Journal of Vacuum Science & Technology B

JVST B

Second Series
Volume 14, Number 4
July/August 1996

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

Papers from the 23rd Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

Papers from the First AVS Topical Conference on Magnetic Surfaces, Interfaces, and Nanostructures

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

19970212 006

An official journal of the American Vacuum Society
Published for the Society by the American Institute of Physics

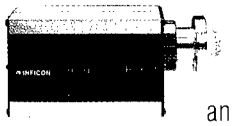
STEP INTO THE FUTURE OF IN SITU GAS ANALYSIS

NO PUMPS REQUIRED

Introducing Transpector® XPR. The first quadrupole-based gas analyzer that operates in PVD processes without complex pumping systems.

Transpector XPR operates at base and PVD process pressures, minus the unwelcome cost and complexity of pressure conversion. A remarkable miniaturized quadrupole, reduced 10 times in size, allows Transpector XPR to operate within a pressure range of 20 mtorr to ultrahigh vacuum. It monitors PVD process, degas, and etch clean for contamination control and photoresist detection—without the cost and maintenance of extra hardware.

When combined with TranspectorWare™ process gas characterization software, Transpector XPR gives you real-time security windows into all stages of your process, for increased uptime and improved wafer yield.



Take the next step in gas analysis—for more information on what Transpector XPR can do for you, call today.

Leybold Inficon, Inc.

Two Technology Place, East Syracuse, NY 13057
Phone: (315) 434-1100 Fax: (315) 437-3803
E-mail: reachus@inficon.com
http://www.inficon.com

REGISTERED ISO 9001 MAY 1993

Innovative Vacuum Technology

LEYBOLD INFICON

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	30 December 1996	Final 1 November 1995 - 31 Oct. 1996

4. TITLE AND SUBTITLE
Final Report:
Twenty -Third Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

5. FUNDING NUMBERS
G: N00014-96-1-0134

6. AUTHOR(S)
C. R. Schulte

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, AZ 85252

8. PERFORMING ORGANIZATION REPORT NUMBER
F1-1996

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
ATTN: ONR 251: Bonnie C. Damanski (703)696-2572
Ballston Tower One
800 North Quincy Street
Arlington, VA 22217-5660 CFDA No. 12.300

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Office of Naval Research position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

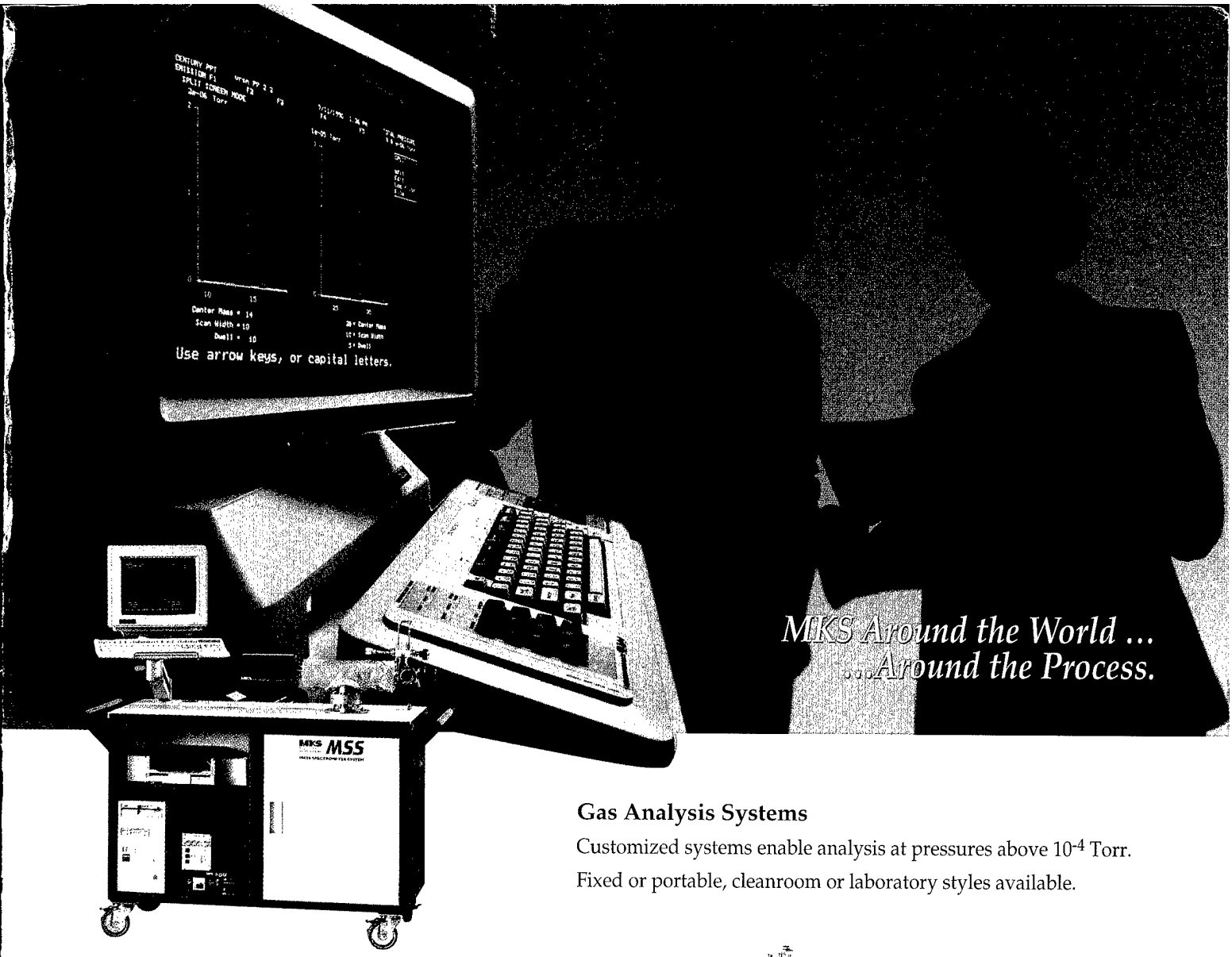
The Twenty-Third Annual Conference on the Physics and Chemistry of Semiconductor Interfaces was held in LaJolla, Calif. and the enclosed report contains papers presented.

14. SUBJECT TERMS
Semiconductor Interfaces

15. NUMBER OF PAGES
1348

16. PRICE CODE

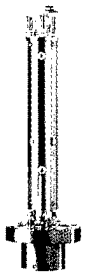
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL



The Full Spectrum Mass Spec Source

Gas Analysis Systems

Customized systems enable analysis at pressures above 10^{-4} Torr. Fixed or portable, cleanroom or laboratory styles available.

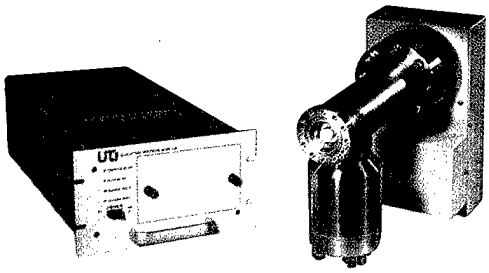


PPT Series Residual Gas Analyzers

Leak test, fingerprint and background analyze with DOS, LabVIEW® and Windows®-based software.

QualiTorr™ Orion® Process Monitors

Application specific solutions for PVD, CVD and Etch. Unmatched stability and accuracy. Proven Windows-based software provides integral alarm and data storage capabilities.



MKS INSTRUMENTS
UTI Division

WEST COAST 2030-C Fortune Drive, San Jose, CA 95131 800-428-9401/408-428-9400 FAX: 408-428-0823
EAST COAST 24 Walpole Park South, Walpole, MA 02081 800-282-1770/508-660-1770 FAX: 508-668-5822
e-mail: uti-east@mksinst.com www: http://www.mksinst.com

QualiTorr™ is trademark, and Orion® is a registered trademark of UTI Division of MKS Instruments, Inc., San Jose, CA. LabView® is a registered trademark of National Instruments, Inc., Austin, TX. Windows® is a registered trademark of Microsoft Corporation, Redmond, WA. 1996 MKS Instruments, Inc. ©. All rights reserved.

Journal of Vacuum Science & Technology B

JVST B

ISSN: 0734-211X
CODEN: JVTBD9

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

Editor: G. E. McGuire,
Microelectronics Center of North Carolina,
Research Triangle Park, NC 27709. Telephone: (919) 248-1910

Supervisor Editorial Office: Rebecca York,
Editorial Assistant: Mary P. Hipsher,
Journal of Vacuum Science and Technology, Microelectronics Center of North Carolina,
3021 Cornwallis Rd., Caller Box 13994, Research Triangle Park, NC 27709-3994.
Telephone: (919) 248-1861 and (919) 248-1884; FAX: (919) 248-1857; E-mail: jvst@mcnc.org

Associate Editors:
Eric Kay, Critical Reviews and Special Issues
C. R. Abernathy (98) University of Florida
E. A. Dobisz (96) Naval Research Lab.
P. C. Fazan (97) Micron Technology
B. Gnade (97) Texas Instruments, Inc.
K. L. Seaward (96) Hewlett-Packard Labs.
J. L. Shaw (98) Naval Research Lab.

JVST Publication Committee:
J. William Rogers, Jr., Chair, University of Washington
John E. Crowell, UC San Diego
Steve Rossnagel, IBM Watson Research Ctr.
Pete Sheldon, NREL

JVST Editorial Board:
Linda M. Cecchi (97) Sandia National Lab.
S. Chambers (96) Pacific Northwest Lab.
R. J. Colton (96) Naval Research Lab.
T. M. Mayer (97) Sandia National Lab.
B. S. Meyerson (96) IBM Watson Research Ctr.
Buddy D. Ratner (97) University of Washington
Gary W. Rubloff (97) NC State University
D. Ruzic (96) University of Illinois
Bruce D. Sartwell (97) Naval Research Lab.

JVST Editorial Staff at AIP: Editorial Supervisor: Deborah McHone; Journal Coordinator: Margaret Reilly; Senior Production Editor: Mary Ellen Mormile

The Journal of Vacuum Science and Technology B is published six times annually (Jan/Feb, Mar/Apr, May/Jun, Jul/Aug, Sep/Oct, Nov/Dec) by the American Institute of Physics (AIP) for and under the editorial management of the American Vacuum Society (AVS). It is an official publication of the AVS and is received by all members of the Society. It is devoted to reports of original research, review articles, and Critical Review articles. The *JVST B* has been established to provide a vehicle for the publication of research dealing with microelectronics and nanometer structures. The emphasis will be on processing, measurement, and phenomena, and will include vacuum processing, plasma processing, materials and structural characterization, microlithography, and the physics and chemistry of submicron and nanometer structures and devices. This journal will publish the proceedings of conferences and symposia that are sponsored by the AVS and its divisions.

Submit Manuscripts to the Editorial Office of the *Journal of Vacuum Science and Technology*, Microelectronics Center of North Carolina, 3021 Cornwallis Rd., Caller Box 13994, Research Triangle Park, North Carolina 27709-3994. Manuscripts of papers presented at AVS-sponsored conferences and symposia and being submitted to *JVST B* should be sent to the Guest Editor appointed for that particular conference. Before preparing a manuscript, authors should read "Information for Contributors," printed in the first issue of each volume of the journal. Submission of a manuscript is a representation that the manuscript has not been published previously nor currently submitted for publication elsewhere. Upon receipt of a manuscript the Editor will send the author a Transfer of Copyright Agreement form. This must be completed by the author and returned *only* to the Editorial Office prior to publication of an accepted paper in the *Journal of Vacuum Science and Technology B*. This written transfer of copyright, which previously was assumed to be implicit in the act of submitting a manuscript, is necessary under the 1978 copyright law in order for the AVS and AIP to continue disseminating research results as widely as possible. Further information may be obtained from AIP.

Publication Charge: To support the cost of wide dissemination of research results through the publication of journal pages and production of a database of articles, the author's institution is requested to pay a *page charge* of $95 per page (with a one-page minimum). The charge (if honored) entitles the author to 100 free reprints. For Errata the minimum page charge is $10, with no free reprints.

Physics Auxiliary Publication Service (PAPS) and/or Electronic PAPS (E-PAPS): For a nominal fee, authors may submit material that is part of and supplemental to a paper, but is too long to be included in the journal. PAPS deposits may be on paper or in electronic media, and can include text, data in graphic or numeric form, computer programs, etc. Retrieval instructions are footnoted in the related published paper. Direct requests to the Editor.

Advertising Rates will be supplied on request from AIP's Advertising Division, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999. Telephone: (516) 576-2440. Fax: (516) 576-2481. E-mail: advtsg@aip.org. All insertion orders and advertising material should be sent to that division.

Copying: Single copies of individual articles may be made for private use or research. Authorization is given (as indicated by the Item Fee Code for this publication) to copy articles beyond the use permitted by Sections 107 and 108 of the U.S. Copyright Law, provided the copying fee of $10 per copy per article is paid to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, USA. Persons desiring to photocopy materials for classroom use should contact the CCC Academic Permissions Service. The Item Fee Code for this publication is 0734-211X/96 $10.00.

Authorization does not extend to systematic or multiple reproduction, to copying for promotional purposes, to electronic storage or distribution, or to republication in any form. In all such cases, specific written permission from AIP must be obtained.

Permission for Other Use: Permission is granted to quote from the journal with the customary acknowledgment of the source. To reprint a figure, table, or other excerpt requires the consent of one of the authors and notification to AIP.

Requests for Permission: Address requests to AIP Office of Rights and Permissions, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Telephone: (516) 576-2268; Fax: (516) 576-2327; Internet: rights@aip.org.

Copyright © 1996 American Vacuum Society. All rights reserved.

American Vacuum Society

Officers

William D. Sproul, *President*
BIRL, Northwestern University

Gary E. McGuire, *President-Elect*
Microelectronics Center of North Carolina

John H. Weaver, *Past-President*
University of Minnesota

William D. Westwood, *Secretary*
Bell-Northern Research, Ltd.

N. Rey Whetten, *Treasurer*
Technical Director
American Vacuum Society

Directors

Leonard J. Brillson
Xerox Webster Research Center

Paula J. Grunthaner
Jet Propulation Laboratory/ CALTECH

Dale C. McIntyre
Optex, Inc.

Gary W. Rubloff
North Carolina State University

Ellen B. Stechel
Sandia National Laboratories

Jerry M. Woodall
Purdue University

JVST

G. Lucovsky, *Editor-in-Chief*
North Carolina State University

AVS Membership Information may be obtained from

Angela Mulligan
AVS Membership and Scholarship Secretary
120 Wall Street
32nd Floor
New York, NY 10005
(212) 248-0200

GET THE EDGE

Del•Seal® High/Ultrahigh Vacuum System Components

Del•Seal flanges and fittings give you the high and ultrahigh vacuum sealing edge that is the industry standard. And you get much more. Like competitive prices, fast shipment from stock, and unsurpassed quality.

Standard *Del•Seal* sizes from 1-1/3" O.D. *mini* through 16-1/2" O.D. in several different styles are available. ASA, Adaptor, Double-Sided, Reducing, ISO *Kwik•Flange* and *Large•Flange*, Wire Seal, and special purpose flanges are also included. Flanges can be ordered separately or with a full range of fittings and accessories.

They're all detailed in a 360 page catalog with prices and available through MDC affiliates around the world. For a catalog and the location of your closest representative, visit our Web site at

http://www.mdc-vacuum.com

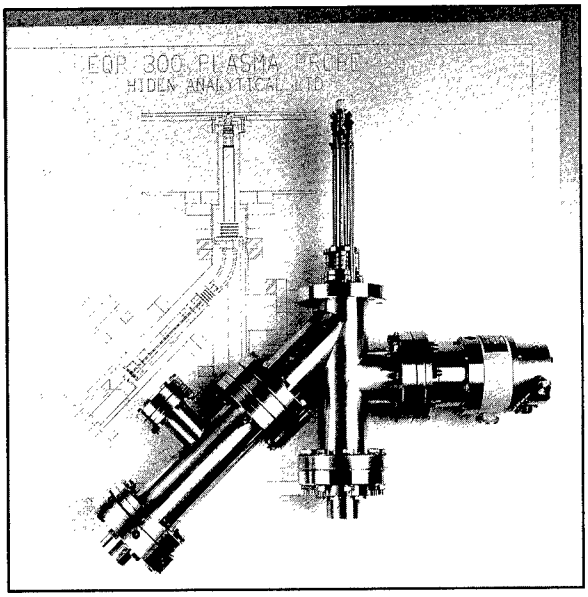
E-mail sales@mdc-vacuum.com
Phone 800-443-8817 or 510-265-3500
Fax 510-887-0626.
Mail 23842 Cabot Blvd.,
Hayward, CA 94545-1651

- **One Source**
- **One Site**
- **Worldwide**

MDC VACUUM PRODUCTS CORPORATION

PLASMA DIAGNOSTICS...
...THE HAL EQP MASS SPECTROMETER

Determine the mass and energy distributions of plasma ions, neutrals and radicals.
Versatile diagnostic tool for deposition, etching and reactive plasmas.
Send for our comprehensive literature pack.



The HAL EQP mass spectrometer uses an electrostatic sector field energy analyser followed by a quadrupole mass filter and pulse counting detector. A comprehensive PC data system provides for data acquisition and display of Mass Spectra, Ion Energy Distributions and Appearance Potential Spectra for radical component studies. Standard gating inputs provide for time resolved plasma analysis.

HIDEN ANALYTICAL

Distribution and Service in the USA by:
Advanced Vacuum Components Inc
2600 Boyce Plaza II, Suite 145
Pittsburgh PA 15241
Tel: (412) 257-9957
Fax: (412) 257-9987

Manufactured in England by:
Hiden Analytical Ltd, Warrington, WA5 5UN
Tel: +44 (0) 1925 445225
Fax: +44 (0) 1925 416518

Instruments and Materials for Surface and Interface Analysis

- Sample Handling & Transfer Systems
- Customized Vacuum Components—Precision Magnetic Manipulators, Loadlocks, Hot/Cold Stages
- Auger Spectrometer System for Defect Review or MBE
- UHV-Compatible Motors
- BEES (Ballistic Electron Emission Spectroscopy)
- Reference Materials

For more information on our portfolio of customized products, call us at (415)965-8205.

SURFACE INTERFACE

110 Pioneer Way, Suite D
Mountain View, CA 94041
Phone: (415)965-8205
Fax: (415)965-8207, E-mail: sii@aip.org

Journal of Vacuum Science & Technology B

JVST B

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

Second Series
Volume 14, Number 4
Jul/Aug 1996

Regular Articles

Single-electron charging of a molecule observed in scanning tunneling scattering experiments
H. Nejo, M. Aono, D. G. Baksheyev, and V. A. Tkachenko **2399**

Current characteristics in near field emission scanning tunneling microscopes
G. Mesa, J. J. Sáenz, and R. García **2403**

Morphological modeling of atomic force microscopy imaging including nanostructure probes and fibrinogen molecules
David L. Wilson, Pranav Dalal, Kenneth S. Kump, William Benard, Ping Xue, Roger E. Marchant, and Steven J. Eppell **2407**

(Continued)

Journal of Vacuum Science & Technology B (ISSN: 0734-211X) is published six times annually (Jan/Feb, Mar/Apr, May/Jun, Jul/Aug, Sep/Oct, Nov/Dec) for the American Vacuum Society by the American Institute of Physics, 500 Sunnyside Blvd., Woodbury, NY 11797-2999. Membership in the American Vacuum Society includes $17.50 from membership dues to be applied towards a subscription to *Journal of Vacuum Science & Technology B.*

Known office of publication is Thorofare, NJ 08086. Periodicals postage paid at Thorofare, NJ 08086, and at additional mailing offices. POSTMASTER: Send address changes to *Journal of Vacuum Science & Technology B,* SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.

Subscription Prices (1996)

	U.S.A. and Poss.	Can., Mex., Central & S. America & Caribbean	Europe, Asia, Africa & Oceania*
JVST B†	$610	$635	$654
JVST B‡	$710	$735	$754
JVST A and B†	$710	$760	$798
JVST A and B#	$710	$735	$735
JVST A and B‡	$810	$860	$898

†Paper only.
‡Paper and CD-ROM.
#CD-ROM only.
*Nonmember subscriptions include air-freight service.

Back-number Prices: 1996 single copy: $115. Prior to 1996 single copies: $95.

Subscriptions, renewals, and address changes should be addressed to *Subscription Fulfillment Division, SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.* Allow at least six weeks advance notice. For address changes please send both old and new addresses, and, if possible, include a label from the plastic mailing wrapper of a recent issue. For your convenience a **change-of-address form is included in every issue of *Physics Today;* please use it.** AIP will honor a request for a missing journal issue only within six months of that issue's actual date of publication. Requests received for missing issues beyond six months of the actual publication date will not be honored. The issue may, however, be purchased at the single-copy price. (Subscription Fulfillment offices are located at SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.)

Single-copy orders (current and back issues) should be addressed to American Institute of Physics, Circulation and Fulfillment Division, 500 Sunnyside Blvd., Woodbury, NY 11797-2999.

Reprints: Reprints can be ordered with or without covers only in multiples of 50 (with a minimum of 100 in each category) from AIP, Circulation & Fulfillment/ Reprints, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Fax: 516-349-9704; Telephone: 800-344-6909 (in U.S. and Canada), or 516-576-2234.

Document Delivery: Copies of articles can be ordered for $15 per copy from the AIP/Member Society document delivery service, "Articles in Physics," 1722 Gilbreth Road, Burlingame, CA 94010-1305; Fax: 415-259-6044; Telephone: 800-480-PHYS (in U.S. and Canada), or 415-259-6002; Internet: articles@aip.org.

Copying fees: The code that appears on the first page of articles in this journal gives the fee for each copy of the article made beyond the free copying permitted by AIP. (See statement under "Copyright" elsewhere in this journal.) If no code appears, no fee applies. The fee for pre-1978 articles is $0.25 per copy. With the exception of copying for advertising and promotional purposes, the express permission of AIP is not required provided the fee is paid through the *Copyright Clearance Center, Inc. (CCC), 222 Rosewood Dr., Danvers, MA 01923.* Contact the CCC for information on how to report copying and remit payment.

Microform: *Journal of Vacuum Society and Technology B* is available on microfiche issued at the same frequency as the printed journal and annually on microfilm. A microform catalog is available from AIP, Circulation & Fulfillment/ Single, Copy Sales, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Fax: 516-349-9704; Telephone: 800-344-6908 (in U.S. and Canada), or 516-576-2277.

Online Availability: Abstracts of journal articles published by AIP and Member Societies (and several other physics publishers) are available in the SPIN database via the AIP online service PINET. Also available on PINET are *Advance Abstracts,* a current awareness service, other physics information resources, as well as internet services (Internet: elecprod@aip.org).

DTIC QUALITY INSPECTED 5

Scanning scattering microscope for surface microtopography and defect imaging
J. Lorincik, D. Marton, R. L. King, and J. Fine 2417

Scanning tunneling microscope study of defect structures on As-terminated Si(001) surfaces
M. D. Jackson, F. M. Leibsle, R. J. Cole, D. A. C. Gregory, D. A. Woolf, and P. Weightman 2424

Simultaneous imaging of Si(111) 7×7 with atomic resolution in scanning tunneling microscopy, atomic force microscopy, and atomic force microscopy noncontact mode
Peter Güthner 2428

Structure imaging by atomic force microscopy and transmission electron microscopy of different light emitting species of porous silicon
R. Massami Sassaki, R. A. Douglas, M. U. Kleinke, and O. Teschke 2432

Field emission characteristics of the scanning tunneling microscope for nanolithography
T. M. Mayer, D. P. Adams, and B. M. Marder 2438

Proximity effect correction for nanolithography
Richard Rau, James H. McClellan, and Timothy J. Drabik 2445

Independent parallel lithography using the atomic force microscope
S. C. Minne, S. R. Manalis, A. Atalar, and C. F. Quate 2456

Monte Carlo simulation of inclined incidence of fast electrons to solids
Y. M. Gueorguiev, G. M. Mladenov, and D. I. Ivanov 2462

Generation mechanism of distortion aberration in a symmetric magnetic doublet for an electron beam projection system
Mamoru Nakasuji and Hiroyasu Shimizu 2467

Reticle fabrication by high acceleration voltage electron beam: Representative figure method for proximity effect correction [VI]
Takayuki Abe 2474

Surface tension, adhesion and wetting of materials for photolithographic process
J. Bauer, G. Drescher, and M. Illig 2485

Polysilicon gate etching in high density plasmas. III. X-ray photoelectron spectroscopy investigation of sidewall passivation of silicon trenches using an oxide hard mask
F. H. Bell and O. Joubert 2493

Influence of patterning in silicon quantum well structures on photoluminescence
Hideo Namatsu, Tomofumi Furuta, Masao Nagase, Kenji Kurihara, Kazumi Iwadate, Katsumi Murase, and Takahiro Makino 2500

Computer simulations of porous silicon formation
Y. M. Weng, J. Y. Qiu, Y. H. Zhou, and X. F. Zong 2505

Comparison of advanced plasma sources for etching applications. V. Polysilicon etching rate, uniformity, profile control, and bulk plasma properties in a helical resonator plasma source
J. T. C. Lee, N. Layadi, K. V. Guinn, H. L. Maynard, F. P. Klemens, D. E. Ibbotson, I. Tepermeister, P. O. Egan, and R. A. Richardson 2510

Surface morphologies for Br-etched Si(100)-2×1: Kinetics of pit growth and step retreat
F. J. Williams, C. M. Aldao, and J. H. Weaver 2519

Deep-etch silicon millimeter waveguide structure for electron acceleration
T. L. Willke and A. D. Feinerman 2524

Minimized response time of optical emission and mass spectrometric signals for optimized endpoint detection
S. Thomas III, H. H. Chen, C. K. Hanish, J. W. Grizzle, and S. W. Pang 2531

Reactive ion etching of sloped sidewalls for surface emitting structures using a shadow mask technique
B. Jacobs and R. Zengerle 2537

Comparison of the physical and electrical properties of electron cyclotron resonance and distributed electron cyclotron resonance SiO_2
M. Firon, M. C. Hugon, B. Agius, Y. Z. Hu, Y. Wang, and E. A. Irene 2543

Reactive-ion etching of WSi_x in CF_4+O_2 and the associated damage in GaAs
Yi-Jen Chan, Chao-Shin Su, and Kuo-Tung Sung 2550

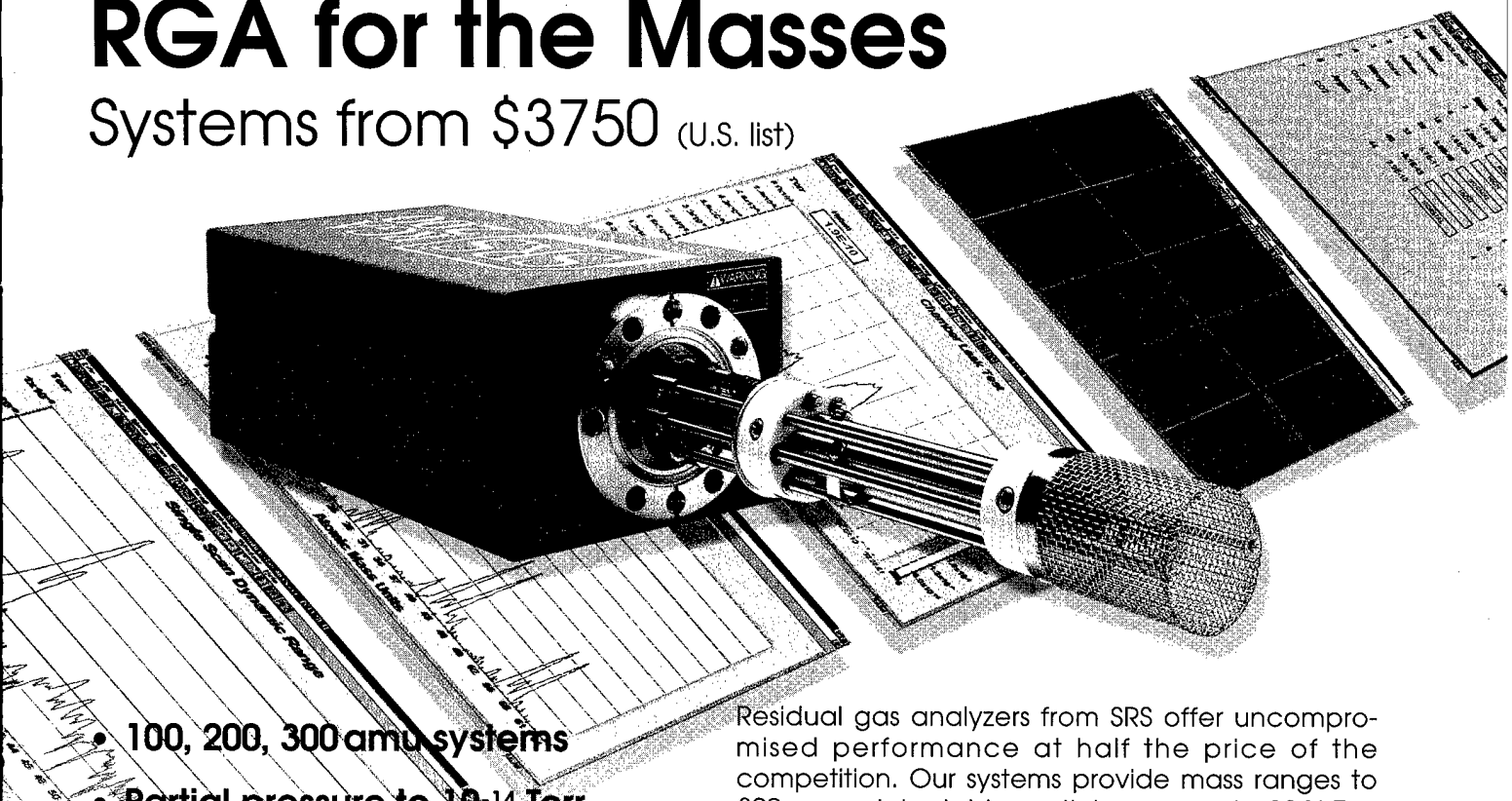
Characterization of electrical damage induced by CH_4/H_2 reactive ion etching of molecular beam epitaxial InAlAs
M. Achouche, A. Clei, and J. C. Harmand 2555

Cl_2/Ar plasma etching of binary, ternary, and quaternary In-based compound semiconductors
J. W. Lee, J. Hong, C. R. Abernathy, E. S. Lambers, S. J. Pearton, W. S. Hobson, and F. Ren 2567

(Continued)

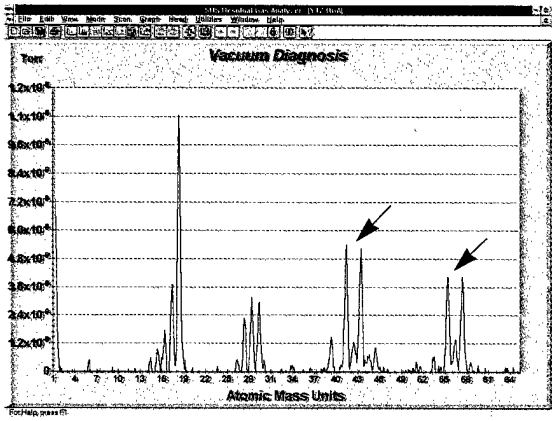
RGA for the Masses

Systems from $3750 (U.S. list)

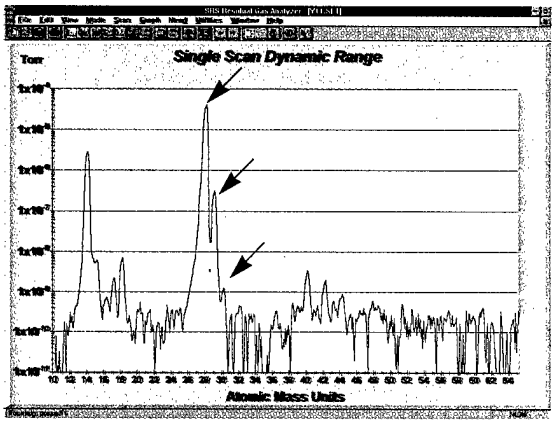


- **100, 200, 300 amu systems**
- **Partial pressure to 10^{-14} Torr**
- **6 orders of magnitude dynamic range in a single scan**
- **Real-time Windows® software**
- **Mass spectra, leak detection and pressure vs. time modes**
- **Multi-head operation**
- **Optional Electron Multiplier**

Residual gas analyzers from SRS offer uncompromised performance at half the price of the competition. Our systems provide mass ranges to 300 amu, detectable partial pressures to 10^{-14} Torr and better than 1 amu resolution. With our interactive Windows® software package you can fully control the RGA sensor, acquire data and analyze your vacuum system in real-time. The simple, rugged quadrupole mass spectrometer attaches directly to a 2 3/4" CF flange. A dual ThO_2Ir filament and a unique continuous dynode electron multiplier provide increased stability and longer life than other designs. Both are field replaceable in a matter of minutes – a first for RGAs! If your application involves residual gas analysis, leak detection or vacuum processing, contact us for a copy of our RGA systems brochure and demo software package.



The RGA is an invaluable vacuum diagnostic tool. Shown above is the mass spectrum of a vacuum chamber contaminated with oil.



With dynamic range that spans 6 orders of magnitude, three isotopes of Nitrogen ($^{14}N_2$, $^{14}N\,^{15}N$, $^{15}N_2$) are clearly detected in a single scan.

Stanford Research Systems
1290-D Reamwood Avenue, Sunnyvale, CA 94089
Telephone (408)744-9040 • FAX (408)744-9049
Email: info@srsys.com • WWW: http://www.srsys.com/srsys

Windows is a registered trademark of Microsoft Corporation

Metal stack etching using a helical resonator plasma
C. B. Labelle, H. L. Maynard, and J. T. C. Lee 2574

Microstructural stability of ohmic contacts to $In_xGa_{1-x}N$
A. Durbha, S. J. Pearton, C. R. Abernathy, J. W. Lee, P. H. Holloway, and F. Ren 2582

Patterned eutectic bonding with Al/Ge thin films for microelectromechanical systems
Bao Vu and Paul M. Zavracky 2588

Simulation of three-dimensional refractory metal step coverage over contact cuts and vias
M. K. Sheergar, T. Smy, S. K. Dew, and M. J. Brett 2595

Liner conformality in ionized magnetron sputter metal deposition processes
S. Hamaguchi and S. M. Rossnagel 2603

High-vacuum versus "environmental" electron beam deposition
Albert Folch, Jordi Servat, Joan Esteve, Javier Tejada, and Miquel Seco 2609

Simulations of metal thin film thermal flow processes
Hung Liao and Timothy S. Cale 2615

Nonlinear analysis of the *I–V* characteristics in Ti/Si and $TiSi_2$/Si Schottky diodes
J. Pérez-Rigueiro, C. Jiménez, R. Pérez-Casero, and J. M. Martínez-Duart 2623

The effect of the *Ti* glue layer in an integrated Ti/TiN/*Ti*/AlSiCu/TiN contact metallization process
L. Ouellet, Y. Tremblay, G. Gagnon, M. Caron, J. F. Currie, S. C. Gujrathi, and M. Biberger 2627

Effects of oxide overlayer on thermal stress and yield behavior of Al alloy films
I.-S. Yeo, S. G. H. Anderson, D. Jawarani, P. S. Ho, A. P. Clarke, S. Saimoto, S. Ramaswami, and R. Cheung 2636

Effects of insulator surface roughness on Al-alloy film properties and electromigration performance in Al-alloy/Ti insulator layered interconnects
Hiroshi Onoda, Tadashi Narita, Kenshin Touchi, and Keiichi Hashimoto 2645

Gap fill dependence of fluorinated polyimide films on solid content, adhesion promoter, spin dwell time, and solvent spray
Shi-Qing Wang and Bin Zhao 2656

Characterization of the Si/SiO_2 interface formed by remote plasma enhanced chemical vapor deposition from SiH_4/N_2O with or without chlorine addition
Young-Bae Park, Xiaodong Li, and Shi-Woo Rhee 2660

Oxide growth on silicon (100) in the plasma phase of dry oxygen using an electron cyclotron resonance source
Keunjoo Kim, M. H. An, Y. G. Shin, M. S. Suh, C. J. Youn, Y. H. Lee, K. B. Lee, and H. J. Lee 2667

Gate quality Si_3N_4 prepared by low temperature remote plasma enhanced chemical vapor deposition for III–V semiconductor-based metal–insulator–semiconductor devices
D. G. Park, M. Tao, D. Li, A. E. Botchkarev, Z. Fan, Z. Wang, S. N. Mohammad, A. Rockett, J. R. Abelson, H. Morkoç, A. R. Heyd, and S. A. Alterovitz 2674

B/Si(100) surface: Atomic structure and epitaxial Si overgrowth
Z. Zhang, M. A. Kulakov, B. Bullemer, I. Eisele, and A. V. Zotov 2684

High phosphorus doping of epitaxial silicon at low temperature and very low pressure
X. D. Huang, P. Han, H. Chen, Y. D. Zheng, L. Q. Hu, R. H. Wang, S. M. Zhu, and D. Feng 2690

Characterization by x-ray photoelectron spectroscopy of the chemical structure of semi-insulating polycrystalline silicon thin films
Fabio Iacona, Salvatore Lombardo, and Salvatore U. Campisano 2693

Inherent possibilities and restrictions of plasma immersion ion implantation systems
S. Mändl, J. Brutscher, R. Günzel, and W. Möller 2701

X-ray photoelectron spectroscopy study on native oxidation of As-implanted Si (100)
Fumiko Yano, Akiko Hiraoka, Toshihiko Itoga, Atsuko Matsubara, Hisao Kojima, Keiichi Kanehori, and Yasuhiro Mitsui 2707

Quantification of metal trace contaminants on Si wafer surfaces by Laser-SNMS and TOF-SIMS using sputter deposited submonolayer standards
A. Schnieders, R. Möllers, M. Terhorst, H.-G. Cramer, E. Niehuis, and A. Benninghoven 2712

Phonon scattering in novel superlattice-asymmetric double barrier resonant tunneling structure
K. Banoo, T. Daniels-Race, C. R. Wallis, and S. W. Teitsworth 2725

Facet generation during molecular beam epitaxy of GaAs/AlGaAs multilayers on GaAs (001) patterned substrates
T. Takebe, M. Fujii, T. Yamamoto, K. Fujita, and T. Watanabe 2731

(Continued)

"That proves it... Same performance, new smaller package - let's go with the MKS 722!"

Device dimensions aren't the only things shrinking! Today's Baratron® process manometers are smaller than ever before. The new Type 722 Baratron pressure transducers offer the same exceptional repeatability, accuracy and zero stability as the industry standard MKS Type 122 in a package that's only 4" high x $1\frac{1}{2}$" in diameter. Same performance, new smaller package, now that's progress.

Call MKS today for product literature or applications assistance at:(800) 227-8766/(508) 975-2350.

MKS
INSTRUMENTS
Baratron® Pressure Products Division

MKS Instruments, Inc. Six Shattuck Road, Andover, MA 01810 Tel: (800)227-8766/(508)975-2350 Fax: (508)975-0093
e-mail: mks@mksinst.com www: http://www.mksinst.com

MKS and Baratron® are registered trademarks of MKS Instruments, Inc., Andover, MA.

1996 MKS Instruments, Inc. ©. All rights reserved.

InP/InGaAs single heterojunction bipolar transistors grown by solid-source molecular beam epitaxy using a phosphorus valved cracker
W. L. Chen, T. P. Chin, J. M. Woodall, and G. I. Haddad 2739

Chemistry of arsenic incorporation during GaAs/GaAs(100) molecular beam epitaxy probed by simultaneous laser flux monitoring and reflection high-energy electron diffraction
Adina K. Ott, Sean M. Casey, April L. Alstrin, and Stephen R. Leone 2742

Solid source molecular beam epitaxy of low threshold 1.55 μm wavelength GaInAs/GaInAsP/InP semiconductor lasers
F. G. Johnson, O. King, F. Seiferth, D. R. Stone, R. D. Whaley, M. Dagenais, and Y. J. Chen 2753

Molecular beam epitaxy grown ZnSe studied by reflectance anisotropy spectroscopy and reflection high-energy electron diffraction
J. -T. Zettler, K. Stahrenberg, W. Richter, H. Wenisch, B. Jobst, and D. Hommel 2757

Solvent effect on the properties of sulfur passivated GaAs
Vasily N. Bessolov, Elena V. Konenkova, and Mikhail V. Lebedev 2761

Rapid Communications

Silicon dioxide chemical vapor deposition using silane and hydrogen peroxide
Frédéric Gaillard, Pascal Brault, and Pierre Brouquet 2767

Shop Notes

Analytical method of gigabit trench doping uniformity by secondary ion mass spectrometry
Naoto Matsuo, Kazuyoshi Tsukamoto, and Tadaki Miyoshi 2770

Errata

Erratum: Measurement of interfacial shear (friction) with an ultrahigh vacuum atomic force microscope [J. Vac. Sci. Technol. B 14, 1289 (1996)]
R. W. Carpick, N. Agraït, D. F. Ogletree, and M. Salmeron 2772

Papers from the 23rd Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

Preface 2774

Transport: Device Physics

Effect of dislocations in strained Si/SiGe on electron mobility
K. Ismail 2776

Quantum transport: Silicon inversion layers and InAlAs–InGaAs heterostructures
D. Vasileska, T. Eldridge, and D. K. Ferry 2780

Ballistic electron emission microscopy of Au–InAs–GaAs system
Mao-long Ke, D. I. Westwood, C. C. Matthai, B. E. Richardson, and R. H. Williams 2786

Interface roughness effects on transport in tunnel structures
D. Z.-Y. Ting and T. C. McGill 2790

Electron transport at Au/InP interface with nanoscopic exclusions
S. Anand, S-B. Carlsson, K. Deppert, L. Montelius, and L. Samuelson 2794

Si Surfaces

Relaxation of the step profile for different microscopic mechanisms
Da-Jiang Liu, Elain S. Fu, M. D. Johnson, John D. Weeks, and Ellen D. Williams 2799

First-principles study of Si 2*p* core-level shifts at water and hydrogen covered Si(001)2×1 surfaces
Alfredo Pasquarello, Mark S. Hybertsen, and Roberto Car 2809

New approach to preparing smooth Si(100) surfaces: Characterization by spectroellipsometry and validation of Si/SiO_2 interfaces properties in metal-oxide-semiconductor devices
D. Schmidt, H. Niimi, B. J. Hinds, D. E. Aspnes, and G. Lucovsky 2812

Schwoebel barriers on Si(111) steps and kinks
S. Kodiyalam, K. E. Khor, and S. Das Sarma 2817

Si Oxidation, Si–SiO_2 Interfaces

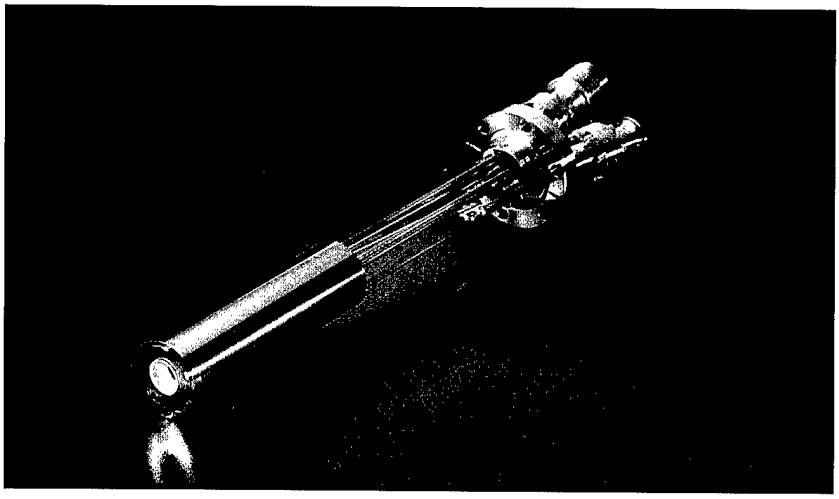
An inquiry concerning the principles of Si 2*p* core-level photoemission shift assignments at the Si/SiO_2 interface
F. R. McFeely, K. Z. Zhang, Mark M. Banaszak Holl, Sunghee Lee, and John E. Bender IV 2824

(Continued)

Attention MBE System Operators

Introducing the New SUMO™ Effusion Cells from EPI

EPI Redefines Effusion Cell Technology



The EPI SUMO™ Effusion Cell – an important advance in MBE equipment technology, available for any production or research MBE system

The new EPI SUMO™ effusion cell is nothing short of revolutionary. It optimizes material quality while maintaining very low levels of morphological defect densities and background doping levels. The beam equivalent pressure remains consistent throughout the life of the material charge, greatly enhancing reproducibility and operator convenience. Shutter activation-related flux transients are virtually undetectable, while providing state-of-the-art uniformities. All models have large useful capacities, and configurations available for both research and multi-wafer systems.

- **Excellent Material Quality**
- **Extremely Low Defect Densities**
- **Increased Flux to Power Ratios**
- **Long-Term Stability - more consistent beam equivalent pressures over the lifetime of the charge**
- **Virtually Undetectable Shutter Activation Related Flux Transients**
- **State-of-the-Art Deposition Uniformity**
- **Significantly Increased Useful Capacity**

If you are looking for a way to maximize your MBE system up-time, increase throughput, improve material quality and reproducibity, and make your life a little easier, contact your local EPI sales office for more information about the EPI SUMO™ Effusion Cell. Be sure to ask for ***The EPI SUMO™ Application Note, February, 1996.***

EPI *MBE Products Group*

EPI
1290 Hammond Road
Saint Paul, MN 55110 U.S.A.
Phone: (1) 612-653-0488
Fax: (1) 612-653-0725
e-mail: info@epimbe.com

EPI Europe
147 Chorley New Road
Horwich, Bolton BL6 5Q
Greater Manchester, U.K.
Phone: (44) 1-204-668366
Fax: (44) 1-204-668466
e-mail: europe@epimbe.com

Japan
Techscience, Ltd.
2-64 Miyamotocho,
Koshigaya-shi
Saitama-ken 343 Japan
Phone: (81) 489-64-3111
Fax: (81) 489-65-1800

Defect properties of Si-, O-, N-, and H-atoms at Si—SiO_2 interfaces
G. Lucovsky, Z. Jing, and D. R. Lee 2832

Oxidation of silicon (100): Experimental data versus a unified chemical model
P. Thanikasalam, T. K. Whidden, and D. K. Ferry 2840

Kinetics of field-induced oxidation of hydrogen-terminated Si(111)
L. Ley, T. Teuschler, K. Mahr, S. Miyazaki, and M. Hundhausen 2845

Photoinduced surface reactions of reverse-biased *n*-type porous Si
Eric J. Lee, Theodore W. Bitner, Anselm P. Hall, and Michael J. Sailor 2850

Si Oxides and Nitrides

Stressing and high field transport studies on device-grade SiO_2 by ballistic electron emission spectroscopy
R. Ludeke, H. J. Wen, and E. Cartier 2855

Ballistic-electron emission microscopy studies of charge trapping in SiO_2
B. Kaczer and J. P. Pelz 2864

Contactless capacitance–voltage and photoluminescence characterization of ultrathin oxide–silicon interfaces formed on hydrogen terminated (111) surfaces
Tamotsu Hashizume, Satoshi Koyanagi, and Hideki Hasegawa 2872

Rapid thermal N_2O oxynitride on Si(100)
Z. H. Lu, R. J. Hussey, M. J. Graham, R. Cao, and S. P. Tay 2882

III–V Surfaces: Passivation

Surface passivation of $In_{0.53}Ga_{0.47}As$ ridge quantum wires using silicon interface control layers
Hajime Fujikura, Satoshi Kodama, Tamotsu Hashizume, and Hideki Hasegawa 2888

Structure of the SiN_x/GaAs (110) interface modified with ultrathin Si and sulfur passivation
L. J. Huang, W. M. Lau, H. T. Tang, W. N. Lennard, I. V. Mitchell, D. Landheer, J.-M. Baribeau, and S. Ingrey 2895

Structural and electronic properties of Sb islands on GaAs (110)
R. Magri, F. Manghi, and C. Calandra 2901

Room-temperature chlorination of As-rich GaAs (110)
William C. Simpson, David K. Shuh, and Jory A. Yarmoff 2909

Interaction of hydrogen ions with oxidized GaAs(100) and AlAs(100) surfaces
Ying-Lan Chang, R. Cao, W. E. Spicer, P. Pianetta, Song Shi, Evelyn Hu, and James Merz 2914

III–V Interfaces

Study on interface abruptness of $In_xGa_{1-x}As/In_yGa_{1-y}As_zP_{1-z}$ heterostructures grown by gas-source molecular beam epitaxy
W. G. Bi and C. W. Tu 2918

Effect of interface defect formation on carrier diffusion and luminescence in $In_{0.2}Ga_{0.8}As/Al_xGa_{1-x}As$ quantum wells
D. H. Rich, K. Rammohan, H. T. Lin, Y. Tang, M. Meshkinpour, and M. S. Goorsky 2922

Role of structural and chemical contributions to valence-band offsets at strained-layer heterojunctions: The GaAs/GaP (001) case
M. Di Ventra, M. Peressi, and A. Baldereschi 2936

Atomic-scale structure of $InAs/InAs_{1-x}Sb_x$ superlattices grown by modulated molecular beam epitaxy
A. Y. Lew, E. T. Yu, and Y. H. Zhang 2940

Evidence for the occupation of DX centers in $In_{0.29}Al_{0.71}As$
A. P. Young and H. H. Wieder 2944

Wide-Bandgap Materials

Microscopic processes during electron cyclotron resonance microwave nitrogen plasma-assisted molecular beam epitaxial growth of GaN/GaAs heterostructures: Experiments and kinetic modeling
Z. Z. Bandić, T. C. McGill, R. J. Hauenstein, and M. L. O'Steen 2948

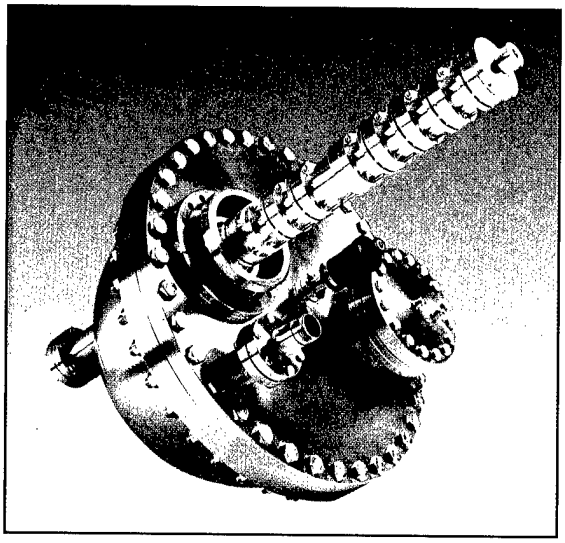
Tensile strain relaxation in GaN_xP_{1-x} ($x \leq 0.1$) grown by chemical beam epitaxy
N. Y. Li, W. S. Wong, D. H. Tomich, K. L. Kavanagh, and C. W. Tu 2952

Nickel doping of boron carbide grown by plasma enhanced chemical vapor deposition
Seong-Don Hwang, N. B. Remmes, P. A. Dowben, and D. N. McIlroy 2957

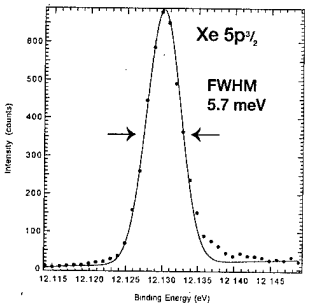
(Continued)

High Resolution Electron Spectroscopy

for XPS-, AES-, UPS- and Synchrotron Applications



Gas phase UPS of Xenon at the synchrotron
(hv = 14 eV, line width of the excitation: 3.3 meV.
Data courtesy of M.R. Siggel, G.C. King, N.J. Mason and I.C. Walker)

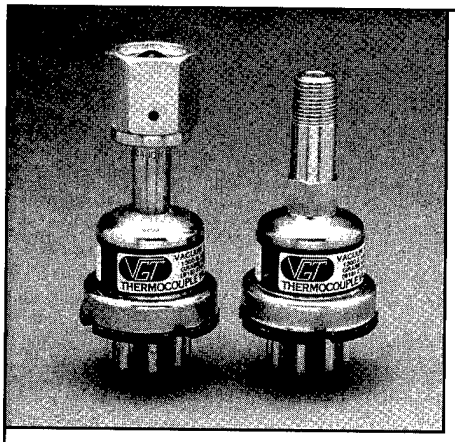


Energy resolution < 5 meV

OMICRON
INSTRUMENTS FOR SURFACE SCIENCE

Headquarters:
OMICRON VAKUUMPHYSIK GmbH
Idsteiner Str. 78 • D-65232 Taunusstein • Germany
Tel. +49 (0) 61 28 / 7 30 44 • Fax +49 (0) 61 28 / 7 35 96

USA:
OMICRON ASSOCIATES
1738 N, Highland Road • Suite G 101 • Pittsburgh, PA 15241
Tel. (412) 831-22 62 • Fax (412) 831-98 28



Why Pay More?

At Vacuum Gauge & Test, we have been manufacturing Thermocouple Tubes since 1982.

We keep all tubes in stock and can ship most orders in 24 hours.

All tubes are enclosed in a contaminate free electroless nickel plated housing. Tubes come with a standard 1/8 IPS mounting port, or upon request with VCR fittings.

Quantity and O.E.M. pricing discounts available, and private labeling is our specialty.

Tubes are available in these equivalences - all MFG's, Type 1518, CVC's GTC-004, Hastings DV-3M, Veeco DV-1M, Hastings DV-6M, Granville Phillips 270 006, Varian/NRC 531, MKS TC-1, CVC-036, Hastings DV-4D, VEECO DV-4M.

Gauge controllers are also available. Call for more information or to request a brochure.

VGT *Vacuum Gauge & Test, Inc.*
200 Litton Dr. #325 • Grass Valley, CA 95945 • (916) 273-9455 • FAX (916) 273-1371

Evolution of deep levels and internal photoemission with annealing temperature at ZnSe/GaAs interfaces
X. Yang, L. J. Brillson, A. D. Raisanen, L. Vanzetti, A. Bonanni, A. Franciosi, M. Grundmann, and D. Bimberg 2961

Interfacial chemical bonds, reactions, and band alignment in ZnSe/GaAs(001) heterojunctions
G. Bratina, T. Ozzello, and A. Franciosi 2967

Interface states of ZnSe/GaAs interface
Z. Yang, I. K. Sou, Y. H. Yeung, G. K. L. Wong, Jie Wang, Cai-xia Jin, and Xiao-yuan Hou 2973

Zn preadsorption on GaAs(100)2×4 prior to ZnSe growth
S. Heun, J. J. Paggel, S. Rubini, and A. Franciosi 2980

Metal–Semiconductor Interfaces

Electronic properties of ideal and interface-modified metal-semiconductor interfaces
Winfried Mönch 2985

Tunable Schottky barriers and the nature of Si interface layers in Al/GaAs(001) diodes
L. Sorba, J. J. Paggel, and A. Franciosi 2994

Schottky barrier tuning at Al/GaAs(100) junctions
C. Berthod, J. Bardi, N. Binggeli, and A. Baldereschi 3000

Surface Influences on Bulk Properties

Near-surface dopant passivation after wet-chemical preparation
L. Ley, J. Ristein, J. Schäfer, and S. Miyazaki 3008

Correlation between surface structure and ordering in GaInP
H. Murata, S. H. Lee, I. H. Ho, and G. B. Stringfellow 3013

Heteroepitaxy

Heteroepitaxy of GaP on Si(100)
K. J. Bachmann, U. Rossow, N. Sukidi, H. Castleberry, and N. Dietz 3019

Characterization of $Si/Si_{1-y}C_y$ superlattices grown by surfactant assisted molecular beam epitaxy
P. O. Pettersson, C. C. Ahn, T. C. McGill, E. T. Croke, and A. T. Hunter 3030

Modulation-doped $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ heterostructures grown on GaAs substrates using step-graded $In_xGa_{1-x}As$ buffers
R. S. Goldman, K. L. Kavanagh, H. H. Wieder, and S. N. Ehrlich 3035

Optical investigations of surface processes in GaP heteroepitaxy on silicon under pulsed chemical beam epitaxy conditions
U. Rossow, N. Dietz, K. J. Bachmann, and D. E. Aspnes 3040

Implications of excess strain in As compound/P compound III–V multilayer superlattices grown by metal-organic vapor-phase epitaxy
A. R. Clawson and C. M. Hanson 3047

Reflectance Anisotropy

***In situ* characterization of ZnSe/GaAs(100) interfaces by reflectance difference spectroscopy**
T. Yasuda, L. H. Kuo, K. Kimura, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao 3052

Optical monitoring of the development of InAs quantum dots on GaAs(001) by reflectance anisotropy spectroscopy
E. Steimetz, J.-T. Zettler, W. Richter, D. I. Westwood, D. A. Woolf, and Z. Sobiesierski 3058

Reflectance anisotropy spectroscopy study of GaAs overlayer growth on submonolayer coverages of Si on the GaAs(001)-*c*(4×4) surface
Z. Sobiesierski, D. I. Westwood, and D. A. Woolf 3065

Interpretation of surface-induced optical anisotropy of clean, hydrogenated, and oxidized vicinal silicon surfaces investigated by reflectance-difference spectroscopy
U. Rossow, L. Mantese, and D. E. Aspnes 3070

***Ab initio* calculations of the reflectance anisotropy spectrum**
J. M. Bass and C. C. Matthai 3075

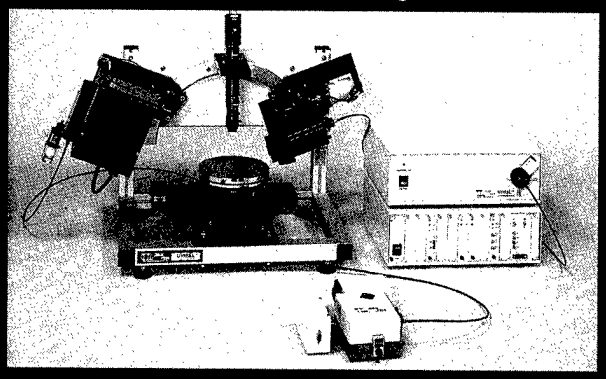
Anisotropic optical reflection by stepped surfaces
P. L. de Boeij, C. M. J. Wijers, and E. Zoethout 3080

(Continued)

No More Experts For Thin Films Characterization

With UVISEL Spectroscopic Ellipsometer

- Ultrafast
- Easy to Use
- High Modularity
- In and Ex-Situ Configuration
- Multiwavelength Capabilities
- Variable Angle of Incidence

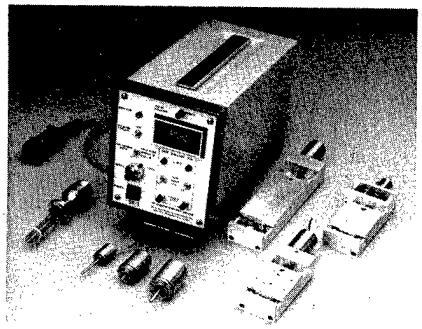


JOBIN YVON - SPEX
Groupe Instruments S.A.

16-18 rue du Canal - 91165 Longjumeau Cedex - France
Tel. : (33) 1/64 54 13 00 - Fax : (33) 1/69 09 93 19
In USA call: (1) 908/494 86 60 - In Netherlands call: (31) 0/172 43 33 23
In Germany call: (49) 89/46 23 17-0 - In Italy call: (39) 2/57 60 30 50
In UK call: (44) 181/204 81 42

Princeton Research Instruments, Inc.

IN-VACUUM MOTION SYSTEMS



Unique in-vacuum stepper motors and accessories provide rotary or translational motion in high and ultra high vacuum without the use of the more traditional mechanical feedthroughs. Features include:

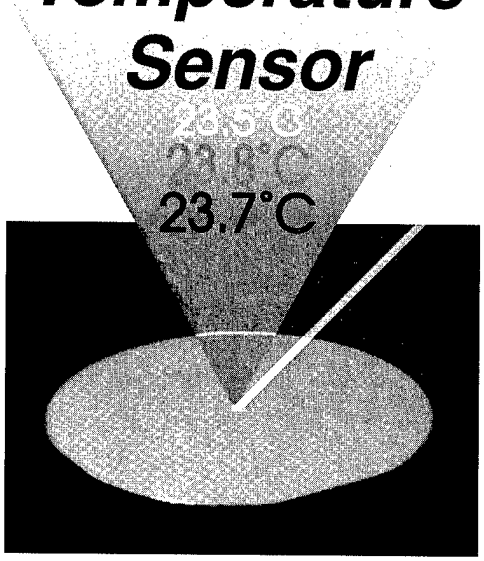
- Motor operation in vacuum as low as 4 X 10^{-10} torr.
- Maximum bakeout and operational temperature of 150° C.
- The motors can be placed anywhere, thus allowing compound motions that are otherwise impossible.
- The motors and motorized slides can be quickly and easily relocated to suit experimental needs.
- The motor driver electronic package is specifically designed for in-vacuum motors.
- The system is ideal for control with a remote controller or computer.
- All products are normally in stock for immediate delivery.
- Our engineers are glad to discuss you. particular application.

PRINCETON RESEARCH INSTRUMENTS

(609) 924-0570
P.O. BOX 1174 PRINCETON, N.J. 08542

DRS1000

Temperature Sensor



Know Your Substrate Temperature During Processing Just by Looking!

±0.2°C
1 second updates

The DRS 1000 is an all-optical, non-pyrometric temperature sensor for use in MBE, CBE, MOCVD, and silicon processing applications.

For more information about the DRS-1000 and other Thermionics vacuum products contact us at:

thermionics northwest, inc.

231-B Otto Street
Port Townsend, WA 98368
Phone: 360/385-7707
FAX: 360/385-6617
email: tnw@thermionics.com

Visit our homepage:
http://www.thermionics.com

Reflection anisotropy spectroscopy, surface photovoltage spectroscopy, and contactless electroreflectance investigation of the InP/$In_{0.53}Ga_{0.47}As$(001) heterojunction system
M. Leibovitch, P. Ram, L. Malikova, Fred H. Pollak, J. L. Freeouf, L. Kronik, B. Mishori, Yoram Shapira, A. R. Clawson, and C. M. Hanson 3089

New Approaches

Physics and chemistry of silicon wafer bonding investigated by infrared absorption spectroscopy
M. K. Weldon, Y. J. Chabal, D. R. Hamann, S. B. Christman, E. E. Chaban, and L. C. Feldman 3095

Nonlinear optical spectroscopy of Si–heterostructure interfaces
C. Meyer, G. Lüpke, Z. G. Lü, A. Gölz, H. Kurz, and G. Lucovsky 3107

Spatial mapping of ordered and disordered domains of GaInP by near-field scanning optical microscopy and scanning capacitance microscopy
J-K. Leong, J. McMurray, C. C. Williams, and G. B. Stringfellow 3113

Papers from the First AVS Topical Conference on Magnetic Surfaces, Interfaces, and Nanostructures

Preface 3118

X-ray magnetic microscopy and spectroscopy using a third generation synchrotron radiation source
A. T. Young, H. A. Padmore, and N. V. Smith 3119

Imaging of ferromagnetic domains using photoelectrons: Photoelectron emission microscopy of neodymium-iron-boron ($Nd_2Fe_{14}B$)
M. Mundschau, J. Romanowicz, J. Y. Wang, D. L. Sun, and H. C. Chen 3126

Effect of silicon processing on giant magnetoresistance
S. L. Burkett, J. Yang, D. Pillai, and M. R. Parker 3131

Electronic magnetic and structural coupling in colossal magnetoresistive (La,Ca)MnO_3
Warren E. Pickett and David J. Singh 3136

Spin-resolved x-ray photoemission studies of ferromagnetic metals
L. E. Klebanoff 3140

Magnetic circular dichroism in reflection electron energy loss spectroscopy?
G. R. Harp, R. F. C. Farrow, and R. F. Marks 3148

Probing surface and thin film magnetic structure with circularly polarized synchrotron radiation
G. D. Waddill, J. G. Tobin, X. Guo, and S. Y. Tong 3152

Theoretical investigation of structural instabilities of Fe layers on face-centered-cubic Cu
Lars Nordström and David J. Singh 3160

Investigation of face-centered-cubic Fe thin films using wedged samples
R. K. Kawakami, Ernesto J. Escorcia-Aparicio, and Z. Q. Qiu 3164

Magnetic x-ray dichroism in the spectroscopy of ultrathin magnetic alloy films
J. G. Tobin, K. W. Goodman, G. J. Mankey, R. F. Willis, J. D. Denlinger, E. Rotenberg, and A. Warwick 3171

Wavelength dependence of the magneto-optical properties of the interfaces of a Au sandwiched (001) Fe film
Wim Geerts, T. Katayama, Y. Suzuki, and J. Childress 3176

Correlation between dynamic magnetic hysteresis loops and nanoscale roughness of ultrathin Co films
Q. Jiang, H.-N. Yang, and G.-C. Wang 3180

Magnetic behavior of $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ pseudomorphic films
S. Z. Wu, F. O. Schumann, G. J. Mankey, and R. F. Willis 3189

Epitaxial growth, structure, and composition of Fe films on GaAs(001)-2×4
E. Kneedler, P. M. Thibado, B. T. Jonker, B. R. Bennett, B. V. Shanabrook, R. J. Wagner, and L. J. Whitman 3193

Mn 3*s* multiplet splitting of pseudomorphic Mn overlayers on Ru(001)
T. K. Sham, M. L. Shek, J. Hrbek, and D. G. Van Campen 3199

Temperature dependent magnetic surface anisotropy in ultrathin Fe films
D. P. Pappas 3203

Distinguishing the close-packed hexagonal and face centered cubic phases of the metallization of diamond by polarization-dependent extended x-ray absorption fine structure
K. M. Kemner, W. T. Elam, V. G. Harris, Y. U. Idzerda, and J. A. Wolf 3207

CUMULATIVE AUTHOR INDEX 3210

Over 7,000 Reasons to *Join* the American Vacuum Society

- *Receipt of 7,000+ pages of technical material*
- *Membership rates on conferences and symposia*
- *Discounts on short courses and educational materials*

All for under a penny a reason! Our programs allow you to:

Interact with Technology Leaders at Symposia and Conferences

Each fall, scientists gather from around the world to attend the American Vacuum Society (AVS) National Symposium —a week-long forum for science and technology exchange in the areas of thin films, microelectronics, nanostructures, processing, surfaces, and vacuum. The symposium features papers from eight technical divisions and two technology groups, short courses, topical conferences on emerging technologies, and an equipment exhibition.

Local chapters host symposia on topics of interest to their local technical communities and equipment exhibitions during the year. Joining the AVS and a local chapter or division is an excellent way to interact with technology leaders and stay informed.

Increase Your Job Skills—More Than 70 Short Courses by Industry Experts

AVS short courses have provided hard-to-find job training for more than 20 years in applied vacuum technology, surface analysis and materials characterization, and materials, thin films, and coatings: processing and properties. The AVS offers over 70 short courses that have been developed by industry experts, and regional programs are held all year long to meet the training needs of local technical communities across the U.S.

Stay Informed—Receive Over 7,000 Pages of AVS Publications Each Year

AVS members receive the *Journal of Vacuum Science and Technology*—the first major physics journal available on CD-ROM. Refereed papers cover surface science, interfaces, nanometer-scale science, STM, plasmas, thin films, and vacuum technology. The *AVS Newsletter* announces meetings and events and includes interesting articles and conference reviews.

Surface Science Spectra is a unique laboratory reference tool published as a journal and on disk. It's an evolving database (including related sample and instrument information) of peer-reviewed XPS, Auger, and SIMS spectra of a wide range of materials.

Maintain Your Technical Edge with a Host of Educational Resources

The AVS has an ongoing program to develop and provide convenient, cost-effective educational and training resources. Formats include monographs (single-subject information); a series of book reprints entitled *AVS Classics;* various instructional materials; recommended practices information; and video tapes that provide industrial training, offer tutorial lectures, or present topical conferences.

Join Us Today—Under a Penny a Reason!

Membership dues are $75 annually ($20, students) and include receipt of *JVST, Physics Today,* and the *AVS Newsletter;* discounts on events and educational materials; and free membership in any of the AVS' chapters or divisions. (There is a separate subscription fee for *Surface Science Spectra.)*

The AVS is a volunteer-based, nonprofit technical society dedicated to advancing the science and technology of thin films, microelectronics, nanostructures, processing, surfaces, and vacuum.

For more information, contact the AVS, 120 Wall St., 32nd Floor, New York, NY 10005, 212-248-0200, fax 212-248-0245, e-mail avsnyc@vacuum.org, http://www.vacuum.org.

Don't Lose Your Technical Edge ...

Attend the premier annual symposium on materials, surfaces, interfaces, and processes!

American Vacuum Society 43rd National Symposium
October 14-18, 1996, Philadelphia, Pennsylvania

Eight technical division programs will be held on issues related to films, microelectronics, nanostructures, processing, surfaces, and vacuum, including:

- Applied Surface Science.
- Electronic Materials and Processing.
- Nanometer-Scale Science and Technology.
- Plasma Science and Technology.
- Surface Science.
- Thin Films.
- Vacuum Metallurgy.
- Vacuum Technology.

In addition, the week-long annual symposium will feature:

Three Topical Conferences

- Flat Panel Displays (FPD).
- Magnetic Surfaces, Interfaces, and Nanostructures.
- Micro-electromechanical Systems (MEMS).

Two Technical Group Programs

- Biomaterial Interfaces.
- Manufacturing Science and Technology.

40+ Short Courses

- Applied Vacuum Technology.
- Surface Analysis and Materials Characterization.
- Materials, Thin Films, and Coatings: Processing and Properties.

150+ Exhibitors

- Vacuum and Deposition Equipment.
- Analysis Systems.
- Vacuum-Related Services.

For more information on the Symposium, visit the AVS Web site at http://www.vacuum.org or contact Angela Mulligan, (212) 248-0200, fax (212) 248-0245, e-mail avsnyc@vacuum.org.

The American Vacuum Society is a volunteer-based, nonprofit organization dedicated to advancing the science and technology of films, microelectronics, nanostructures, processing, surfaces, and vacuum and to providing a variety of educational opportunities. There are eight technology-based divisions, three technical groups, 21 local-area chapters, and about 6,000 members.

Single-electron charging of a molecule observed in scanning tunneling scattering experiments

H. Nejo[a)]
National Research Institute for Metals, 1-2 Sengen, Tsukuba 305, Japan

M. Aono
The Physical and Chemical Institute, 2-1 Hirosawa, Wako 351-01, Japan

D. G. Baksheyev and V. A. Tkachenko
Novosibirsk State University, Novosibirsk 630090, Russia

(Received 25 July 1995; accepted 9 April 1996)

Two extreme cases can be solved for electron charging effects where the first one ignores the electron correlation and the second one treats only one-electron scattering. So far, Coulomb blockade has been treated by the orthodox theory [Likharev, IBM J. Res. Dev. **32**, 144 (1988)], where the charging energy is expressed as $e^2/2C$ using the capacitance C of a tunnel junction. On the other hand, one-electron scattering can be solved exactly by using the Schrödinger equation. But it is difficult to solve for a many-body system by the Schrödinger equation. In this article, stress is placed on the case where the electric field cannot be ignored anymore. A treatment of the charging energy is proposed using a screening model or an electromagnetic field model by employing a cavity model in the field of photon. © *1996 American Vacuum Society.*

I. EXPERIMENTS

It is evident that the small central electrodes of a double tunnel junction whose sizes are on the atomic scale show a Coulomb staircase, as discussed below. The liquid crystal (LC) molecule 4′-7-alkyl-4-cyanobiphenyl (7CB) was used as the central electrode of a double tunnel junction (Fig. 1), where a scanning tunneling microscope (STM) tip acts as one outer electrode and shows Coulomb blockade [Figs. 2(a) and 2(b)].[1–3] Also, a fullerene molecule was used as a central electrode of another double tunnel junction, and it also shows Coulomb blockade [Fig. 2(a)].[4] Other molecules, 1,7-$(CH_3)_2$-1,2-$C_2B_{10}H_9Tl(OCOCF_3)_2$ in a Langmuir–Blodgett (LB) film [Fig. 2(c)],[5] and octasubstituted metallophtalocyanine (PcPd) between electrodes, also show a Coulomb staircase [Fig. 2(d)].[6]

II. ORTHODOX COULOMB BLOCKADE MODEL

Figures 2(a)–2(d) shows the calculated $i-v$ characteristics (solid lines). They were calculated based on the orthodox theory[7] where the charging energy is described as

$$\frac{n^2e^2}{2C}, \tag{1}$$

where C is the sum of the capacitances of a double tunnel junction. Coulomb blockade is observed when the tunnel resistance R is greater than the quantum resistance $R_0=\hbar/2e^2$ and also $e^2/2C$ is greater than the thermal energy k_BT. The orthodox theory treats the case where the number of electrons in the central electrode is so great that each electron in the central electrode does not interact and the extent of the charging can be treated as a capacitor.

[a)]Electronic mail: nejo@nrim.go.jp

III. ONE-ELECTRON SOLUTION (RESONANT TUNNELING)

Another extreme case is treating only one electron; then scattering of one electron by a potential can be solved exactly by the Schrödinger equation. In this approach, the one-particle one-dimensional Schrödinger equation is solved with piecewise constant potentials $U(x)$ and $V(x)$ and the effective mass $m(x)$ (for example, see Fig. 3):

$$i\hbar\frac{d\Psi}{dt}=-\frac{\hbar^2}{2}\frac{d}{dx}\left(\frac{1}{m}\frac{d\Psi}{dx}\right)+(U+V\cos\omega t)\Psi. \tag{2}$$

The solution of this equation obeys the condition of time periodicity of $|\Psi|^2$:

$$\Psi(\epsilon,t+\tau)=\exp\begin{pmatrix} & \epsilon & \\ -i & -\tau & \\ & \hbar & \end{pmatrix}\Psi(\epsilon,t),\quad \tau=\frac{2\pi}{\omega},$$

$$0<\epsilon<\hbar\omega, \tag{3}$$

where ϵ is the quasienergy. In every interval of constancy of U, V, and m, this solution looks like

$$\Psi_a(\epsilon,x,t)=\sum_{m=-\infty}^{\infty}\Psi_a(\epsilon_m,x)\exp\begin{pmatrix} & \epsilon_m & \\ -i & -t & \\ & \hbar & \end{pmatrix}$$

$$\times\exp\begin{pmatrix} & V_a & \\ -i & - & \\ & \hbar\omega & \end{pmatrix}\sin\omega t, \tag{4}$$

where a is the index of the interval, $\epsilon_m=\epsilon+m\hbar\omega$,

$$\Psi_a(\epsilon_m,x)=A_a(\epsilon_m)e^{ik_a(\epsilon_m)x}+B_a(\epsilon_m)e^{-ik_a(\epsilon_m)x} \tag{5}$$

and the wave vector

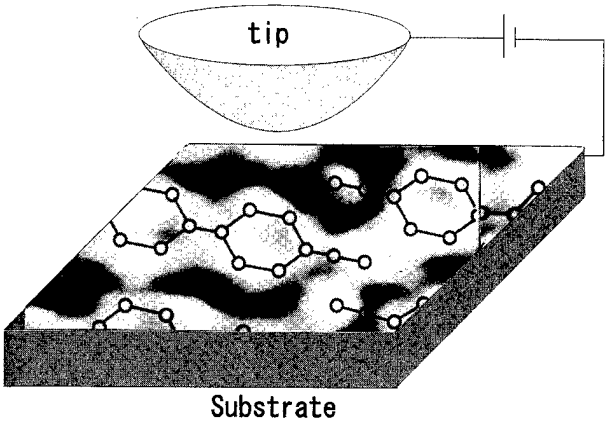


FIG. 1. Schematic configuration of a 7CB LC molecule on a substrate and under an STM tip which forms a double tunnel junction.

$$k_a(\epsilon_m)=\sqrt{\frac{2m_a(\epsilon_m-U_a)}{\hbar^2}}. \tag{6}$$

Figure 4 shows $|\Psi|^2$ at quasilevel ϵ_1, i.e., for incident electron energy $E_0=1$. It is seen that this quasilevel is the second level in the well (in the limit of infinitely thick barriers). Figure 5 shows $|\Psi|^2$ at quasilevel ϵ_1, when electrons flow from the side of thinner barrier (horizontally flipped potential of Fig. 3). The probability density in the well becomes so large that the accounting of electron–electron repulsion becomes important.

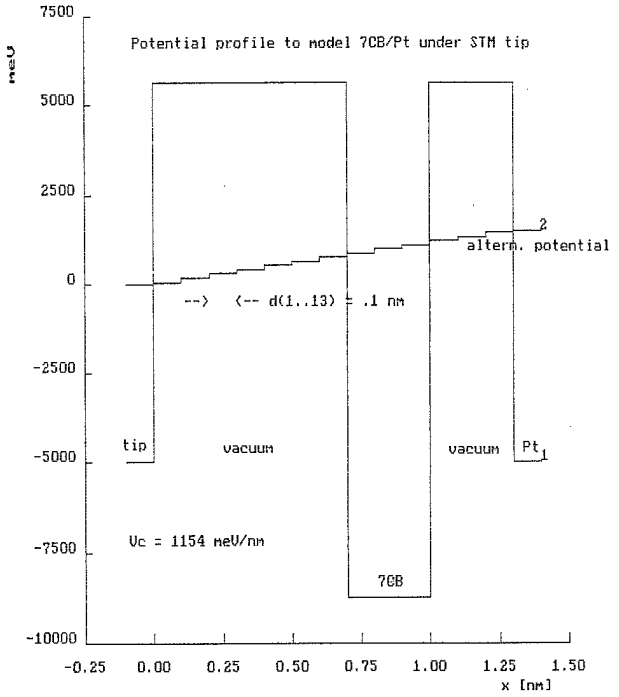


FIG. 3. One-dimensional potential profile to model an LC molecule sketched in Fig. 1.

IV. SMALL TUNNEL JUNCTIONS

When the size of the central electrode becomes small, the charging energy of a double tunnel junction can be converted

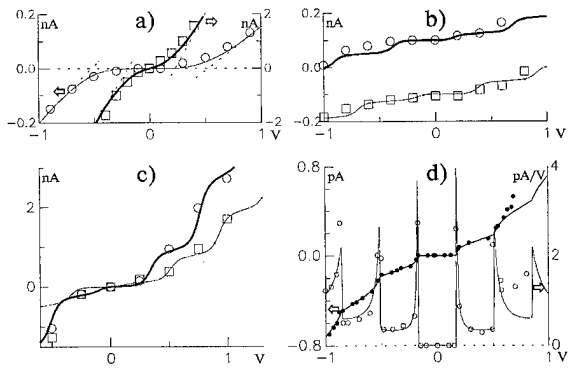


FIG. 2. $I-V$ curves calculated within the orthodox Coulomb blockade model (Ref. 7) with use of our programs (Ref. 10). (a) For a C_{60} molecule on Pd (left axis, Coulomb gap 1 eV) and for a monolayer of 7CB LC molecules on graphite (right axis, Coulomb gap 0.4 eV) under an STM tip at 300 K. Experimental values are marked with circles [from Ref. 4, Fig. 1(a)] and squares (from Ref. 2, Fig. 1). (b) For a monolayer of 7CB LC molecules on a ZrO_2/Pt substrate at 300 K with corresponding experimental values from Ref. 3, Figs. 5A (circles) and 5D (squares). Curves are shifted for clarity. (c) For 1,7-$(CH_3)_2$-1,2-$C_2B_{10}H_9Tl$ $(OCOCF_3)_2$ clusters in a LB film on graphite under an STM tip at 300 K [from Ref. 5, Fig. 2, curves 1 (circles) and 2 (squares)]. (d) $I-V$ and dI/dV curves for a monolayer of PcPd molecules between two crossed gold stripes at 4.2 K. Corresponding experimental values from Ref. 6, Figs. 5(a) and 5(b), are marked with empty (dI/dV) or filled ($I-V$) circles.

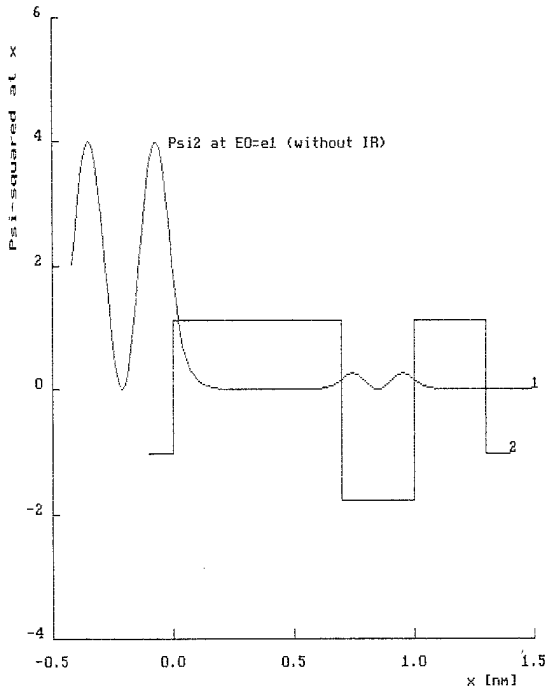


FIG. 4. $|\Psi|^2$ at quasilevel ϵ_1, i.e., for incident electron energy $E_0=1$. It is seen that this quasilevel is the second level in the well (in the limit of infinitely thick barriers).

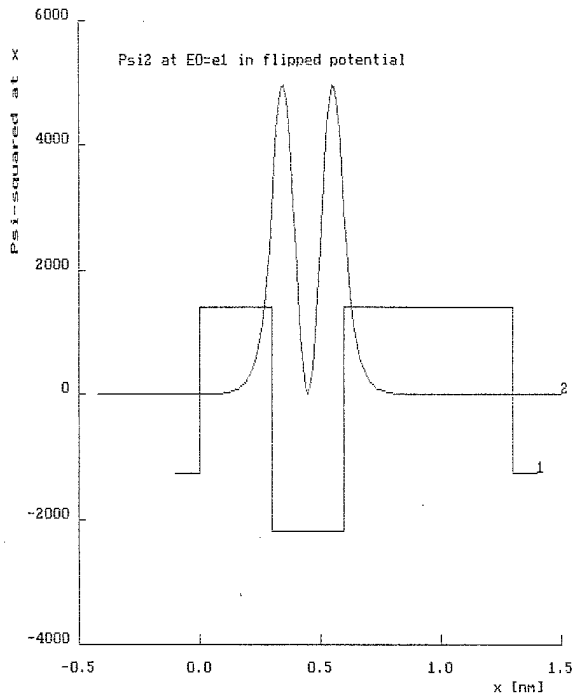


FIG. 5. $|\Psi|^2$ at quasilevel ϵ_1, when electrons flow from the side of thinner barrier (horizontally flipped potential of Fig. 3). The probability density in the well becomes so large that the accounting of the electron–electron repulsion becomes important.

to the charging energy of the central electrode. The charging energy of a double tunnel junction can be described as the charging energy of the central electrode:

$$\frac{(n_1-n_2)^2e^2}{2C}+\text{c.c.}, \tag{7}$$

where n_1 and n_2 are the number of electrons through tunnel junctions 1 and 2, respectively, and $C=C_1+C_2$, where C_1 and C_2 are the capacitances of the tunnel junctions 1 and 2, respectively.

When the battery works as the energy difference between n and $n+1$ electrons in the central electrode, the $(n+1)$th electron can tunnel to the central electrode:

$$eV=\frac{(n+1)^2e^2}{2C}-\frac{n^2e^2}{2C}. \tag{8}$$

A. Screening model

When the central electrode consists of a molecule or a small metal droplet, they are dielectric materials themselves, or all the materials are polarized at least under the high electric field (typically 10^9 V/m). The lines of electric force start from the tip (tip positive) and once they reach the negative pole of the molecule, they start again from the positive pole and reach the substrate (substrate negative) (Fig. 6). The electron tunnels from the substrate to the positive area of the molecule. When the polarity of the tip and the substrate changes, again the electron tunnels to the positive area of the molecule.

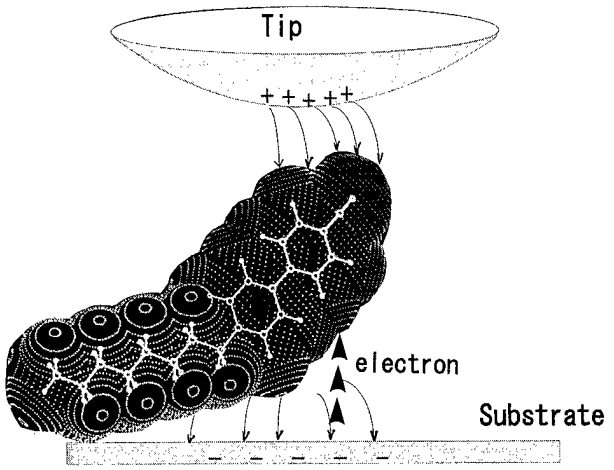


FIG. 6. The dielectric material in the tunnel gap. The lines of electric force start from the tip (tip positive) and once they reach the negative pole of the molecule, they start from the positive pole again and reach the substrate (substrate negative). The LC molecular surface potential image (without including the electric field effect) is shown. The nitrogen atom and the center of the biphenyl group have negative potential, and hydrogen atoms and carbon atoms of the biphenyl group have positive potential. A numerical potential value (P) vs color is green, $P<-15.0$ kcal/mol; blue, $-15.0<P<-9.0$; light blue, $-9.0<P<-3.0$; yellow, $-3.0<P<3.0$; violet, $3.0<P<9.0$; red, $9.0<P$.

The potential of n electrons $\varphi_0(\mathbf{r})$ without any screening is expressed as

$$\varphi_0(\mathbf{r})=\frac{ne}{r}. \tag{9}$$

In the positive gas, the potential of n electrons is screened and the effective potential $\varphi(\mathbf{r})$ is written as

$$\nabla^2\varphi(\mathbf{r})=-4\pi ne\,\delta(\mathbf{r})+4\pi e\langle\delta\rho(\mathbf{r})\rangle, \tag{10}$$

where $\delta(\mathbf{r})$ is a delta function and $-e\langle\delta\rho(\mathbf{r})\rangle$ is the induced charge density. Then the charging energy $(n+1)^2e^2/2C$ of the capacitor model should be replaced by $\varphi(r)e/2$, and this quantity reflects the screening effect so the energy may change by the number of electrons in the central electrode.

Alternatively, the charging energy can be expressed as the electromagnetic energy of tunnel junctions as

$$\int_V \mathbf{E}\,\delta\mathbf{D}, \tag{11}$$

where $\mathbf{E}$ is electric field, $\mathbf{D}$ is electric displacement, and the integral is taken over the tunnel gap space V. From the viewpoint of electromagnetic energy, until the electromagnetic field is charged up, the next electron cannot enter the central electrode. Also, the preceding electron should not escape the central electrode before the electromagnetic field is charged up. The former condition is equivalent to the condition of $e^2/2C>eV$ and the latter of $R\gg R_Q$ for observing Coulomb blockade.

This electromagnetic energy varies with increasing number of electrons in the central electrode, but it does not increase by the unit of e/C, since the electromagnetic field is not quantized by e/C.

B. Hartree–Fock calculation

Furthermore, the exact solution of a many-body system is obtained by *ab initio* calculation. By ignoring electron correlation, the total energy increase of the central electrode by an increment of one electron can be approximated using the Hartree–Fock method, and the condition for making $(n+1)$th electron be able to enter is

$$eV=E(n+1)-E(n). \tag{12}$$

But the Hartree–Fock calculation[8] of an isolated molecule gives the energy difference $E(n+1)-E(n)$ as 10 eV for 7CB,[1] and it is much larger than the experimental values of 0.1–0.2 eV. This brings us to the consideration of correlation energy.

The number of electrons (N_{TF}) in the circle of a radius of the Thomas–Fermi length is

$$N_{\mathrm{TF}}=\frac{4\pi n}{3k_{\mathrm{TF}}^3}=\frac{1}{12r_s^{3/2}}, \tag{13}$$

where n is number density of electrons, k_{TF} is the Thomas–Fermi wave number, and

$$r_s=\left(\frac{3}{4\pi n}\right)^{1/3}\frac{me^2}{\hbar^2}. \tag{14}$$

Since $r_s>1$ for a small metallic central electrode, Thomas–Fermi screening does not work anymore. Now, we have to consider the strong correlation case. But since the extent of correlation energy is 1 eV at most,[9] the calculated energy with correlation energy still does not match the experimental values. This would show that the most important factor for increasing the charging of a molecule between two outer electrodes is the screening as described Sec. IV A.

V. CONCLUSION

We and other groups have shown that a Coulomb staircase can be observed even for atomic scale central electrodes or molecules. Application of the orthodox theory is too crude for such a small number of electrons. On the other hand, one-electron scattering by potentials can be solved exactly, but it cannot include the charging effects. Up to now, one of the suitable approximations considers the screening effect. This way of treatment should be further improved.

ACKNOWLEDGMENT

The authors thank H. Hori for useful discussions.

[1]H. Nejoh, Nature **353**, 640 (1991).
[2]H. Nejoh, S. Watanabe, M. Aono, and N. Shima, JJAP Series **9**, 122 (1993).
[3]H. Nejoh, M. Ueda, and M. Aono, Jpn. J. Appl. Phys. **32**, 1480 (1993).
[4]H. Nejoh, S. Watanabe, M. Nakamura, and M. Aono, in Proceedings of the 3rd International Conference on High Technology, Chiba, 1992 (unpublished), p. 115.
[5]A. A. Zubilov *et al.*, JETP Lett. **20**, 41 (1994).
[6]C. M. Fisher, M. Burghard, S. Roth, and K. von Klitzing, Europhys. Lett. **28**, 129 (1994).
[7]K. K. Likharev, IBM J. Res. Dev. **32**, 144 (1988).
[8]M. J. Fish *et al.*, GAUSSIAN 90 (Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburgh, PA).
[9]S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982).
[10]D. G. Baksheyev and V. A. Tkachenko, in Proceedings of the Second International Conference NANO-II, Moscow, August, 1993 [Herald of Russian Acad. Technol. Sci. **1**, 723 (1994)].

Current characteristics in near field emission scanning tunneling microscopes

G. Mesa and J. J. Sáenz[a)]
Departamento de Física de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain

R. García
Centro Nacional de Microelectrónica, CSIC, Serrano 144, E-28049 Madrid, Spain

(Received 25 July 1995; accepted 18 April 1996)

The operation of the scanning tunneling microscope (STM) in the near field emission regime provides a direct, noninvasive approach for investigating surfaces at nanometer scale. Here, we present a study of the current characteristics in a near field emission STM. The influence of tip's geometry is analyzed. We show that the electron field emission from the sample is stable against tip-shape changes due to adsorbate diffusion or atomic rearrangements. © *1996 American Vacuum Society.*

I. INTRODUCTION

The scanning tunneling microscope (STM) has become the most versatile tool for imaging and manipulation of surfaces of conducting materials.[1,2] In the most common operation mode, the STM tip is scanned at a constant current of the order of nA and at applied voltages well below the tip and sample work functions. This implies typical tip–sample distances of the order or smaller than 1 nm. While this has the advantage of high lateral resolution, the close proximity of probe and sample may induce unwanted modifications of the interface. This phenomena is particularly severe when working with soft materials, such as molecular films or biomolecules. One way to overcome these problems is to operate the STM in the near field emission regime.[3] By using relatively high voltages and very small currents, the working distance can be of the order of a few nanometers. This minimizes the probe–sample interaction forces. Biological membranes adsorbed on conducting surfaces have been successfully observed[4,5] following this approach. On the other hand, low voltage and high current conditions can be used to modify and manipulate biological material.[5]

The performance of a near field emission STM as an imaging tool has been discussed recently.[3] By using a simple model for the probe–sample interface, we showed that it is possible to obtain subatomic vertical resolution and a few nanometers lateral resolution at working intensities of the order of pA. The best performance is obtained when the electrons are emitted from the sample. This is in contrast with earlier field emission approaches like the "topografiner" [6] where electrons were always emitted from a sharp tip. The high energy of the electrons impinging the sample could modify it. While the sample damage induced by electron impact is a severe drawback in the imaging mode, it has, in contrast, important technological applications. The low energy electron beam generated in the near field emission regime can be used for high resolution lithographic patterning in radiation sensitive materials.[7–10] On the other hand, atomic-scale desorption induced by electrons field emitted from a STM tip has been reported recently.[11]

Three regimes can be defined in the current (I) versus voltage (V) curves in a metal–insulator–metal junction. Tunneling when the applied voltage is lower than the work function φ; the intermediate region of voltages ($eV \gtrsim \varphi$) called the near field emission regime and the standard field emission or Fowler–Nordheim regime for high voltages ($eV \gg \varphi$).

In this work, we describe the current characteristics in near field emission STM. We study the dependence of the distance–voltage ($S-V$) curves on the tip's geometry. The stability of the current is also analyzed. When the emission comes from the tip, adsorbates that may appear on the tip would produce instabilities in the emission characteristics. This is due to changes in the work function and in the field very close to the protrusion. However, the changes in the field decay very fast with the distance from the protrusion (at least in the far field regime). It has been argued[3] that these have a negligible effect on the field at the sample surface. As we will see, electric field calculations show that this is indeed the case. The instabilities could then be avoided if the polarity is reversed, i.e., if the electrons are field emitted from the sample.

II. DISTANCE VERSUS VOLTAGE IN THE CONSTANT CURRENT MODE

To calculate the total current flowing between tip and sample, it is necessary to determine the current densities on the surface of the tip. For this we need an accurate evaluation of the potential and field distributions around the emitter surface. Sphere-plane[7] and hyperboloidal analytical models[3,6] have been used to estimate the electrostatic field in the near field emission regime. The numerical solution of the Laplace equation for different tip shapes and tip–sample distances[12] show that these estimations are only valid for tip–sample distances smaller than $\approx 1/2$ the tip radius. To develop a more realistic approach, we have modeled the tip as a truncated cone with an spherical apex characterized by

[a)]Author to whom correspondence should be addressed; Electronic mail: juanjo@dune.fmc.uam.es

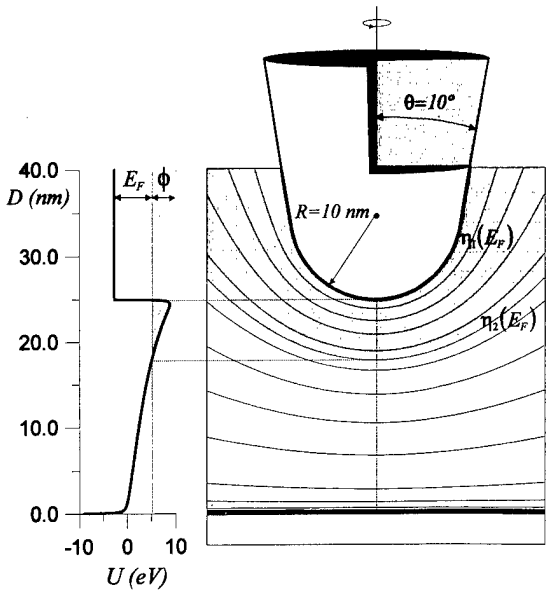


FIG. 1. Schematic view of the potential energy configuration. The upper part shows typical electrostatic equipotential lines. The shadowed region corresponds to the three-dimensional tunneling barrier. The corresponding energy diagram along the tip axis is also shown.

the curvature radius R at the tip apex and the cone half angle θ_C (see Fig. 1). For each tip–sample distance S, the electrostatic potential $V(\mathbf{r})$ as well as the classical image potential $V_{\text{im}}(\mathbf{r})$ are calculated numerically[12] (typical equipotential lines are shown in Fig. 1).

Once the potential distribution is known, we calculate the total current intensity I by using the following semiclassical approach.[13,14] The current density $J(\mathbf{r}_S)$ at each point of the emitter surface, $\mathbf{r}_S$, is obtained from a one-dimensional model,[15]

$$J(\mathbf{r}_S)\approx\frac{em}{2\pi^2\hbar^3}\left[\int_0^{E_F-eV_0}T(E_n,\mathbf{r}_S)dE_n + \int_{E_F-eV_0}^{E_F}(E_F-E_n)T(E_n,\mathbf{r}_S)dE_n\right], \quad (1)$$

where E_n is the normal energy in the tunneling direction, E_F is the Fermi energy, V_0 is the applied voltage, and $T(E_n,\mathbf{r}_S)$ is the tunneling probability. For simplicity, we assume that both tip and sample have the same Fermi energies E_F and work functions $\varphi(E_F=8$ eV and $\varphi=4.7)$. The $T(E_n,\mathbf{r}_S)$ associated with each emitting point is calculated within the Wentzel–Kramers–Brillouin approximation

$$T(E_n,\mathbf{r}_S)\approx\exp\left[-2\int_{\eta_1}^{\eta_2}\left\{\frac{2m}{\hbar^2}(U(\eta)-E_n)\right\}^{1/2}d\eta\right], \quad (2)$$

where $d\eta$ is the differential length along the field line normal to the surface at $\mathbf{r}_S$, and the tunneling barrier $U(\eta)$ is the exact potential calculated along this field line. $\eta_1(E_n)$ and $\eta_2(E_n)$ are the classical turning points at a given energy E_n. In Fig. 1, we show schematically the potential energy configuration used in the calculations. In a previous work,[3] we had used a mean barrier approximation[15] for the tunneling probability. However, this is only valid for relatively thick barriers.[16] Here we have calculated numerically $T(E_n)$ given by Eq. (2). Finally, the total intensity is obtained after integration of $J(\mathbf{r}_S)$ over the emitter surface.

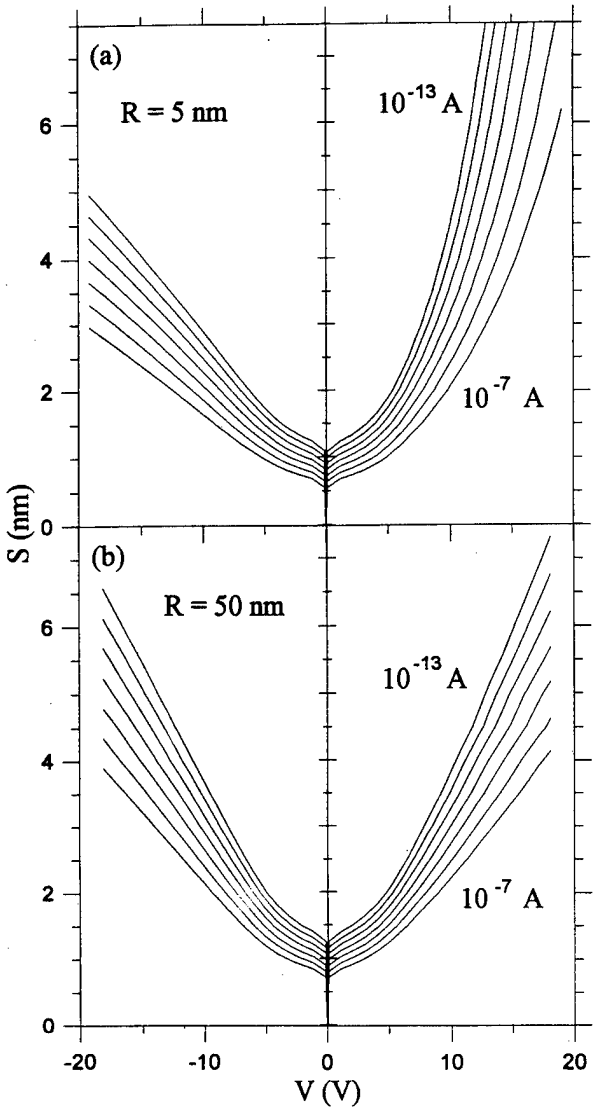


FIG. 2. Tip sample distances vs applied voltage at constant current I for different I values. In the field emission regime the current is controlled by the electric field. Consequently, the current increases with decreasing the tip's radius if the emission is from the tip (positive bias). The opposite happens when polarity is reversed. Intensity range is between 0.1 μA and 0.1 pA.

The dependence of the tip–sample distance S with the applied voltage V for two different tip radii and different set currents is presented in Fig. 2. Positive bias corresponds to electron emission from the tip. The qualitative behavior of the $S-V$ plot at constant current can be understood by using a simple argument. A constant current experiment in near field emission $(V>\varphi)$ implies an approximately constant field at the emitter surface. When the gap distance becomes comparable to the tip radius, there is a clear asymmetry in the curves because of the field is always higher on the tip

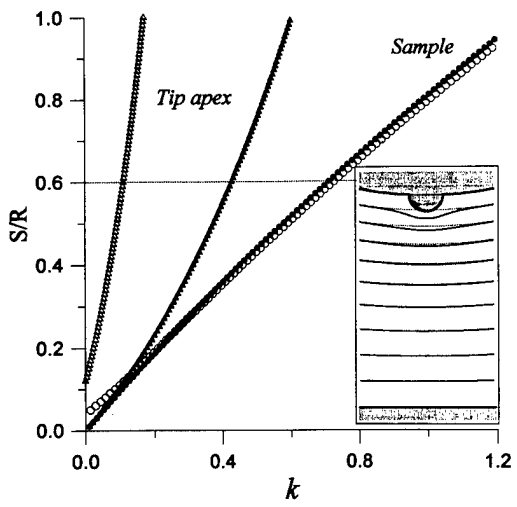


FIG. 3. S/R vs the k factor$=V/(FR)$, corresponding to the emission from the tip and from the sample. Full symbols correspond to a smooth tip. Open symbols reflect the changes in k due to the appearance of a small hemispherical protrusion ($R_b=0.05R$) on the tip apex. The inset shows the influence of a bump in the equipotential lines for a distance of $0.6R$. As can be seen, the equipotentials close to the sample are almost unaffected by the change of the tip apex geometry.

than on the sample [see, e.g., the results for $R=5$ nm, Fig. 2(a)]. When the distance is much smaller than the tip radius, the field drops linearly in the gap region. As a consequence, the results are almost independent of the polarity [see Fig. 2(b)]. It is worth noticing that, because of the semiclassical approach used, these results do not take into account the influence of field emission resonances.[17] These should appear as small quantum oscillations overimposed to the semiclassical result.[17]

III. CURRENT STABILITY IN NEAR FIELD EMISSION STM

In this section, we discuss the current stability against adsorbates or atomic rearrangements on the tip apex. The electric field F at the tip apex can be written as $F=V/(kR)$, where R is the radius of curvature of the tip and k is a geometric factor.[18] Since $k=V/(FR)$ is proportional to the applied voltage, the S vs k plot gives the dependence of the tip–sample distance with the applied voltage in a constant current (constant field) near field emission experiment. In Fig. 3, we present a comparison between the $S-k$ curves for emission from the tip (full triangles) and from the sample (full circles) surfaces. For distances $S/R \ll 1$, the voltage drops linearly within the tip–sample gap, and the k factor is linear with S. At distances of the order of $\approx 0.2R$, k already deviates from the linear behavior. This is the same behavior we have obtained in Fig. 2.

The presence of a small bump on the tip can produce a dramatic change in the current characteristics. In Fig. 3, we present the changes in k on the tip (open circles) and on the sample (open triangles) due to the sudden appearance of a small protrusion of radius $R_b=0.05R$. The field is strongly enhanced in the vicinity of the bump while it is almost unaffected at the sample surface (see inset in Fig. 3). This effect has important consequences in the performance of the STM in the near field emission regime. In a STM experiment in the constant current mode at fixed applied voltage, the feedback loop would try to keep the field constant by changing the tip–sample distance. Let us assume a typical STM tip of radius, say, ≈ 10 nm and a working distance of ≈ 3 nm. When electrons are field emitted *from the tip*, the appearance of an atomic-size bump ($R_b \approx 0.5$ nm) on the tip apex would lead the tip to retract more than 10 nm in order to keep constant field and voltage (i.e., constant k). On the other hand, the emission *from the sample* is almost unaffected.

IV. CONCLUSIONS

We have studied the current characteristics in near field emission STM by using a realistic model of the tip–sample interface. We have also studied the effects due to the presence of a small protrusion on the tip apex. We have shown that the field on the sample surface is almost unaffected by the presence of a small protrusion on the tip apex. As a consequence the electron emission from the sample is stable against tip-shape changes due to adsorbate diffusion or atomic rearrangements. We think that one of the main applications of our calculations is the optimization of the experimental conditions for imaging and modification of molecular films and thin oxide layers.[19]

ACKNOWLEDGMENTS

The authors would like to thank P. H. Cutler for interesting suggestions about the calculation of tunneling probabilities. They gratefully acknowledge support from the DGICYT under Contract Nos. PB92-0081 and PB94-0016.

[1] *Scanning Tunneling Microscopy (STM) and Related Methods*, NATO ASI Series E: Applied Sciences, Vol. 184, edited by R. J. Behm, N. García, and H. Rohrer (Kluwer, Dordrecht, The Netherlands, 1990).
[2] *Atomic and Nanometer-Scale Modification of Materials: Fundamentals and Applications*, NATO ASI Series E: Applied Sciences, Vol. 239, edited by P. Avouris (Kluwer, Dordrecht, The Netherlands, 1993).
[3] J. J. Sáenz and R. García, Appl. Phys. Lett. **65**, 3022 (1994).
[4] R. Guckenberger, W. Wiegrbe, A. Hillerbrand, T. Hartmann, Z. Wang, and W. Baumeister, Ultramicroscopy **31**, 327 (1989); R. Guckenberger, B. Hacker, T. Hartmann, T. Scheybani, W. Wang, W. Wiegrbe, and W. Baumeister, J. Vac. Sci. Technol. B **9**, 1227 (1991).
[5] R. García, Appl. Phys. Lett. **64**, 1162 (1994); J. Tamayo, J. J. Sáenz, and R. García, J. Vac. Sci. Technol. A **13**, 1737 (1995); R. García, J. Tamayo, J. M. Soler, and C. Bustamante, Langmuir **11**, 2109 (1995).
[6] R. Young, J. Ward, and F. Scire, Rev. Sci. Instrum. **43**, 999 (1972).
[7] M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **3**, 198 (1985).
[8] C. R. K. Marrian, F. K. Perkins, S. L. Brandow, T. S. Koloski, E. A. Dobisz, and J. M. Calvert, Appl. Phys. Lett. **64**, 390 (1994); C. R. K. Marrian, E. A. Dobisz, and J. M. Calvert, p. 139 in Ref. 2.
[9] M. J. Lercel, G. F. Redinbo, H. G. Craighead, C. W. Sheen, and D. L. Allara, Appl. Phys. Lett. **65**, 974 (1994).
[10] L. Stockman, G. Neuttiens, C. Van Haesendonck, and Y. Bruynseraede, Appl. Phys. Lett. **62**, 2935 (1993); C. Van Haesendonck, L. Stockman, Y. Bruynseraede, L. Langer, V. Bayot, J.-P. Issi, J. P. Heremans, and C. H. Olk, in *Proceedings of the 14th General Conference of CMD of the EPS*, Phys. Scr. T **55**, 86 (1994).

[11]T.-C. Shen, C. Wang, G. C. Abeln, J. R. Tucker, J. W. Lyding, Ph. Avouris, and R. E. Walkup, Science **268** 1590 (1995).
[12]G. Mesa, E. Dobado-Fuentes, and J. J. Sáenz, J. Appl. Phys. **79**, 39 (1996); G. Mesa and J. J. Sáenz, Appl. Phys. Lett. (to be published).
[13]P. A. Serena, L. Escapa, J. J. Sáenz, N. García, and H. Rohrer, J. Microsc. **152**, 43 (1988).
[14]J. J. Sáenz, N. García, Vu Thien Binh, and H. de Raedt, p. 409 in Ref. 1; J. Phys. Colloq. (France) **50**, C8-73 (1989).
[15]J. G. Simmons, J. Appl. Phys. **34**, 1793 (1963).
[16]N. M. Miskovsky, P. H. Cutler, and T. E. Feuchtwang, Appl. Phys. A **27**, 139 (1982); H. Q. Nguyen, P. H. Cutler, T. E. Feuchtwang, and N. Miskovsky, Surf. Sci. **160**, 331 (1985).
[17]K. H. Gundlach, Solid State Electron. **9**, 949 (1966); R. García, J. J. Sáenz, J. M. Soler, and N. García, J. Phys. C **19**, L131 (1986).
[18]R. Gomer, *Field Emission and Field Ionization* (Harvard University Press, Cambridge, MA, 1961).
[19]T. G. Ruskell, R. K. Workman, D. Chen, D. Sarid, S. Dahl, and S. Gilbert, Appl. Phys. Lett. **68**, 93 (1996).

Morphological modeling of atomic force microscopy imaging including nanostructure probes and fibrinogen molecules

David L. Wilson[a)]
Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106 and Department of Radiology, University Hospitals of Cleveland, Cleveland, Ohio 44106

Pranav Dalal
Department of Rhysiology and Biophysics, Case Western Reserve University, Cleveland, Ohio 44106

Kenneth S. Kump
Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106

William Benard
Department of Physiology and Biophysics, Case Western Reserve University, Cleveland, Ohio 44106

Ping Xue, Roger E. Marchant, and Steven J. Eppell
Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106

(Received 8 November 1995; accepted 19 March 1996)

Due to the finite size of the probe tip, atomic force microscopy (AFM) images of biomolecules, and other structures similar in size, are laterally enlarged. We use mathematical morphology, a non-linear image processing method, to model the interaction between probe tip and sample. In a typical imaging situation, baseline dimensions are most affected by the probe and widths can be 80% tip and 20% molecule. Using the morphological model and a known tip, we can restore the image so that it more closely resembles the actual surface. Morphological restoration is ideal in some regions, giving the exact sample surface, and improved in others. In the case of a carbon probe, restoration increases the perfectly obtained surface area by as much as 160 times. Following restoration, lateral widths at fixed heights are improved by as much as 75%. Restoration greatly improves image resolution even if one uses probes consisting of very small candidate structures, e.g., nanotubes and Bucky balls. The tip imaging process is also modeled, and we find that calibration spheres should be larger than the molecules of interest and that for many tips, there is little or no advantage to using smaller spheres. A blood plasma protein, fibrinogen, is modeled, and AFM and restored images of single molecules are computed. © *1996 American Vacuum Society.*

I. INTRODUCTION

The atomic force microscope (AFM) is a valuable tool in the three-dimensional analysis of biopolymers.[1–4] Using AFM, heights are measured accurately, but lateral dimensions are enlarged due to the finite size of the tip. This is particularly problematic with most biopolymers which fall in the size range of 4 to 10 nm in lateral dimensions by 1 to 5 nm in height.[5–10] A more recent application of AFM is the measurement of intermolecular forces.[11,12] It has become increasingly obvious that accurate information on tip and sample size and shape is needed to equate theory with measured forces.[13] These issues require one to understand and to possibly correct the distortion in AFM images.

In contact and tapping mode imaging of structures such as globular proteins with relatively high (a "high" object can contact the probe at a point above the apex where the width of the probe is comparable to the lateral size of the object) rounded sides, there is normally "contact" between the sample and an extended region of the probe as it travels up, over, and down the structure. The interaction is non-linear–contact is either all or nothing. Unlike a linear imaging system, AFM can exactly measure the height of a very small object but greatly enlarge its lateral dimensions. With the assumption of a hard-surface, sliding-contact, we and others model the interaction between sample and probe and propose restoration techniques to remove the tip-induced distortion and improve image resolution.[14–22] The interaction between tip and sample is not properly modeled by convolution, and AFM restoration is not deconvolution although some refer to the non-linear process by this name.[18]

In a previous report, we used mathematical morphology to model the probe-sample interaction, developed the AFM morphological restoration technique, measured tip surfaces by imaging known objects, and created restored images with improved resolution.[20] In the restored images, some regions were equal to the true sample surface before restoration, some were equal following restoration, and others were improved by restoration but not equal to the sample surface. Although these three regions were identified, their relative sizes were unknown.

In a recent paper, Westra and Thompson modeled AFM imaging of thin film surfaces.[23] They measured errors in elevation, related them to what they considered a perceivable

[a)]Author to whom correspondence should be addressed; Electronic mail: dlw@po.cwru.edu

change in grey-scale, and determined an image perception measure of distortion. In simulated AFM images without restoration, they found that the radius of curvature of the probe should be 1/10 the size of the imaged structures in order to have no perceivable difference in 90% of the pixels. In the case of biomolecules, adoption of this rule-of-thumb would create an impossible situation requiring probe tips with an apex one or two atoms wide by ten atoms high. We will investigate the issue of probe size and include an analysis of morphological restoration.

In the present report, we use morphological modeling to explore the characteristics and limitations of AFM imaging. Using spherical models, we quantify the amount of distortion before and after restoration. We analyze the resolution limit of AFM using very small candidate probe tips such as nanotubes and Bucky balls. Using models of single fibrinogen molecules, we predict the distortion introduced when one images a biomolecule. Although fibrinogen is used as an example, results will aid interpretation of AFM measurements from all molecules.

II. THEORY

We briefly describe the theory of morphological modeling and restoration that is presented in more detail by Wilson *et al.*[20]

Morphological image processing is processing with regard to shape. It is helpful to consider a grey-scale image as a two-dimensional surface with positive (white) grey values corresponding to hills or ridges in the surface terrain. The structuring element (SE) incorporates the shape of interest, and, in our application, it is a two-dimensional gray-scale surface equal to the AFM probe tip. There are two elementary operations, erosion ($\ominus$) and dilation ($\oplus$). An opening ($\bigcirc$) is obtained by eroding the input image and then dilating the resultant image using the same SE in both operations, $M \bigcirc L = M \ominus L \oplus L$. If the input image is first dilated and then eroded, a closing ($\bullet$) is obtained, $M \bullet L = M \oplus L \ominus L$. An opening can be interpreted physically as taking an object whose top surface is described by the SE, pressing it up against the underside intensity surface of the image, and moving it about while marking the highest point reached by the object at each pixel in the image. Positive excursions smaller than the SE are eliminated with an opening. In a closing, the SE is reflected left–right and flipped upside down. The resulting structure is then pressed down from the top onto the grey-scale image surface and moved around. The minimum gray-scale values reached by the reflected, upside down SE are recorded and become $M \bullet L$. Negative excursions smaller than the SE are removed by a closing.

Between the probe and sample, we assume a hard-surface interaction with no lateral forces. In Figure 1, the original sample surface is s; the AFM image surface is z; and the restored surface with improved resolution is r. Given a tip, t, and a model sample, s, one can calculate the AFM image surface, z,[20]

$$z = s \oplus t. \tag{1}$$

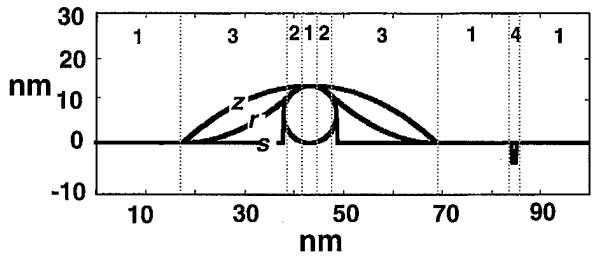


FIG. 1. Morphological restoration is illustrated. The sample surface is imaged by the probe tip. As compared to the measured AFM image, z, the restored image, r, more closely follows the sample surface, s, and has improved lateral dimension. For the morphological operations, the probe tip surface is considered to point upward with the tip center placed at the origin (not shown). In region <1>, $r=z=s$; in region <2>, $r=s \neq z$; and in region <3>, $r \neq s \neq z$. In regions <1> and <2>, ideal restoration is obtained and the restored surface equals the sample surface. In region <3> restoration gives an image surface having improved resolution. In region <4>, $r=z \neq s$, and there is no improvement with restoration.

Given the AFM image surface and a tip, one can obtain a restored image having improved resolution,

$$r = z \ominus t. \tag{2}$$

Presently, we model interactions and use a known sample surface, s; hence, we can calculate r from a closing ($r=s \oplus t \ominus t = s \bullet t$). In most situations, one restores a *measured* image of an unknown sample with a known tip. In this case, we compute $r=z \ominus t$.

We can analytically investigate the tip–sample interaction using simplified, one-dimensional structures consisting of a parabolic tip and a circular sample. A closing with a parabola removes small negative excursions and bounds the image from above. We can analyze the interaction in terms of derivatives.[24,25] The equation of a tip-down parabola is $y=Bx^2$ with $B>0$, and this gives a constant second derivative, $2B$, that bounds the second derivative of a closed image. In the case of the top of a circle $(x^2+y^2=R^2; y>0, -R \leq x \leq R)$, we obtain $[d^2y/dx^2 = (-1/R)/[1-(x/R)^2]^{3/2}; -R \leq x \leq R]$ and a second derivative that is always negative. Hence, with regard to second derivatives, the parabola always exceeds the circle, and we anticipate that a closing of the circle top with a parabola will leave the circle top unchanged. A closing also bounds the first derivative. The first derivative of a circle approaches $\pm\infty$ as $x \rightarrow \mp R$; i.e., tangents at the edges of the circle are vertical. Closing with a parabola will bound this limit. From these analyses of continuous functions, we anticipate that with the exception of the left and right edges, closing with a parabola of any size will not change the surface of the sample and the circle will be perfectly obtained. In practice, we are concerned with discrete computations using digitally sampled AFM images of tips and samples. More important, the spherical object always rests on a planar substrate, and this greatly limits the closing. These issues will be investigated later.

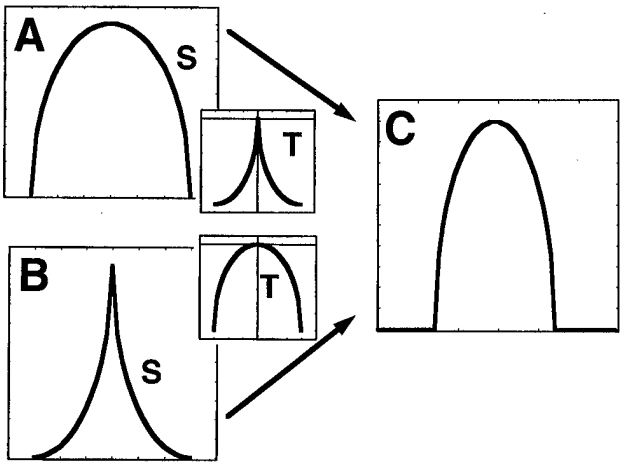


FIG. 2. One can switch sample and tip surfaces and obtain identical AFM images. In A, there is a rounded sample, s, and a pointed tip, t. In B, there is a pointed sample and a rounded tip. In both cases, identical image surfaces, z, are obtained in C.

Dilation is commutative, $M \oplus L = L \oplus M$;[26] hence, identical AFM images are obtained by dilating the sample by the tip, or the tip by the sample:

$$z = s \oplus t = t \oplus s. \tag{3}$$

It follows that we obtain an identical AFM image surface by interchanging the tip and sample surfaces.[20] In Figures 2A and 2B, very different samples are used yet identical AFM images are obtained. Obviously, an AFM image can very much depend upon the shape of the probe tip.

Using a known, "calibration" sample such as a gold or polystyrene sphere,[20] one can exploit the commutative property to obtain an image of the probe tip. We can substantially improve the resolution of the tip image by restoration. Recall that $r = z \ominus t$; switching s and t gives a restored tip surface,

$$r_{tip} = z \ominus s. \tag{4}$$

The generated tip surface can then be used to restore an image of unknown objects such as proteins and nanostructures.

With regard to restoration, a potential complication is that in a closing the dilation and erosion must be done with the same SE. In the case of AFM imaging, we have a good, but not ideal, estimate of the probe tip, and this may limit the ability to create a restored image. That is, we image with a real tip and restore with a slightly different, estimated tip. This creates an interesting error propagation problem. Even if we measure the probe tip exactly, we are uncertain about its effective origin. Fortunately, a change in origin of the tip in the restoration (erosion) process gives a true restored image that is simply shifted spatially; that is, $z \ominus t_{(x_s,y_s)} = (z \ominus t)_{-(x_s,y_s)} = r_{-(x_s,y_s)}$ where $t_{(x_s,y_s)}$ is t translated by (x_s, y_s). Spatial translation of the restored image should not affect interpretation.

There are at least four important regions associated with a restoration which are shown in Figure 1. In region <1>, both the raw AFM image and the restored AFM image are equal to the sample surface ($r = z = s$). In region <2>, the restoration recovers the sample surface ($r = s$), and the raw AFM image is degraded ($z \neq s$). In region <3>, the restoration improves image resolution but does not exactly recover the sample surface ($r \neq s$). In region <3>, restoration error in the lateral dimension progressively increases as one approaches the baseline, and the restored shape in region <3> reflects the probe shape. In region <4>, the probe is so large that no portion of it enters a negative excursion in the sample. In this case, the restoration does not improve resolution and $r = z$ but $r \neq s$. Restoration in regions <1> and <2> is *ideal*; it obtains the sample surface. In region <3>, restoration gives an *improved* version of the image. Finally, it should be understood that restoration does not allow one to image a region on the undersurface of a sample not reached by AFM.

III. METHODS

A. Computations and visualizations

In IDL (Interactive Data Language, Research Systems Inc., Boulder, CO), we construct mathematical models of probe tips and of sample objects including fibrinogen molecules. The models are written to image files and morphologically processed using IDL and/or DIPStation (Perceptive Systems Inc., Boulder, CO). In the case of measurements, DIPStation is used to properly prepare the data. For example, we scan images and cutout regions of interest for further processing.[20] Grey-scale and surface visualizations are done in DIPStation and IDL.

B. Models of AFM probe tips

We use a paraboloid to model the probe tip. A paraboloid well describes the carbon probes that we have imaged;[20] it is easily parametrized; and it has desirable mathematical properties with regard to second derivatives as described in Section II. The probe becomes a structuring element in the morphological calculations, and it is defined to be apex-up at the origin,

$$z = -B_v(x^2 + y^2). \tag{5}$$

For ease of analysis, some describe the probe tip as a sphere, and it may reasonably describe the region very near the apex.[5] However, in the case of biomolecules with high sides, a paraboloid will better describe contacts between the probe and sample. It may also be a more accurate model for future force calculations where interactions with some layers above the apex of the probe are required.

We are interested in the limits of imaging obtained with very narrow probes. For the case of a one-dimensional (1D) parabolic tip ($y = Bx^2$), the radius of curvature is $1/(2B)$ at the apex. We use this to compute the parabola size that will just enclose various objects of interest. One possibility is a single atom with diameter of 2 Å giving $B = 5\ \text{nm}^{-1}$. Another structure is the mono-atomic tip described by Fink.[27] Field ion micrograph tip images show that the first three

layers consist of one, three, and seven atoms. Assuming 2 Å diameter atoms and hexagonal close packing, we create a model of the first three layers (≈ 7 Å) of the probe tip. We find that a paraboloid with $B_v = 5$ nm^{-1} encloses all but a very small portion of the third layer of this model. Other potential probe tips are carbon Bucky balls and nanotubes. A nanotube consists of a cylinder with spherical caps resembling Bucky balls at the ends; nanotubes are significantly stiffer than any other presently known materials[28] making them ideal candidates for AFM probes. The diameter of a nanotube cap is ≈20 Å giving $B = 0.5$ nm^{-1}.[28] Recently reported measurements on currently used tips yields radius of curvature measurements that range from 30 nm[29] down to 3–6 nm,[30] and this corresponds to $0.017 < B < 0.17$ nm^{-1}. For measured electron-beam deposited carbon probes,[20] we determined $B \approx 0.015$ nm^{-1}.

These structures give us some probe sizes. On graphs, we place marks for a carbon probe ($B = 0.015$ nm^{-1}), a Bucky ball ($B = 0.5$ nm^{-1}), and a single atom ($B = 5$ nm^{-1}). It should be understood that we model a parabola which envelops either the Bucky ball or the single atom. Different results may be obtained if one models a cylindrical nanotube with a Bucky ball cap. In the case of fibrinogen imaging, we use three realistically dimensioned carbon tips: narrow ($B_v = 0.030$ nm^{-1}), medium ($B_v = 0.015$ nm^{-1}), and wide ($B_v = 0.0075$ nm^{-1}). The medium probe mimics our measurements of carbon probes.

C. Fibrinogen model

We create a molecular model of the blood plasma glycoprotein, fibrinogen using lateral dimensions from electron microscopy studies[31] and heights from our own preliminary AFM measurements. The fibrinogen model consists of two globular D domains (ellipsoids with equal lateral diameters of 6.5 nm and a height of 1.28 nm) and a central E domain (an ellipsoid with equal lateral diameters of 5.0 nm and a height of 1.0 nm). The globular domains are linked by coiled-coil structures which we model as cylinders. We link the three ellipsoids with two cylinders of 14.75 nm length and 1.2 nm diameter. The height of the E domain is set so that the ratio of E to D heights is the same as the ratio of E to D lateral dimensions.

We use simple structures to create the fibrinogen model, and equations are listed below where x and y refer to positions along the planar substrate and z refers to height. The top of an ellipsoidal volume (Equation 6) provides a model of a globular domain, and a cylinder of length l oriented along the x axis (Equation 7) provides connections between globular domains,

$$\frac{(x-x_0)^2}{E_x^2}+\frac{(y-y_0)^2}{E_y^2}+\frac{(z-z_0)^2}{E_z^2}=1, \tag{6}$$

$$\frac{(y-y_0)}{C^2}+\frac{(z-z_0)^2}{C^2}=1; \quad x_0 \leqslant x \leqslant x_0+l. \tag{7}$$

To construct a molecular model, we link ellipsoids with cylinders. Ellipsoids always rest on a planar substrate. Cylinders are parallel to the substrate and attached at the center of the facing surface of the smaller, central ellipsoid. In the case of AFM imaging, only the top of a three-dimensional object is measured, and we create a two-dimensional surface using the top of all objects. Each globular domain has lateral dimensions (E_x and E_y) and a height (E_z).

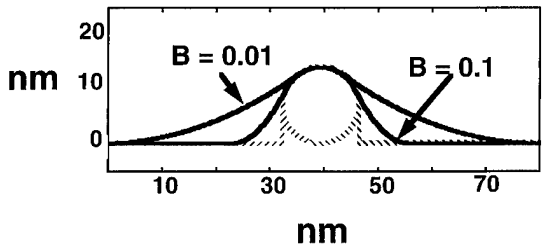


FIG. 3. Effect of tip size on restoration. The sample, dashed curve, is a circle, and the two restored, solid curves, are plotted. The raw AFM image is not shown. As B increases, the tip gets sharper and the restored image improves.

IV. RESULTS

We first investigate the limits of AFM imaging and the properties of AFM restoration using simplified, one-dimensional models. Probe models vary from hypothetical, synthetic nanostructure probes to conventional carbon and silicon nitride probes.[20] Later, we investigate three-dimensional models of fibrinogen.

A. Properties of morphological restoration

Using discrete calculations, we investigate the interaction of a parabolic tip with a circular sample.

To test the analytical analysis in Section II, we investigate a simplified arrangement where a circle is raised on a narrow, arbitrarily high pedestal above the substrate in order to eliminate the effect of floor contact. Using a range of parabolas ($0.01 < B < 10$) and circles ($1 < R < 100$), we determine that the restored curve (r) exactly equals the sample circle (s) at all points on the top surface. These results agree with the analytical, continuous analysis in Section II. Extending to a two-dimensional analysis and ignoring the substrate, a sphere can be restored exactly with any reasonably-sized paraboloid tip.

With the substrate present, we obtain restored images of a circle using two different parabolas (Figure 3). In both cases, the restored curve is much improved as compared to the AFM curve, z (not shown), and it is exactly equal to the sample circle as long as the tip does not contact the substrate. On either side of the circle, we obtain an image of the tip parabola where the parabola contacts the planar substrate and the edge of the circle. The more narrow probe yields a better restoration.

For continuous analytical functions, we derive an expression for the point of contact on a circle when the tip rests on the baseline substrate (see the Appendix). Following restoration, above this contact, the circle is imaged perfectly. We compute the angle of perfect imaging using Equation A6,

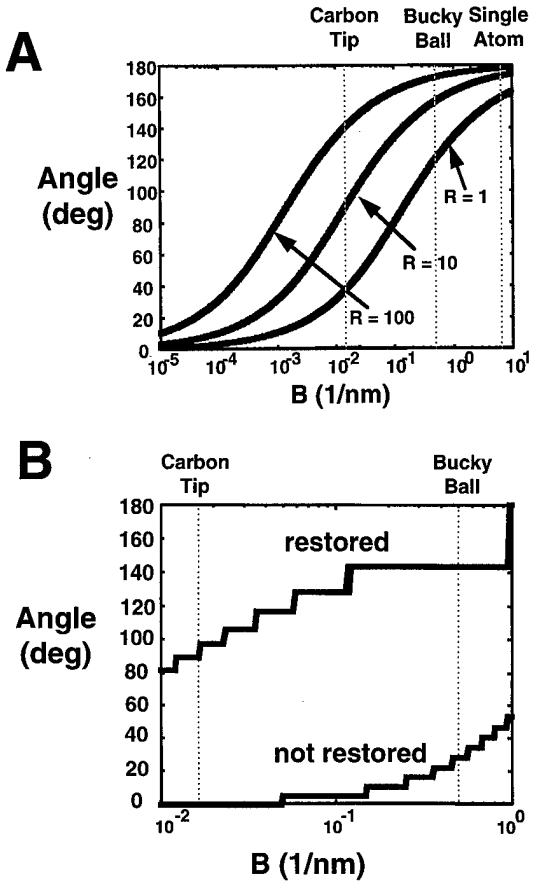


FIG. 4. The angle of perfect imaging is investigated for continuous (A) and discrete (B) computations. In A, we use the expression derived in the Appendix for continuous functions (Eq. (A6)) and plot the angle of perfect imaging as a function of independent variables, R, the radius of the imaged circle, and B, the parabola size. Prior to restoration, only an infinitesimal point is perfectly imaged. Using discrete calculations, a much different result is obtained, and the angle of perfect imaging is shown before and after restoration (B). Vertical lines mark B values obtained from a carbon probe, a Bucky ball, and a single atom. In the discrete calculations, the pixel size is 0.5 nm and the elevation calibration is 0.076 nm/gray-value, typical values used in our biomolecule imaging experiments.

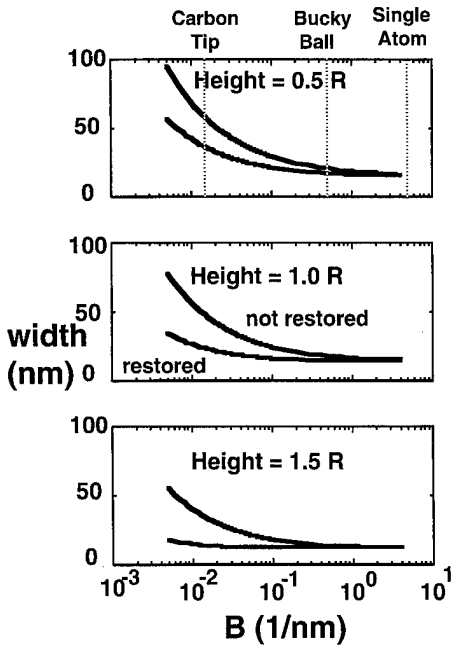


FIG. 5. Given a circle of diameter 14 nm, we compute the widths of the AFM curve, z (upper curve in each panel) and the restored curve, r (lower curve in each panel) at different heights above the planar substrate where 1.5 R refers to a height of 1.5 times the radius of 7 nm. Widths are plotted as a function of the parabola parameter, B. As B increases, the curves converge to the actual width of the sample. Vertical lines mark B values obtained from a carbon probe, a Bucky ball, and a single atom.

and in Figure 4A, we plot the result as a function of the parabola parameter, B, on the x-axis and the circle diameter, R. As expected, as B increases and the tip "sharpens," the initial parabola-circle contact moves down the side of the circle and the angle approaches the ideal result of 180°. The angle also depends upon the size of the sphere, and in the case of the smallest sphere ($R=1$ nm), it is still well below 180° at $B=10$ nm^{-1}. Extending to a two-dimensional surface, the angle measurement is proportional to surface area. Hence, for the case of a 10 nm sphere and a carbon probe, the restored angle is 90° out of 180° and 50% of the surface area is restored exactly. Prior to restoration, only a single infinitesimal point at the top of the sphere is imaged perfectly.

Images consist of pixels which are discrete samples in space and grey level, and discrete results will sometimes differ from analytical results on continuous functions. For this reason, in discrete simulations, we compute the angle of perfect imaging with and without restoration (Figure 4B). We determine the number of pixels that equal the sample pixels and convert this number to an angle using basic geometry. The angle of perfect imaging is obtained as a function of parabola size, B, and the radius of the circle, R. In this discrete simulation, the pixel size is 0.5 nm and the elevation calibration is 0.076 nm/gray-value, typical values used in our biomolecule imaging experiments. For the case of the carbon probe and a circle ($R=10$ nm), a single pixel is perfectly obtained prior to restoration, and this corresponds to an angle of $<5.6°$, the angle contained within ± 1 discrete samples at the top. After restoration, we get 90°, or greater than a 16-fold increase in the exactly restored surface area. Prior to restoration, in order to obtain more than one perfect pixel, a much narrower probe must be used ($B = 0.05$ nm^{-1} as observed in Figure 4B). Results prior to restoration depend greatly on pixel size, and we also investigate a much smaller pixel size of 0.05 nm (not shown). In this case, the "not restored" curve is shifted down the y-axis. Once again, for a carbon probe, we get 90° following restoration and one pixel before restoration. In this case of finer sampling, the angle contained within one pixel is reduced by an order of magnitude to $<0.56°$ and the perfectly obtained surface area is increased at least 160 times by restoration.

The improvement in lateral measurements following restoration is shown in Figure 5. For both the AFM and restored

curves, we plot the width as a function of B on the x-axis and the height above the planar substrate. There is significant improvement with restoration at all heights shown. As expected, as B increases and the tip becomes sharper, the curves converge to the actual width. For the case of a carbon probe and a circle of 14 nm in diameter, at baseline, the AFM and restored widths are both 69 nm (not shown). Hence, at baseline, the width measurement is 80% probe and 20% sample. Half-way up at a height of $1.0R$, the widths are 50.5 nm and 25 nm before and after restoration, respectively. These values give 72% probe and 28% sample prior to restoration and 44% probe and 56% sample following restoration. At a height of $1.5R$, there is more improvement. In this case, the widths are 37.5 nm and 15.6 nm before and after restoration, respectively. At this height of $1.5R$, the width of the circle is 12 nm, and the widths give 68% probe and 32% sample before restoration and 22% probe and 78% sample following restoration.

We have seen that restoration greatly improves AFM resolution when realistic probes and sample objects are used. However, with some unusual tip shapes, there is no improvement following restoration (Figure 6). Over a broad range of sample surfaces, following restoration, images from a convex tip (A) improve but a straight-edge tip (B) and a concave tip (C) do not. The requirement for improvement is that the tip be convex; for a point-down tip this means that the second derivative is positive.

B. Tip imaging

As described in Section II, one can create a restored tip surface using a known sample object. A parabolic tip is imaged using circle samples of various sizes. In Figure 7, we plot the *restored* curve as an overlay on the actual tip curve. We find that the tip is perfectly imaged until the probe reaches the floor, at which point, an image of the circle is obtained. If one creates a restored tip image using a small sphere of radius R, one obtains a truncated tip image which flattens at a distance $2R$ from the apex.

C. Application to fibrinogen

In Figure 8A, the fibrinogen model (center) is imaged using a carbon probe (left) and a silicon nitride probe (right). As determined previously, silicon nitride tips are significantly more lumpy and asymmetrical than carbon probes.[20] As a result, the image obtained with the silicon nitride probe is quite lumpy. Both AFM images are significantly enlarged.

The fibrinogen model (Figure 8B, middle) is dilated to create the AFM image (top) using the medium paraboloid tip model ($B_v = 0.015\ \text{nm}^{-1}$). The AFM image is morphologically restored (bottom). The restored image more closely approximates the model; that is, the globular domain linkages and the globular domains are more clearly defined. We get the correct height at the center of the domains both in the case of the AFM and restored images.

Next, we will quantify the improvement of the restored image over the AFM image. In Figure 8C, a fibrinogen

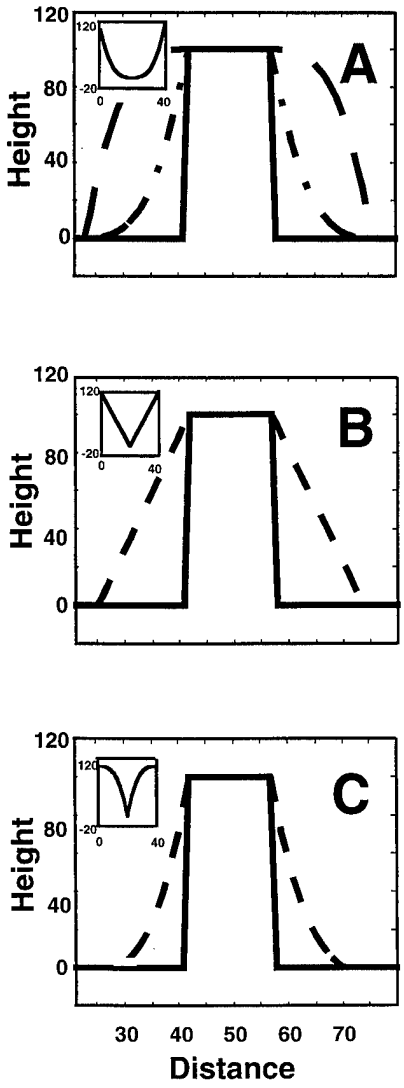


FIG. 6. The effect of morphological processing is investigated using some unusual tip shapes. In panel A, a realistic, convex tip is shown (inset). The sample curve (solid), AFM curve (outer dashed), and restored curve (inner dashed) are plotted. In this case, there is significant improvement. With a straight-edge tip (B) and with a concave tip (C), there is no improvement with restoration. The restored curve and AFM curve are identical. Tips in the insets are shown point down for illustration purposes, but in the morphological restoration process, they are defined to be point up.

model is imaged and restored using the wide (top), medium (middle), and narrow (bottom) model tips. Using grey-scales, we label regions defined previously. They are: white (<1>), black (<2>), and grey (<3>). Note that the relatively uninteresting region at the baseline is white (<1>). Because of the success of the restoration process, the area of region <2> is much greater than region <1>. In addition, following restoration, the area of region <2> is substantially increased with the smaller probe. The size of region <3> is decreased substantially when we use the smaller probe because this region reflecting the probe image at the boundary is smaller with the narrow tip.

We count labeled pixels, and report them as triplets (<1>, <2>, <3>), We neglect the uninteresting region

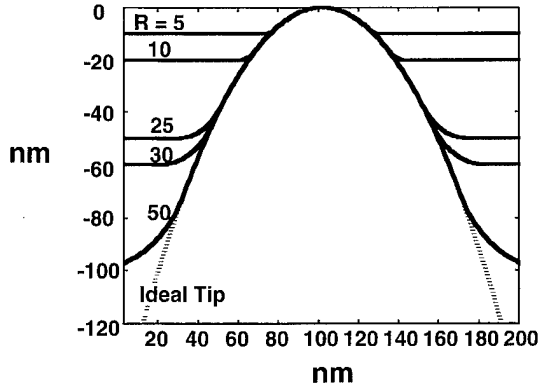


FIG. 7. A parabolic tip is imaged using various circular samples. The original tip is shown as a dashed curve. Other curves are the restored imaged tip using samples having radii from 5 to 50 nm as indicated in the figure. The restoration flattens out on either side because the tip contacts the floor.

<1> pixels at the baseline. The number of pixels are as follows: wide (3, 84, 10066), medium (19, 134, 6122), and narrow (33, 210, 3806). Ratios of perfectly obtained pixels (restored/not-restored) are wide (29), medium (8.05), and narrow (7.36). Note that these ratios are greatly affected by the pixels in the cylindrical linkages. Region <3> pixels are greatly affected by the size of the tip. One way to compare the size of region <3> is to compute ratios of numbers of pixels. The ratios are 0.37 (narrow/wide) and 0.6 (medium/wide). Yet another way to compare the effect of probe is the ratio of perfectly obtained pixels following restoration. The ratios are 1.7 (medium/wide) and 2.8 (narrow/wide). Even with restoration, significant improvement is obtained using the narrow probe tip.

V. DISCUSSION

A. Properties of morphological restoration

The one-dimensional modeling studies illustrate the importance of the restoration process. In the case of continuous functions, an infinitesimal point is perfectly obtained prior to restoration. For a carbon probe and a 10 nm diameter circle, 90° is perfectly obtained following restoration (Figure 4A) giving an infinite improvement. In the case of discrete computations, more than one pixel can be perfectly sampled prior to restoration especially for very sharp tips, wide samples, and/or large pixels (Figure 4B). However, in the case of a realistic carbon probe, only one pixel is obtained perfectly prior to restoration. In the case of a carbon probe and a 0.05 nm pixel, the angle contained within ± 1 pixel is 0.56°, and the perfectly obtained surface area is increased by over 160 times following restoration. To put it another way, prior to restoration, the perfectly obtained surface area is 0.3% of the total; following restoration, it is 50%. Restoration also improves measured lateral widths by as much as 75% (Figure 5).

We find that both in the case of AFM and restored surfaces, the width at the bottom of an object is most affected by the presence of the tip. In the case of currently realistic tips such as carbon probes used to image spherical objects of approximately 14 nm diameter, a width at baseline is approximately 80% probe and 20% molecule. Thus, the common practice of reporting the width at baseline very much depends upon the size of a tip. By reporting other widths above the substrate, AFM researchers can better compare results which are less dependent on tip widths and shape (Figure 5). For example, at 1/4 below the maximum height, before restoration, the width measurement consists of 68% probe and 32% molecule. Following restoration, the numbers improve to 22% probe and 78% molecule.

As described in Section I, Westra and Thompson modeled AFM imaging of thin film surfaces[23] and determined a rule-of-thumb that the radius of curvature of the probe should be 1/10 the size of the imaged structures. We have extended their analysis to include a range of object sizes (1 to 100 nm radius), a range of tip sizes (radii of curvature 0.1 nm to 50 μm), and morphological restoration. We require pixels to be perfectly obtained following restoration; Westra and Thompson use a relaxed condition requiring a grey-scale error less than that perceived by the human eye on 90% of pixels. If one requires 90% of the pixels to be perfectly obtained, our restoration analysis also shows that one needs a tip radius $\approx$ 1/10 the radius of the object (Figure 4A). However, if one relaxes the condition of 90% to 50% perfectly restored area, then even objects as small as 10 nm in radius can be obtained adequately in a restored image using a conventional carbon tip. In addition, objects as small as 1 nm in radius can be imaged with potentially obtainable tips larger than a Bucky ball. These are conservative analyses, and an even larger proportion of pixels have relatively little elevation distortion. Also, we find that many features in our fibrinogen model are plainly seen following restoration. We can determine the number of globular domains and their spacing. Hence, for globular biomolecules, useful information is obtained with relatively large probes.

Given the tremendous potential improvement, we recommend restoration of all AFM images. This is especially important for biomolecules consisting of globular domains in the range of 1–10 nm. Remarkably, even with very narrow, candidate probes such as a Bucky ball, there is significant improvement with restoration (Figure 4B). In fact, there is almost a nine-fold improvement in the correctly obtained surface area.

B. Tip shape and calibration

With some unusual, non-convex probe tip shapes, there is no improvement with restoration (Figure 6). All carbon probes that we have measured using AFM are convex within 15 nm of the apex.[20] Also, high resolution, scanning electron microscopy images of these tips show them to be convex within several hundred nanometers of the apex. On the other hand, silicon nitride tips can show local concavities within 15 nm of the apex.[20] A convex shape is implicit in reports of tips having a radius from 30 nm[29] down to 3–6 nm.[30] Atomic-scale tips are desirable, and in the case of carefully constructed tips containing a single atom at the end, the first

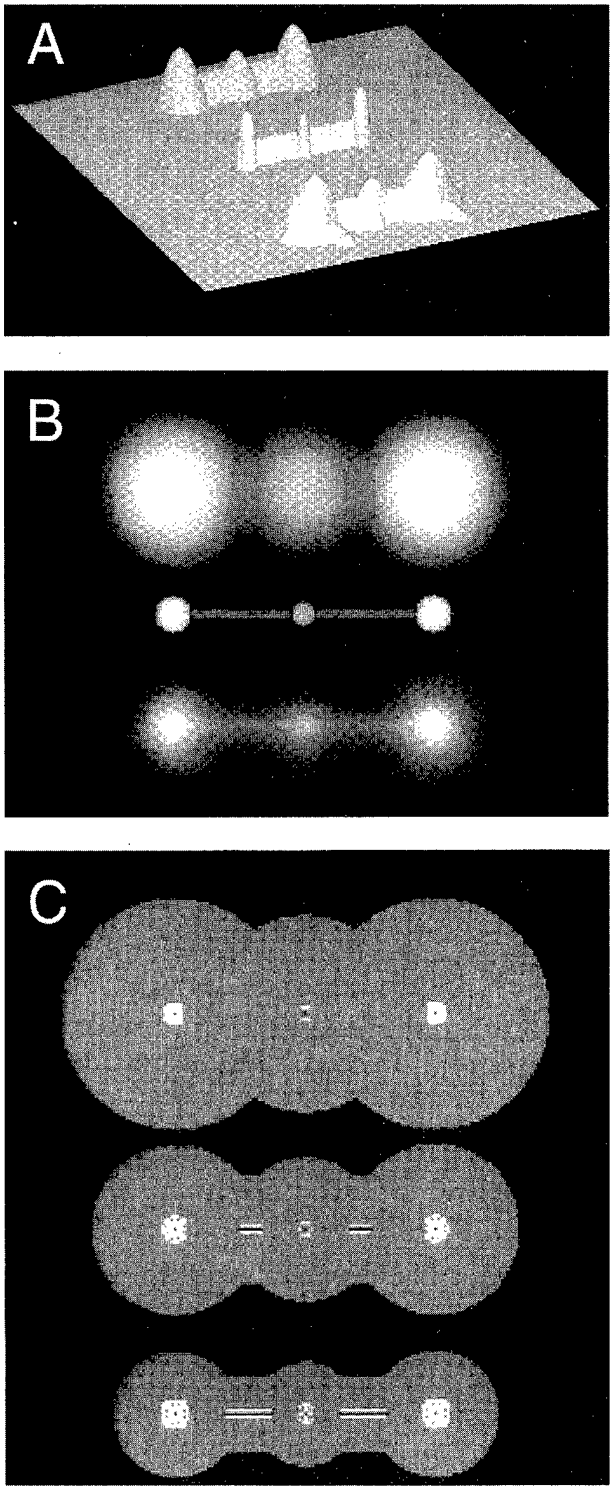


FIG. 8. Fibrinogen models are investigated. In A, a fibrinogen model (center) is imaged using a carbon probe (left) and a silicon nitride probe (right). As described in Section III C, the model is the top surface of a structure consisting of ellipses and cylinders. Surfaces are rendered using shading. The image obtained with the silicon nitride probe is quite lumpy due to the shape of the probe tip. In B, we demonstrate the restoration process. The fibrinogen model (center) is dilated with a paraboloid tip (top) and this is eroded to create the restored image (bottom). Heights are rendered as a grey-scale image with light values being high elevations. In C, the fibrinogen model is imaged and restored using a wide (top), medium (middle), and wide (bottom) tip. Regions obtained following restoration are identified in shades of gray. They are: white (region <1>), black (region <2>), and grey (region <3>).

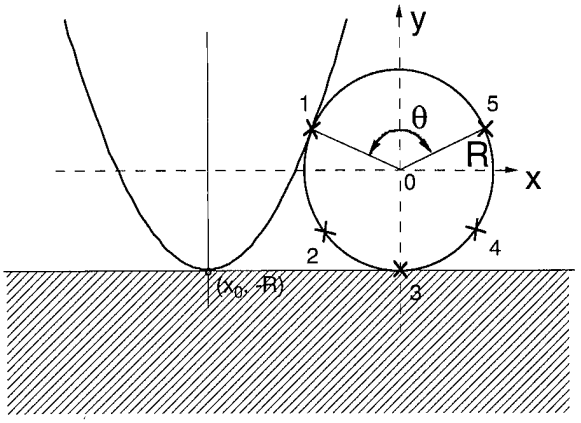


FIG. 9. A parabola tip will contact a circle as shown. Both rest on the planar substrate which is the top of the hatched area. Mathematically, there are five potential contact points that are marked with crosses and enumerated. In the text, we derive an expression for the physically meaningful contact points (1) and (5) in terms of variables defined in the figure. The region above the contact points marked by the angle θ is obtained perfectly following restoration.

three layers contain one, three, and seven atoms.[27] An enveloping surface over an atomic model consisting of these three layers has fairly straight surfaces with local convexities and concavities. However, we have never seen atomic scale features resulting from atoms on the side of our tips riding over a sharp edge (e.g., a step on mica or highly oriented pyrolytic graphite). Thus, we do not expect atomic scale convexities and concavities to greatly limit the usefulness of restoration.

One of our interests is to determine optimal calibration objects for imaging tips. Contrary to our initial expectation, with smooth, paraboloid-shaped tips, we find no need to use small spheres to get an accurate representation of the tip (Figure 7), and this is a demonstration of our arguments in Section II. Moreover, the downside with a small sphere is that one obtains a truncated tip image (Figure 9). If one uses a truncated tip to restore larger objects, unacceptable distortions will result. One rational, but imperfect, solution is to terminate the tip image near the "baseline" by placing the useful portion of the tip on a pedestal or marking SE pixels in the unusable region as ones not to be computed. A better solution is to measure the tip surface with a bigger calibration sphere. Thus, the results show that there is little benefit to measuring a smooth tip with a small sphere and that the sphere size should be larger than any unknown object of interest. It is understood that irregularly shaped tips with small protrusions and indentations must be sampled with a smaller sphere.

Accumulation or removal of debris on the probe tip during AFM scanning is a potential source of error. By changing the size of the SE, one can model the effect of this phenomenon on AFM and/or restored images. Obviously, if one uses an improper tip for restoration, this will propagate to an error in the restored image. Vesenka *et al.*[32] suggest dispersing calibration standards on the surface containing the objects of interest. In this way, the accumulation and removal of debris

from the tip can be monitored within an image. In principle, morphological modeling can then be used to dynamically alter the tip shape to reflect contaminants.

C. Imaging of biomolecules

Figure 8B illustrates the dramatic improvement of fibrinogen images following restoration. With restoration, conventional probes can be used to obtain useful images containing recognizable domain structures. Molecular modeling also helps identify the limitations imposed by the probe tip. AFM images obtained with a silicon nitride tip are very distorted and asymmetrical (Figure 8A). This effect is reduced, but not eliminated by morphological restoration (not shown). In the case of carbon probe tips, there are fewer small excursion distortions, but there is significant lateral enlargement. Following restoration of the medium tip, the perfectly obtained pixels are increased by a factor of 8. Note that with molecular modeling we count pixels rather than measure surface area. This helps explain the somewhat less dramatic effects than obtained with one-dimensional modeling. Comparing results with different probe sizes in Figure 8C, we determine that small probes are desirable. Even in this case, restoration is extremely important for imaging biomolecules.

VI. CONCLUSION

We have quantified the improvement obtained when using morphological restoration on AFM images. Using currently available tips to image molecularly sized objects, the number of perfectly obtained pixels is increased at least 7 times by restoration. For optimal imaging, restoration is not enough, and significant improvement is obtained with sharper probe tips. However, even with the sharpest tips that are likely to be available in the foreseeable future, morphological restoration provides significant improvement in AFM imaging. This is especially true when imaging globular structures like biomolecules. We conclude that morphological modeling and processing greatly aid interpretation of AFM measurements.

ACKNOWLEDGMENTS

Support for S.J.E. and R.E.M. is provided by NIH HL40047, the Whitaker Foundation, and the Center for Cardiovascular Biomaterials. P.D. is supported by the Biophysics/Bioengineering training program in the Department of Physiology and Biophysics at CWRU.

APPENDIX: ANGLE OF IDEAL IMAGING FOR A SPHERICAL SAMPLE AND A PARABOLOID TIP

We derive equations describing the contact point between a paraboloid tip and a spherical sample both of which are resting on a floor produced by the planar substrate. Because of the spherical symmetry, we can simplify the geometry and investigate the problem of a parabola and circle (Figure 9). In the figure, the top of the hatched region indicates the floor, and the origin is judiciously chosen in the center of the circle. The contacts (x,y) must lie on both curves at a point where the tangents are equal. Under these conditions, there are mathematically five possible solutions which are marked with crosses and enumerated on the circle (Figure 9). Only points (1) and (5) are physically meaningful in our case.

Given the coordinate system in the figure, the equation of a circle of radius R is

$$x^2+y^2=R^2 \tag{A1}$$

and the parabola is given in terms of a parameter B and an offset, x_0,

$$y+R=B(x-x_0)^2. \tag{A2}$$

The tangent of the circle is $dy/dx=-x/y$ and the tangent of the parabola is $dy/dx=2B(x-x_0)$. Equating tangents and rearranging for x_0, we obtain the equation below,

$$x_0=\frac{x}{2By}+x. \tag{A3}$$

Substituting into Equation (A2) and further eliminating x using Equation (A1), we obtain an equation containing y to the third power,

$$y+R=\frac{R^2-y^2}{4By^2}. \tag{A4}$$

There are three solutions: $y_1=(-1+\sqrt{1+16BR})/8B$, $y_2=(-1-\sqrt{1+16BR})/8B$, and $y_3=-R$. In Figure 9, these three solutions correspond to points (1) and (5); (2) and (4); and (3), respectively. We use y_1 which corresponds to the physically interesting solutions at (1) and (5). Substituting y_1 into Equation (A1), we obtain two x values corresponding to points (1) and (5), respectively.

We are interested in the angle of perfect restoration (θ) above the contact points (Figure 9). Using the y_1 solution, we obtain the final result below,

$$\theta=180^\circ-2\sin^{-1}\left(\frac{y}{R}\right), \tag{A5}$$

$$=180^\circ-2\sin^{-1}\left(\frac{-1+\sqrt{1+16BR}}{8BR}\right). \tag{A6}$$

[1]W. A. Rees *et al.*, Science **260**, 1646 (1993).
[2]H. G. Hansma *et al.*, Scanning **15**, 296 (1993).
[3]S. Ohnishi, M. Hara, T. Furuno, and H. Sasabe, Biophys. J. **65**, 573 (1993).
[4]B. G. De-Grooth and C. A. Putman, J. Microsc. **168**, 239 (1992).
[5]S. J. Eppell, F. R. Zypman, and R. E. Marchant, Langmuir **9**, 2281 (1993).
[6]H. Slayter, J. Loscalzo, P. Bockenstedt, and R. I. Handin, J. Biol. Chem. **260**, 8559 (1985).
[7]T. Thundat *et al.*, J. Vac. Sci. Technol. A **10**, 630 (1992).
[8]J. Yang and Z. Shao, Ultramicroscopy **50**, 157 (1993).
[9]Y. L. Lyubchenko *et al.*, Nucleic Acid Res. **21**, 1117 (1993).
[10]T. Thundat *et al.*, Appl. Phys. Lett. **63**, 2150 (1993).
[11]E. L. Florin, V. T. Moy, and H. E. Gaub, Science **264**, 415 (1994).
[12]G. U. Lee, D. A. Kidwell, and R. J. Colton, Langmuir **10**, 354 (1994).
[13]D. D. Leckband, J. N. Israelachvili, F. J. Schmitt, and W. Knoll, Science **255**, 1419 (1992).
[14]D. Keller, Surf. Sci. **253**, 353 (1991).
[15]J. Vesenka, R. Miller, and E. Henderson, Rev. Sci. Instrum. **65**, 2249 (1994).
[16]D. J. Keller and F. S. Franke, Surf. Sci. **294**, 409 (1993).
[17]P. Markiewicz and M. C. Goh, Langmuir **10**, 5 (1994).
[18]P. Markiewicz and M. C. Goh, J. Vac. Sci. Technol. B **13**, 1115 (1995).

[19]P. Markiewicz and M. C. Goh, Rev. Sci. Instrum. **66**, 3186 (1995).
[20]D. L. Wilson, K. S. Kump, S. J. Eppell, and R. E. Marchant, Langmuir **11**, 265 (1995).
[21]J. S. Villarrubia, Surf. Sci. **321**, 287 (1994).
[22]G. S. Pingali and R. Jain, *Proceedings IEEE Workshop Applications of Computer Vision* (IEEE, Piscataway, NJ, 1992), pp. 282–289.
[23]K. L. Westra and D. J. Thomson, J. Vac. Sci. Technol. B **12**, 3176 (1994).
[24]S. R. Sternberg, Comput. Vision Graphics Image Proccess. **35**, 333 (1986).
[25]R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision* (Addison-Wesley, New York, NY, 1992), Vol. 1.
[26]R. M. Haralick, S. R. Sternberg, and X. Zhuang, IEEE Trans. Pat. Analys. Mach. Intell. **PAMI-9(4)**, 532 (1987).
[27]H.-W. Fink, IBM J. Res. Dev. **30**, 460 (1986).
[28]S. Wang and D. Zhou, Chem. Phys. Lett. **225**, 165 (1994).
[29]T. R. Albrecht, S. Akamine, T. E. Carver, and C. F. Quate, J. Vac. Sci. Technol. B **8**, 3386 (1990).
[30]H. G. Hansma, K. A. Browne, M. Bezanilla, and T. C. Bruice, Biochemistry **33**, 8436 (1994).
[31]W. E. Fowler and H. P. Erickson, J. Mol. Biol. **134**, 241 (1979).
[32]J. Vesenka, S. Manne, and R. Giberson, Biophys. J. **65**, 992 (1993).

Scanning scattering microscope for surface microtopography and defect imaging

J. Lorincik,[a] D. Marton,[b] R. L. King, and J. Fine[c]
Surface and Microanalysis Science Division, National Institute of Standard and Technology, Gaithersburg, Maryland 20899-0001

(Received 1 April 1996; accepted 13 May 1996)

A recently developed optical-scatter instrument, the scanning scattering microscope (SSM) produces two-dimensional images of very small surface features and of variations in surface topography. In its present configuration, the lateral resolution is better than 10 μm and its sensitivity to surface roughness is in the angstrom (rms) range. The performance of the SSM has been demonstrated using calibration gratings, Si(100) surfaces, and Ge thin films grown on silicon; intercomparison also was made with atomic force microscope measurements. These results indicate that this scanned optical technique is a very useful noncontact method for evaluating surface microtopography.

I. INTRODUCTION

The continuing impact of miniaturization and shrinking feature size on semiconductor device performance has greatly increased the need for noncontact methods for measuring the topography and defect structure of surfaces and buried interfaces.[1,2] In the development of ultrathin (~5 nm) gate insulators, for example, where microroughness and defects directly affect electrical characteristics,[2] it is critical to monitor and control both the surface and interface topography as well as defect size and density. Precise on-line measurement methods are needed not only of the average surface roughness but, even more so, of the lateral changes in topography, often with a lateral resolution of less than one micrometer. The scanning scattering microscope (SSM), which we describe here, has the capability of filling many of these needs; we have used it to produce two-dimensional maps of surface microroughness and defect structure with subnanometer roughness sensitivity and with micrometer lateral resolution. The images generated by the SSM allow us to visualize very subtle variations in surface microroughness and defect structure.

The SSM is a recent development of one of the earliest optical techniques for the surface roughness measurement of finely finished surfaces, the total integrating scattering (TIS) method.[3] This method is based on the concept that rough surfaces scatter more light (into off-specular directions) than smooth surfaces. The information content of the off-specular scattered light, however, is quite different from that of the specular beam so that optical techniques based on scattered light measurement (such as TIS) are characterized by the rms roughness of the surface and not by the shape or actual height of surface features. In its original implementation, TIS measurements are done in a static mode using a relatively large (~1 mm^2) beam of light to obtain an average value of surface roughness for the illuminated area. The concept behind the SSM technique is to make use of the information contained in the scattered light (TIS) but to use a microfocused, scanned light beam in order to produce two-dimensional maps of surface roughness and defect distributions.

The usefulness of the SSM technique for imaging lateral variations in surface roughness was initially demonstrated with a prototype instrument[4–6] operating at atmospheric pressure. Earlier measurements obtained in a static, unscanned mode (i.e., the TIS method) showed that sensitive microroughness measurements also could be made, *in situ*, under ultrahigh vacuum conditions.[7] In this article we describe a newly designed, substantially improved version of the original SSM concept.

The general form of a TIS scatterometer is shown in Fig. 1. The laser beam impinges on the sample at normal incidence; the specular beam is reflected back to a photodetector to measure I_0—the specular-light intensity. The scattered off-specular light I_s is focused by a hemispherical mirror (sometimes called a Coblenz sphere) onto a photodetector. This photodetector is located in the basal plane of the mirror, spaced equally distant from the center of the mirror as the illuminated spot on the sample, i.e., the photodetector and the illuminated spot are at conjugate focal points of the hemisphere. The ratio of the scattered-light intensity to the specularly reflected light intensity is defined as the TIS. Quantitatively, this technique is based on an approximate theory[8–10] which, in the smooth-surface limit ($R_q < 0.1\ \lambda$), directly relates TIS to rms roughness R_q (in optics also known as σ)

$$\mathrm{TIS} = I_s/I_0 = (4\pi R_q/\lambda)^2, \tag{1}$$

where I_s is the off-specular light intensity, I_0 the specularly reflected intensity, and λ the wavelength of the laser light source. With certain limitations,[1,10] TIS measurements are now widely used for surface roughness metrology and are

[a]Permanent address: J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Prague.
[b]Permanent address: Chemistry Department, University of Houston, Texas.
[c]Electronic mail: JFINE@ENH.NIST.GOV

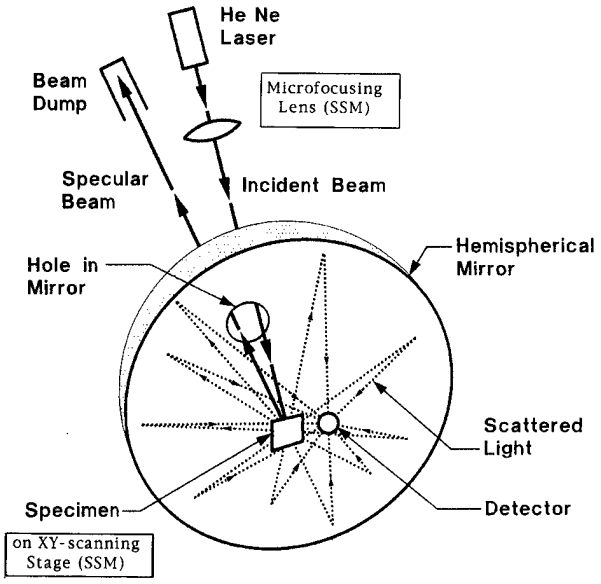


FIG. 1. Optical scattering concept of TIS and the SSM.

also used to identify surface defects as well as monitor thin-film fabrication. The TIS method has been adopted by the American Society for Testing and Materials as a standard test method (F1048-87)[11] for surface roughness.

The SSM technique which we are describing in this article combines the features of TIS with those of the scanned probe techniques. This has resulted in two major modifications to the general TIS setup in Fig. 1.

(1) XY raster scanning of the sample, in a point-by-point manner, so that TIS [or rms roughness calculated using Eq. (1)] can be plotted as a function of sample position to create two-dimensional roughness maps.
(2) Use of a focused laser beam to produce an illumination spot size of micrometer dimensions for improved lateral resolution. As a result, the power density of the microfocused spot is much higher and, therefore, provides an increased sensitivity to surface microroughness.

II. THE SSM APPARATUS

A schematic diagram of the improved SSM is shown in Fig. 2. It is composed of four basic elements: an illumination system, a beam analyzer system, a detection system, and a sample scanning system.

The illumination system consists of a highly stabilized He–Ne laser (3 mW, single mode, polarized, power stability better than 0.1%), a Faraday isolator, a spatial filter with beam expander, an illumination lens, a cube beam splitter, a

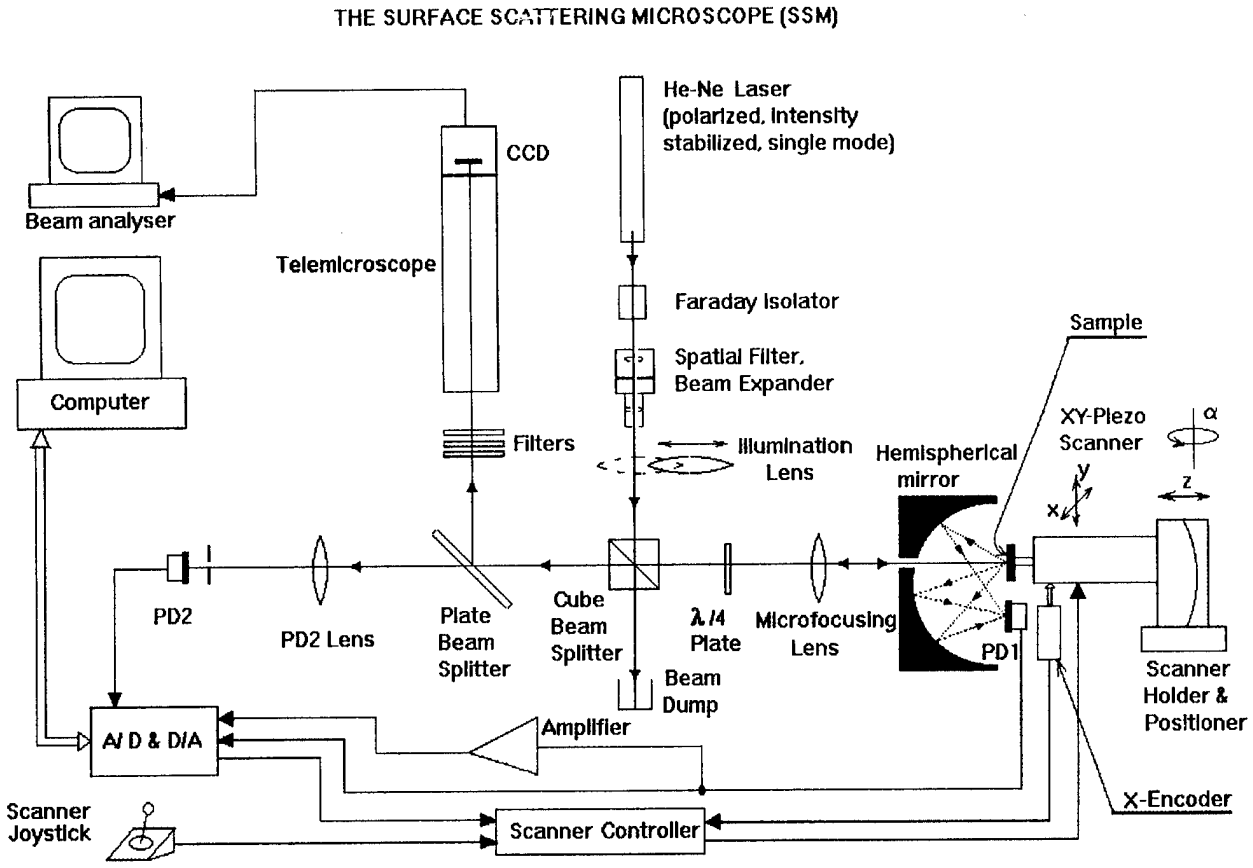


FIG. 2. Experimental configuration of the SSM.

quarter-wave plate, and a microfocusing lens ($f=110$ mm). The sensitivity of the stabilized laser to retroreflectance necessitated the use of a Faraday isolator. The choice of spatial filter, beam expander, and focusing lens was made so as to minimize the size of the beam spot on the sample.

The beam analysis system consists of a telemicroscope, a CCD camera, a plate beam splitter, and a beam profile analyzer. A part of the specular beam is reflected by the plate beam splitter into the telemicroscope and CCD camera of the beam analysis system; a magnified image of the beam can then be observed and its beam profile analyzed.

The detection system consists of a hemispherical mirror (inner diameter=132 mm), the detector for the scattered light (PD1), the detector for the specularly reflected light (PD2), a quarter-wave plate, and a focusing lens for PD2. The relative sensitivity of PD1 and PD2 was calibrated. The hemispherical mirror has a hole (diameter=8.0 mm) positioned 8.0 mm off axis. The beam enters the hole parallel to the axis of the mirror. The hemispherical mirror focuses the scattered light onto the photodetector PD1 placed at the conjugate focal point of the mirror (Fig. 1). For collection of the off-specular light, this geometry imposes a lower limit of ~3.5° on the scattering angles. From the geometry of this instrument, the upper-limit scattering angle would be close to 90°. Using a one dimensional grating equation,[10] we estimate the bandwidth to be between 0.095 and 1.6 μm^{-1} (which correspond to surface wavelengths between 0.6 and 10.5 μm). It is difficult, however, to define the upper bandwidth limit (for a 90° scattering angle) from the geometry alone as the light scattered to large angles may be reflected by the photodetector surface and, therefore, would not contribute to the electronic signal for I_s; the upper bandwidth limit may thus be less than we have estimated. The quarter-wave plate changes the polarization of the specular beam from circular to linear p polarized so that the cube beam splitter is nearly transparent to it. The specular light is then focused by the PD2 lens onto the second photodetector (PD2). The signals from both photodetectors are fed into an A/D converter and processed by the computer. For the scattered light, there is an option to amplify the signal 50× so that there is adequate dynamic range for measuring weak signals.

The scanning system consists of XY-piezo drivers, a linear position encoder of submicrometer resolution, and the scanner holder with tilt and rotation. The piezo drivers and the encoder operate in an open loop setup for the y direction and in a closed loop for the x direction. They, as well as the whole SSM system, are controlled, by the SUPERSCOPE II software[12]* installed and customized on our computer. We have found the scan linearity to be better than 1 μm over a 512×512 μm^2 scan.

The illumination lens ($f=1$ m) (Fig. 2), when inserted in the light path, serves to illuminate an area of about 1 mm^2 of the specimen so that the specimen can be viewed directly on the beam analyzer display and regions of interest located. The specimen can be moved by the joystick in both x and y directions to the selected region.

An alignment procedure has been devised which, once it is initially established, allows for routine specimen installation and focusing. Once the SSM is aligned (this includes focusing the telemicroscope on the image of the microfocusing lens), optimal focusing of the laser beam on the sample can be achieved by moving the sample in the z direction, so as to obtain the smallest round spot on the beam analyzer display.

Data acquisition is controlled by the SUPERSCOPE II system,[12]* customized for our SSM. The maximum data acquisition rate is 30 pixels per second. If, for each pixel, we average 100 voltage measurements from each photodetector, the acquisition time for taking parallel specular and off-specular images (256×256 pixels with 1 μm step size) is about 1 h. After collecting the data, some processing, statistics, and wavefront manipulation can be immediately done within the SUPERSCOPE II software.[12]* The data are then loaded into an image analysis program and further processed; "rms roughness maps" also can be created by our application programs.

III. RESULTS AND DISCUSSION

A. Lateral resolution test

To evaluate the lateral resolution of the SSM (for the laser focusing optics described here), we used a VLSI standards STM180A z-calibration mate-rial.[13]* This specimen has three gratings with 18 nm high steps made by VLSI technology. The gratings are platinum-coated silicon-dioxide patterns on a silicon die. We used gratings with periodicities of 20 and 10 μm. In Fig. 3(a), we present the scattered light image of the 20 μm grating, i.e., 10 μm square pits (18 nm deep) separated from each other by a 10 μm wide step. In Fig. 3(b) an AFM image of the same area is presented; the darker regions are deeper than the lighter ones by 18 nm. This scan was taken in the contact mode using an Ultralever* tip from Park Scientific.* All of the AFM measurements in this work were done in air with a Park Scientific Autoprobe LS system.[14]*

If we overlay the two pictures in Fig. 3, we can clearly interpret the SSM image. The laser light is scattered mainly from the edges of the square pits in the VLSI pattern in Fig. 3(b). If we assume the same reflectivity from the bottom of the pits as from top of the pattern and that the profile of the beam is Gaussian, then the contrast obtained from the square features with the darker dots in their centers [Fig. 3(a)] corresponds to a beam diameter [full width at half-maximum intensity (FWHM) of the intensity profile] of 8–10 μm. The beam radius at the focal point of the microfocusing lens can also be estimated[9,15] from

$$r=A\lambda f/(\pi r_0), \tag{2}$$

where λ (632.8 nm) is the laser light wavelength, f (110 mm) is the focal length of the microfocusing lens, r_0 (7.5 mm) is the radius of the expanded beam, and A is a constant correc-

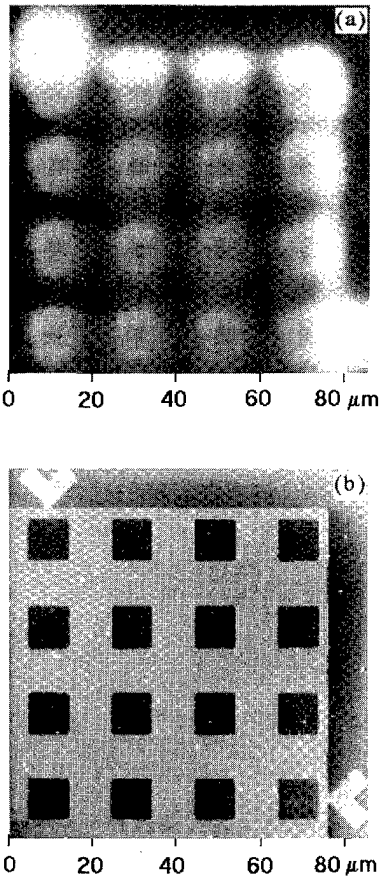


FIG. 3. Images of a part of a 20 μm grating (VLSI standards STM180A*). (a) The SSM scattered-light image (scan size is 86×86 μm^2). The brighter areas are from light scattered by the edges of the VLSI pattern. (b) An AFM image of the same area as taken with the SSM. The dark areas are deeper than the brighter ones. The small white spots are from particulates which remain on the surface after cleaning.

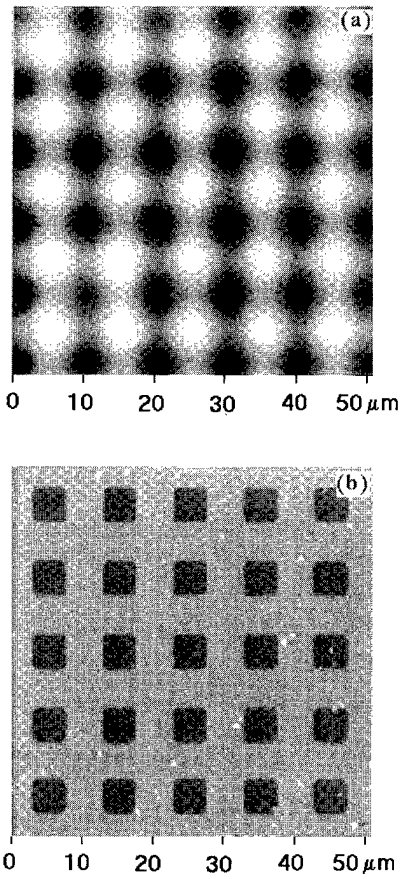


FIG. 4. Images of a part of a 10 μm grating (VLSI standards STM180A*). (a) The SSM scattered-light image (scan size is 51×51 μm^2). The brighter areas are from light scattered by the edges of the VLSI pattern. (b) An AFM image of the same area as taken with the SSM. The dark areas are deeper than the brighter ones. The small white spots are from particulates.

tion factor ($\sim$1.00) balancing spherical aberration and diffraction of the lens. Equation (2) gives a beam spot diameter $2r=6-8$ μm, which is in reasonable agreement with our estimates from Fig. 3(a). The brighter spots on the edge of the grating in Fig. 3(a) are due to the close proximity ($\sim$2 μm) of two or more scattering features (edges) that contribute additional scattered light to the edge regions of the grating.

The small bright spots in the AFM image in Fig. 3(b) are due to particulates on the sample surface. Many of these particulates, however, are not detected by the SSM; they are not evident in Fig. 3(a). One possible explanation as to why these particles are not strong optical scatterers may be related to their optical constants—they may be essentially transparent. Both the SSM and the AFM now operate in an air environment containing many dust particles. In the former case, the sample is mounted in a vertical plane, but in the latter case it is horizontal and is more prone to dust contamination. It is also possible that the sample was contaminated with dust particles during the sample transfer from the SSM to the AFM and also during the actual AFM measurements.

Figure 4(a) is a scattered light image of the 10 μm grating, i.e., 5 μm square pits separated from each other by 5 μm terraces that are 18 nm high. For comparison and direct interpretation, an AFM image of the same area is displayed in Fig. 4(b). The bright areas in Fig. 4(a) come from the superposed scattering of the four edges of the square pits, i.e., we cannot resolve two strongly scattering features (the edges of the pits) 5 μm apart or closer.

From the SSM measurements of the VLSI standards* sample and the above comments, we conclude that the present lateral resolution of our instrument is between 5 and 10 μm and is mainly due to the imaging properties (long focal length, small numerical aperture) of our microfocusing lens.

B. Test of surface roughness sensitivity

It is possible to convert the scattered light intensity data into rms roughness by using the familiar relationship for a perpendicular incident beam

$$R_q=\lambda/(4\pi)\sqrt{(I_s/I_0)}, \tag{3}$$

as given by Eq. (1).[9,10]

From Eq. (3) and measurements of the noise level of the SSM data acquisition system, we estimate the lower limit of detectable rms roughness to be $\sim$0.1 nm.

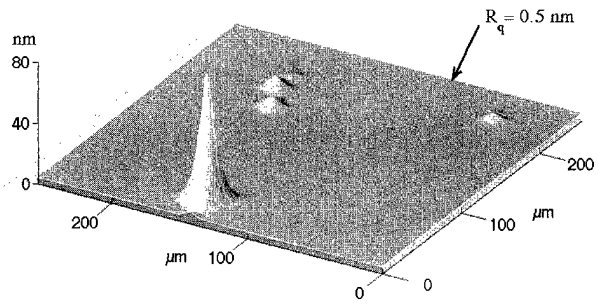


FIG. 5. An rms roughness map of the Si(100) surface in a three-dimensional representation (scan size is 256×256 μm^2). This image was constructed from the ratio of the scattered light to the specular light, both taken simultaneously by the SSM; the value of R_q for each pixel was calculated using Eq. (3). The calculated rms roughness of the Si surface is 0.5 nm. The five peaks are due to surface particles.

We tested the surface roughness sensitivity of our SSM experimentally with a very smooth Si(100) sample. The rms roughness (R_q) of this sample as measured by our AFM was 0.2–0.3 nm. This rms roughness was obtained by scanning over an area of the surface which was free of particulates.

In Fig. 5, we present a rms roughness map of the Si(100) surface created from the ratio of the scattered off-specular light intensity and the corresponding specular intensity using Eq. (3); the size of the measured area is 256×256 μm^2. The R_q value for the dark regions in Fig. 5 is 0.5 nm. The five bright spots most probably come from particulates sticking to the silicon surface. The assumptions under which Eq. (3) was derived[8–10] are not applicable to the brighter areas so in this case the R_q values cannot be interpreted in a similar way to the case of the dark area.

We remeasured the same area as in Fig. 5 with the scattered light signal amplified 50 times and obtained an identical image. The R_q value, however, was found to be 0.8 nm for the dark area.

Both R_q values for the Si(100) specimen obtained by the SSM are larger than those obtained by the AFM. There are several possible explanations for this difference.

(1) These two instruments, basically, measure different characteristics associated with surface roughness. For the AFM, the tip radius and geometry greatly affect the range of surface-wavelength features that can be accurately determined. Surface wavelengths less than four tip radii are not accurately traced and thus affect the AFM "rms roughness" results. Similar restrictions also apply to the SSM where the laser light wavelength λ affects the roughness sensitivity for surface wavelengths shorter than λ. In essence, these two techniques differently sample the roughness features of a surface (i.e., the bandwidth of the two techniques are different) and thus there is no a priori reason to expect that the roughness data will be the same.

(2) A small fraction of the specular beam (after reflection from the sample) may be scattered from the interfaces of the optical elements, primarily from the first interface of the microfocusing lens (i.e., back towards the sample and PD1), and thereby could increase the signal collected by PD1. Since the specular light intensity is typically three orders of magnitude greater than that of the scattered light, even a very small portion of the specular beam reflected back towards the sample could account for the above difference in R_q values. This effect might be reduced by using antireflection coatings on the microfocusing lens.

(3) PD2 in Fig. 2 may not provide an accurate value of I_0 since the specular light intensity is attenuated by each optical interface on its path away from the sample. We have measured this attenuation and obtained an attenuation coefficient which was used to correct the intensity measured by PD2 in order to obtain I_0. The above procedure was complicated by the fact that the laser beam is polarized and thus the transmission characteristics of the optical components depend on the polarization state. Since different specimens, or different areas of one specimen, can depolarize the incoming light differently, the attenuation coefficient for I_0 may be specimen dependent. This dependence of the attenuation factor on specimen characteristics may account for a part of the difference between R_q values obtained with the SSM and the AFM. A more accurate technique for determining I_0 is under development.

(4) There may have been submicrometer low-scattering particulates (too close to each other to be resolved with our laser beam) which can increase the rms value over the scanned area.

From the above measurements, we conclude that the sensitivity of the SSM for quantitative measurements of rms surface roughness is in the Angstrom range. Additional measurements on ultra smooth silicon surfaces are in progress and should provide us with a better estimate of the rms roughness sensitivity.

C. Defects in a Ge thin film evaporated on SiO_2/Si

A germanium thin film evaporated onto the natural silicon dioxide layer on silicon was measured by both the SSM and the AFM. Due to the difference in lattice parameters, the thin Ge films grown on SiO_2/Si were expected to have many structural defects that could be seen with the SSM. The same areas of these samples were measured with the SSM and the AFM in order to compare directly the two techniques.

We first looked at a 24.8 nm germanium thin film evaporated on SiO_2/Si. For direct comparison with the AFM, a sample navigation procedure was devised:

(1) A mark was made on the specimen of ~1 mm length with the tip of a scalpel.
(2) We then examined the specimen with the auxiliary optical microscope of the AFM to find a unique feature of ~100 μm size—a scratch, microcrack, dot pattern, etc.—and measured its coordinates with respect to the manually made mark.
(3) This unique feature was imaged with the telemicroscope of the SSM and SSM scans were made over an area close to this feature. Figure 6(a) is a SSM scan over an

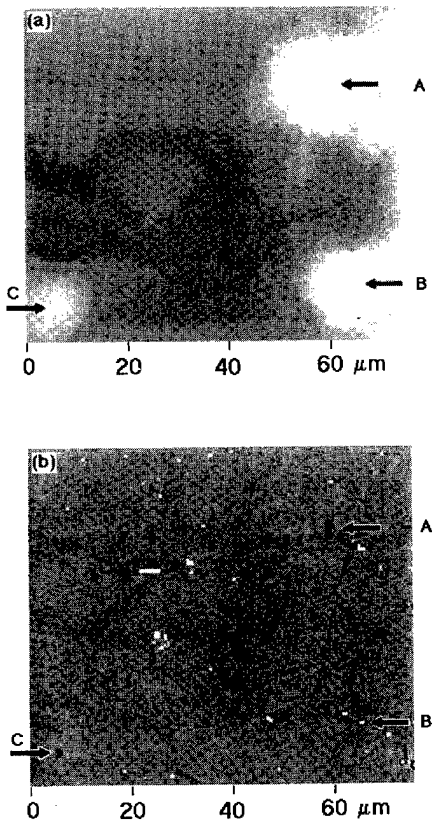


FIG. 6. Images of the surface of a 24.8 nm Ge thin film evaporated onto a SiO_2/Si substrate. The arrows, labeled by capital letters, point to defects which give a strong scattered-light signal. (a) The SSM scattered-light image (scan size is 75×65 μm^2). The brighter areas are from light scattered by defects in the Ge thin film. (b) An AFM image of the same area taken after the SSM measurements.

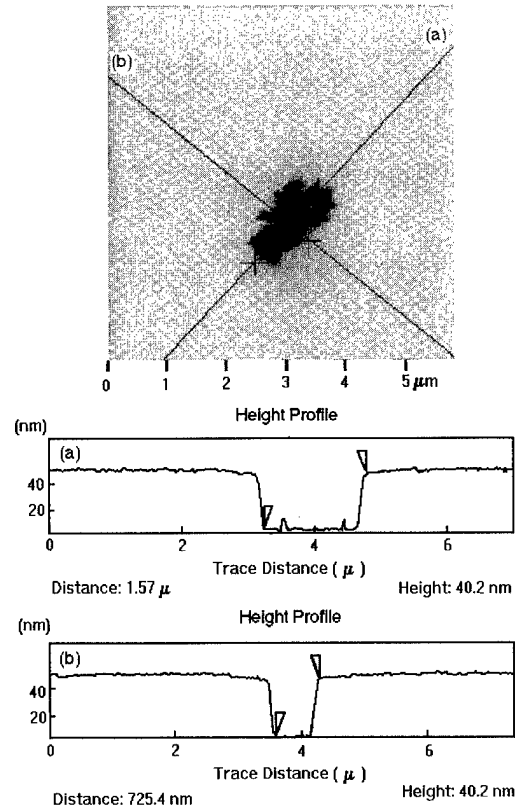


FIG. 7. Small area scan over the (A) defect in Fig. 6(b). Scans (a) and (b) are the results of line measurements of defect (A).

area of 75×65 μm^2 close to a "natural" scratch—a unique feature — on the specimen. A small fraction of the scattered light signal from this scratch is on the right side of Fig. 6(a); the actual scratch is about 10–20 μm further to the right. For comparison with the AFM, we choose three spots marked by arrows and labeled by capital letters in Fig. 6(a) and measured their positions with respect to the natural scratch.

(4) We were then able to precisely locate the region of interest in the AFM. Measurements were made in both contact and noncontact modes; reasonable agreement was found between the two modes so that only the contact mode AFM images will be presented here.

Figure 6(b) is an AFM image of the identical area to that shown in the SSM measurement of Fig. 6(a). There are more features — bright spots and dark ones than in Fig. 6(a). Since we located the position of the bright spot centers with respect to the natural scratch in Fig. 6(a) (with a precision of ~1 μm), we were able to correlate unambiguously the morphological defects in Fig. 6(b) with the scattering centers in Fig. 6(a). The features in Fig. 6(b) corresponding to the bright spots in Fig. 6(a) are marked identically in both images. The A-defect and the C-defect—the dark ones—are clearly identifiable. The B-defect—the bright one—looks like a feature on top of the surface as opposed to A and C. To get a more detailed picture of the scattering centers, we took small area scans over them. As an example, we present a picture of the A defect and the corresponding line scans in Fig. 7. This defect is a pit ~40 nm deep and ~1.5 μm diam. The C defect is a pit of the same depth but of ~0.5 μm diam. The B defect is a pit of ~1 μm diam, but we could not measure its depth since there was a particle sticking to the edge and covering ~75% of its opening. As a result, this defect looks like a bright spot in Fig. 6(b), and in this case the intense light scattering is from a feature that is underneath the particle, i.e., a pit or other subsurface defect. It seems clear that the three features labeled in Fig. 6(a) are due to submicrometer pits at the surface; the features responsible for the bright spots seen by the AFM [Fig. 6(b)] but not by the SSM [Fig. 6(a)], however, do not seem to be strong optical scatterers, probably because they are nearly transparent.

The SSM measurements of Ge/SiO_2/Si samples indicate that this technique is very sensitive to surface defects such as pits and cracks and that may, therefore, be very useful for solving technological problems associated with thin film growth.

IV. CONCLUSION

An optical scanning instrument, the SSM, has been developed specifically for imaging lateral microroughness variations at surfaces and interfaces. The capabilities of this instrument were evaluated using three different test materials:

TABLE I. Scanning scattering microscope performance parameters.

Light source	He–Ne laser, single-mode, intensity stabilized, power 3 mW, λ=632.8 nm
Laser beam stability (intensity)	~0.1%
Laser beam angle of incidence (from the surface normal)	0°
Focused laser beam spot size (FWHM)	$\leq$10 μm
Lateral resolution	$\leq$10 μm (limited by beam spot size)
Surface roughness sensitivity (minimal detectable roughness) rms roughness range (estimated)	in the range of 0.1–0.8 nm
Data acquisition speed	0.1–30 nm
Scanning linearity in x and y	256×256 pixels/h
Working distance	$\leq$1.0 μm
Bandwidth (estimated from the geometry of the SSM)	centimeters
	0.095–1.6 μm^{-1}
Surface wavelength range (estimated)	(0.06/λ–1/λ)
	0.6 μm–10.5 μm

a VLSI standards* patterned sample, a very smooth Si(100) sample, and a deposited Ge thin film containing various surface defects. Results obtained with these materials demonstrate the very high sensitivity of this technique for evaluating changes in microroughness at the subnanometer level. This sensitivity is especially useful since it can be realized in an instrument that operates at a large working distance (cm) from the sample surface. A summary of the performance parameters of the SSM is given in Table I. The SSM has a unique capability for noncontact measurements of variations in surface microroughness; in its current configuration the spatial resolution is about 8 μm and the rms roughness sensitivity is approximately a few tenths of a nanometer.

[1]T. V. Vorburger and E. C. Teague, Precis. Eng. 61 (1981).
[2]P. O. Hahn, M. Grundner, A. Schnegg, and H. Jacob, Appl. Surf. Sci. **39**, 436 (1989).
[3]H. E. Bennett and J. O. Porteus, J. Opt. Soc. Am. **51**, 123 (1961).
[4]J. Fine and D. Marton, U.S. Patent No. 4,954,722 (1990).
[5]D. Marton and J. Fine, *Microphysics of Surfaces: Beam Induced Processes* (Optical Society of America, Washington, DC, 1991), Vol. 3, p. 146.
[6]R&D 100 Award for the development of the Scanning Scattering Microscope, R&D Magazine (1990).
[7]D. Marton and J. Fine, Thin Solid Films **185**, 79 (1990).
[8]H. Davies, Proc. Inst. Elect. Eng. **101**, 209 (1954).
[9]P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (Pergamon, New York, 1963).
[10]J. C. Stover, *Optical Scattering, Measurement and Analysis* (SPIE, Washington, DC, 1995).
[11]F1048-87, Annual Book of ASTM standards, 06.01, 992 (1995).
[12]**SUPERSCOPE II Instruction Manual* (GWI Instruments Inc., Somerville, MA, 1994).
[13]*STM 180A Surface Topography Standard Application Note, VLSI Standards Inc., Mountain View, CA, 1992.
[14]**Autoprobe LS Manual* (Park Scientific Instruments, Sunnyvale, CA, 1994).
[15]**Optics Guide 5* (Melles Griot, Irvine, CA, 1990), pp. 8–18.

*In this article certain commercial equipment, instruments or materials are identified (by an asterisk) in order to specify experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standard and Technology, nor is it intended to imply that the materials or equipment identified are necessary the best available for the purpose.

Scanning tunneling microscope study of defect structures on As-terminated Si(001) surfaces

M. D. Jackson,[a] F. M. Leibsle, R. J. Cole, D. A. C. Gregory, D. A. Woolf,[b] and P. Weightman
I.R.C. in Surface Science and Department of Physics, Liverpool University, P.O. Box 147, Liverpool, United Kingdom L69 3BX

(Received 24 January 1996; accepted 13 May 1996)

As-terminated Si surfaces are model semiconductor interface systems with a wide range of technological applications. We have studied As-terminated Si (001) surfaces with scanning tunneling microscopy which reveals that it is possible to produce a well ordered surface. Some characteristic types of defects do occur on this surface, however, which are found to be of particular interest. These include antiphase domain boundaries, step edges, and long, atomically straight trenches running perpendicular to the dimer rows across the surface. The nature and origin of these defects are discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

Arsenic termination of the Si(001)2×1 surface forms a well-ordered passivating layer. With increasing miniaturization of semiconductor devices, the depletion zones associated with the surface become a significant problem. A passivating layer such as As has the potential to remove such a depletion region and enable further miniaturization. It is also desirable to integrate the optical properties of GaAs with the industrial capabilities of Si growth by growing GaAs structures on Si substrates. Device structures including bipolar transistors,[1] solar cells,[2] and heterojunction dipole lasers[3] have been grown successfully, but unfortunately a high density of dislocation defects, which may hinder device operation, are found to occur in the GaAs and attributed to originate from strain occurring from the lattice mismatch.[1–3] In these studies, it was found necessary to grow a large buffer layer to reduce the strain from the lattice mismatch. These buffer layers can constitute over 50% of the thickness of the device and with the drive for smaller devices; the origins of this strain is a matter of considerable interest. Growth of GaAs is usually performed on a Si surface after formation of a monolayer of As,[4] and it is therefore crucial to the understanding of the heteroepitaxial growth of GaAs on Si to understand the structure of the As terminated Si surface, paying particular attention to the defect structure on this surface which may seed defects in crystal growth on top.

II. As TERMINATED Si(001) SURFACE

The (001) surface of Si forms a (2×1) reconstruction by producing Si dimers in the [1$\bar{1}$0] direction which form rows in the [110] direction. Initial scanning tunnel microscopy (STM) studies suggested symmetric dimers[5,6] with the dimers arranged in rows which are observed to be of the order of 50 Å long. The surface has quite a high level of defects and buckled dimers were observed with the STM near to vacancy type defects on the surface[5,6] which gave rise to areas of local $c(4\times2)$ and $p(2\times2)$ symmetry. Later, low-temperature STM studies supported asymmetric dimers[7] while photoemission studies have supported both symmetric, covalent dimers[8] and asymmetric, ionic dimers.[9] Much work has been done since to answer the question of whether the dimers are symmetric or asymmetric.

When the surface is terminated with arsenic, the As atoms form symmetric As dimers on the surface as demonstrated by angle resolved photoemission studies,[10,11] grazing incidence x-ray diffraction investigations[12] and core-level spectroscopy which indicates one site, i.e., symmetric As dimers.[13] The reconstruction is shown in Fig. 1 and because the As atoms are fully coordinated, they create a stable, passivating layer.[10,11,14] The lowered surface energy of arsenic termination of Si, therefore, inhibits interdiffusion and produces a well-ordered, chemically inert, model surface.

An As-terminated surface can be produced by molecular beam epitaxy (MBE) growth of arsenic on MBE grown Si(001)(2×1). For ~monolayer (ML) coverages, STM studies have reported very well-ordered rows of symmetric As dimers with few defects.[15,16] The vicinal surface has also been observed with STM.[4,15–17] These authors[15,17] report well-ordered rows of As dimers perpendicular to step edges that are *not smooth* for samples grown in conditions similar to ours. Though it is possible to produce surfaces with the dimer rows parallel to step edges by using specific growth conditions.[17] Total energy calculations of GaAs growth on this surface[18] indicate that growth of one or two epitaxial layers of GaAs on ideal, atomically flat Si(001) would be hindered by rehybridisation of Ga and As in the first layer, indeed high resolution transition electron microscope (HRTEM) and STM observations indicate that GaAs growth occurs on the Si(001):As surface by 3*d* nucleation and island growth.[19] The fact that growth is observed to be easier on the vicinal surface suggests that step edges play an important role in growth, and this, in turn, suggests that defects on the flat surface might be of importance in the nucleation and growth process.

[a]Electronic mail: phineas@if.ssci.liu.ac.uk
[b]Present addresss: Department of Physics, University College of Wales, Cardiff, United Kingdom CF2 3YB.

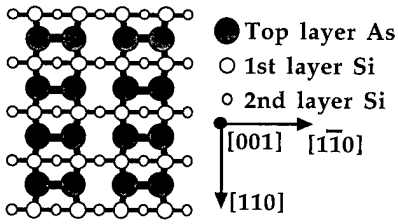


FIG. 1. Ideal reconstructed surface of Si(001)2×1:As the arsenic dimerizes and forms rows.

III. EXPERIMENT

The details of the sample growth have been specified elsewhere.[20] The *n*-type (0.006–0.015 Ω/cm, Sb doped) Si(001) substrates were thoroughly cleaned at a 850 °C before being allowed to cool under an As_4 flux to 120 °C. At this point, reflection high energy electron diffraction (RHEED) observations indicated a (2×1) pattern characteristic of the As terminated Si(001) surface. At a temperature of 120 °C the As_4 flux was shuttered and the sample was allowed to cool. Arsenic was then deposited on the sample to cap it so that it might be transported in air between laboratories. Post-growth RHEED observations confirmed the As overlayer to be of an amorphous nature, and Auger analysis later confirmed the cap to be greater than 280 Å thick.[20]

The STM measurements were carried out in an Omicron STM instrument at UHV, operating at a base pressure of $<10^{10}$ mbar. The system is fitted with a preparation chamber and a fast entry port for rapid introduction of samples. The main chamber is equipped with a retractable four-grid rear-view low-energy electron diffraction (LEED) optics. The STM is designed for room-temperature measurements with the tungsten tip held at ground potential and the sample biased. Imaging was carried out in constant current mode at positive and negative bias (tunneling into empty/out of filled states in the As atoms, respectively).

Following introduction to the STM apparatus, the sample was prepared by thermally desorbing the protective As cap by heating to 400 °C for 60 min in ultrahigh vacuum (UHV). A very sharp 2×1 LEED pattern was observed; this process is known to leave a single ML of As terminating the surface.[11] The sample was then placed in the STM chamber and imaging proceeded.

IV. OBSERVATIONS AND DISCUSSION

Figure 2 shows a large 1000×700 Å area of the sample surface and shows rows of symmetric As dimers that are often atomically straight for over a hundred ångstroms across the surface. There are islands of undesorped arsenic left on the surface but very few vacancy or interstitial defects, which represents much better ordering on this surface than the clean Si(001) (2×1) surface, in agreement with previous work.[16] The As islands cover about 20% of the surface, which is considerably more than in other work[15,17] where As was deposited on the vicinal surface using a *K*-cell *in situ*. We attribute our larger island coverage to our different preparation technique in which there is a lot of As present on the surface to start. The islands may have been desorbed more completely if a longer anneal was used. Step edges are observed to be ragged and the dimer rows are ordered perpendicular to each other on adjacent terraces indicating single steps.

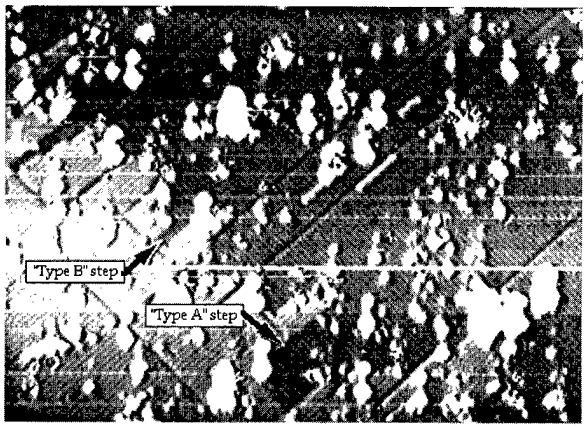


FIG. 2. Area of sample 1000×700 Å imaged with bias +2 V and tunneling current of 1 nA. Image shows long straight dimer rows, a few undesorbed arsenic islands, and *trench* defects. Most of the length of the step edge that is visible can be seen to be *type A* like our model (Fig. 3) discussed in the text with only a few areas of *type B* step.

Following Chadi's[21] classification of steps on the clean Si(001)2×1 surface as types A and B, we see, in Fig. 2, many examples of steps which are similar to type A; in that rows of arsenic dimers are perpendicular to steps up and dimer rows run parallel to the edge of steps down. We measure the steps to be 1.8 Å high on average which agrees with a single step model and find that the horizontal distance, perpendicular to a step, between the step edge and the start of a row of dimers on the lower terrace is about 11.4 Å. These dimensions agree with a simple model of a single step with missing dimers at the base of the step which we show in Fig. 3 with the appropriate experimentally measured distances indicated. In contrast, we see proportionally much fewer examples of steps that are like B-type steps on the surface and

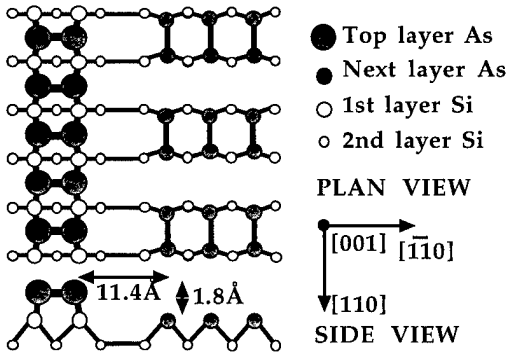


FIG. 3. Suggested model of the single steps found predominantly on this Si(001)2×1:As surface. The experimentally determined distances discussed in the text are indicated in their appropriate places.

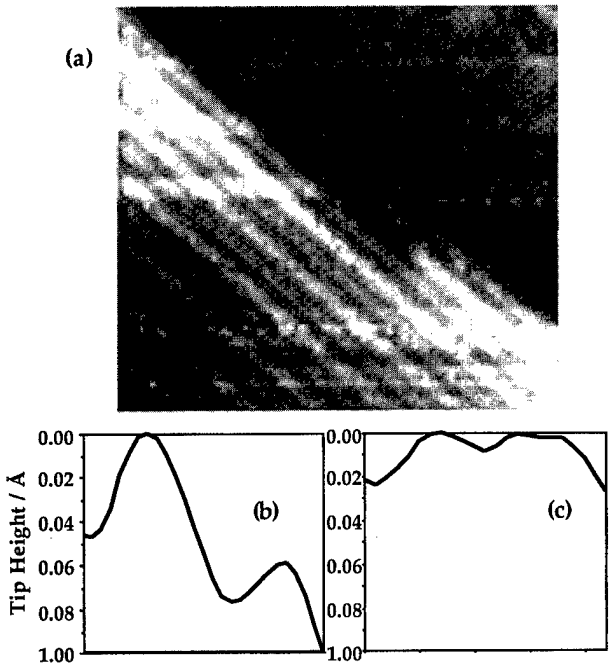


FIG. 4. (a) STM image of step edge taken with bias −2 V and tunneling current of 1 nA. Line profiles taken across the dimers on the (b) upper and (c) lower terrace, shown to the same vertical scale. The dimers appear to be asymmetric on the upper terrace but appear symmetric on the lower.

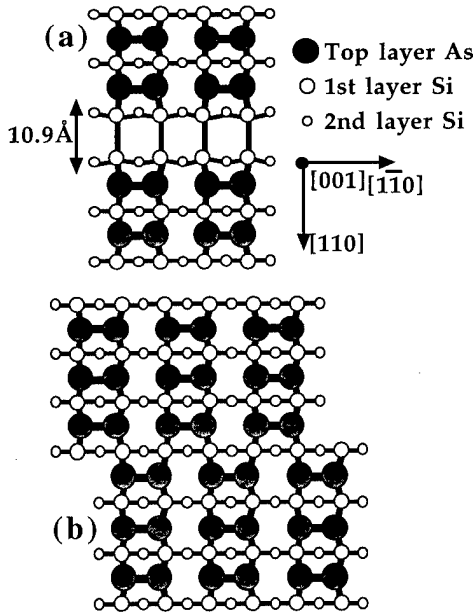


FIG. 5. (a) Ball and stick type model of an in-phase missing-dimer trench. Assuming that the Si atoms in the trench dimerize to reduce the energy of their dangling bonds. The experimentally determined width of the trench is shown. (b) Ball and stick model of an antiphase domain boundary or *scar* type defect.

find that they are often accompanied by nearby trench defects or As islands. Our results then indicate that this surface preferentially forms steps that are similar to A-type single steps on the clean Si(001)2×1 surface which is in agreement with previous work.[15,17]

We observe a low density of steps overall on the surface which leads to very large, atomically flat, areas. These terraces are observed to be generally about 1500 nm^2 in extent, with some terraces being over 2100 nm^2. The step edges allow us to make an interesting observation about the STM as a tool. Figure 4 shows a poor resolution image of a step edge, the image is so poor that the dimer rows on the lower terrace are difficult to distinguish but are seen more clearly when line profiles are taken. The line profiles show apparently asymmetric dimers on the upper terrace, on the left, while on the adjacent terrace separated by one step down; the dimers exhibit no asymmetry. This only occurred on a small set of scans, and we attribute it to an asymmetry in the tip. Although similar effects, seen in an STM study of Si(001) 2×1,[22] were attributed to asymmetric and symmetric dimers existing on the surface due to the near-degeneracy of their energies,[23] we do not believe that to be the explanation here due to the strong theoretical and experimental evidence that the As dimers are symmetric.

The defects that are most obvious on the surfaces are the atomically straight trenches which are perpendicular to the dimer rows and are extremely long, extending over 850 Å in some cases. Similar trenches have been observed before on the vicinal surface where the dimer rows are parallel to the step edges.[17] In our experiment, the absence of steps on the samples allows us to see the extraordinary length of the trenches. We observed the trenches to be about 0.6 Å deep and about 10.9 Å wide; dimensions which are consistent with a single missing dimer type defect [in Fig. 5(a)]. In Fig. 6, we show an example of a trench, on the right, along with an antiphase domain boundary on the left which is consistent with a simple model shown in Fig. 5(b). In the trench shown in Fig. 6, the dimer rows are in-phase on either side of the trench, but we also observe examples where the rows are out-of-phase on either side of the trench. Since we find both in-phase and out-of-phase trenches on the surface in equal amounts, and we see simple antiphase domain boundaries such as that in Fig. 6, we conclude that the trenches do not form to accommodate antiphase domains. Nor are these trench defects simply steps down and steps up close together since, as we have demonstrated above, the surface has a preference for the dimer rows to be aligned along the step edges which the opposite to the trench edges. Furthermore, the

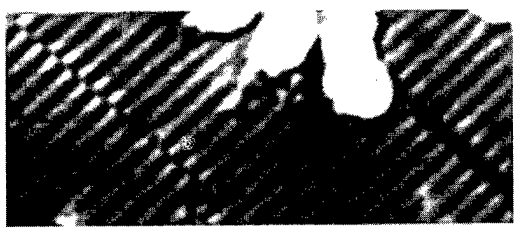


FIG. 6. STM image of area approximately 300×150 Å taken at sample bias of +2 V with tunneling current of 1 nA. Image shows two types of defect; an antiphase domain boundary on the left and an in-phase trench on the right.

trenches do not resemble the disordered dimer channels induced by Ni on Si(001)[24] by virtue of their extreme order.

Similar trench-like structures have been reported on *N*-terminated Cu(001) surfaces[25] and on other semiconductor systems. Trenches up to 200 Å long have been observed on the Ge(111):As surface,[16] trenches were reported on the Si(001)2×1:*P* surface,[26,27] well ordered trenches separated by 8 to 15 dimers have been reported on Si(001):Ge[28] and in the study of Si(001):Bi[29] it was also found that very well ordered trenches form. In that study, the samples gave LEED patterns of $(2\times n)$; examination by STM revealed trenches separated by n dimers, and n was found to vary from 5 to 12 depending on the annealing temperature.[29] We note that in our specimens, the unequal separation of the trenches seen in Fig. 2 is confirmed by the (2×1) LEED pattern. We note that the As-terminated surface only forms trenches perpendicular to the dimer rows in contrast to the *P*-terminated surface where both perpendicular and parallel trenches are seen.[27] Furthermore, on the *P*-terminated surface the trenches were often observed to continue over step edges for short distances.[27] This only seems to happen occasionally on the As-terminated surface with just a couple of examples to be seen in Fig. 2. The proposed missing dimer model for the perpendicular trenches on the Si(001)2×1:*P* surface suggested two missing dimers[27] whereas our measurements have indicated that a single row of missing dimers is most likely. Any differences between the P- and As-terminated surfaces are probably due to the different atomic sizes of phosphorus and arsenic.

In all these previous examples, the trenches were believed to be strain relief mechanisms. We therefore conclude that the trenches seen here form as a mechanism for strain relief due to the incommensurate nature of the As overlayer. We note that no such defects have been reported on studies on 4° vicinal Si(001) where the dimer rows run perpendicular to the step edges,[15,17] and attribute this to the terraces on this surface being only 40–50 Å wide allowing strain to be relieved on the terrace by relaxation over the edge of a step. In our study, the separation of the trenches varies from 10 to 35 dimers. This is larger than reported on the other systems[28,29] showing that the trench separation depends on preparation and not only the extent of lattice mismatch.

Since the trenches are only found in the $[1\bar{1}0]$ direction, this means that the structure is strained only in the [110] direction. This is physically reasonable since to dimerize the surface layer, atoms have to move towards each other to achieve a bonding distance, the energy of repulsion of the sp^3 bonds on the Si atoms being offset by the free-energy loss of the dimer atom's dangling bonds being removed. Thus, when the silicon surface has to accommodate the relatively large arsenic atom, it is quite easy for it to relax towards the preferred bond angle of 109° in the direction along the bond, i.e., the $[1\bar{1}0]$ direction (see Fig. 1). In the [110] direction, however, no such relaxation is possible which means that after a certain number of dimers have been added to a row, the surface must relieve the strain by the formation of trenches. This hypothesis also explains the missing As dimers along the bottom of the step in the model shown in Fig. 3, there being both strain and steric hindrance effects preventing arsenic adatoms from sticking at the base of the step.

V. CONCLUSION

We have shown a number of defects that occur on the otherwise well-ordered, Si(001)2×1:As surface. We suggest models of step edges, antiphase boundaries, and trenches from our STM images of these defects, and we propose that the trenches are formed as a means of strain relief due to the incommensurate nature of the arsenic overlayer.

ACKNOWLEDGMENTS

The authors would like to acknowledge Trevor Morrison and Steve Barrett for work on the STM analysis software. The project was funded by the U.K. EPSRC and the PESSI network of the HCMP of the E.U.

[1]R. Fischer, N. Chand, W. Kopp, H. Morkoc, L. P. Erickson, and R. Youngman, Appl. Phys. Lett. **47**, 397 (1985).
[2]Y. Itoh, T. Nishioka, A. Yamamoto, and M. Yamaguchi, Appl. Phys. Lett. **49**, 1614 (1986).
[3]T. H. Windhorn and G. M. Metze, Appl. Phys. Lett. **47**, 1031 (1985).
[4]D. K. Beigelsen, R. D. Bringans, J. E. Northrup, and L.-E. Swartz, Appl. Phys. Lett. **57**, 2319 (1990).
[5]R. M. Tromp, R. J. Hamers, and J. E. Demuth, Phys. Rev. Lett. **55**, 1303 (1985).
[6]R. J. Hamers, R. M. Tromp, and J. E. Demuth, Phys. Rev. B **34**, 5343 (1986).
[7]R. A. Wolkow, Phys. Rev. Lett. **68**, 2636 (1992).
[8]D.-S. Lin, T. Miller, and T.-C. Chiang, Phys. Rev. Lett. **67**, 2187 (1991).
[9]G. K. Wertheim, D. M. Riffe, J. E. Rowe, and P. H. Citrin, Phys. Rev. Lett. **67**, 120 (1991).
[10]R. I. G. Uhrberg, R. D. Bringans, R. Z. Bachrach, and J. E. Northrup, Phys. Rev. Lett. **56**, 520 (1986).
[11]R. I. G. Uhrberg, R. D. Bringans, R. Z. Bachrach, and J. E. Northrup, J. Vac. Sci. Technol. A **4**, 1259 (1986).
[12]N. Jedrecy, M. Sauvage-Simkin, R. Pinchaux, J. Massies, N. Greiser, and V. H. Etgens, Surf. Sci. **230**, 197 (1990).
[13]R. D. Bringans, M. A. Olmstead, R. I. G. Uhrberg, and R. Z. Bachrach, Phys. Rev. B **36**, 9569 (1987).
[14]T-H. Shen and C. C. Matthai, J. Phys. Condensed Matter **3**, 6169 (1991).
[15]D. K. Biegelsen, L.-E. Swartz, and R. D. Bringans, J. Vac. Sci. Technol. A **8**, 280 (1990).
[16]R. S. Becker, T. Klitsner, and J. S. Vickers, J. Microsc. **152**, 157 (1988).
[17]R. D. Bringans, D. K. Biegelsen, and L.-E. Swartz, Phys. Rev. B **44**, 3054 (1991).
[18]E. Kaxiras and J. D. Joannopoulos, Surf. Sci. **224**, 515 (1989).
[19]D. K. Beigelsen, F. A. Ponce, A. J. Smith, and Tramontana, J. Appl. Phys. **61**, 1856 (1987).
[20]J. A. Evans, A. D. Laine, P. Weightman, J. A. D. Matthew, D. A. Woolf, D. I. Westwood, and R. H. Williams, Phys. Rev. B **46**, 1513 (1992).
[21]D. J. Chadi, Phys. Rev. Lett. **59**, 1691 (1987).
[22]R. Wiesendanger, D. Burgler, G. Tarrach, H.-J. Guntherodt, I. V. Shvets, and J. M. D. Coey, Surf. Sci. **274**, 93 (1992).
[23]I. P. Batra, Phys. Rev. B **41**, 5048 (1990).
[24]H. Niehus, U. K. Kohler, M. Copel, and J. E. Demuth, J. Microsc. **152**, 735 (1988).
[25]F. M. Leibsle, S. S. Dhesi, S. D. Barrett, and A. W. Robinson, Surf. Sci. **317**, 309 (1994).
[26]Y. Wang, X. Chen, and R. J. Hamers, Phys. Rev. B **50**, 4534 (1994).
[27]L. Kipp, R. D. Bringans, D. K. Beigelsen, J. E. Northrup, A. Garcia, and L.-E. Swartz, Phys. Rev. B **52**, 5843 (1995).
[28]J. Knall and J. B. Pethica, Surf. Sci. **265**, 156–167 (1992).
[29]Ch. Park, R. Z. Bakhizin, T. Hashizume, and T. Sakurai, J. Vac. Sci. Technol. B **12**, 2049 (1994).

Simultaneous imaging of Si(111) 7×7 with atomic resolution in scanning tunneling microscopy, atomic force microscopy, and atomic force microscopy noncontact mode

Peter Güthner[a)]
Omicron Vakuumphysik GmbH D-65232, Taunusstein, Germany

(Received 25 October 1995; accepted 10 May 1996)

The reconstructed Si (111) 7×7 surface was imaged in several operation modes of the combined ultrahigh vacuum atomic force microscope/scanning tunnel microscope. By imaging single atom defects on the sample surface a clear proof of the atomic resolution in noncontact mode of the force microscope was possible. By simultaneous measurements of several interaction parameters and by the investigation of force-distance curves, it was possible to explain the origin of the interaction. © *1996 American Vacuum Society.*

I. INTRODUCTION

The 7×7 reconstruction of the Si (111) surface was analyzed by scanning tunneling microscope (STM) measurements in ultrahigh vacuum (UHV).[1] This structure often shows single atom defects. By imaging these defects the true atomic resolution of a microscope can be proven. In the STM mode, the atomic resolution on Si (111) can be achieved easily. With an atomic force microscope (AFM) it is much more difficult to get similar resolution. Most "atomic resolution images" by AFM do not show single atom defects or clearly resolved steps, so it is pseudo atomic resolution by multitip imaging. AFM images of a monoatomic cleavage step on NaF[2] showed a lateral step width of 1 nm although the regular lattice pattern could be imaged on the terraces. This step width indicates that no true atomic resolution was achieved.

First AFM measurements on Si (111) were done in contact mode. The adhesion force of the point contact between the clean Si surface and the silicon tip is up to 10^3 nN. [3] This large force destroys the surface reconstruction. The adhesion forces could be reduced by coating the cantilever with polytetrafluorethylene (PTFE). [3] With such a tip, it was possible to image the atomic periodicity and the corner holes of the Si surface, but, due to friction effects, no true atomic resolution was possible.

First measurements in AFM noncontact mode [4] on the Si (111) surface were able to resolve the 7×7 reconstruction for a very short period of time, [5] but the quality of the image was much worse than typical STM images. Also by AFM noncontact mode, it was possible to image the InP (110) surface with atomic resolution. [6]

II. EXPERIMENTAL SETUP

The Si (111) 7×7 surface can only be prepared in UHV. Therefore, a two chamber UHV system was used (pumped by turbo pump, ion getter pump, and titanium sublimation pump). The pressure in the analysis chamber was below 5×10^{-11} mbar during the experiments. For sample and tip preparation, the system was equipped with sputter ion source and sample heater for direct or resistive heating. To check the sample preparation a low-energy electron diffraction (LEED) system was available.

The experiments were made with a standard commercial instrument (OMICRON UHV AFM/STM) which is based on a design of Howald and co-workers. [7] It uses the widely acknowledged beam deflection technique. The crucial alignment of the light beam onto cantilever and position sensitive detector (PSD) has been achieved using two mirrors mounted on ball motors. [8] In this setup, the sample is scanned and the force detection is fixed. For the noncontact mode operation an oscillator circuit is used to stimulate the cantilever to vibrate at its resonance frequency which is detected by a frequency modulation detector. The principle of frequency detection was described by Albrecht and coworkers. [9] A snap-in protection measures the damping of the cantilever and retracts the tip before it is too close to the surface.

The control electronics (OMICRON Scala System) allows a choice of several operation modes. It is possible to select tunneling current, cantilever deflection, or frequency shift as the input source for the distance regulation. A topography image of the sample surface can be recorded by measuring the z position of the tip with active feedback. If a slow feedback is selected the tip cannot follow the atomic corrugation and the z position of the tip stays nearly constant. In both operation modes, several input signals such as tunneling current, cantilever deflection, and/or frequency shift can be recorded simultaneously. In spectroscopy mode, the scanning is stopped at each scan point and a spectroscopy curve is measured.

For the experiments, single crystal silicon cantilevers were used. They have a force constant in the range of 10 N/m and a resonance frequency of about 300 kHz, [10] the doping of the silicon is sufficient to use the cantilevers for STM experiments. The cantilevers have been cleaned *in situ* by sputtering with an argon ion beam.

[a)]Electronic mail: Peter_ Guethner@omihqger.ccmail.compuserve.com

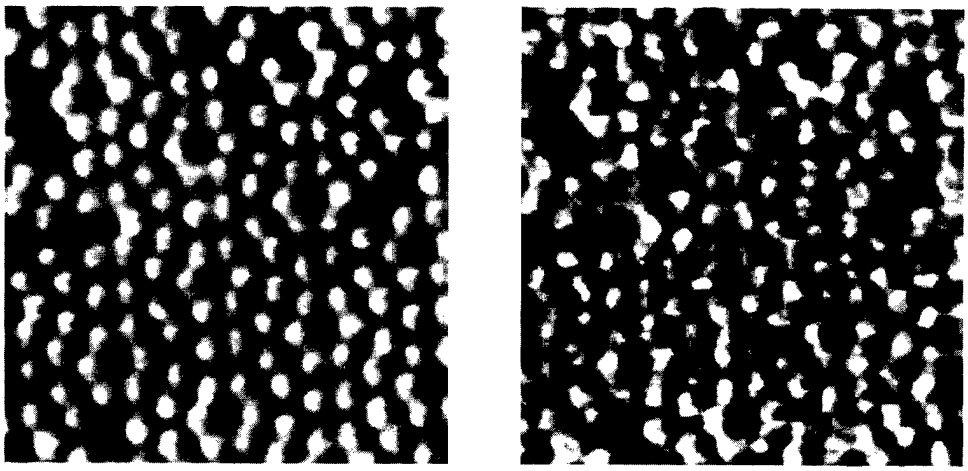


FIG. 1. Left: STM image (10×10 nm) of a Si (111) 7×7 surface measured with a conductive cantilever. Right: simultaneous measurement of the interaction force between tip and surface during the tunneling process (white=lower attractive force).

III. RESULTS

The combined AFM/STM was calibrated by STM measurements with a tungsten tip on Si (111) 7×7. The measurement was also used to proof the quality of the surface preparation. For all other experiments described in this article, conductive silicon cantilevers were used. Scanning tunneling microscopy (STM) measurements on Si (111) using these conductive cantilevers as a tunneling tip resolved the atomic structure of the 7×7 surface reconstruction. These images showed an atomic corrugation of up to 0.5 nm which is much larger than known from normal STM images with tungsten tips. This large corrugation can be explained by interaction force bending the cantilever. To proof this effect, the deflection of the cantilever was recorded simultaneously while measuring a cantilever STM image. The STM image was recorded in constant current mode (active feedback) at a tunneling current of 0.5 nA and a gap voltage of 2 V (sample positive). The cantilever was bent towards the surface which indicates that the interaction force was attractive. At the positions of the atoms the interaction force was about 0.8 nN smaller than at the position of the holes. In Fig. 1 both images are shown.

The same silicon sample was imaged later on in AFM noncontact mode. It was possible to image the atomic reconstruction with similar measurement parameters as described in the literature.[5] For the experiments, a cantilever with 10 N/m and a resonance frequency of 290 kHz was used. The feedback setpoint was adjusted to −95 Hz and the peak-to-peak oscillation was approximately 20 nm. The atomic structure can be seen over the entire image, but due to the large tip surface separation with these parameters, the lateral resolution of such images was not satisfying.

Another set of experiments was started by combining STM and AFM noncontact mode. The Si (111) surface was imaged first with a conducting cantilever in STM mode. It was possible to detect the resonance frequency of the cantilever during the tunneling process although the oscillation was not excited by the piezo. This thermal oscillation of the cantilever with an amplitude of approximately 0.03 nm was not visible in the force signal, but the high sensitivity of the FM detector allows to measure the frequency of this small vibration (see Fig. 2). The interaction force was attractive resulting in a decrease of the resonance frequency. At the position of the atoms the attractive force was stronger ($\Delta f = -145$ Hz) than at the position of the holes ($\Delta f = -12$ Hz).

Now the cantilever was excited to an oscillation of about 20 nm_{pp} at the resonance frequency. If the tip is close enough to the surface and a voltage is applied between tip and sample a current can be detected. This current can be explained by the tip touching the surface at the closest point of the oscillation. The current is increasing for closer distances which indicates that the tip is touching for a longer period of the oscillation. The expected periodic maxima in the tunneling current were not detected. Only the average current was measured due to the limited bandwidth (30 kHz) of the current amplifier. Figure 3 shows the distance dependence of the point contact current and the simultaneously measured shift of the resonance frequency of the cantilever. The data were taken during a tip approach.

A voltage of 2 V (sample positive) was applied. The current increases linearly up to 50 nA which is the saturation of the current amplifier. The resonance frequency is first decreasing due to the attractive van der Waals and electrostatic forces. However, as soon as the tip starts to touch the surface, the frequency increases again due to the repulsive van der Waals forces.

It is possible to image the Si (111) 7×7 surface reconstruction by measuring this point contact current. Furthermore the spatial display of the frequency shift leads to a high resolution image if the tunneling current feedback is virtually zero. The frequency is shifted to lower values by the attractive van der Waals force and the electrostatic force. Figure 4 shows a smaller decrease of the frequency for the position of the atoms than for the position of the corner holes. The difference in resonance frequency between atoms and holes is 20 Hz, the average frequency shift was −680 Hz. This effect

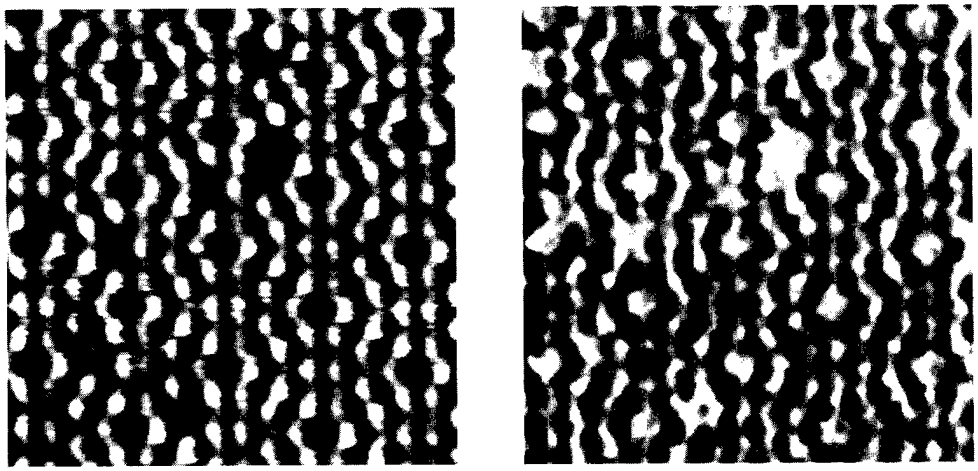


FIG. 2. Left: STM image (11×11 nm) of a Si (111) 7×7 surface measured with a conducting cantilever. Right: simultaneous measurement of the frequency shift of the thermal oscillation of the cantilever (white=lower attractive force).

can be explained by a repulsive van der Waals force between the surface atom underneath the tip and the very end of the tip. This force is added to the attractive van der Waals forces and the Coulomb force showing no or only very small local variations. The repulsive van der Waals forces indicate that the tip is very close to the surface and therefore the quality of the image is similar to STM images of the Si (111) 7×7 surface.

A high quality image of the Si (111) 7×7 reconstruction in AFM noncontact mode is possible by choosing similar parameters as for the image described above. Again, a voltage (1.9 V) was applied between tip and sample, but now the frequency shift was used as an input to the distance regulation. The setpoint was adjusted to a frequency shift of −870 Hz and the feedback was too slow to follow the atomic corrugation. An image taken with these parameters is shown in Fig. 5. At the position of the atoms ($\Delta f = -862$ Hz) the attractive force is again lower than at the position of the holes ($\Delta f = -872$ Hz).

The tip should follow the atomic structure if the feedback speed is increased. However, due to the repulsive forces at the position of the atoms this is not the case. If the feedback loop tries to keep the frequency shift constant the tip has to come closer to the surface at the position of the tip. This effect inverts the image of the atomic structure which could be seen in the measurements. Because of the inversion of the slope in the $\Delta f(z)$ curve (see Fig. 3), it was not possible to get a stable image in constant frequency shift mode. The instability of the distance regulation often leads to a snap-in of the cantilever which then destroys the tip.

IV. CONCLUSION

For the first time, it was possible to image the Si (111) 7×7 surface reconstruction simultaneously in several operation modes of the combined AFM/STM, such as STM com-

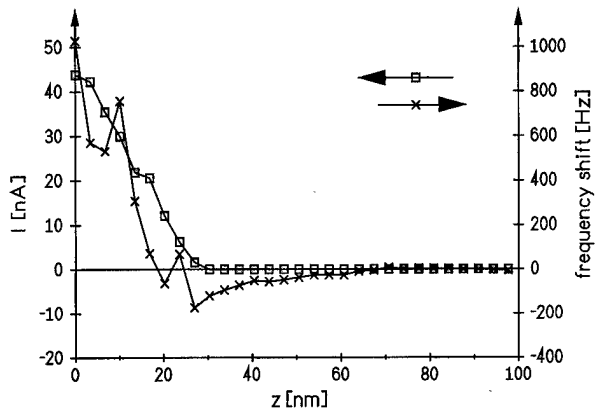


FIG. 3. Point contact current and frequency shift as a function of tip height (x=frequency shift, square=point contact current).

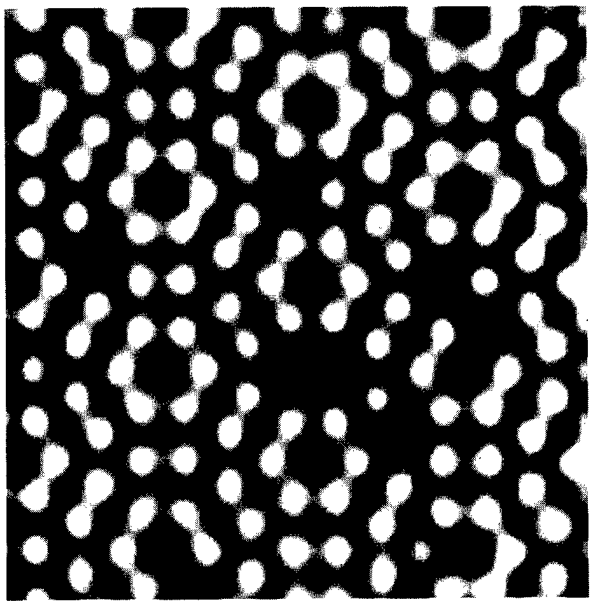


FIG. 4. AFM noncontact mode image (9×9 nm) of Si (111) 7×7 with slow STM feedback (white=lower attractive force).

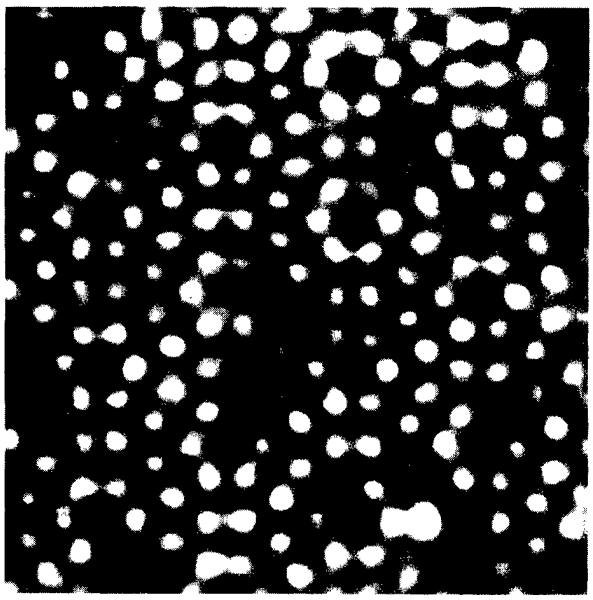


FIG. 5. AFM noncontact mode image (10×10 nm) of Si (111) 7×7 with slow feedback on frequency shift (white=lower attractive force).

bined with measurement of cantilever deflection or resonance frequency. The images demonstrate that the atomic force microscope is capable of achieving similar resolution as a STM.

The quality of the images were very reproducible after finding the right parameters. It was possible to image the Si (111) surface for several times and with several tips with similar quality of the images. The tip preparation is important to get these results. Even more crucial is the small gap distance between tip and sample surface which can be controlled perfectly by the tunneling current. It was already possible to repeat the experiment with the same tip preparation and the same imaging parameters in another laboratory. [11]

The results can be explained qualitatively by superimposing van der Waals forces and electrostatic forces. The AFM noncontact mode (constant frequency shift imaging) is only possible if the resonance frequency decreases with decreasing gap distance. For typical interaction potentials this is only true for larger distances; for smaller distances the resonance frequency increases again due to repulsive forces. If an attractive electrostatic force is superimposed, the turnover point is moved to smaller distances, therefore, noncontact mode images are possible for smaller gap width than without an applied voltage which leads to better lateral resolution.

The electrostatic forces have a long interaction range so that this interaction is not modulated on an atomic scale. The local variations of the resonance frequency are only caused by van der Waals forces. By varying the tip sample distance, it is possible to get a higher frequency on top of the atoms if the tip is very close or to get a lower frequency on top of the atoms if the tip is further away. A more quantitative explanation of the results can be made by measuring spectroscopic data with spatial resolution such as $I(z)$, $I(V)$, $F(z)$, $F(V)$, $Df(z)$, or $Df(V)$ curves. These experiments are planned for the future. To understand the interaction in more detail, a theoretical simulation of the measurements would be very helpful.

ACKNOWLEDGMENTS

The technical help of R. Weiler with the design of the FM detector electronics is gratefully acknowledged. For many stimulating discussions, the author thanks M. Sander and T. Berghaus.

[1]G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **50**, 120 (1983).
[2]L. Howald, H. Haefke, R. Lüthi, E. Meyer, G. Gerth, H. Rudin, and H. J. Güntherodt, Phys. Rev. B **49**, 5651 (1994).
[3]L. Howald, R. Lüthi, E. Meyer, P. Güthner, and H. J. Güntherodt, Z. Phys. B **93**, 267 (1994).
[4]Y. Martin. C. C. Williams, and H. K. Wickramasinghe, J. Appl. Phys. **61**, 4723 (1987).
[5]F. J. Giessibl, Science **267**, 68 (1995).
[6]H. Ueyama, M. Ohta, Y. Sugawara, and S. Morita, Jpn. J. Appl. Phys. **34**, 1086 (1995).
[7]L. Howald, E. Meyer, R. Lüthi, H. Haefke, R. Overney, H. Rudin, and H. J. Güntherodt, Appl. Phys. Lett. **63**, 117 (1993).
[8]L. Howald, H. Rudin, and H. J. Güntherodt, Rev. Sci. Instrum. **63**, 3909 (1992).
[9]T. R. Albrecht, P. Grütter, D. Horne, and D. Rugar, J. Appl. Phys. **69**, 668 (1991).
[10]Nanosensor GmbH, Wetzlar, Germany.
[11]R. Lüthi and E. Meyer (private communication).

Structure imaging by atomic force microscopy and transmission electron microscopy of different light emitting species of porous silicon

R. Massami Sassaki
Instituto de Quimica, Universidade Estadual de Campinas, 13083-970 Campinas, São Paulo, Brazil

R. A. Douglas,[a] M. U. Kleinke, and O. Teschke
Instituto de Fisica Gleb Wataghin, Departamento de Fisica Aplicada, Universidade Estadual de Campinas, 13083-970 Campinas, São Paulo, Brazil

(Received 6 February 1996; accepted 1 May 1996)

The complex pattern of the nanowire skeletons of different light emitting porous silicon structures is investigated by transmission electron microscopy (TEM) and atomic force microscopy (AFM). Diffraction lines and dark field images are used to identify and determine the crystallite specimen long range order. TEM images give the size and particle orientation, and AFM images show a three-dimensional pattern formed by an interconnecting skeleton of particles. Near infrared photoluminescent porous silicon (0.006 Ω cm) structures show a skeleton of nanosized silicon aggregates which form domains of spatially oriented crystallites. For red photoluminescent samples (4.9 Ω cm) the electron diffraction spots are discontinuously split into tiny intensity maxima. The diameter of the wire structure forming porous silicon as measured by TEM allows us to estimate the distortion of the AFM images due to the finite size of the tip radius. A critical angle $\alpha_0 = 2 \arctan[K/(1-K)]^{1/2}$, where K is the ratio of the height of the structure to the tip diameter was defined and it was shown that for structure walls steeper than α_0 the distortion may be substantial.
© *1996 American Vacuum Society.*

I. INTRODUCTION

A debate exists on the precise processes responsible for porous silicon (PS) photoluminescence.[1–4] A better understanding of the complex structure of skeletons consisting of nanosized particles in different light emitting PS species is required. Increased control over formation morphologies is also of fundamental importance for many applications proposed for these materials. The most adequate techniques to investigate the atomic structure of metal and semiconductor surfaces are provided by scanning tunneling microscopy (STM) and TEM.[5,6] STM measurements on films thicker than 500 nm revealed instabilities in the tunnel current so that surface topography could not be definitely determined.[7] In a recent paper, we reported TEM results which reveal the nanostructure of luminescent PS.[8–11] PS surfaces have been previously investigated by using AFM and STM.[12–14] Theories of image formation have been reported, the most prevalent being the tip sample convolution which is based on purely geometrical considerations assuming a strictly stiff cantilever, with only one degree of freedom in the vertical direction.[15] The essence of the theory is that in each observed image there exists some information about the tip in addition to the information about the sample.

In the experiments described below, we exploit the capabilities of AFM combined with TEM to examine the nature of the structure of PS layers. Moreover, in order to gain a better understanding of the image formation mechanism in AFM, a comparison of AFM and TEM images is also performed. To this end, we report AFM studies of the three-dimensional pattern formed by the PS layer, while TEM studies give the size and orientation of the nanoparticles.

[a]Electronic mail: ross@ifi.unicamp.br

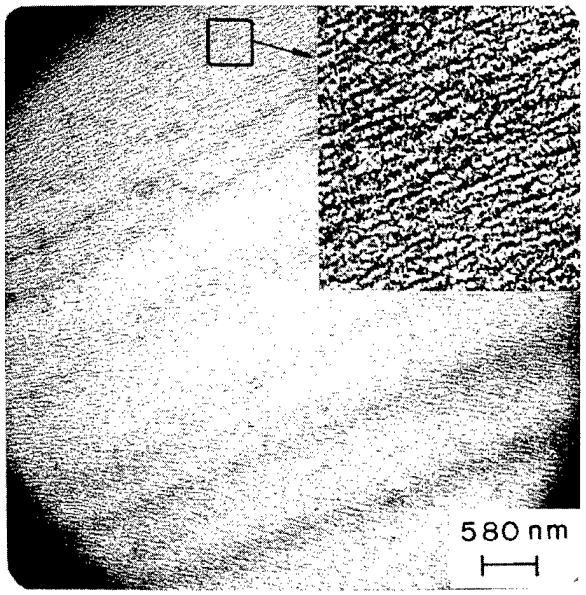
II. EXPERIMENT

Single-crystal, polished ⟨100⟩ and ⟨111⟩ wafers of doped (p- and n-type) silicon of 0.006, 0.4, and 4.9 Ω cm resistivities were cut into rectangles with areas of approximately 1.0 cm^2. The etching bath was a 50:50 (by volume) solution of 48% HF (Merck) and 95% ethanol (Merck). The samples were etched using a model 273A potentiostat/galvanostat (Princeton Applied Research). After anodization the silicon slab edge was inserted into the TEM for study, as previously described.[8] TEM examinations were carried out in a Zeiss CEM902 microscope using 80 keV electrons and equipped with an electron energy loss spectrometer (EELS) and an image intensifying camera.

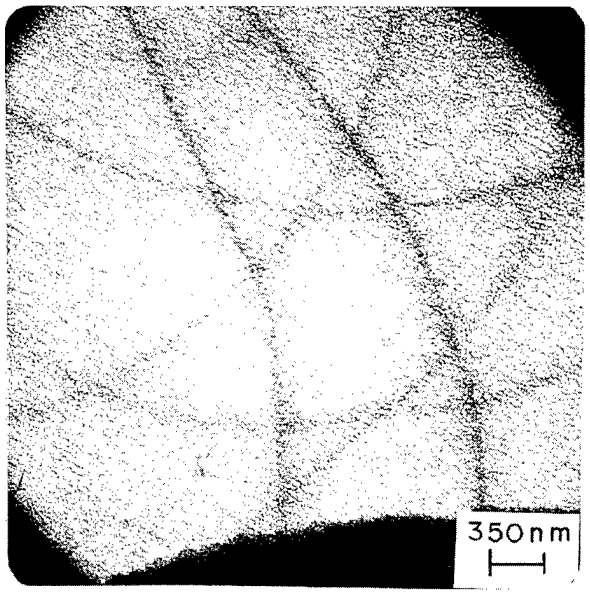
The silicon samples for AFM observations were removed from the electrolyte without removing the liquid adhered to its surface, placed in a glass container, and immersed in liquid nitrogen. The frozen samples were then placed under vacuum (10^{-3} Torr), and their glass containers were immersed in liquid nitrogen. Solution sublimation from the samples took about 2 h. AFM observations were carried out in a TopoMetrix TMX 2000 AFM instrument, using the variable force mode in the repulsive force region. Images were obtained in air on freshly produced surfaces. The room temperature photoluminescence system has been described previously.[9]

III. RESULTS

Images obtained by TEM from a 0.009 Ω cm, n-type, ⟨100⟩ oriented silicon sample, anodized at current density of



(a)



(b)

FIG. 1. Transmission electron micrographs of the porous silicon structure formed on ⟨100⟩ oriented silicon, n-type, 0.009 Ω cm sample. (a) Inset shows a region with 4× magnification. X indicates a 20-nm-diam particle. Diffraction contrast lines form a pattern of almost straight lines. (b) Curved diffraction contrast lines and microscopic structure.

100 mA/cm^2, are displayed in Figs. 1(a) and 1(b). This sample shows a near infrared photoluminescence peak at 1.45 eV. Two distinct types of structures can be seen in Fig. 1(a): a pattern of columnar dark spots ~4 nm in diameter over the entire PS surface and an array of holes (whitish areas) ~25 nm across, forming a structure of pores perpendicular to the plane of the silicon slab and exhibiting wall dimensions of the order of ~30 nm. In addition, Fig. 1(a) shows diffraction contrast lines on the surface of the PS films. These contrast lines arise from the etching action when the crystallites forming the crystal are slightly tilted. It is also

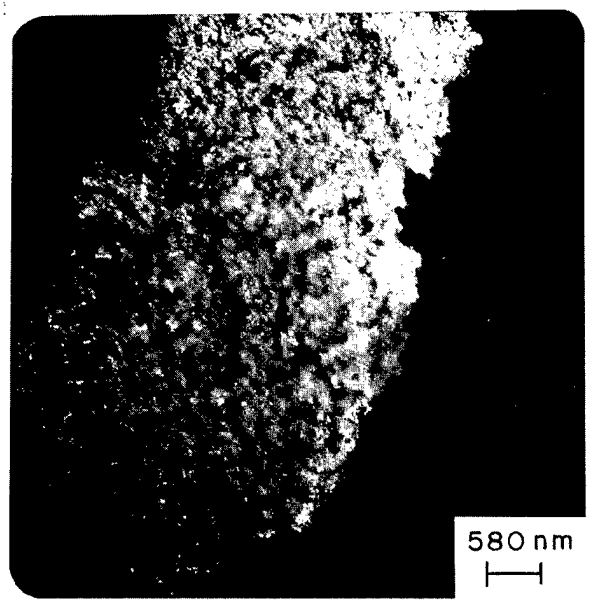
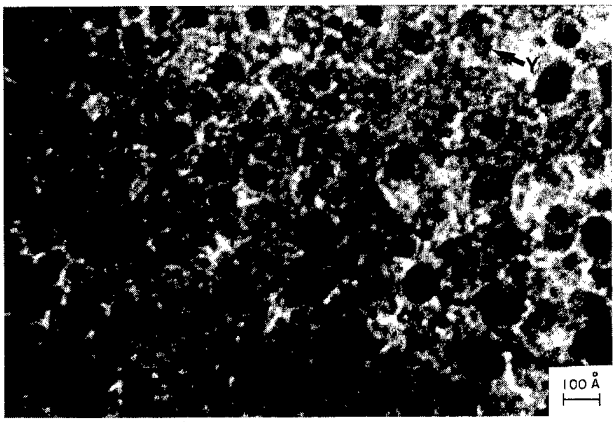


FIG. 2. Dark field images obtained by TEM in the {311} directions of a porous silicon structure formed in ⟨100⟩ oriented silicon n-type, 0.009 Ω cm. The long-range ordering resulting from the original crystalline structure is given by the width of the brightest region (~500 nm).

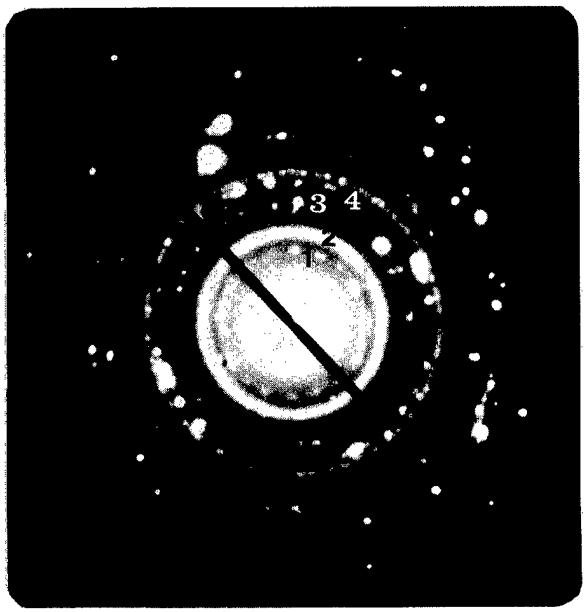
evident that when the misorientation of the crystallites distorts the original grid into curved lines, the dark fringes are curved, as shown in Fig. 1(b).

In order to check the crystallites spatially variable orientation in the etched samples, dark field images, as shown in Fig. 2, were obtained by TEM. Spots corresponding to the {111}, {220}, {311}, {400}, {331}, and {422} diffraction were observed. The dark field image for the {311} direction (Fig. 2) clearly shows that the nanocrystals in the brightest area of the photograph are preferentially aligned along this orientation. The long-range ordering resulting from the original crystalline structure in the sample is of the order of ~500 nm.

Figure 3(a) shows the TEM image of a 4.9 Ω cm, p-type, ⟨100⟩ oriented silicon sample, anodized at a current density of 5 mA/cm^2. This sample shows a photoluminescence peak at 1.8 eV. The shape of the etched parts of the sample are emphasized by a whitish color. The structure of the anodized silicon is seen to be composed of highly isolated columns [marked Y in Fig. 3(a)] as small as 15 Å cross-section diameter. Figure 3(b) shows that the asterism of the electron diffraction spots is discontinuously split into tiny intensity maxima. Spots corresponding to the {311}, {331}, and {533} diffraction were observed [indicated by 1, 3, and 4 in Fig. 3(b)]. This observation indicates that segmented crystallites are produced, which are relatively undistorted among themselves although differing slightly in orientation from their neighbors. The etched lattice is formed by an aggregate of comparatively strain-free subgrains separated from one another and having nearly the same orientation as the original crystal.



(a)



(b)

FIG. 3. (a) TEM images of porous silicon structure formed by anodizing a 4.9 Ω cm ⟨100⟩ oriented, *p*-type silicon samples. Y indicates a particle with a diameter of ~15 Å. (b) Electron anodized diffraction pattern of porous silicon obtained by TEM showing spots corresponding to the ⟨311⟩, ⟨331⟩, and ⟨533⟩ diffraction, indicated by 1, 3, and 4.

The surface topography of PS samples was then investigated by AFM. An image corresponding to a central position in the 1 cm diameter, 0.009 Ω cm samples, anodized with a current of 100 mA/cm^2 during 20 min, is shown in Fig. 4. Etched holes are emphasized by dark color. A skeleton structure is observed, which is formed by a pattern of large holes having diameters of ~1 μm (dark regions, indicated by Y). The walls of these pores have dimensions of this order as well.

Silicon tips from Park Instruments [Ultralever™(Si) and Microlever™ (SiN)] were used to obtain improved spatial resolution. An image obtained by AFM (Microlever™ tip) from a *p*-type 4.9 Ω cm ⟨100⟩ oriented silicon sample, anodized during 30 min at 5 mA/cm^2 current density, is displayed in Fig. 5. These samples show a red photoluminescent peak at 1.8 eV.[10] The image shows a pattern of columns (light isolated spots) over the entire PS surface. A more detailed mapping of the PS formed structure is shown by its profile. Typical profiles of this sample observed by AFM show valleys and hills as small as 55 Å wide [see Fig. 6(a) obtained with the Microlever™ tip and Fig. 6(b) with the Ultralever™ tip].

IV. DISCUSSION

A. TEM observations, diffraction lines, and dark field images

The results above demonstrate that PS layers formed during silicon anodization of different light emitting species also have different structures (see Figs. 1–3). Diffraction lines and dark field images are used to identify and define the crystallite specimen long-range order. The long-range order shown in Fig. 2 is ~500 nm. Figure 3(b) shows that PS formed on 4.9 Ω cm samples is composed of small loosely connected crystallites, as previously reported[16] PS layers are therefore formed by slightly disoriented lattice segments or subgrains, separated by etched holes. TEM images give the sizes and particle orientations (Figs. 1–3).

B. Observation by AFM and TEM

A comparison of TEM and AFM images of samples was then performed. Figures 1(a) and 1(b) are TEM images and Fig. 4 depicts an AFM image of the 0.009 Ω cm, *n*-type, ⟨100⟩ oriented silicon substrate each showing different aspects of the porous nanosized structure. The AFM images give a three-dimensional view of the skeleton structure of PS which cannot be observed by TEM. This three-dimensionality allows the visualization of the interconnecting structure.

An AFM image corresponding to a red-photoluminescence sample is presented in Figs. 5 and 6. Note that it was not possible to resolve features smaller than those shown in this figure. The discrepancy between the PS surface shown by TEM and the one shown by AFM is due to the response of the probe tip to the force exerted by the surface structure. AFM images are a reconstruction from digital data obtained from a probe tip response to a sample's topography as the tip is scanned across the surface. Studies involving structures highly three-dimensional, as PS layers, can be complicated by the fact that although the vertical resolution is usually excellent, the AFM produces distorted lateral features on the steeply inclined sides of nanoparticles. The effect of the finite curvature of the tip in AFM imaging has been noted by several authors.[17–21] Therefore a knowledge of the tip geometry is of fundamental importance.

C. AFM image distortions and geometrical resolution limitations

The distortion of an AFM image is illustrated by the response of a conical tip with an apex tip radius R_T to various structures as shown in Figs. 7(a) and 7(b). Two situations can be distinguished, namely $\alpha < \alpha_0$ or $\alpha \geq \alpha_0$, where

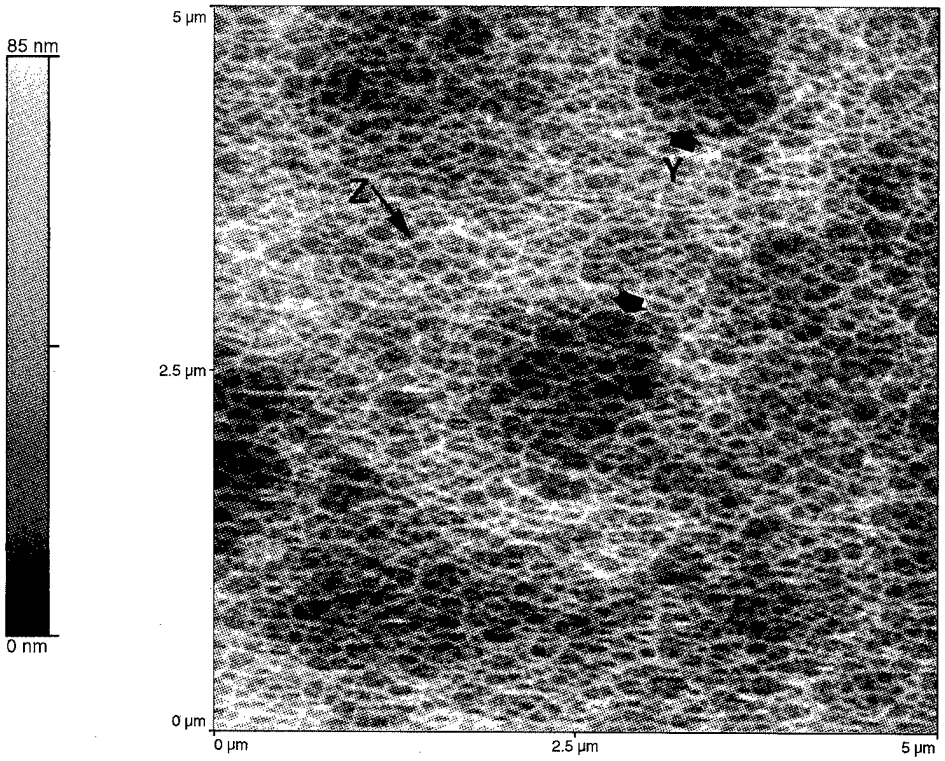


FIG. 4. An AFM image of a porous silicon etched in ⟨100⟩ oriented silicon, n-type, 0.009 Ω cm. It shows a skeleton structure of holes and walls formed by a substructure of ~40-nm-thick walls, indicated by Z.

$\alpha_0=2\arctan[K/(1-K)]^{1/2}$, $K=h/2R_T$, h is the height of the object, R_T is the apex radius, and α is angle of rise or fall of the structure. Figure 7(b) shows that for $\alpha<\alpha_0$, the increase in the width of a structure in the sample is given by $R_T \sin(\alpha/2)$ for each rise or fall of the sample structure. This value does not depend on the height h of the structure.

For $\alpha\geqslant\alpha_0$ and small values of h ($h/R_T<0.3$) only the spherical apex surveys the structure. Figure 7(a) shows the observed profiles for this situation, while Fig. 8(a) displays the unobserved region d or "hole region" and Fig. 8(b) the critical angle α_0 as a function of the object height (h) for the tip apex radii of 200, 50, 30 and 15 Å. The distortions introduced do not affect the height, but only the width and the increase in the width will depend on the height (h) and angles (α's) of the structure. Table I gives the values of the calculated full width at half height for objects of various heights h and apex tip radii R_T. The object is a symmetrical trapezoid with a flat horizontal top of 15 Å width, and α is

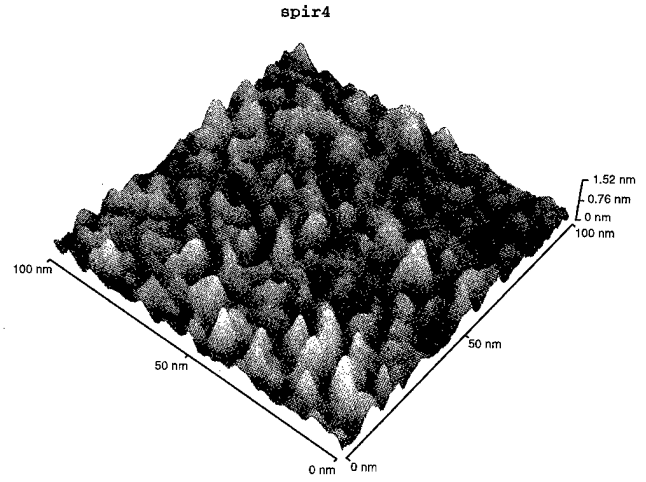


FIG. 5. An AFM image of p-type, 4.9 Ω cm ⟨100⟩ oriented silicon sample anodized during 30 min at 5 mA/cm^2 current density. It shows a pattern of columns over the entire surface.

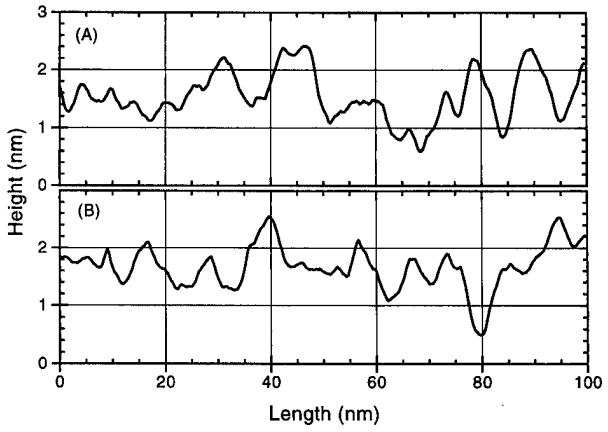


FIG. 6. Typical profiles of the sample of Fig. 5 with (a) a Microlever™ SiN tip; (b) an Ultralever™ Si tip.

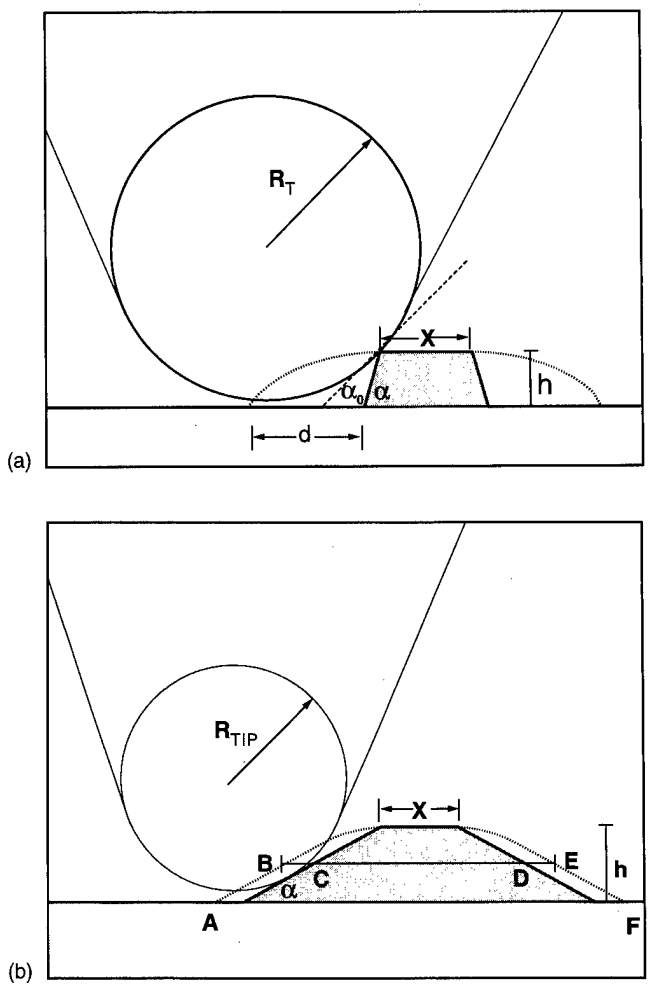


FIG. 7. Distortion of an AFM image illustrated by the response of a tip of radius R_T to a step structure which rises at an angle α to a vertical height h. (a) α larger than the critical angle α_0; (b) α smaller than the critical angle α_0.

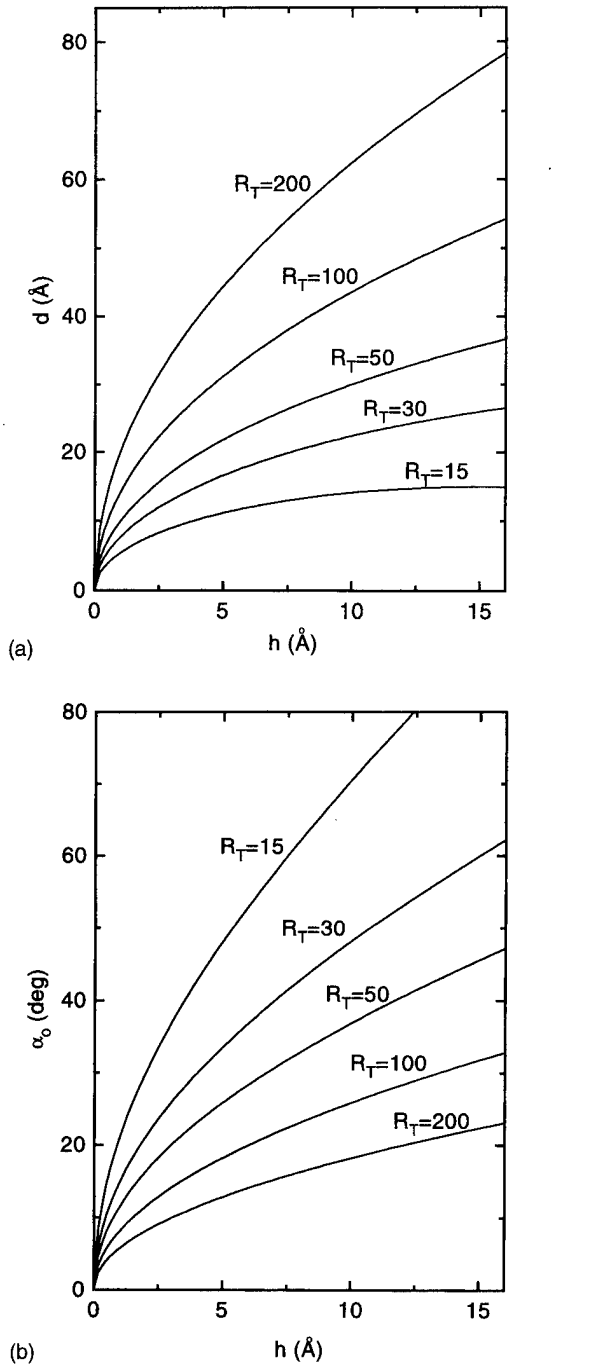


FIG. 8. (a) Value of the unobserved region length d as a function of the vertical height for various tip radii R_T. (b) Critical angle value α_0 as a function of the vertical height for various tip radii R_T.

the angle of rise (or fall) of the structure. Observe that the measured particle width depends critically on both the particle height and the tip radius.

D. Porous silicon AFM formed images

The spatial resolution of the TEM images is 15 Å, while that of the AFM depends critically on the profile of the probe tip and sample. When the tip apex radius is of a magnitude similar to the one of the sample structure, the distances between adjacent peaks in the samples are well reproduced in the observed image. Also, there is less distortion of peaks than of valleys. The depth of valleys in the image is strongly dependent on the radius of the tip. The radii of tips were determinated from the profiles shows in Fig. 6. For the Microlever™ tips the estimated radius obtained from the profiles shown in Fig. 6(a) is 48±10 Å and for the Ultralever™ tips [Fig. 6(b)] the radius is equal to 30±6 Å.

The porous silicon layers are formed by a skeleton structure of ~15 Å diameter cylinders as shown in Fig. 3(a). The profiles present peaks with heights around 10 Å. Using this height value and the estimated radius of the tips, the critical angle for the Ultralever™ and Microlever™ tips are equal to 60° and 42°, respectively, which are larger than the observed α's of 10° to 15°, as illustrated in Fig. 7(b). Thus the true width is the measured width minus $R\sin(\alpha/2)$ for both the rise and the fall. Isolated peak profiles shown in Figs. 6(a) and 6(b) were measured and the ascending and the descending angles calculated as well as the radii of the peaks. The smallest particle width was estimated to be 15±4 Å (for the Microlever™ tip) and 15±8 Å (for the Ultralever™ tip).

TABLE I. The observed full widths at half-maximum (FWHM) of structures with $\alpha > \alpha_0$, where the width is 15 Å at the top for various heights from 2 to 15 Å and for tips with apex radii of 200, 100, 50, 30, and 15 Å. No specification is made of the half angle of the cone, so that no values of FWHM are given for $h > 0.3R_T$ where this parameter would be necessary.

h (Å) \ R (Å)	200	100	50	30	15
15	124	91	68	...	...
10	104	77	59	48	...
05	78	59	46	39	32
02	55	43	35	30	26

V. CONCLUSIONS

In conclusion, this work has shown that different structures are responsible for near infrared and red luminescent PS. Diffraction lines and dark field images are used to identify and define the crystallite specimen long-range order. TEM images give the size and particle orientation, and AFM images show a three-dimensional pattern formed by an interconnecting skeleton of particles. The spatial resolution of the TEM images is 15 Å, while that of the AFM depends critically on the profile of the probe tip and the height and slope of the observed structure.

The measured diameter wire structure forming porous silicon structure by AFM and TEM allows us to calculate the distortion of the AFM images due to the finite size of the tip radius. A critical angle $\alpha_0 = 2 \arctan[K/(1-K)]^{1/2}$, where K is the ratio of the height of the structure to the tip diameter, was defined and substantial distortions were observed for structure walls that are steeper than the critical angle.

ACKNOWLEDGMENTS

The authors are grateful to L. O. Bonugli and J. R. Castro for technical assistance and FAPESP 93/0961-5, CNPq 520060/93-8, and FAEP/UNICAMP 0726/94 for financial support.

[1]L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[2]F. Buda, J. Kohanoff, and M. Parrinello, Phys. Rev. Lett. **69**, 1272 (1992).
[3]J. P. Proot, C. Delerue, and G. Alan, Appl. Phys. Lett. **61**, 1948 (1992).
[4]S. Furukawa and T. Miyasato, Jpn. J. Appl. Phys. **27**, L-2207 (1988).
[5]D. Sarid, *Scanning Force Microscopy* (Oxford University, Oxford, 1991), and references cited therein.
[6]G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, Appl. Phys. Lett. **40**, 178 (1982).
[7]M. Enachescu, E. Hartmann, A. Kux, and F. Koch, J. Lumin. **57**, 191 (1993).
[8]O. Teschke, M. C. Gonçalves, and F. Galembeck, Appl. Phys. Lett. **63**, 1348 (1993).
[9]O. Teschke, F. Alvarez, L. R. Tessler, and M. U. Kleinke, Appl. Phys. Lett. **63**, 1927 (1993).
[10]O. Teschke, Appl. Phys. Lett. **64**, 1986 (1994).
[11]O. Teschke, F. Galembeck, M. C. Gonçalves, and C. V. Davanzo, Appl. Phys. Lett. **64**, 3590 (1994).
[12]J. M. Gomez-Rodriguez, A. M. Baro, and V. P. Parkhutik, Appl. Surf. Sci. **44**, 185 (1990).
[13]G. B. Amisola, R. Behrensmeier, J. M. Galligan, F. A. Otter, F. Namavar, and N. M. Kalkoran, Appl. Phys. Lett. **61**, 2595 (1992).
[14]V. P. Parkhutik, J. M. Abella, J. M. Martinez-Duart, J. M. Gomez-Rodriguez, A. M. Baro, and V. I. Shershulsky, Appl. Phys. Lett. **62**, 366 (1993).
[15]C. Odin, J. P. Aimé, Z. El Kaakour, and T. Bouhacina, Surf. Sci. **317**, 321 (1994).
[16]K. Barla, R. Herino, G. Bomchil, J. C. Pfister, and A. Freund, J. Cryst. Growth **68**, 727 (1984).
[17]Ph. Niedermann and O. Fisher, J. Microsc. **152**, 93 (1988).
[18]M. Stedman, J. Microsc. **152**, 611 (1988).
[19]G. Reiss, J. Vancea, H. Wittman, J. Zweeck, and H. Hoffmann, J. Appl. Phys. **67**, 1156 (1990).
[20]K. L. Westra, A. W. Mitchell, and D. J. Thomson, J. Appl. Phys. **74**, 3608 (1993).
[21]J. E. Griffith and D. A. Grigg, J. Appl. Phys. **74**, R 83 (1993).

Field emission characteristics of the scanning tunneling microscope for nanolithography

T. M. Mayer,[a] D. P. Adams, and B. M. Marder
Sandia National Laboratories, Albuquerque, New Mexico 87185-1413

(Received 16 January 1996; accepted 29 May 1996)

We present a systematic study of the performance of scanning tunneling microscope (STM)-based, low energy electron beam lithography, using simulations of field emission from STM tips, emphasizing realistic conditions of tip geometry and operation. We calculate the potentials and electric field for a hemispherical model emitter in an axially symmetric system. Emission current density at the tip is calculated using the Fowler–Nordheim equation, and current density at the sample is obtained by calculating trajectories of emitted electrons. We characterize the beam diameter at the sample as a function of emitter radius, tip–sample bias, emission current, resist thickness, and tip work function. The beam diameter is primarily affected by the tip–sample gap, increasing at larger gaps, characteristic of high bias and large tip curvature. For optimal tip radius the beam diameter increases linearly with bias from approximately 2 nm at 5 V to 25 nm at 50 V. Beam diameter is nearly independent of emission current over the range 0.05–50 nA. Dielectric resist films cause an increase in beam diameter due to increased tip–substrate gap. Beam diameter is very sensitive to tip work function, increasing dramatically for low work function tips. Tips comprised of asperities on flat surfaces produce significantly smaller beams compared to "standard" tips of the same emitter radius. However, for low bias (<15 V) beam diameter becomes insensitive to tip geometry. We compare these simulations to selected experimental results to evaluate the limitations to performance and assess the feasibility of routine sub-10 nm structure fabrication using STM-based low energy electron beam lithography. © *1996 American Vacuum Society.*

I. INTRODUCTION

Low energy electron beam lithography is an attractive option for nanometer-scale device fabrication because of the virtual elimination of scattering effects which limit resolution in conventional electron beam lithography.[1,2] The scanning tunneling microscope (STM) operating in field emission mode is widely employed to produce electron beams with energy <100 eV for this purpose. The relative simplicity of the system, small beam size, and very high current density at low energy makes the STM an ideal research tool, and a potentially useful manufacturing tool for nanometer-scale structures.

There are a number of possible approaches to nanostructure fabrication using the STM,[3,4] but the most promising for widespread utility is low energy electron exposure of an imaging material (resist), which is used to transfer a pattern into a substrate or film. A variety of resist materials has been investigated using STM-based lithography, including PMMA and other conventional resists,[2,5,6] thin siloxane films,[7] molecular films such as Langmuir–Blodgett and self-assembled monolayer films,[8–10] thin inorganic materials, such as CaF_2,[11] and adsorbed atomic layers, particularly hydrogen on Si.[12–14]

The limitations to feature size and resolution obtainable by this process are functions of both the resist material performance and the ability to form a small diameter beam of low energy electrons. However, there have been few investigations of the performance of a STM tip as a low energy electron beam source. McCord and Pease[15] performed simulations of electron emission from an isolated sphere in the field emission regime (5 eV<E<100 eV), and identified many of the major features of STM-based lithography. Beam diameter at the sample was shown to be a sensitive function of the tip radius, gap, and tip–sample bias.

Since this early modeling effort, many experimental demonstrations of nanometer-scale lithography have appeared. In particular, we have shown that it is possible to directly measure the current distribution from STM tips used in lithographic processes.[14] These results give impetus to more extensive modeling using realistic emitter structures and operating conditions that may be compared directly to experiments. In this article we present results of field emission simulations of STM tips as low energy (<50 eV) electron beam sources for nanolithography. We emphasize comparison to realistic conditions where a tip is operated in constant voltage and current mode, using feedback control of the gap to maintain a fixed current. We investigate the effects of bias voltage and tip radius on beam diameter for idealized hemispherical tips. We also investigate the dependence of beam diameter on total emission current, tip work function, and the presence of a dielectric (resist) layer between the tip and substrate. Alternate tip configurations in which the electric field at the tip is tailored to confine the trajectories of emitted electrons into a small beam diameter are evaluated. These results help establish reasonable expectations for performance of STM-based nanolithography.

Finally, we critically examine selected experimental systems to evaluate their performance compared to the expected

[a]Electronic mail: tmmayer@sandia.gov

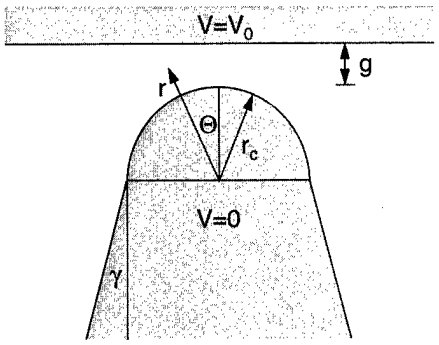


FIG. 1. Geometry of the field emission calculation. The cathode (tip) and anode (sample) are assumed to be conducting.

emitter characteristics. We conclude that routine sub-10 nm lithography is feasible using low energy (<20 V) exposures and very thin, atomic or molecular layer resists.

II. FIELD EMISSION SIMULATIONS

The field emission simulations are carried out in the geometry shown in Fig. 1. We define an axially symmetric system, in which the cathode (emitter) is a conducting hemisphere of radius r_c. The emitter is attached to a shank with a variable cone angle, γ. The anode (sample) forms a conducting plane at some distance z_a, where the gap between emitter and sample is $g=(z_a-r_c)$. A dielectric film (resist) can be interposed between the tip and sample.

The potential in the space between the anode and cathode obeys Laplace's equation:

$$\nabla\epsilon\nabla V=0, \tag{1}$$

where V is the potential at any point in the vacuum or resist film and ϵ is the dielectric constant of the medium. In spherical polar coordinates Eq. (1) is[16]

$$\frac{\partial}{\partial r}\epsilon r^2\frac{\partial V}{\partial r}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\epsilon\sin\theta\frac{\partial V}{\partial\theta}=0. \tag{2}$$

After specifying the emitter geometry and the emitter–sample gap, we define a set of grid points in r and θ, and solve Eq. (2) for $V(r,\theta)$ using a finite difference method with the boundary conditions that $V=0$ at the cathode (tip) and $V=V_0$ at the anode (sample).

The emission current density at the tip is given by the Fowler–Nordheim equation:[17]

$$J=\frac{AE^2}{\phi}\exp\left(\frac{B}{\phi^{1/2}}-\frac{C\phi^{3/2}}{E}\right), \tag{3}$$

where E is the electric field, ϕ is the work function of the tip, and A, B, and C are constants. The electric field at the emitter surface is the gradient of $V(r_c,\theta)$. We integrate the emission over the entire tip surface and adjust the anode distance, z_a, to obtain the desired total emission current at a fixed bias.

Sample current density is obtained by calculating trajectories of electrons leaving the tip at different angles, and mapping the emission current density from the tip to the sample. We assume that electrons leave the tip normal to its surface, with zero kinetic energy, and we ignore scattering of the electrons in the resist film (if present).

Most calculations are carried out using a tip work function, ϕ=4.5 eV, appropriate for a W tip.[18] The shank angle, γ, has minimal effect on the electric field at the tip and the trajectories of electrons traveling to the sample, except for very sharp tips and large tip–sample distances. All "standard" tip simulations are carried out using a shank angle of 20°. We also explore some alternate tip geometries in an attempt to confine or focus the electron current at the sample by tailoring the electric field in the vicinity of the tip.

III. RESULTS AND DISCUSSION

A. Characteristics of standard tips

We have examined emission characteristics of standard tips as a function of tip radius, anode voltage, total emission current, resist thickness, and tip work function. In typical operation of an STM-based lithography system, the anode voltage and emission current are held constant, and the gap is adjusted by a feedback circuit to maintain constant electric field at the tip and thus constant emission. This is a very important aspect, since the electric field configuration in the gap and trajectories of electrons to the sample are very sensitive to the tip–sample gap. For direct comparison to this mode of STM operation, all simulations presented here treat the gap as a free parameter. We have previously shown that these simulations agree well with the gap measured as a function of anode voltage in an operating STM.[14] Since we are primarily interested in the current density distribution at the sample, results are presented in terms of the beam diameter at the sample surface. We define the beam diameter as that which contains 60% of the total current. (This is approximately equivalent to the standard method of measuring beam diameter in electron beam lithography systems.)

Considering standard tips with radii of 5–50 nm, we calculate the beam current density distributions for biases from 5 to 50 V. The current density distribution at the sample follows a functional form of approximately $\exp(-r^{1.8}/\alpha)$, where the width parameter, α, depends on the bias and the tip–sample gap. The beam diameter as a function of bias voltage is shown in Fig. 2(a). The increase in beam diameter with increasing bias, for all tip radii, is primarily due to increased gap at increased bias, and resulting radial spreading of electron trajectories. This is demonstrated in Fig. 2(b), where the beam diameter is seen to increase approximately linearly with the gap that is required to produce 1 nA current at the given bias.

Tips with higher curvature (small r_c) develop higher electric fields at a given V_0 and gap, so that the feedback circuit responds by retracting the tip to maintain the preset emission. For sharper tips emission also occurs over a larger range of angle, θ, at the apex of the tip. This results in the phenomenon shown in Fig. 2(a) in which sharp tips give smaller beam diameters at low bias but larger beam diameters at increasing bias, due to the larger gap. These results are in good agreement with the earlier simulations of Mc-

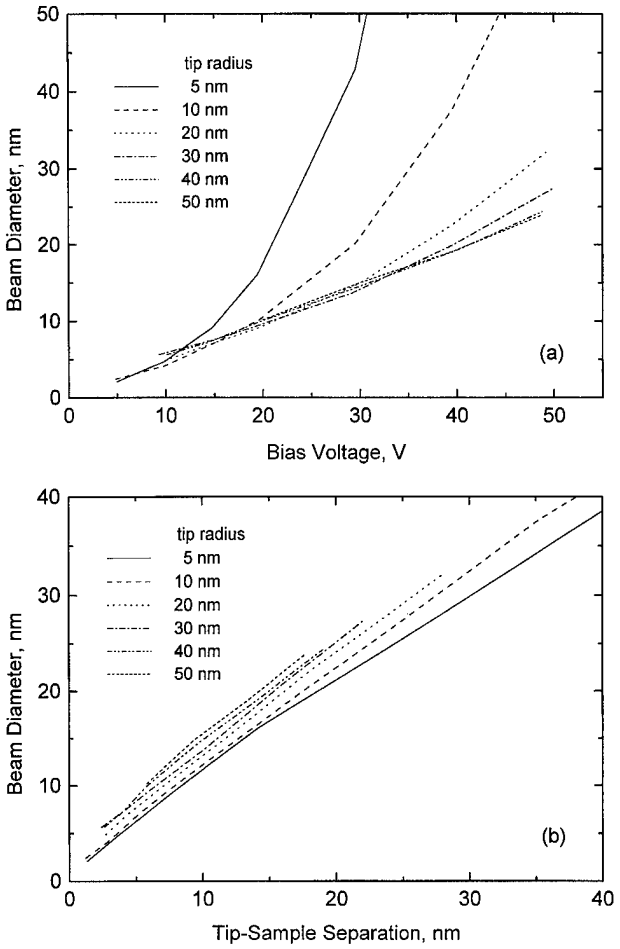


FIG. 2. (a) Beam diameter vs sample bias for various tip radii at 1 nA emission current; (b) beam diameter vs tip sample separation for the calculations shown in (a).

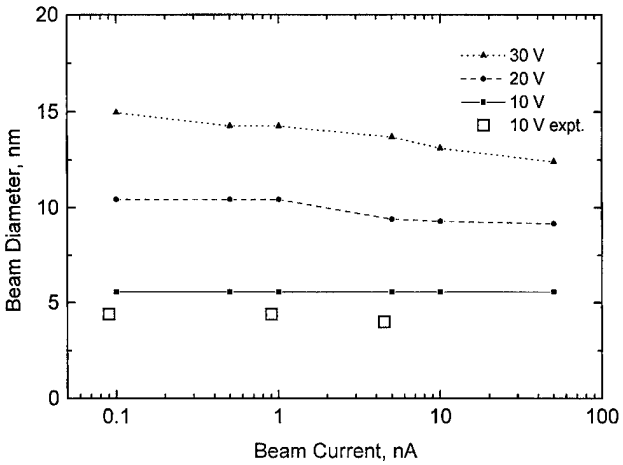


FIG. 3. Beam diameter vs emission current for a 30-nm-radius tip at 10–30 V sample bias. Experimental data (from Ref. 14) are for 10 V bias.

Cord and Pease,[15] who demonstrated that there is an optimum ratio of tip radius to gap which gives minimum beam diameter. In general, the beam diameter increases approximately linearly with bias for optimum tip radius, and best performance is achieved at low bias for all tip radii.

We have examined other influences on beam diameter for the standard tip geometry. Dependence of beam diameter on beam current is shown in Fig. 3, for a tip radius of 30 nm, and for several biases. The beam diameter decreases slightly at high current due the decreased gap required for higher fields. But the dependence is fairly weak over three orders of magnitude of current due to the steep dependence of emission on electric field characteristic of the Fowler–Nordheim emission process. Only very small changes in gap are necessary to alter the field sufficiently for large variations in current. We have ignored space charge effects, which can substantially alter the trajectories of low energy electrons in high current density beams. For currents <1 μA, at typical bias and tip–sample gap, there is, on average, only one electron in the gap at a time.

The presence of a resist film will alter the electric field at the tip and in the gap. The primary effect of interposing a resist film will be to increase the tip–sample separation, causing increased beam diameter. A resist will also have a higher dielectric constant than vacuum (or air). Presuming that the tip is not allowed to penetrate the resist film, the electric field in the vacuum gap will be approximately $E(\text{gap}) \approx \epsilon E(\text{resist})$, where ϵ is the resist dielectric constant. Because $E(\text{resist})$ decreases with increasing resist thickness, $d(\text{gap})$ must decrease in order to produce adequate electric field at the tip to maintain emission. For thick resists, the tip must be very close to the resist surface in order to maintain emission [$d(\text{gap})<1$ nm for resist thickness of 50 nm]. In general, particularly for resists with low dielectric constant, as resist thickness increases, it is difficult to maintain an adequate electric field at the tip to provide the desired emission current. The tip will then be forced to penetrate the resist in order to establish the required electric field.

Beam diameter at the resist/vacuum interface is small, because of the reduced $d(\text{gap})$, however the beam spreads traversing the resist film. Calculations for a tip radius of 30 nm, 20 V bias, and $\epsilon=3$ show that the beam diameter at the resist/substrate interface increases with resist thickness with a slope of ≈ 0.3. For this calculation, we assume the electron travels through the resist film without scattering. Of course electrons will scatter in the resist, as they must to deposit energy, and the volume of exposed material will be further increased by the scattering of these low energy electrons, comparable to the scattering of secondary electrons produced by high energy electron beam exposure.

Alternate tip materials with lower work function or tips with low work function adsorbates[19] can be employed. This leads to higher emission current at a given electric field, as seen from Eq. (3). This is generally desirable for other field emission devices, such as emitters for electron microscopes, vacuum microelectronics, or display devices. However in a lithographic application, a low work function tip will achieve a given current density at a larger tip–sample gap, resulting in larger beam diameter at a given current and voltage than for the same geometry tip of a higher work function. The effect is quite dramatic, and is shown in Fig. 4. Plotted are

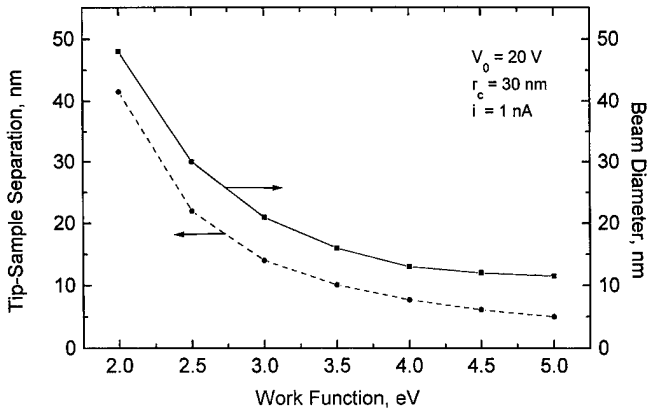


FIG. 4. Effect of tip work function on tip–sample separation and beam diameter for a 30-nm-radius tip, 20 V bias, and 1 nA emission.

tip–sample gap and beam diameter at the sample as a function of tip work function for a 30-nm-radius tip at 20 V bias and 1 nA current. The direct correlation of tip–sample gap and beam diameter is well demonstrated here. Additional effects of tip geometry and bias behave in a similar manner to Fig. 2.

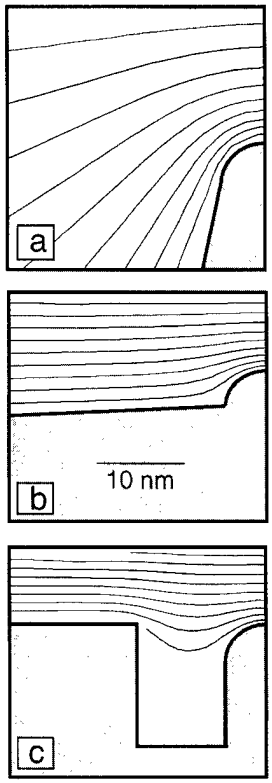


FIG. 5. Equipotential contours for (a) standard tip, (b) asperity tip, and (c) recessed tip. Contour interval is $0.1V_0$. Tip–sample gap is that required for 1 nA emission at 20 V bias.

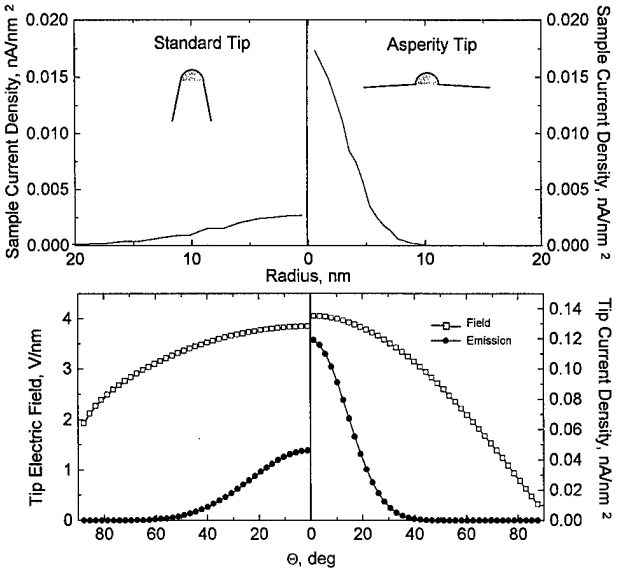


FIG. 6. Tip electric field, tip emission current density, and sample current density for standard and super tips. Emitter radius is 5 nm, bias is 20 V, and emission is 1 nA for each.

B. Alternate tip configurations

The idealized standard tip configuration shown in Fig. 1 is unlikely to be representative of real electron emitters for a number of reasons. While tips may appear round and smooth on a macroscopic scale, invariably, microscopic details of the tip surface will include local asperities due to tip preparation, tip alternation, contamination, or even discontinuities due to atomic steps and defects on the surface. These will all lead to a local alteration of the electric field and emission characteristics that deviate from the ideal spherical tip.[20] It is generally assumed in scanning tunneling microscopy that tunneling occurs primarily through only a few atoms at a local asperity, and many tip preparation procedures are effectively intended to produce small asperities.

Asperities can also be intentionally fabricated in a controlled fashion to produce a desired electric field configuration at the tip. "Super tips"[21,22] have been constructed by producing controlled asperities on the end of relatively blunt tips for use as electron and ion emitters with a reduced angular emission profile. While a very small asperity indicates a high curvature emitter, the electric field is suppressed at large polar angles by the proximity of a flat surface, and emission is restricted to a small range of polar angles at the apex of the asperity. The result is a beam that is focused in the forward direction.

This focusing effect can also be useful in lithographic applications. We have examined to alternate tip configurations intended to produce this focusing effect, shown in Fig. 5 along with equipotential contours for a tip–sample gap that would produce 1 nA current at a sample bias of 20 V. The origin of the effect is clear from the shapes of the equipotential contours and the recognition that field lines are perpendicular to the potential contours. The standard tip in Fig. 5(a)

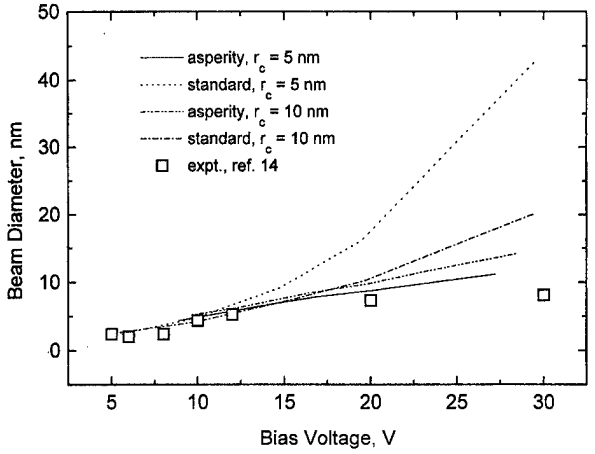


FIG. 7. Beam diameter vs sample bias for standard and asperity tips with radii of 5 and 10 nm at 1 nA emission. Experimental data are from Ref. 14.

with a 5 nm radius exhibits large curvature of the potential contours, leading to a large degree of radial spreading of the electron trajectories. Figure 5(b) depicts an asperity tip consisting of a hemisphere with 5 nm radius on a nearly flat surface (shank angle $=85°$). The potential contours are comparatively much flatter in the region of the emitter, giving electron trajectories focused in the forward direction. Figure 6 shows a comparison of the electric field and emission profiles at the tip and the current profile at the surface for the standard tip and the asperity tip in Figs. 5(a) and 5(b). For the standard tip, at 20 V bias and 1 nA emission, the tip–sample gap is 17 nm, and emission occurs out to $>40°$ polar angle on the tip. This results in a diffuse beam with a diameter of 21 nm. The asperity tip, however, shows high field and significant emission only out to 30° polar angle, and the tip–sample gap is reduced to 10 nm. This gives a beam diameter of 9 nm, or less than half that of the standard tip.

Beam diameter versus sample bias for asperity tips and standard tips of 5 and 10 nm tip radius are shown in Fig. 7. The characteristics of the asperity tip emission follow the same trends as those for the standard tip, only with decreased beam diameter due to the focusing effect and reduced tip–sample separation. We note from Fig. 7 that the effect is most pronounced for small asperities. For the larger (10 nm) asperity the alteration of the potential near the tip apex is less pronounced and the focusing effect is minimized. We also note that the focusing effect is most pronounced at higher bias, where the tip–sample separation is large. Small tip–sample separation at low bias also results in a minimal alteration of the potential near the apex, and minimal focusing.

Comparison of Fig. 7 with Fig. 2(a) shows that the 5 nm asperity tip produces a beam slightly smaller than the "optimum" standard tip at all biases above 15 V. However, at biases of less than about 15 V, the beam diameter is relatively insensitive to the tip geometry, whether it is an asperity or standard tip. This insensitivity suggests that for low voltage lithography one need not be very concerned about the nature of the tip.

Further focusing of electron trajectories can be obtained using an emitter structure similar to that shown in Fig. 5(c). This is a recessed emitter than could potentially be fabricated by ion sputtering techniques similar to those used to make extremely sharp tips for metrology applications,[23] and recessed emitters for vacuum microelectronics.[24,25] In this schematic example the recessed tip is a hemisphere with a diameter of 5 nm. The top of the tip is at the same height as the planar cathode, and the tip is separated from the rest of the cathode by a 10-nm-wide trench. The intention here is to use the planar cathode as a Wehnelt cylinder to focus electron trajectories in the forward direction. The focusing effect is evident, producing a beam diameter of 7.2 nm for the conditions of Fig. 6, compared to 18 nm for the standard tip and 8.8 nm for the asperity tip. The performance of this tip depends primarily on the width of the trench between the tip and the planar cathode. Structures with trench width of >20 nm perform less well than the tip of Fig. 5(b). Performance could be increased by narrowing the trench or by recessing the emitter below the surface of the planar cathode. Of course the corners of the planar cathode would have to be suitably rounded to prevent development of an excessive field at this corner. Fabrication of such a tip becomes problematic, and operation of a recessed tip in close proximity to a sample may result in the planar cathode being too close to the sample (with very high electric field at the planar cathode surface). This type of emitter structure may be useful in other field emission devices that require a focused electron beam.[26]

We have not attempted to model emitters consisting of only a single or very few atoms. Such ultrasharp tips have been fabricated and used as point source emitters for electron microscopes.[19,27] As expected, they demonstrate very high fields and emission current densities at the apex. Cutler and co-workers,[28] however, have shown that ultrasharp tips do not obey the usual formulation of the Fowler–Nordheim emission theory, and suggest that Eq. (3) may be in error by orders of magnitude for tip radii <10 nm. It appears that a more sophisticated treatment of the field emission process is required for modeling atomically sharp tips.

C. Comparison to experiments

While most experimental demonstrations of STM-based lithography have not specified tip geometry nor directly measured electron current distributions, we can compare these simulations to a number of experimental results demonstrating feature sizes on the order of 10 nm that have appeared over the past few years.

We have directly measured the electron current distribution at the sample by imaging the distribution of desorbed H atoms from Si(001) surfaces in the STM, reported elsewhere.[14] Above a threshold energy of 6–8 eV, this electron stimulated desorption process provides a direct measure of the electron flux distribution with minimal electron scattering effects. We also measured the tip–sample separation as a function of sample bias. All of these measurements were made with a W tip, prepared by electrochemical etching, that was not evaluated prior to introduction into the STM. Over the range of 5–30 V bias, we observed a beam diameter that

varied from 2 to 8 nm. The measured tip–sample gap ranged from, nominally, 1 nm at tunneling conditions of 2 V bias to 16 nm at 40 V. The tip–sample gap agrees well with standard tip simulations for hemispherical emitters with $r=30-50$ nm.

The observed beam diameters, however, were smaller than those calculated for standard tips of optimum radius, particularly at higher bias, but agree well with the simulations including focusing effects associated with emission from asperities on a blunt tip, shown in Fig. 7. We do not know if emission is occurring from a single site or from a number of asperities. (We have seen experimental evidence for multiple site emission, on occasion.) Beam diameter was observed to be essentially constant over more than two orders of magnitude of emission current at fixed bias, in good agreement with the calculations, as shown in Fig. 4. The estimated functional form of the beam profile is also in reasonable agreement with the calculations. The experimental data are best fit with an exponential form, decaying radially as r^1 to $r^{1.5}$, while the calculations indicate an $r^{1.8}$ dependence. Feature sizes produced as a function of exposure dose show broadening, consistent with the beam profile measured at low dose, with some evidence of electron reflection effects at high bias and dose. This system, by eliminating scattering effects and using an ideally thin, self-developing resist appears limited only by electron current distribution from the tip.

Effects of electron scattering, large tip–sample separation with thick resists, and postexposure processing can all limit the size of features that can be fabricated independently of the performance of a given tip. Exposure of self-assembled monolayer resists should show minimal electron scattering effects and minimal beam spreading due to increased tip–sample gap, and most accurately reflect the beam diameter at the surface. Marrian *et al.*[10] have exposed monolayers of a phenylethylenediaminemethoxysilane resist, approximately, 1 nm thick and achieved linewidths of <10 nm for low voltage (<10 V) exposures. Exposures of 4-chloromethylphenyltrichlorosilane at 8 V produced features <20 nm (measurement limited by the grain size of the Ni subsequently plated on the sample). Direct comparison to our simulations is not possible, but these studies appear to be approaching the limit of patterning ability determined by beam diameter.

Zhang *et al.*[8] have examined polymethylmethacrylate (PMMA) and polyvinylcinnamate films prepared by the Langmuir–Blodgett technique that give 4.5–14-nm-thick films. Linewidths of 30–100 nm were observed after exposure at 20–30 V sample bias, and various development and pattern transfer procedures. These dimensions are considerably larger than the expected beam diameter at these conditions and probably reflect the effects of resist development and subsequent etching processes on broadening of the feature dimension.

A number of investigators have examined thicker films of conventional electron beam resists prepared by spin-on techniques. McCord and Pease[5] report fabrication of 20 nm linewidths in 20-nm-thick films of PMMA. Dobisz and Marrian[2] and Marian *et al.*[6] report linewidths from 30 to 50 nm for 50–80-nm-thick resist (Shipley SAL-601) films. These results apparently show little broadening due to resist processing and are consistent with broadening of the electron beam profile by the presence of a thick resist film.

While a number of studies appear to be approaching the beam-limited performance, definition of sub-10 nm features has only been demonstrated for atomic or molecular layer resists. As suggested by the simulations, beam spreading (and ultimately electron scattering) in thicker (>10 nm) resist films leads to larger feature sizes. Routine sub-10 nm lithography may be limited to very thin resists, such as self-assembled monolayer films, or the adsorbed atomic or molecular layer.

IV. SUMMARY

We have presented a systematic study of the performance of STM-based, low energy electron beam lithography based on simulations of field emission from STM tips in close proximity to a substrate surface. The major factor in determining the beam diameter is the tip–sample gap, which is set to obtain the necessary field strength at the tip for the desired emission. The beam diameter increases as a function of bias because of the larger gap at higher voltage and subsequent spreading of the electron trajectories. At a given bias there is a trade-off between tip radius and tip–sample gap that leads to an optimum tip radius for minimum beam size. At low bias (<10 V) sharp tips give smallest beam size, whereas at high bias (>40 V) more blunt tips result in smaller beam diameter. For optimal tip radius the beam diameter increases linearly with bias from approximately 2 nm at 5 V to 25 nm at 50 V. Beam diameter is nearly independent of emission current over the range 0.05–50 nA.

Dielectric resist films interposed between tip and sample cause an increase in beam diameter due to increased tip–substrate gap. Reduction of the tip work function results in a dramatic increase in beam diameter due to the larger gap required to obtain the desired emission from a low work function tip.

We have examined alternative tip structures intended to tailor the electric field in the vicinity of the tip and focus the electron beam to a smaller diameter. Asperities consisting of a small radius hemisphere on a nearly flat surface produce beams with significantly smaller diameter than those produced by a standard tip of the same emitter radius. However, for sample bias <15 V the beam diameter becomes relatively insensitive to tip geometry, converging on a beam diameter of ~5 nm at 10 V bias for all tips.

The simulations agree quite well with experiments that directly measure the current distribution at the sample. These results, using adsorbed atomic H as a resist, display feature sizes limited only by the characteristics of the electron emitter. Experiments using molecular layer resists, such as self-assembled monolayer films, also approach resolution limits set by the electron beam diameter. Given the results of the simulations presented here and the beam-diameter-limited

performance of a number of experimental systems, we feel confidence that routine sub-10 nm structure fabrication using STM-based low energy electron beam lithography is feasible with current techniques.

ACKNOWLEDGMENTS

The authors have enjoyed fruitful discussions with B. S. Swartzentruber, G. L. Kellogg, and H. Koops. This work was supported by the U.S. Department of Energy under Contract No. De-AC04-94AL850000.

[1]Y. W. Yau, R. F. W. Pease, A. A. Iranamesh, and K. J. Polasko, J. Vac. Sci. Technol. **19**, 1048 (1981).
[2]E. A. Dobisz and C. R. K. Marrian, J. Vac. Sci. Technol. B **9**, 3024 (1991).
[3]G. M. Shedd and P. M. Russell, Nanotechnology **1**, 67 (1990).
[4]*Technology of Proximal Probe Lithography*, edited by C. R. K. Marrian (SPIE Institutes for Advanced Optical Technologies, Bellingham, WA, 1993), Vol. IS10.
[5]M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **4**, 86 (1986).
[6]C. R. K. Marrian, E. Dobisz, and J. A. Dagata, J. Vac. Sci. Technol. B **10**, 2877 (1992).
[7]S. W. Park, H. T. Soh, C. F. Quate, and S. I. Park, Appl. Phys. Lett. **67**, 2415 (1995).
[8]H. Zhang, L. S. Hordon, S. W. J. Kuan, P. Maccagno, and R. F. W. Pease, J. Vac. Sci. Technol. B **7**, 1717 (1989).
[9]L. Stockman, G. Neuttiens, C. van Haesendonck, and Y. Bruynsereade, Appl. Phys. Lett. **62**, 2935 (1993).
[10]C. R. K. Marrian, F. K. Perkins, S. L. Brandow, T. S. Koloski, E. A. Dobisz, and J. M. Calvert, Appl. Phys. Lett. **64**, 390 (1994); F. K. Perkins, E. A. Dobisz, C. R. K. Marrian, and S. L. Brandow, J. Vac. Sci. Technol. B **13**, 2841 (1995).
[11]M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **5**, 430 (1987).
[12]R. S. Becker, G. S. Higashi, Y. J. Chabal, and A. J. Becker, Phys. Rev. Lett. **65**, 1917 (1990).
[13]J. W. Lyding, G. C. Abeln, T. C. Shen, C. Wang, and J. R. Tucker, J. Vac. Sci. Technol. B **12**, 3735 (1994).
[14]D. P. Adams, T. M. Mayer, and B. S. Swartzentruber, J. Vac. Sci. Technol. B **14**, 1642 (1996).
[15]M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **3**, 198 (1985).
[16]P. L. Lorrain, D. P. Corson, and F. Lorrain, *Electromagnetic Fields and Waves* (Freeman, New York, 1988).
[17]R. H. Fowler and L. W. Nordheim, Proc. R. Soc. London Ser. A **119**, 173 (1928); C. A. Spindt, I. Brodie, L. Humphrey, and E. R. Westerberg, J. Appl. Phys. **47**, 5248 (1976).
[18]H. B. Michaelson, J. Appl. Phys. **48**, 4731 (1977).
[19]R. Morin and H.-W. Fink, Appl. Phys. Lett. **65**, 2362 (1994).
[20]R. Gomer, *Field Emission and Field Ionization* (Harvard University Press, Cambridge, 1961).
[21]K. Jousten, K. Bohringer, R. Borret, and S. Kalbitzer, Ultramicroscopy **26**, 301 (1988).
[22]R. Borret, K. Bohringer, and S. Kalbitzer, J. Phys. D: Appl. Phys. **23**, 1271 (1990).
[23]M. J. Vasile, D. A. Grigg, J. E. Griffith, E. A. Fitzgerald, and P. E. Russell, Rev. Sci. Instrum. **62**, 2167 (1991).
[24]J. Ishikawa, T. Ohtake, Y. Goto, H. Tsuji, N. Fukuyama, K. Inoue, S. Nagamachi, Y. Yamakage, M. Ueda, H. Maruno, and M. Asari, J. Vac. Sci. Technol. B **13**, 452 (1995).
[25]W. K. Lo, M. Skvarla, C. W. Lo, H. G. Craighead, and M. S. Isaacson, J. Vac. Sci. Technol. B **13**, 2441 (1995).
[26]M. G. R. Thompson, J. Vac. Sci. Technol. B **13**, 2455 (1995).
[27]H.-W. Fink, IBM J. Res. Devel. **30**, 460 (1986); H.-W. Fink, Phys. Scr. **38**, 260 (1988).
[28]P. H. Cutler, J. He, N. M. Miskovsky, T. E. Sullivan, and B. Weiss, Prog. Surf. Sci. **42**, 169 (1993).

Proximity effect correction for nanolithography

Richard Rau, James H. McClellan,[a] and Timothy J. Drabik
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0250

(Received 12 July 1995; accepted 10 May 1996)

The resolution of electron-beam lithography is limited by the proximity effect, which is due to the scattering of incident electrons. This scattering leads to exposure of neighboring areas and, therefore, to pattern degradations. It can be modeled by convolving the incident dosage distribution with a point spread function and the subsequent developing process by a pointwise nonlinear function. This article presents an iterative algorithm that exploits this overall model to compute corrections of the proximity effect in the nanometer range. A convex error function is derived that satisfies all physical constraints and the problem is set up as a convex, constrained, and nonlinear minimization problem. The correction algorithm is a hybrid of the conjugate gradient and gradient projection algorithms. Its performance for a test pattern is evaluated and then a detailed analysis of the effect of inaccuracies of the model parameters and of the necessary quantization of the corrected incident dosage distribution is presented. © *1996 American Vacuum Society.*

I. INTRODUCTION

The resolution of electron-beam lithography is limited by the so-called proximity effect, which describes degradations of the pattern during exposure due to forward scattering of electrons on their path through the resist and due to backward scattering of electrons from the substrate. The scattered electrons change locally the chemical properties of the resist, which in turn causes different solution rates of the developer, so that areas exposed by scattered electrons are also partly removed.

In 1975 Chang[1] suggested modeling the scattering as a convolution of the incident dosage distribution with a proximity function. The proximity function describes the distribution of the received exposure at some resist depth for a pointlike incident electron beam. His proximity function consists of a sum of two Gaussian terms, a narrow one for the forward scattering and a wide one for the backward scattering. This model has undergone some extensions since 1975, but is widely used to design correction algorithms for the proximity effect.

The idea of many correction techniques is to modify the incident dosage distribution so that the degradations in the developed pattern due to the proximity effect are reduced. Specifications on the accuracy of the correction depend strongly on several conditions, such as the actual physical system, the feature sizes of interest, the basic properties and tolerances on the geometry of the features, etc. Therefore, a diverse number of approaches and techniques have been developed during the last 20 years.[2–16] An issue of major concern in all the published work is the trade-off between accuracy and computational efficiency in designing the correction. However, with decreasing feature sizes more sophisticated algorithms are necessary to achieve the required resolution.

We present here an algorithm that allows correction for features in the nanometer range. It incorporates two important ideas to obtain high resolution. First, it is based on an extended mathematical model that includes the nonlinear effects of the developing process. Second, it computes an individual incident dosage value for every incident beam location to achieve the highest possible resolution.

In Sec. II the assumed model used is described in detail. For an efficient implementation of proximity effect correction, fast Gaussian filters are necessary. This issue is addressed in Sec. III. Section IV then defines a convex error function which reduces the problem of proximity effect correction to a constrained nonlinear optimization problem. The final algorithm presented in Sec. V is a hybrid of gradient projection and conjugate gradient algorithms. Simulation results and an error analysis are given in Secs. VI and VII.

In the following we use the term exposure for the received energy in the resist due to incident electrons and scattering effects.

II. DESCRIPTION OF THE MODEL

The model for the overall lithography process consists of two steps, the scattering of electrons described by the proximity function and the chemical developing step leading to a three-dimensional resist relief. Recent results[17–20] showed that Chang's double Gaussian model for the proximity function $F(x,y)$ is not accurate enough if very small feature dimensions are approached. Since additional Gaussian terms[19,20] diminish this problem, we define $F(x,y)$ to be the general sum of n Gaussians:

$$F(x,y)=\frac{1}{\pi\sum_{i=1}^{n}\eta_i}\sum_{i=1}^{n}\frac{\eta_i}{\sigma_i^2}\,e^{-(x^2+y^2)/\sigma_i^2}. \tag{1}$$

The η_i define the contribution of each term. In this form $F(x,y)$ is normalized:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(x,y)\,dx\,dy=1. \tag{2}$$

In all devised proximity correction algorithms, it is implicitly assumed that $F(x,y)$ is space invariant and that at every

[a]Electronic mail: jim.mcclellan@ee.gatech.edu

pixel linear superposition holds for the exposure. Thus, the exposure distribution in the resist $E(x,y)$ can be determined by a two-dimensional convolution of the applied incident dose distribution $D(x,y)$ and the proximity function $F(x,y)$:

$$E(x,y)=cD(x,y)**F(x,y), \tag{3}$$

where c is a constant factor and ** denotes a two-dimensional convolution.

The lithographical process is stepped in horizontal directions, so that the beam is applied at discrete rectangular grid points with an equal spacing Δ. Since the incident beam has a Gaussian distribution for most lithographical equipment, it can simply be incorporated into $F(x,y)$ by adjusting the σ_i in Eq. (1). The incident dose distribution $D(x,y)$ is then

$$D(x,y)=\sum_{k=0}^{M-1}\sum_{l=0}^{N-1}\tilde{d}(k,l)\,\delta(x-k\Delta,y-l\Delta), \tag{4}$$

where M and N are the number of grid points in the x and y direction and $\tilde{d}(k,l)$ is the dose applied at the grid point $(k\Delta,l\Delta)$. Thus, $E(x,y)$ can be determined by the superposition of a finite number of continuous functions. In particular, if we are only interested in discrete points on a rectangular grid with spacing Δ/v with v being an integer, then $E[n_1(\Delta/v),n_2(\Delta/v)]$ can be determined by a discrete convolution.

If we define

$$d(n_1,n_2)=\begin{cases}\dfrac{1}{D_0}\tilde{d}\left(\dfrac{n_1}{v},\dfrac{n_2}{v}\right) & \text{if } \dfrac{n_1}{v} \text{ and } \dfrac{n_2}{v} \text{ are integers}\\ 0 \quad \text{otherwise}\end{cases}, \tag{5}$$

where D_0 is a normalization dose, and

$$f(n_1,n_2)=\Delta^2 F\left(n_1\frac{\Delta}{v},n_2\frac{\Delta}{v}\right) \tag{6}$$

then a discrete convolution equivalent to Eq. (3) can be written in a unit-free form:

$$e(n_1,n_2)=\frac{cD_0}{E_0\Delta^2}\,d(n_1,n_2)**f(n_1,n_2), \tag{7}$$

where E_0 is a normalization factor for the exposure. It is important to point out that Eq. (7) is exact. It is not the result of discretization by approximating a continuous integral by a discrete sum, as it is often done in modeling of continuous systems. For this research, we always take $v=1$ so that $d(n_1,n_2)=(1/D_0)\tilde{d}(n_1,n_2)$.

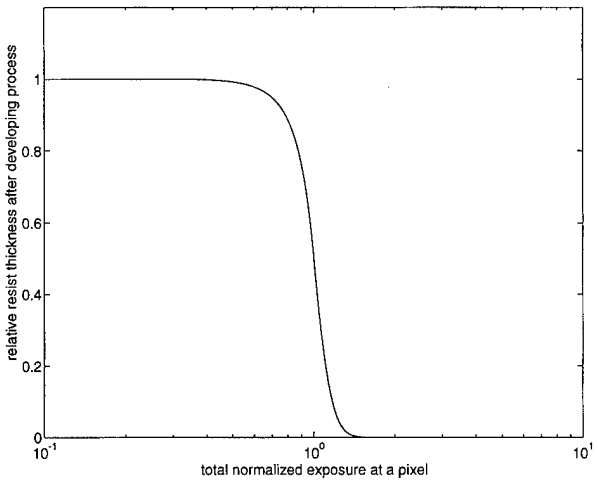


FIG. 1. Characteristic curve for a positive resist.

In the subsequent developing step, the exposed areas are removed in the case of a positive resist. The process itself depends on many parameters such as resist composition, thickness of the resist, type and application of the developer, bake conditions, etc. However, for a fixed set of these conditions, a resist can be characterized by a nonlinear curve (Fig. 1) giving the amount of resist removed as a function of the exposure.[21] The exposure corresponding to a resist removal of 50% is normalized to 1. Hence, if we neglect effects due to horizontal etching and a nonuniform energy distribution in the vertical direction, the developing process can be described by a nonlinear, point-wise function $s(\cdot)$ which applies to every pixel. The two intervals of zero slope of $s(\cdot)$ in Fig. 1 indicate regions where the resist stays completely unchanged or is totally removed. They can be exploited to achieve better corrections.

The overall two-step discrete pattern formation process can now be written as

$$p(n_1,n_2)=s[d(n_1,n_2)**f(n_1,n_2)], \tag{8}$$

where $p(n_1,n_2)$ is the final resist pattern. A block diagram is shown in Fig. 2. The normalization constants D_0 and E_0 were incorporated into the incident dosage distributions $d(n_1,n_2)$.

The objective is to compute a $d(n_1,n_2)$ such that $p(n_1,n_2)$ approximates a desired pattern as closely as possible. The only desired resist patterns considered here are those which are strictly binary and which can be represented

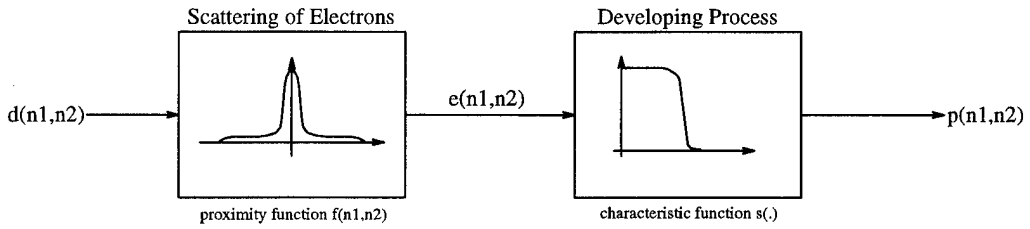


FIG. 2. Model for the overall proximity effect.

accurately by a discrete pattern $p_{\text{des}}(n_1,n_2)$ on the grid introduced above. Furthermore, we assume in our approach that there is no additional exposure from outside the pattern. For convenience, pixels of the desired pattern with completely removed resist will be called zero pixels or zero regions, and pixels with the original resist thickness one pixels or one regions, respectively.

To obtain a more compact representation of the following mathematical derivations, a matrix notation widely used in image processing is adopted here. We use lexicographical ordering of the two-dimensional signals to represent them as column vectors (bold small letter). The convolution with the proximity function $f(n_1,n_2)$, can be written as a matrix multiplication with a block Toeplitz matrix $\mathbf{F}$ (bold capital letter). The point-wise nonlinearity $s(\mathbf{v})$ would be applied to every element of the vector $\mathbf{v}$.

III. EFFICIENT GAUSSIAN FILTERS

For high-energy electron beams, the spread of the backscattered electrons can be several micrometers, which translates for a grid spacing of a few nanometers into a two-dimensional proximity function of several hundred pixels in radius. Since our proposed proximity correction technique will be iterative, numerous convolutions of the pattern with the proximity function will be necessary. Thus, very efficient Gaussian filters[22,23] are needed to compute corrections in a reasonable time.

The most efficient way to implement a good approximation of a Gaussian filter is to exploit the Central Limit Theorem, which states, in a general form, that the convolution of a large number of positive functions with finite first and second moments converges under mild constraints in the limit to a Gaussian function. Thus, a Gaussian filter can be implemented by iteratively convolving a signal with an arbitrary, but simple, positive kernel. The uniform box filter is very well suited for that purpose.[24,25] A box filter of length N and gain N can be implemented by the recursive equation

$$y[n]=y[n-1]+x[n]-x[n-N] \tag{9}$$

with only one addition and one subtraction per output pixel.

This idea can be used to implement good approximations of two-dimensional Gaussian filters with only 17 floating point operations per output pixel by applying two box filters along each axis and along each 45° diagonal. Furthermore, the amount of computation can be reduced to less than eight floating point operations for wide Gaussian filters, because methods from multirate signal processing can be applied. If less accuracy is acceptable, implementations with fewer than two floating point operations per pixel are possible. A detailed discussion of the resulting filter structures and a stability analysis for cascaded box filters can be found elsewhere.[22,23] The important result here is that Gaussian filters of any size can be implemented with only a few floating point operations per pixel. Thus, there is no computational difference between narrow Gaussians for the forward scattered and wide Gaussians for the backward scattered electrons. These can be treated simultaneously, in the 2D convolution of Eq. (8).

IV. DEFINITION OF A CONVEX ERROR FUNCTION

A specific error measure has to be defined in order to compute an optimal incident dose distribution $d(n_1,n_2)$. We consider the squared Euclidean norm of the distance between the actual developed pattern and the desired pattern. The proximity effect correction problem is then

$$\min_{\mathbf{d}} \sum_i (\mathbf{p}_{\text{des}_i}-s(\mathbf{Fd})_i)^2, \quad \mathbf{d}\geqslant\mathbf{0}, \tag{10}$$

where $s(\mathbf{Fd})_i$ denotes the ith element of the vector $s(\mathbf{Fd})$. The non-negativity of $\mathbf{d}$ is significant, because a negative value would imply the removal of deposited energy. This definition of the error function (10) has the disadvantage of not being convex, because $s(\cdot)$ is a nonconvex function. Therefore, general algorithms widely used in nonlinear programming may converge only to suboptimal solutions. This problem can be overcome by modifying Eq. (10) so that there is a different error function $e_i(\cdot)$ for every pixel. Let $e_i(\cdot)$ be the error function for the exposure at vector element $(\mathbf{Fd})_i$, then we can write the total squared error as

$$E(\mathbf{d})=\sum_i e_i^2[(\mathbf{Fd})_i]. \tag{11}$$

Due to the binary nature of $\mathbf{p}_{\text{des}}$, it is sufficient to introduce only two different error functions, one for the zero pixels and the other for the one pixels. To incorporate the nonlinearity $s(\cdot)$ we define two new functions, $s_1(\cdot)$ and $s_0(\cdot)$:

$$s_0(x)=\begin{cases} s(x) & \text{if } x\in[1,\infty) \\ s'(1)(x-1)+s(1) & \text{if } x\in[0,1) \end{cases}, \tag{12}$$

$$s_1(x)=\begin{cases} s(x) & \text{if } x\in[0,1) \\ s'(1)(x-1)+s(1) & x\in[1,\infty) \end{cases}, \tag{13}$$

where $s'(\cdot)$ is the first derivative of $s(\cdot)$. Both $s_1(\cdot)$ and $s_0(\cdot)$ are continuously differentiable and they allow a straightforward definition of $e_0(\cdot)$ and $e_1(\cdot)$ for zero or one pixels, respectively,

$$e_0(\cdot)=s_0(\cdot), \tag{14}$$

$$e_1(\cdot)=1-s_1(\cdot). \tag{15}$$

It can easily be shown that the resulting $E(\mathbf{d})$ is convex,[22] but not strictly convex. Hence, there may be more than one global minimum point, but the set of all global minima is convex.[25]

In the case of a positive resist, a zero region of a pattern will develop correctly, if the remaining resist thickness is zero. On the other hand, a one region of a pattern will develop correctly if the remaining resist layer is sufficiently thick. This requirement may even be met when about half of the thickness of the resist layer is removed. Thus, the deviations in one and zero regions from the desired pattern are of very different significance in terms of a good correction. This

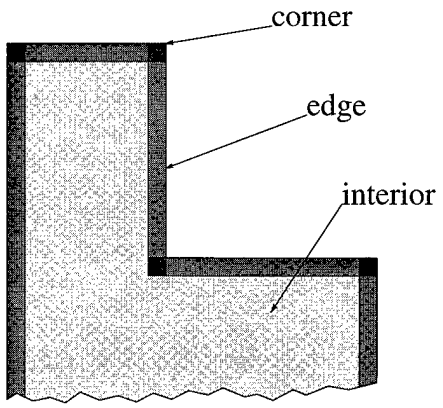


FIG. 3. Distribution of the weighting factors in the weighting matrix **W**.

observation can be incorporated into our correction technique by using an individual weighting factor $\mathbf{w}_i$ for every pixel. In general, the larger $\mathbf{w}_i$ is, the smaller will be the local error. The best results are achieved by using different weights at edges, corners, and in the interior of the zero regions (see Fig. 3) and unit weight in one regions. The weights should be largest for the interior part of zero regions, less at the edges, and smallest at corners. Such a distribution allows us to change the location of transition regions to a certain extent.

Putting all this together, we can formulate the following minimization problem. Let $e(\cdot,\cdot)$ be a vector element-wise function of two vector arguments:

$$e(\mathbf{x}_i,\mathbf{y}_i)=\begin{cases} e_1(\mathbf{x}_i) & \text{if } \mathbf{y}_i=1 \\ e_0(\mathbf{x}_i) & \text{if } \mathbf{y}_i=0 \end{cases}. \tag{16}$$

The proximity correction problem becomes a nonlinear, convex, constrained optimization problem

$$\min_{\mathbf{d}}[e(\mathbf{Fd},\mathbf{p}_{\text{des}})^T\mathbf{W}e(\mathbf{Fd},\mathbf{p}_{\text{des}})], \quad \mathbf{d}\geqslant\mathbf{0}, \tag{17}$$

where **W** is a diagonal matrix with the weighting factors $\mathbf{W}_{ii}=\mathbf{w}_i$. We will use the notation $E(\mathbf{d})$ for the overall error term in Eq. (17).

V. PROXIMITY EFFECT CORRECTION ALGORITHM

Since the problem of proximity effect correction is a constrained minimization problem, the well established framework of nonlinear optimization theory can be applied. A hybrid of the Fletcher–Reeves conjugate gradient algorithm,[25] Rosen's projected gradient algorithm,[25] and the Goldstein–Levitin–Polyak algorithm[27] showed best performance for this particular problem. Here we show only the central ideas and concepts behind our new algorithm. A detailed listing in pseudocode of the algorithm can be found in the Appendix.

The algorithm is a descent method and has, therefore, the underlying basic form

$$\mathbf{d}_{k+1}=\mathbf{d}_k+\alpha_k\mathbf{h}_k, \tag{18}$$

where $\mathbf{d}_k$ is the incident dosage vector after k iterations. The sequence $\{\mathbf{d}_k\}$ constitutes the vertices of a polygon in the multidimensional search space, which converges to the solution vector $\mathbf{d}^*$. Every segment of the polygon is determined in two steps. First, a descent direction $\mathbf{h}_k$ is determined and second, an appropriate step length α_k is chosen such that

$$E(\mathbf{d}_{k+1})\leqslant E(\mathbf{d}_k). \tag{19}$$

The non-negativity constraint in Eq. (17) can be incorporated by applying a closest point projection P into the set of non-negative points after every iteration. To ensure convergence, more stringent requirements than in the unconstrained case have to be placed on α_k.[4] For the new proximity effect correction algorithm, diverse experiments showed that α_k can be determined by solving the unconstrained minimization

$$\alpha_k=\arg\min_{\alpha} E(\mathbf{d}_k+\alpha\mathbf{h}_k). \tag{20}$$

This yields an appropriate α_k with a minimum of computations. The total error $E(\mathbf{d})$ is locally approximately quadratic, so that Eq. (20) can be reduced to minimizing a parabola. A detailed step size rule for α_k can be found in Appendix 2.

If an element $\mathbf{d}_{k_i}$ of the vector $\mathbf{d}_k$ is zero, then $\mathbf{d}_{k_i}$ lies on the border of the set of non-negative points, and the corresponding constraint is said to be active. For $\mathbf{h}_k$ we use a conjugate direction vector determined in the same way as in the Fletcher–Reeves conjugate gradient algorithm.[25] However, we restrict $\mathbf{h}_k$ to the space induced by the current set of nonactive constraints by setting all $\mathbf{h}_{k_i}$ to zero for which $\mathbf{d}_{k_i}=0$. This implies that once a constraint becomes active, it stays active at least until a local minimum is found. Hence, particular care must be taken in the choice of the initial incident dosage distribution $\mathbf{d}_0$.

Experiments have shown that the convergence rate during the first few iterations is higher when those constraints active in $\mathbf{d}^*$ are already active in $\mathbf{d}_0$. In several publications it is assumed that $\mathbf{d}^*$ is zero wherever $\mathbf{p}_{\text{des}}$ is zero. However, it turns out that admitting incident dosage in zero regions around convex corners of the pattern (see Fig. 4) improves the correction results. We, therefore, use a starting distribution of the form

$$\mathbf{d}_0=\beta\mathbf{p}_{\text{des}}+\epsilon\mathbf{c}, \tag{21}$$

where **c** is a binary pattern. Figure 5 shows **c** as gray regions around the convex corners of the black zero regions. Since the term $\epsilon\,\mathbf{c}$ is used only to ensure that the constraints around the convex corners are not active, ϵ can be chosen very

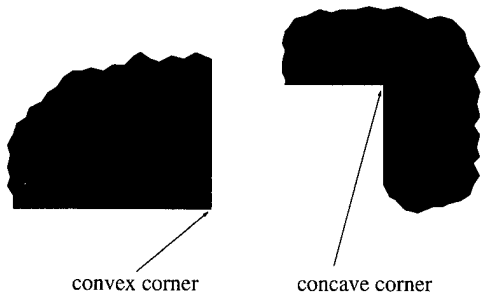


FIG. 4. Explanation of convex and concave corners.

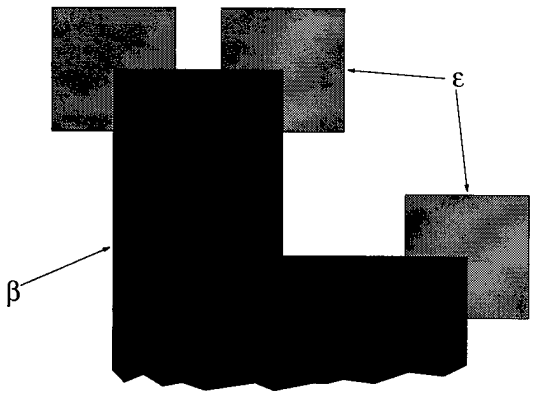


FIG. 5. General form of the starting incident dosage distribution for the iterative algorithm. Black regions have value β and gray regions have value ϵ.

small. The scaling factor β should minimize $E(\beta \mathbf{p}_{\text{des}})$. It takes into account that the required incident dosage depends among other things on the density of the pattern, because the scattering of electrons directs some of the incident dosage to the zero regions of $\mathbf{p}_{\text{des}}$. Hence, it holds in general that the smaller the pattern density is, the higher the incident dosage has to be, since a higher portion of the dosage is directed to insignificant areas.

VI. PERFORMANCE OF THE ITERATIVE ALGORITHM

A. Test data

For test purposes a set of parameters for the model was defined using the data for a 30 kV electron beam onto silicon substrate given by Gentilli *et al.*[28] The parameters for the proximity function $F(x,y)$ in Eq. (1) are $n=3$, $\eta_1=1$, $\sigma_1=0.053$ μm, $\eta_2=0.587$, $\sigma_2=3.55$ μm, $\eta_3=0.087$, and $\sigma_3=0.27$ μm. We use in Eq. (4) a grid of $\Delta=10$ nm spacing. The Gaussian with σ_1 was implemented as a truncated, separable finite impulse response (FIR) filter, and the Gaussians with σ_2 and σ_3 were computed with the efficient Gaussian approximation outlined in Sec. III. We use the data for a high contrast, positive polymethylmethacrylate (PMMA) developer with contrast $\gamma=9.4$ (see Fig. 1).[28]

A 4×4 μm^2 test pattern was designed. In Fig. 6 the zero regions of the pattern are black and the one regions are white. The test pattern contains primarily small features with sizes down to 50 nm, which is less than σ_1. The most difficult features to correct are the square holes with adjacent corners and sidelengths of 50, 100, 200, and 300 nm (lower right corner of Fig. 6), the small square with a sidelength of 50 nm (upper left corner of Fig. 6), the double cross structure with linewidths of 50 and 100 nm (upper left part of Fig. 6), the pine tree structure with a buried inner line (upper right part of Fig. 6), and closely spaced narrow lines with a linewidth and a line spacing of 50 nm each (outer left part of Fig. 6).

It is important to point out that only minor effects due to the backscattering of electrons for the test pattern are encountered, because the test pattern is small compared to the range of the backscatter Gaussian. Thus, the background dosage due to this effect is small. As this work is primarily dedicated to the correction of the proximity effect in nanolithography, we are mainly interested in the ability of the proposed algorithm to correct small features and details primarily degraded by the forward scattering. For that purpose the size of the test pattern in Fig. 6 is sufficient.

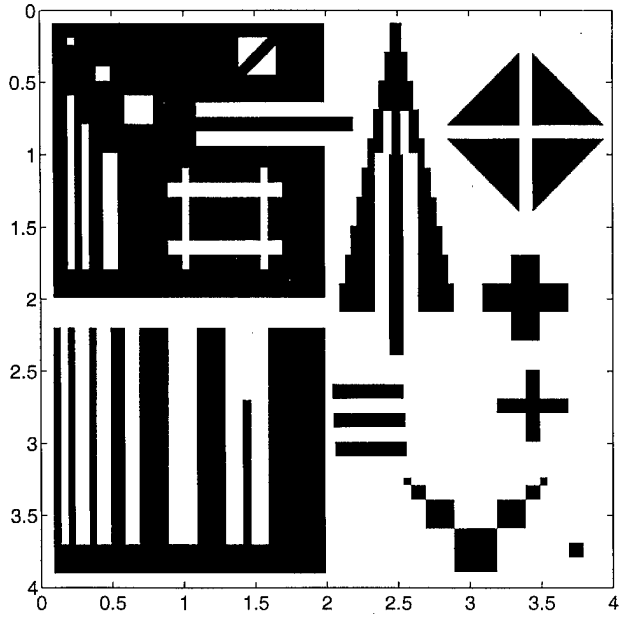


FIG. 6. Test pattern of size 4×4 μm^2.

B. Correction of the test pattern

We evaluate the performance of the algorithm within the model defined in Sec. II and under the various assumptions made about the actual physical process. Within the first few iterations of the correction algorithm the error decreases rapidly. For larger numbers of iterations the decrease slows down. Therefore, it seems reasonable not to run the algorithm until it converges to an optimal solution, but to use the result after about 100–200 iterations as the proximity corrected input pattern for the lithographical process. Furthermore, we have observed that the correction becomes more sensitive to inaccuracies in the model parameters when the number of iterations is high.

To evaluate the quality of the correction we create gray slice plots of the developed corrected pattern for three distinct regions. A zero region is assumed to be completely developed out if the remaining resist thickness is less than 2% of the original thickness and will be marked black. On the other hand, a one region of the pattern is developed correctly if the remaining resist layer is sufficiently thick (we require at least 60% of the original thickness). The corresponding regions will be marked white. The remaining pixels with intermediate values specify transition regions and are marked in gray. Figure 7 shows this kind of gray slice plot

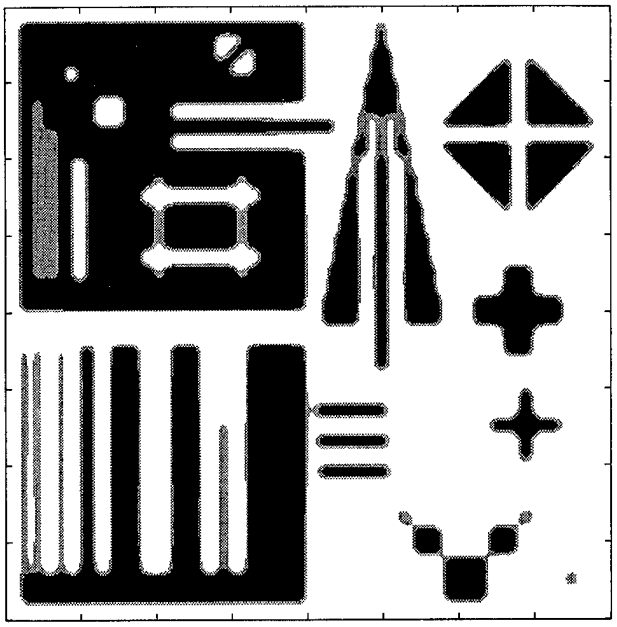


FIG. 7. Gray slice image of the developed pattern without proximity effect correction.

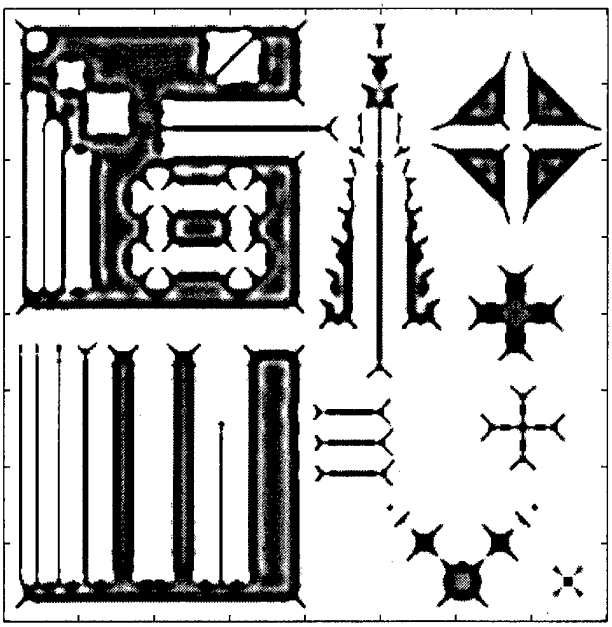


FIG. 9. Proximity effect corrected incident dosage after 150 iterations. Black denotes the highest incident dosage.

for the developed, uncorrected test pattern. Small features of 100 nm width and less exhibit significant deviations.

In our implementation of the correction algorithm, we used for the weighting matrix $\mathbf{W}$ the weights 350:5:1 for the interior, the edges, and the corners of the zero regions, respectively. Figure 8 shows the gray slice plot for the developed corrected pattern after 150 iterations. Nearly all features are developed out correctly. The calculated incident dosage distribution can be seen in Fig. 9. Obviously, the optimal incident dosage distribution is a combination of dose modulation and edge displacement.

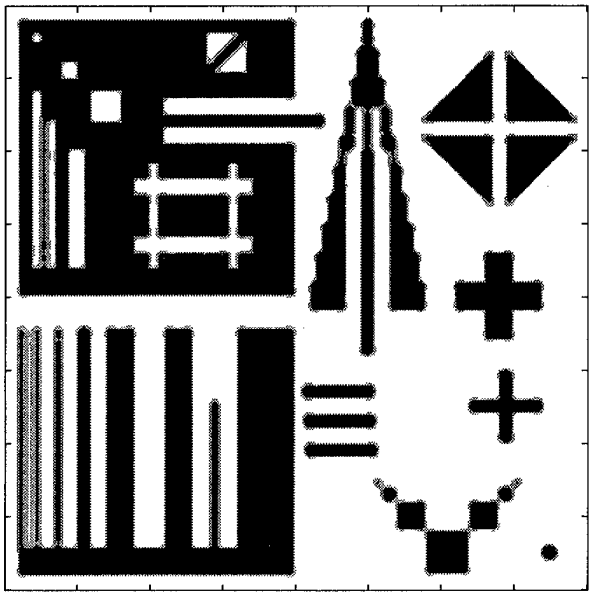


FIG. 8. Gray slice image of the developed pattern after 150 iterations.

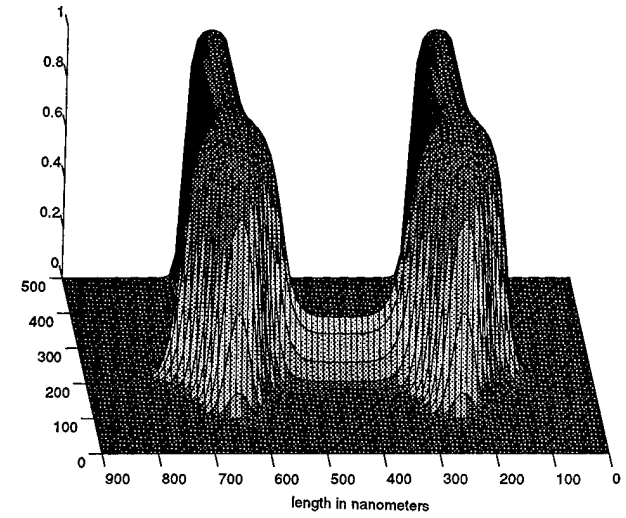
A better understanding of the improvements in the pattern obtained by our algorithm can be seen by examining small or complicated features more closely. Two such regions in Fig. 6 are the double cross and the narrow adjacent lines in the comb structure. Figures 10(a) and 10(b) show a part of the developed double cross structure without correction and after 150 iterations, respectively. The improvements are remarkable. In the case of the narrow adjacent lines (left lower corner of Fig. 6) in Figs. 11(a) and 11(b), the separating resist wall between the two leftmost lines has reduced height and wide transition regions. Here the physical resolution limits of the system are reached.

VII. ERROR ANALYSIS

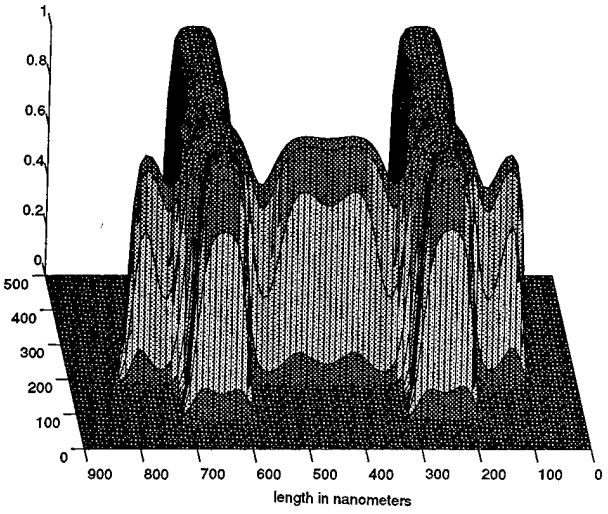
In the evaluation of the algorithm it was assumed that the model parameters were known exactly. The determination of the precise proximity function is, however, difficult and measured parameters for similar systems vary significantly in the literature. Furthermore, the lithographical equipment might only be able to change the incident dosage at a point in certain discrete steps, whereas the corrected incident dosage distribution can contain any positive value. Thus, the actual applied incident dosage distribution would be a quantized version of the corrected one. In this section, we will assess the influence of model parameter inaccuracies and of quantization effects on the quality of the developed pattern.

A. Inaccuracies of the proximity function

For our examinations we use corrected incident dosage distributions $\hat{\mathbf{d}}_k$ after k iterations for the test pattern and the model parameters of Sec. VI. We then determine $\tilde{E}(\hat{\mathbf{d}}_k)$ for perturbed proximity functions $\tilde{F}$ to simulate the usage of inaccurate model parameters. In every experiment only one

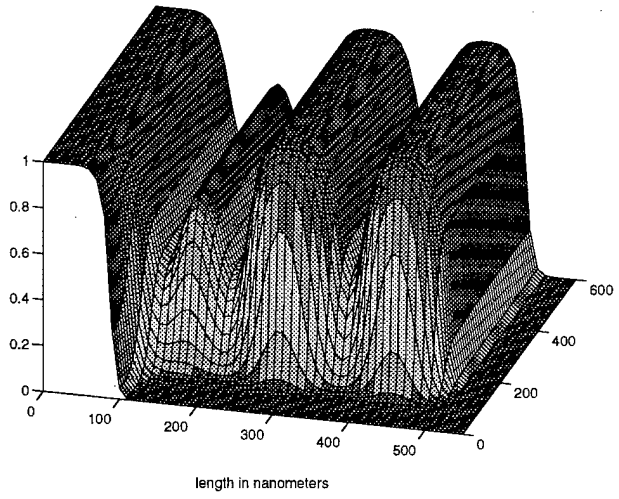


(a)

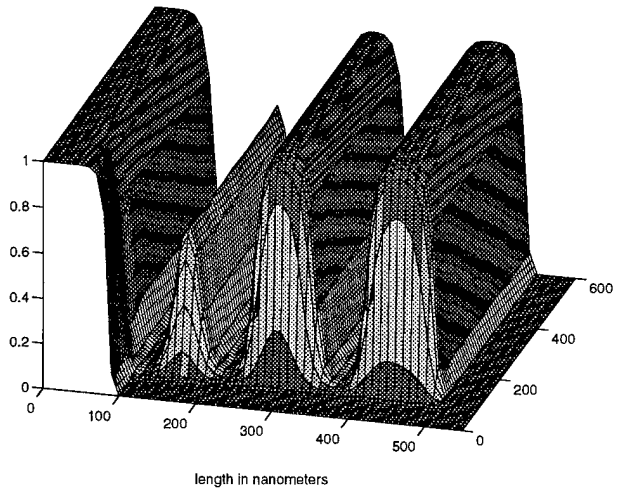


(b)

FIG. 10. A part of the developed double cross structure (a) without correction and (b) after 150 iterations.



(a)



(b)

FIG. 11. Developed 50 nm narrow lines in comb structure (a) without correction and (b) after 150 iterations.

Gaussian was modified by varying σ_i but the general form of a triple Gaussian proximity function was maintained. We denote the standard deviation of the modified Gaussian as $\tilde{\sigma}_i$. It represents the proximity function during the actual lithographical process. In the experiments $\tilde{\sigma}_i$ was varied by 20% around σ_i. Thus, a $\tilde{\sigma}_i$ larger than σ_i was used to simulate too small model parameters, and vice versa. Figure 12 shows the results for the wide Gaussian with standard deviation σ_2 modeling the backward scattering of the electrons. Obviously $\tilde{E}(\hat{\mathbf{d}}_k)$ is more sensitive to too large model parameters than to too small ones. The effect of varying σ_2 slightly decreases with the number of iterations. The main conclusions for the wide Gaussian σ_2 are that small deviations of σ_2 have only little influence on $\tilde{E}(\hat{\mathbf{d}}_k)$, that this characteristic is fairly independent of the number of iterations, and that it is preferable to use a σ_2 in the model which is too small rather than too large.

Variations of the narrow Gaussian with σ_1 modeling the forward scattering of the electrons (see Fig. 13) reveal a completely different behavior. With an increasing number of iterations the sensitivity of $\tilde{E}(\hat{\mathbf{d}}_k)$ for $\tilde{\sigma}/\sigma<1$ increases dramatically. Of particular importance are the intersections of error curves for different numbers of iterations. They show that inaccuracies in the parameters for the forward scattering Gaussian can actually deteriorate the correction results when the number of iterations is increased. This effect can easily be explained. The corrected incident dosage distribution after 150 iterations (see Fig. 9) compensates for the proximity effect not only by modulation of the incident dosage but also

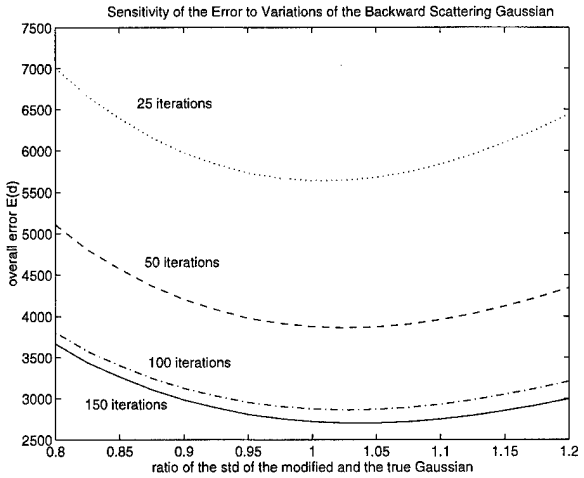


FIG. 12. Sensitivity of $\tilde{E}(\hat{\mathbf{d}}_k)$ to variations of the standard deviation σ_2. The x axis is given in $\tilde{\sigma}_2/\sigma_2$.

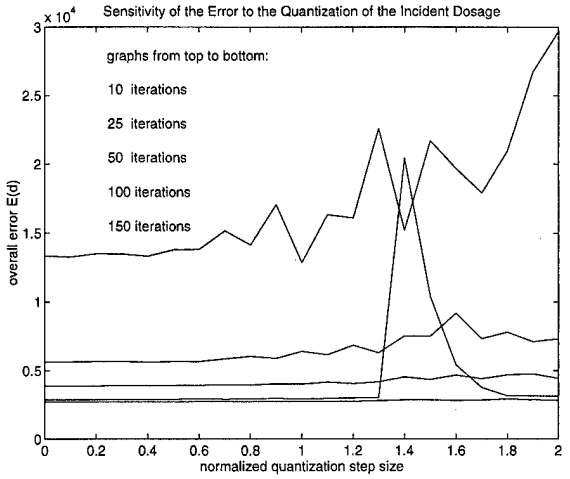


FIG. 14. Sensitivity of $E(Q\hat{\mathbf{d}}_k)$ to uniform quantization for different quantization step sizes.

by exploiting the forward scattering of the electrons itself. Some regions, especially along edges and at concave corners, are not exposed to any incident electrons, only to scattered ones. The correct exposure in these regions depends completely on the accuracy of the narrow model Gaussian. If it is larger than the Gaussian in the lithographical process, significant pattern deteriorations occur, as some areas do not receive any exposure. In general it is advantageous to use σ_1 too small rather than too large.

The influence of the third Gaussian term is insignificant. For the same kind of experiments the increase of $\tilde{E}(\hat{\mathbf{d}}_k)$ due to variations of σ_3 never exceeded 1%. The main result of this analysis is that the narrow Gaussian is the most critical part of the proximity function. Its accuracy determines the achievable quality in the corrected developed pattern. Since the experimental determination of the exact proximity function in general and of the narrow Gaussian in particular is very difficult, the best set of model parameters probably has to be determined by several test corrections for slightly modified, experimentally determined sets of parameters.

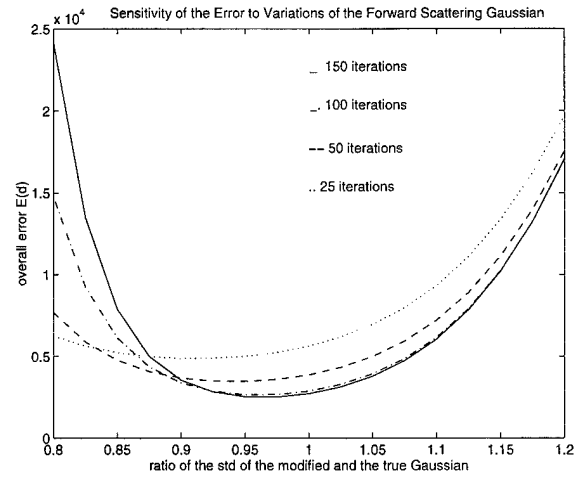
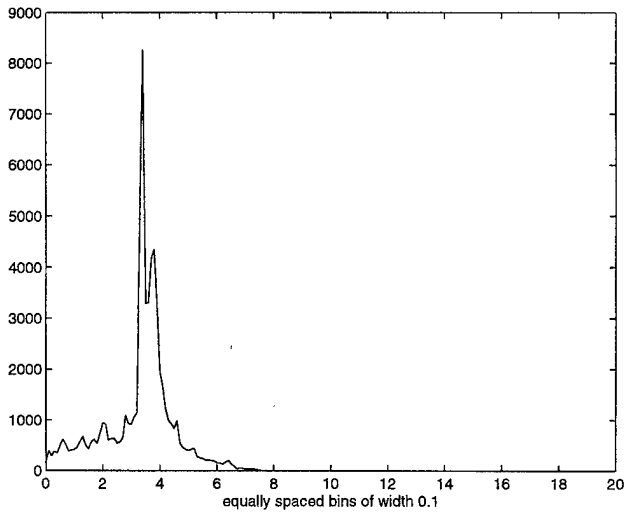


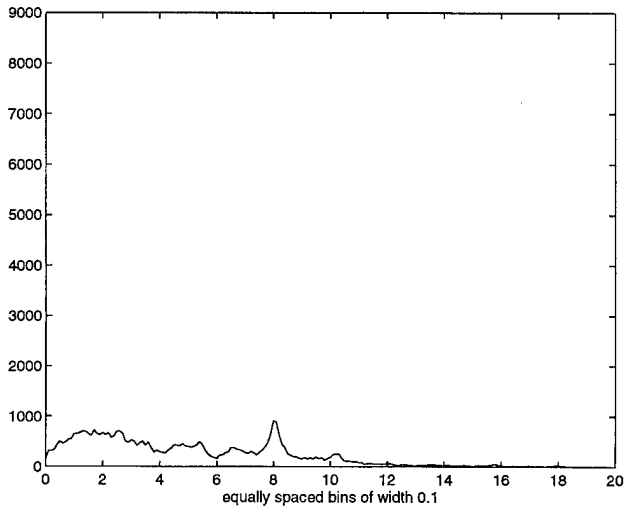
FIG. 13. Sensitivity of $\tilde{E}(\hat{\mathbf{d}}_k)$ to variations of the standard deviation σ_1. The x axis is given in $\tilde{\sigma}_2/\sigma_2$.

B. Errors due to quantization

To examine the influence of the dosage quantization, $E[Q(\mathbf{d}_k)]$ was computed for uniform midtread quantizers $Q(\cdot)$ with different quantization step sizes. In lithography systems the minimum quantization step size is given by some minimum shutter time. Figure 14 shows the results for the corrected incident dosage after 10, 25, 50, 100, and 150 iterations and normalized quantization step sizes from 0 to 2. The normalization constant is the same as introduced in Sec. II, i.e., a normalized incident dosage of magnitude one over a large area leads to 50% removal of the resist during the developing process. The sensitivity of the error to quantization seems to decrease in general with an increasing number of iterations of the correction algorithm. This can be explained by examining the development of the histogram of nonzero values of the vector elements of $\hat{\mathbf{d}}_k$ for increasing k. The histogram of the uncorrected incident dosage values consists of one main peak at β and a small peak at ϵ in Eq. (21). Thus, the quantization introduces the same quantization error at all points of $\hat{\mathbf{d}}_0$ with value β or ϵ, respectively. Hence, $\hat{\mathbf{d}}_0$ is very sensitive to quantization. After a few iterations the histogram of $\hat{\mathbf{d}}_k$ is still mainly concentrated around β [see Fig. 15(a)]. However, for an increasing number of iterations the histogram of $\hat{\mathbf{d}}_k$ widens and loses its distinct peak [see Fig. 15(b)]. A uniform quantizer is optimal in the mean square sense for uniform distributions. Thus, with an increasing number of iterations the sensitivity of the overall error to quantization decreases in general. However, as we can see for the case of 100 iterations, extraordinarily large errors can be introduced under certain conditions. In this particular case, the incident dosage of a large part of a zero region was mapped to a too small value. Due to the high weighting factor for zero regions in the error norm (17), the increase in the overall error becomes very significant. It is



(a)



(b)

FIG. 15. Histograms for nonzero values of the incident distributions after (a) 50 and (b) 150 iterations.

interesting that this effect occurs only for a certain number of iterations and disappears again for higher k, as the distribution of $\hat{\mathbf{d}}_k$ changes. Figure 16 shows the gray slice plot for $\hat{\mathbf{d}}_{150}$ developed with subsequent quantization of step size one and a modified $\tilde{\mathbf{F}}$ in the lithographical process with a $\tilde{\sigma}_2$ that is 5% smaller than the σ_2 in the correction algorithm. The quality of the result is still very good and only little additional artifacts appear.

VIII. CONCLUSIONS

A new algorithm to compute corrections to the proximity effect in the nanometer range was presented. It exploits properties of the proximity effect and the subsequent chemical developing process. The devised mathematical model includes a convex error norm, so that the problem of proximity effect correction can be reduced to a nonlinear, convex, and

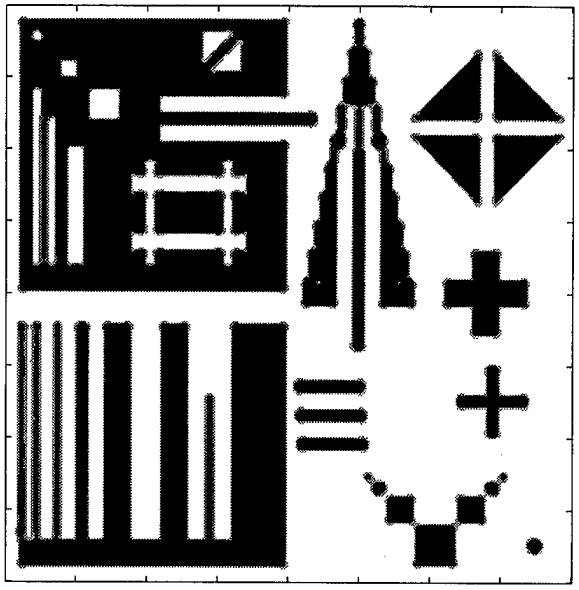


FIG. 16. Gray slice image of the developed pattern with corrected and quantized incident dosage distribution after 150 iterations and a modified proximity function in the lithographical process.

constrained optimization problem. The suggested correction algorithm is a hybrid of different conjugate gradient and gradient projection algorithms. Since the incident dosage distribution has to be convolved several times with the proximity function during the correction, very efficient Gaussian filters were used. The performance of the algorithm was demonstrated for a test pattern with features as small as 50 nm. Finally, the effect of inaccuracies in the model parameters and of quantization of the corrected incident dosage distribution was analyzed. The Gaussian term modeling the forward scattering of the electrons has to be known precisely in order to achieve the high resolution of this correction technique. The quantization effects, on the other hand, are not that significant, so that good results should be achievable on most lithographical equipment.

ACKNOWLEDGMENT

J.H.M. is supported by the Joint Services Electronics Program under Contract No. DAAH-04-93-G-0027.

APPENDIX

1. The algorithm

First of all, some operators have to be defined. The closest point projection P into the set of non-negative points is

$$P\mathbf{d}=\max(\mathbf{0},\mathbf{d}), \tag{A1}$$

where $\max(\mathbf{0},\mathbf{d})$ is the vector, whose ith element is the maximum of the set $\{0,d_i\}$. The set of active constraints at $\mathbf{d}$ is given as

$$A(\mathbf{d})=\{i:d_i=0\}. \tag{A2}$$

Of interest is the subspace S_Q of $\mathbf{R}^N$, which is induced by a subset $Q\subseteq\{1,2,\ldots,N\}$ as

$$S_Q=\{\mathbf{x}\in\mathbf{R}^N: x_i=0 \quad \text{if } i\in Q\}. \tag{A3}$$

Particularly important is $S_{A(\mathbf{d})}$, the linear subspace defined by the set of active constraints at $\mathbf{d}$. C_Q denotes the linear, orthogonal projection onto S_Q. It is

$$C_Q\mathbf{u}=\mathbf{u}; \quad \forall\mathbf{u}\in S_Q, \tag{A4}$$

$$\langle(I-C_Q)\mathbf{u},C_Q\mathbf{u}\rangle=0; \quad \forall\mathbf{u}\in\mathbf{R}^N, \tag{A5}$$

where I is the identity operator, $\langle\cdot,\cdot\rangle$ denotes the inner product. C_Q can be implemented very efficiently, as the ith vector element of $C_Q\mathbf{u}$ has the form

$$(C_Q\mathbf{u})_i=\begin{cases}0 & \text{if } i\in Q\\ u_i & \text{if } i\notin Q\end{cases}. \tag{A6}$$

The conditions for a global solution $\mathbf{d}^*$ of Eq. (17) are

$$C_{A(\mathbf{d}^*)}\nabla E(\mathbf{d}^*)=\mathbf{0}, \tag{A7}$$

$$(I-C_{A(\mathbf{d}^*)})\nabla E(\mathbf{d}^*)\geq\mathbf{0}. \tag{A8}$$

Using these definitions, the algorithm can be written in the following form:

(1) Use Eq. (21) as the incident dosage $\mathbf{d}_0$.
(2) Determine the set of active constraints $Q=A(\mathbf{d}_0)$.
(3) Compute the projected gradient $\mathbf{g}_0=\mathbf{C}_Q\nabla E(\mathbf{d}_0)$ and set $\mathbf{h}_0=-\mathbf{g}_0$.
(4) For $k=0,1,\ldots,n-1$:
  (a) Use the step size rule of Appendix 2 to determine α_k.
  (b) Set $\mathbf{d}_{k+1}=P(\mathbf{d}_k+\alpha_k\mathbf{h}_k)$.
  (c) Determine the set of active constraints $Q=A(\mathbf{d}_{k+1})$.
  (d) If $k<n-1$:
  (e) Compute the projected gradient $\mathbf{g}_{k+1}=C_Q\nabla E(\mathbf{d}_{k+1})$.
  (f) If $\mathbf{g}_{k+1}\neq\mathbf{0}$:
  (g) —Compute $\beta_k=\langle\mathbf{g}_{k+1},\mathbf{g}_{k+1}\rangle/\langle\mathbf{g}_k,\mathbf{g}_k\rangle$;
  (h) —set $\mathbf{h}_{k+1}=-\mathbf{g}_{k+1}+\beta_k C_Q\mathbf{h}_k$.
  (i) Else $\mathbf{g}_{k+1}=\mathbf{0}$:
  (j) —Determine $\mathbf{g}^{\perp}_{k+1}=(I-C_Q)\nabla E(\mathbf{d}_{k+1})$;
  (k) —if $\mathbf{g}^{\perp}_{k+1}\geq\mathbf{0}$, $\mathbf{d}_{k+1}$ is the global minimum. Stop;
  (l) —else, set Q to $Q\backslash\{i: g^{\perp}_i<0\}$ and go to 3.
  (m) Next k.
(5) Replace $\mathbf{d}_0$ by $\mathbf{d}_n$ and go to 3.

In practice, n can be set to ∞. A more comprehensive discussion of the algorithm itself, its properties, and its performance in comparison to other algorithms can be found in the thesis.[22]

2. The step size rule

As the point-wise error functions $e_0(\cdot)$ in Eq. (14) and $e_1(\cdot)$ in Eq. (15) are nonlinear only over a small interval, a local linearization of $e_0(\cdot)$ and $e_1(\cdot)$ will introduce very little error, if the step length α to the local minimum of the line search problem $E(\mathbf{d}_k+\alpha\mathbf{h}_k)$ is small and if many of the points $(\mathbf{Fd})_i$ for which either $e_0(\cdot)$ or $e_1(\cdot)$ is evaluated lie in the linear regions of $e_0(\cdot)$ or $e_1(\cdot)$, respectively. In this case, $E(\cdot)$ in Eq. (17) is well approximated locally by a quadratic functional. Hence, $E(\mathbf{d}_k+\alpha\mathbf{h}_k)$ is for small α well approximated by a parabola and the minimization of the function can be replaced by the determination of the minimum of the approximating parabola. A parabola is specified uniquely by three parameters, e.g., from two values and from the derivative at one point:

$$y_0=E(\mathbf{d}_k+0\mathbf{h}_k),$$

$$y_0'=\left.\frac{dE(\mathbf{d}_k+\alpha\mathbf{h}_k)}{d\alpha}\right|_{\alpha=0}=\langle\nabla E(\mathbf{d}_k),\mathbf{h}_k\rangle,$$

$$y_1=E(\mathbf{d}_k+\alpha_{k-\eta}\mathbf{h}_k),$$

where η is some small integer. The first two values can be computed very efficiently as byproducts of step 4e so that only one extra evaluation of $e(\cdot)$ is necessary. With this data α_k is

$$\alpha_k=-\frac{\alpha^2_{k-\eta}y_0'}{2(y_1-y_0-\alpha_{k-\eta}y_0')}. \tag{A9}$$

We have to take special precautions, if Eq. (A9) provides a negative α or leads to an α which does not decrease the error norm. This problem is attacked in the following step size rule:

(1) Determine α_k with Eq. (A9).
(2) If $\alpha_k<0$:
  (a) Set $\beta_1=\alpha_{k-\eta}$.
  (b) Set $\beta_2=2\beta_1$.
  (c) Set $y_2=E(\mathbf{d}_k+\beta_2\mathbf{h}_k)$.
  (d) Determine
  $$\alpha_k=-\frac{1}{2}\frac{\beta_1^2(y_0-y_2)+\beta_2^2(y_1-y_0)}{\beta_1(y_2-y_0)+\beta_2(y_0-y_1)}.$$
  (e)If $\alpha_k<0$
  (f) Set $y_1=y_2$, $\beta_1=\beta_2$.
  (g) Go back to 2b.
(3) If $E(\mathbf{d}_k+\alpha_k\mathbf{h}_k)\geq E(\mathbf{d}_k)$
  (a) Set $\beta_1=\alpha_{k-\eta}$.
  (b) Set $\beta_2=\alpha_k$.
  (c) Set $y_2=E(\mathbf{d}_k+\beta_2\mathbf{h}_k)$.
  (d) Determine
  $$\alpha_k=-\frac{1}{2}\frac{\beta_1^2(y_0-y_2)+\beta_2^2(y_1-y_0)}{\beta_1(y_2-y_0)+\beta_2(y_0-y_1)}.$$
  (e)If $E(\mathbf{d}_k+\alpha_k\mathbf{h}_k)\geq E(\mathbf{d}_k)$
  (f) If $y_1>y_2$,
  $-y_1=y_2$, $\beta_1=\beta_2$.
  (g) Set $y_2=E(\mathbf{d}_k+\alpha_k\mathbf{h}_k)$, $\beta_2=\alpha_k$.
  (h) Go back to 3d.

Step (2) repeats parabolic fits as long as the resulting α_k is negative. The way the successive β_2 are picked ensures that this procedure needs only a finite number of steps. Step (3) also repeats parabolic fits until the overall error is decreased.

In all numerical experiments with several test patterns and hundreds of iterations, the cases step (2) or step (3) were required only very few times. Thus steps (2) and (3) are security measures rather than essential parts of the step size

rule. In almost every case, the result of Eq. (A9) is sufficient. Hence, the proposed step size rule needs in general only two additional evaluations of $E(\cdot)$.

[1]T. H. P. Chang, J. Vac. Sci. Technol. **12**, 1271 (1975).
[2]Y. C. Pati, A. Teolis, D. Park, R. Bass, K. Rhee, B. Bradie, and M. C. Peckerar, J. Vac. Sci. Technol. B **8**, 1882 (1990).
[3]P. Vermeulen, R. Jonckheere, and L. Van Den Hove, J. Vac. Sci. Technol. B **7**, 1556 (1989).
[4]C.-Y. Chang, G. Owen, F. R. Pease, and T. Kailath, SPIE **1671**, 208 (1992).
[5]F. M. H. Y. S. Okazaki, N. Saitou, and Y. Sakitani, J. Vac. Sci. Technol. B **10**, 3272 (1992).
[6]S.-Y. Lee, J. C. Jacob, C.-M. Chen, J. A. McMillan, and N. C. MacDonald, J. Vac. Sci. Technol. B **9**, 3048 (1991).
[7]K. Harafuji, A. Misaka, N. Nomura, M. Kawamoto, and H. Yamashita, IEEE Trans. Comput.-Aided Design Integr. Circuits Systems **12**, 1508 (1993).
[8]R. C. Frye, K. D. Cummings, and E. A. Rietman, SPIE **1263**, 175 (1990).
[9]H. Hübner, Microelectron. Eng. **18**, 275 (1992).
[10]J. M. Pavkovich, J. Vac. Sci. Technol. B **4**, 159 (1986).
[11]H. Eismann, T. Waas, and H. Hartmann, J. Vac. Sci. Technol. B **11**, 2741 (1993).
[12]A. M. Carroll, J. Appl. Phys. **52**, 434 (1981).
[13]E. Barouch, R. Coifman, J. Ma, M. C. Peckerar, and V. Rokhlin, SPIE **1924**, 150 (1993).
[14]M. Parikh, J. Appl. Phys. **50**, 4371 (1979).
[15]T. R. Groves, J. Vac. Sci. Technol. B **11**, 2746 (1993).
[16]M. E. Haslam and J. F. McDonald, J. Vac. Sci. Technol. B **6**, 436 (1988).
[17]T. H. P. Chang, D. P. Kern, E. Kratschmer, K. Y. Lee, H. E. Luhn, M. A. McCord, S. A. Rishton, and Y. Vladimirdky, IBM J. Res. Devel. **32**, 462 (1988).
[18]S. A. Rishton and D. P. Kern, J. Vac. Sci. Technol. B **5**, 135 (1987).
[19]M. G. Rosenfield, S. J. Wind, W. W. Molzen, and P. D. Gerber, Microelectron. Eng. **11**, 617 (1990).
[20]S. J. Wind, M. G. Rosenfield, G. Pepper, W. W. Molzen, and P. D. Gerber, J. Vac. Sci. Technol. B **7**, 1507 (1989).
[21]G. R. Brewer, *Electron-Beam Technology in Microelectronic Fabrication* (Academic, New York, 1980).
[22]R. Rau, M.S. thesis, Georgia Institute of Technology (1995).
[23]R. Rau and J. H. McClellan (to be published).
[24]W. M. Wells, III, IEEE Trans. Pattern Anal. Machine Intelligence **8**, 234 (1986).
[25]R. Andonie and E. Carai, Comput. Artificial Intelligence **11**, 363 (1992).
[26]D. G. Luenberger, *Linear and Nonlinear Programming* (Addison-Wesley, Reading, MA, 1984).
[27]D. P. Bertsekas, IEEE Trans. Autom. Control, **AC-21**, 174 (1976).
[28]M. Gentilli, L. Grella, A. Lucchesini, L. Luciani, L. Mastrogiacomo, and P. Musumeci, J. Vac. Sci. Technol. B **8**, 1867 (1990).
[29]G. H. Bernstein, D. A. Hill, and Wen-Ping Liu, J. Appl. Phys. **71**, 4066 (1992).

Independent parallel lithography using the atomic force microscope

S. C. Minne,[a] S. R. Manalis, A. Atalar,[b] and C. F. Quate[c]
E. L. Ginzton Laboratory, Stanford University, Stanford, California 94305

(Received 15 March 1996; accepted 13 May 1996)

Independent parallel features have been lithographically patterned with a 2×1 array of individually controlled cantilevers using an atomic force microscope. Control of the individual cantilevers was achieved with an integrated piezoelectric actuator in feedback with a piezoresistive sensor. Patterns were formed on ⟨100⟩ single crystal silicon by using a computer controlled tip voltage to locally enhance the oxidation of the silicon. Using the piezoresistor directly as a force sensor, parallel images can be simultaneously acquired in the constant force mode. A discussion of electrostatic forces due to applied tip voltages, hysteresis characteristics of the actuator, and the cantilever system is also presented. © *1996 American Vacuum Society.*

I. INTRODUCTION

The scanning probe microscope (SPM) has demonstrated itself to be a versatile and effective tool for patterning surfaces at the nanometer scale. Two common methods for modifying surfaces using probe microscopes are direct physical patterning and electric field assisted patterning. While both methods of surface modification are quite different, they both require that a sharp tip interacts with the surface to be patterned.

Physical patterning consists of scribing or indenting a sample using the tip of the SPM. Jung[1] has used this process to scribe patterns into polymer surfaces, and Mamin[2] has used the physical indentation process in conjunction with laser heating to store 100 nm bits at 100 kHz in a polymer surface. This approach has the advantage that the sample is typically much softer and generally unreactive with the tip, thereby reducing tip wear. The literature on electrical modification of surfaces with probe microscopes is much more extensive. Scanning probe lithography was pioneered by Dagata,[3] who patterned ⟨111⟩ silicon with the scanning tunneling microscope (STM), and Lyding[4] has used this same technique in ultrahigh vacuum (UHV) to pattern features less than a few nanometers. Snow and Campbell have modified this technique and patterned Si[5] and GaAs[6] with the atomic force microscope.

This form of lithography can be used to pattern a variety of surfaces.[7] Using the atomic force microscope (AFM) or STM, various materials have been patterned, including chromium by Song,[8] titanium by Sugimura,[9] and amorphous silicon by Kramer.[10] Recently, we have demonstrated the use of amorphous silicon as a general resist for lithography.[11] Also, organic materials have been modified by electrical means, as demonstrated by Majumdar[12] using PMMA. Finally, Sugimura[13] has used the STM to pattern amorphous silicon and then globally modify the surface by gold electroplating.

Further work in this area has yielded features more complex than simple lines on substrates. Previously we demonstrated the compatibility of scanning probe lithography with semiconductor processing by fabricating a 0.1 μm n-metal–oxide–semiconductor field-effect transistor (nMOSFET[14]) with the AFM. Campbell[15] has fabricated a side-gated transistor on a silicon-on-insulator wafer, and Matsumoto[16] has fabricated a single electron transistor by modifying titanium. More recently, Snow[17] has demonstrated real time lithographic monitoring using AFM lithography in order to fabricate metal–oxide–metal junctions.

While scanning probe lithography has revealed its potential value by fabricating operational devices, it suffers the drawback of being very slow since the scanning probe is an inherently serial device. For this form of lithography to be accepted beyond the single device level, the throughput of this approach must be greatly improved. We have addressed this problem in two ways, by increasing the tip velocity and by devising a parallel scheme.

In general, the tip velocity of scanning probes is governed by the speed of the feedback control of the microscope, which is itself limited by the resonant frequency of the actuator (typically less than 1 kHz). By integrating the actuator onto the cantilever, it is now possible to work up to near the cantilever resonant frequency (20–100 kHz), allowing the tip velocity to be significantly increased. With our integrated actuator, we have shown tip velocities in feedback operation of up to 3 mm/s with a piezoresistive sensor[18] and up to 1 cm/s with optical lever detection.[19]

Increased tip velocity reduces the time that the tip will take to cover a given area, but, even with the increased speed, the time necessary for a single tip to cover a wafer area is unacceptably long. Parallelism is an alternative approach to circumvent this problem. Previously we have used parallel arrays of five cantilevers for imaging and lithography,[11] but in that work, we did not have individual control of the cantilevers and therefore could not operate each cantilever in feedback. We have since fabricated an array of cantilevers with integrated sensors and actuators and used them for parallel imaging in feedback.[20] In this article, we advance our previous work on the 2×1 arrays by modifying their design to show parallel imaging without the need for force correction. We also demonstrate that the areas

[a]Electronic mail: sminne@leland.stanford.edu
[b]On leave from: Bilkent University, Ankara, Turkey.
[c]Electronic mail: quate@ee.stanford.edu

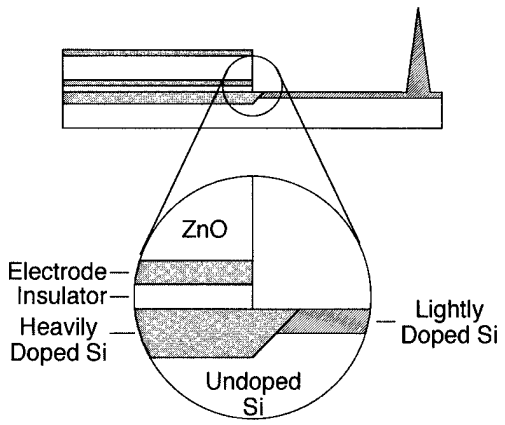


FIG. 1. Schematic cross-sectional view of a piezoresistive cantilever with integrated piezoelectric actuator. The piezoresistor lies on the surface of the silicon, separated from the ZnO actuator by silicon dioxide and silicon nitride. The piezoresistor doping extends to the apex of the tip.

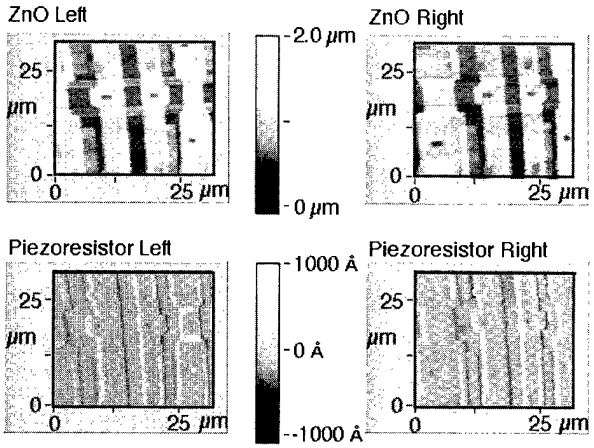


FIG. 2. Parallel constant force AFM image of a microcircuit. The ZnO maintains constant force over 2 μm of topography. Beneath the images (ZnO signal) are the corresponding error signals. The reduced error signal scale indicates constant force is being maintained.

scanned by the two cantilevers can be simultaneously patterned using a system with feedback and compatible with high speed AFM.

II. CANTILEVERS

The cantilevers used in this work are silicon piezoresistive cantilevers with integrated single crystal silicon tips developed by Tortonese.[21] We have modified the cantilever structure to include a zinc oxide (ZnO) actuator, which is fabricated on top of the silicon cantilever (Fig. 1). Since ZnO is piezoelectric, an applied electric field will cause the ZnO film to expand in the direction of the electric field and contract in the lateral direction. Due to the bimorph nature of the films, this contraction results in bending.

When imaging, a feedback loop monitors the cantilever deflection with the piezoresistor to determine the voltage that the ZnO actuator needs to maintain constant force between the tip and the sample. The fabrication process and initial results using this type of AFM cantilever with an integrated sensor and actuator are presented in Ref. 20.

The main disadvantage of the system described in Ref. 20 is that the sensor and actuator were not completely mechanically decoupled. Since the piezoresistive sensor extended over the entire length of the cantilever, the portion of the sensor beneath the ZnO deflects with the ZnO, generating a signal unrelated to the force on the tip. In order to take a feedback image, it is necessary to electronically compensate this mechanical coupling which limited the useful range of the actuator.

In this work, we have modified the cantilever by doping the area beneath the ZnO[22] to eliminate sensor-actuator coupling and the need for the correction circuit. A schematic diagram of the side view of the cantilever including the new implant is shown in Fig. 1. The increased doping serves two purposes. The piezoresistive effect[23] can be described by the following equation:

$$\frac{\Delta\rho}{\rho}(x,y,z)=\pi_L S(x,y,z), \tag{1}$$

where ρ is the resistivity of the cantilever material, π_L is the longitudinal piezoresistive coefficient, and S is the stress. Assuming a constant piezoresistor width and depth, $\Delta R/R$ is found by integrating and normalizing Eq. (1) over the length of the resistor. Our sensing circuitry measures changes in resistance of the entire cantilever (both the area under the ZnO and the tip area, which contains the active piezoresistive sensor). By increasing the doping under the ZnO region, we decrease the piezoresistive coefficient (π_L) by 80%, a reduction that directly affects the contribution to the sensor signal induced by the actuator. Another improvement comes from the nature of the fractional change. As the doping increases, the resistance decreases, and as a result, the absolute change in resistance from a given stress is reduced.

The small coupling permits imaging without the need for correction circuitry. An uncorrected parallel constant force image of an integrated circuit containing vertical topography of 2 μm is presented in Fig. 2. The top images are the voltages controlling the ZnO actuator and represent the surface topography, while the lower images are the piezoresistor signals which represent the error signals. For constant force imaging the error signals should be zero. The reduced scale on the error signal images shows that constant force is effectively maintained.

The frequency response of the cantilevers has been presented in detail in previous work.[18–20] The hysteresis and linearity measurements of the cantilever are presented in Fig. 3. Linearity and hysteresis measurements of the actuator were obtained by measuring the cantilever deflection with an optical level system. The optical system is calibrated to the out of contact movement of the cantilever in the following manner: The tip is brought close to the surface of a calibrated piezotube. The voltage to the ZnO is ramped until the tip to strikes the sample. The piezotube is then retracted from the tip by a known amount. The amount of additional voltage to

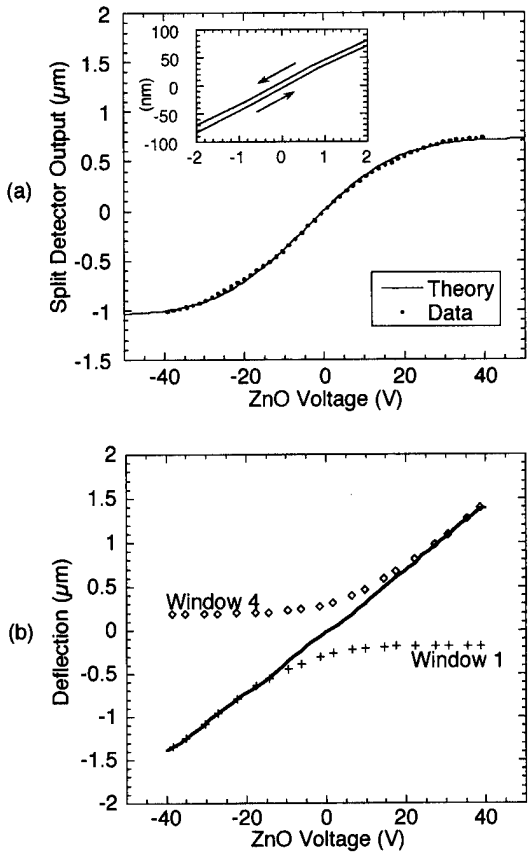


FIG. 3. (a) Hysteresis curve for the ZnO actuator. The points represent the data, while the solid line is a simulation of a Gaussian beam passing over a split photodiode. Inset: Expanded view of the data around zero applied voltage showing 20 nm maximum hysteresis. (b) Linearity plot for the ZnO actuator. The solid line is a composite curve of four photodiode positions representing the overall linearity curve of the actuator. The points of the first and final windows used to construct the overall plot are included. This method was required because the cantilever displaced the beam beyond the linear range of the photodiode.

the ZnO required to bring the tip back into contact with the surface provides a calibration for tip movement in air for a change in ZnO voltage. The cantilever is then deflected by the ZnO in air while the output of the laser is monitored. The slope of the response combined with the previous result gives the change in laser output for a given displacement of a freely suspended tip.

Figure 3(a) shows the output of a split photodetector when the ZnO voltage is swept from -40 to 40 V and then back to -40 V. The inset of Fig. 3(a) shows an expanded view of the hysteresis loop around zero volts applied voltage. The maximum hysteresis of the ZnO in this voltage range is 20 nm. Unfortunately, it is not possible to determine the linearity of the actuator from Fig. 3(a). The split photodiode detector used to measure the position of the beam reflected off the cantilever has a limited range for which the output is linear with respect to cantilever deflection. This limitation occurs because the size of the reflected beam is finite and must cover an adequate portion of both sides of the detector. In our microscope, the extent of the linear range covers roughly 1 μm of deflection and is insufficient to characterize the entire range of the ZnO actuator. It is interesting to note that the sigmoidal shape of Fig. 3(a) reflects the Gaussian distribution of our laser diode. We have also simulated a Gaussian beam traversing a split photodiode and plotted the results with the data in Fig. 3(a).

In order to measure linearity of the actuator, we manually translated the linear window of the detector in discrete steps so that it covered the entire range of the reflected laser beam. This provided us with a series of curves, each of which has only a small linear range. We then extracted and shifted the linear data from each curve to construct the overall linearity curve. The data was shifted such that overlapping voltages in adjacent windows had the same deflection.

Two such windows, which represent the initial and final data curves used for constructing the overall plot, along with the final construction, are plotted in Fig. 3(b). In these windows, the defector position was adjusted so that response is linear at the maximum excursion of the cantilever. Once the beam is completely deflected to one side of the photodiode, the detector cannot provide information about the cantilever deflection and the curves flatten. The composite curve of Fig. 3(b) shows excellent linearity of the device over its entire range of 3 μm.

III. ELECTROSTATIC FORCES

Electrical AFM lithography involves applying a voltage between the tip and the substrate to be modified. The force produced by the applied voltage is the negative spatial derivative of the energy, $U=\frac{1}{2}CV^2$, where C is the capacitance and V is the voltage. The capacitance is primarily a function of the distance between the tip and the sample, and the effective area of the tip.

When the AFM operates in a repulsive mode (contact imaging), the electrostatic attraction produced by an applied voltage between the tip and sample increases the contact force but does not cause the cantilever to deflect. If the sample is soft, the applied voltage provides a method for physical modification using an electronic signal. If the sample is hard, the increased forces accelerate the wear on the delicate tip. Since the detector is sensitive only to cantilever deflection, it is difficult to accurately monitor this attractive force while the tip remains in constant contact with the surface.

In order to measure the contact force produced by the electrostatic attraction, we retracted and extended the sample while monitoring the cantilever deflection with the piezoresistor. A force curve was obtained by plotting the deflection versus vertical position. Figure 4(a) shows a typical force curve where the sample is biased to -7.5 V and being retracted from the tip. An additional bias is provided by the piezoresistor measurement process. For this data, the extra bias is 0.6 V at the tip. The electrostatic attraction can be determined by multiplying the cantilever spring constant by the distance for which the tip is electrostatically fixed to the sample during retraction. The electrostatic force is labeled "*EF*" on Fig. 4(a). In our experiments, we find that the elec-

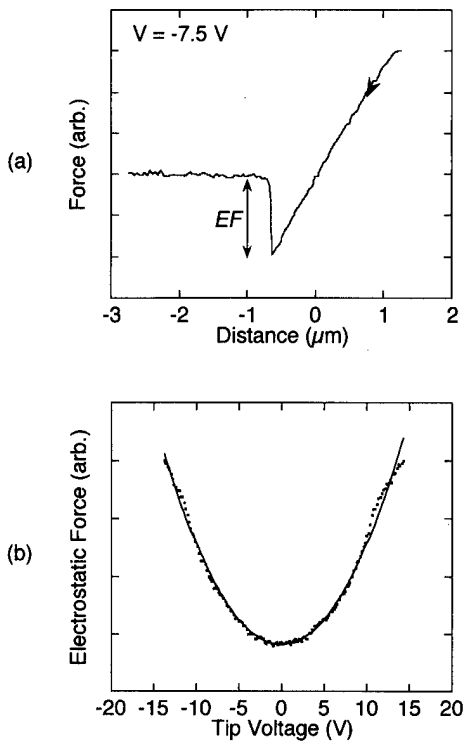


FIG. 4. (a) Force curve trace as the cantilever is retracted from the sample. The sticking is due to the electrostatic attraction of the applied voltage between the tip and sample. (b) The electrostatic attraction between the tip and the sample as a function of applied voltage.

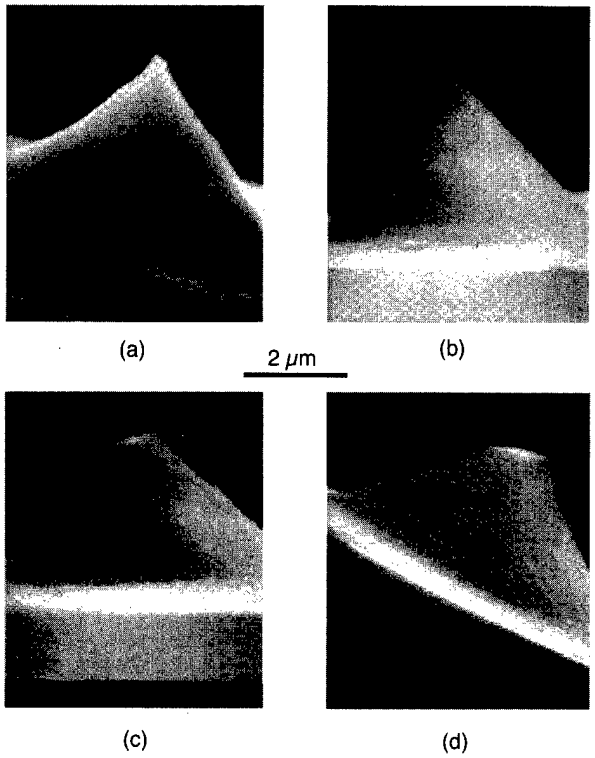


FIG. 5. SEM micrographs of (a) a new tip, (b) a slightly worn tip which is still useful for silicon lithography, (c) a heavily worn tip which is no longer useful for silicon lithography, and (d) the tip used for the measurement of Fig. 4. In (d), the heavy damage is due to the measurement process described in Fig. 4.

trostatic force is significantly greater than the meniscus force. The electrostatic force can be measured as a function of applied voltage by reproducing the data in Fig. 4(a) for many voltages and is shown in Fig. 4(b). A second order polynomial describes the data well in Fig. 4(b), indicating that the force is proportional to the square of the applied voltage.

Figure 5(a) shows a scanning electron micrograph of the tip used to obtain the data in Fig. 4. This tip initially had a radius of curvature of ~300 Å, but due to the process we used to measure the contact force versus voltage, the tip sharpness was reduced to ~1.5 μm, as shown in Fig. 5(d). During initial measurements made with this tip, we measured a contact force of 20 nN for a voltage of 5 V. As the measurement process was repeated, this force continued to increase until stabilizing at a force of 1.25 μN. Presumably the tip had dulled to a point that no further damage was occurring. The results of Fig. 4 and 5 indicate that the contact force between the tip and sample under an applied voltage is dependent on the magnitude of the voltage and the degree of tip wear.

Preliminary results show that tip wear can be a significant problem. In our lithographic experiments, we worked with single crystal silicon tips patterning single crystal silicon substrates. Due to the hardness of the substrate and the chemical interaction between silicon tip and silicon sample, and because of the increased force due to the applied voltage, the impact on the tip is substantial. Tips showing various states of wear are shown in Fig. 5. In Fig. 5(a), we show a new tip which has a radius of curvature of ~300 Å. Figure 5(b) shows a tip that has performed and is still capable of performing high quality lithography, although the radius of curvature has increased to ~2000 Å. Figure 5(c) shows a tip that was originally effective in performing lithography but has been worn down to the point where lithography is no longer possible. The surface diameter of the tip is about 1 μm. It can be seen that the tip wear is due to the scanning process and is exacerbated by the electrostatic forces from the voltage used for lithography—the straight angle on the worn tip corresponds to the scanning angle of our AFM. We believe that lithography fails as the tip wears because of the decrease in electric field due to the dulling geometry of the tip and because most of the conductive implant in the tip region has worn off.

IV. LITHOGRAPHY

Our preferred method for performing lithography is to use the conductive path in the piezoresistor to apply appropriate voltages to the conducting tip. Initially the ZnO was used to control the lithography by biasing the sample and pulling the tip off the surface when lithography was not desired, but this method showed limited success due to the large electrostatic forces. In other attempts, we tried to use the ZnO as both the sensor and actuator,[24] and dedicated the piezoresistor for

lithographic voltages only. While this method provided good imaging capabilities, the lithographic signal would often interfere with the feedback loop. We believe that this approach is still attractive, but at this time our cantilevers are not optimized for the use of the ZnO as a sensor.

In order to use the piezoresistor as both a sensor and an electrical conductor for lithography, it is necessary to separate the two signals. By using an ac bridge, we can sense piezoresistor deflections at high frequency (typically 100 kHz) while sending the electrical pulses for lithography through the cantilever at lower frequencies (filtered to 100 Hz). A benefit of this system is that very high scan speeds can be obtained using the ac bridge (see Ref. 18). Another consideration when using the piezoresistor as a path for lithography is electrical breakdown between the piezoresistor and the bottom ZnO electrode. Referring back to Fig. 1, the piezoresistor is separated from the bottom electrode by a thin film of dielectric. Since the bottom electrode is held at ground, a lithography signal can cause breakdown between the films. It is therefore necessary to bias both ZnO electrodes and the piezoresistor with the lithography voltage used for writing the lithographic patterns.

Also, for high resolution imaging it is necessary to use a high gain low noise amplifier for sensing the cantilever deflection. Unfortunately, the large voltage needed for lithography normally precludes the use of such amplifiers in our circuitry.

To address the problems of tip speed, dielectric breakdown, and low noise amplification, two transformers and a battery powered amplifier were used in order to decouple the cantilever from ground. By biasing the entire cantilever (both ZnO electrodes and the piezoresistor) to the lithography voltage, the piezoresistor could be measured at ac without being influenced by the high voltage lithographic signals. Also, by biasing the entire cantilever, breakdown between the films is no longer a concern. Duplicating the electronics for each of the cantilevers within the array, parallel independent lithography in feedback is readily accomplished.

In general, we have found that more reliable pattern generation is obtained by using a combination of open and closed loop feedback. The open loop portion occurs only when the lithography pulse is on. When writing, the ZnO exerts a force to partially oppose the electrostatic forces in order to reduce tip wear and increase tip lifetime. In this way, cantilevers are maintained in constant force mode as they profile the surface. Only when writing voltage is applied do they pull back from the surface to compensate for the electrostatic force. Once the lithography signal is turned off, the tips return to the normal feedback mode. Lithography has been accomplished using both this mode as well as regular closed loop feedback.

Figure 6(a) shows an optical image of a large scale independent parallel lithographic pattern using the open and closed feedback system. The substrate is ⟨100⟩ silicon and is patterned with a lithographic voltage of 20 V. The presumed mechanism for the lithography is that the large electric field desorbes hydrogen and increases the oxidation rate on the

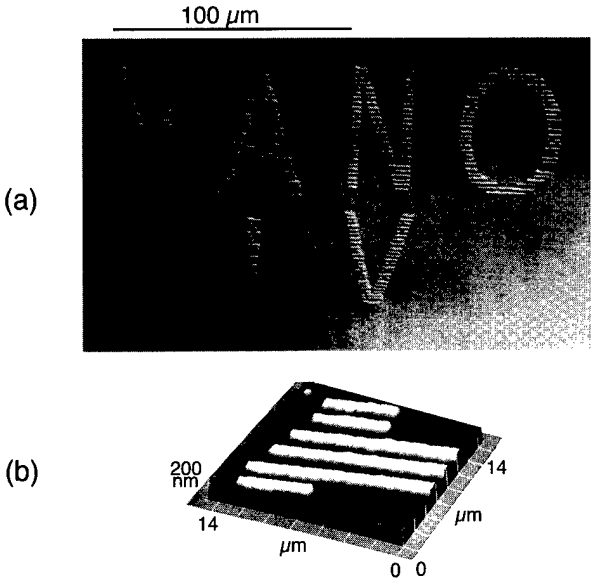


FIG. 6. (a) Independent parallel lithography using cantilevers with integrated sensors and actuators. The lithography covers 200 μm×100 μm with one tip writing "N A" over "I" and the other simultaneously writing "N O" over "V" in a single pass. (b) Magnified AFM image of (a).

exposed silicon. Details of this process are described elsewhere.[3,5] The latent image is transferred to the substrate with a KOH etch. In Fig. 6(a), each tip has traversed a 100 μm×100 μm area of the 200 μm×100 μm total area in a raster pattern. A computer program controlled the scanning and coordinated the lithography voltages. With one pass, the left tip patterned the letters "N A" over the "I" and the right tip patterned the letters "N O" over the "V". Figure 6(b) is an enlarged AFM image of the pattern in the square on Fig. 6(a). The modulation in the lines is caused by the digitization from computer control.

V. CONCLUSIONS

Expansion of scanning probe microscope lithography through parallelism, combined with the high scan speeds enabled by integrated actuators, are important steps that begin to address the issue of throughput for a probe based lithography system. Using this approach, we have achieved the first demonstration of parallel, independent AFM lithography using cantilevers with independent control spanning an area of 20 000 μm^2.

We believe further improvement in cantilever design can be implemented. Optimization of the ZnO geometry to improve its simultaneous sensing and actuating capabilities would allow the elimination of the piezoresistor leaving the silicon exclusively for lithographic signal path. Alternatively, adding a separate conducting path to the tip would serve the same function, while allowing us to maintain the current design.

ACKNOWLEDGMENTS

The authors would like to thank Babur Hadimioglu and Jim Zesch at Xerox PARC for the ZnO deposition and

Charles Ahn and Nick Ulman for insightful discussions. Primary support for this work came from the Joint Services Electronics Program N0014-91-J-1050 of the Office of Naval Research with partial support from the NSF and CMR at Stanford. S. C. M. acknowledges the support of the Leland T. Edwards Fellowship and S. R. M. acknowledges the support of the Urbanek Fellowship.

[1]T. A. Jung, A. Moser, H. J. Hug, D. Brodbeck, R. Hofer, H. R. Hidber, and U. D. Schwarz, Ultramicroscopy **42**, 1446 (1992).
[2]H. J. Mamin and D. Rugar, Appl. Phys. Lett. **61**, 1003 (1992).
[3]J. A. Dagata, J. Schneir, H. H. Harary, C. J. Evans, M. T. Postek, and J. Bennett, Appl. Phys. Lett. **56**, 2001 (1990).
[4]J. W. Lyding, T. C. Chen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, Appl. Phys. Lett. **64**, 2010 (1994).
[5]E. S. Snow and P. M. Campbell, Appl. Phys. Lett. **64**, 1932 (1994).
[6]E. S. Snow, P. M. Campbell, and B. V. Shanabrook, Appl. Phys. Lett. **63**, 3488 (1993).
[7]C. R. K. Marrian, *Technology of Proximal Probe Lithography* (SPIE Optical Engineering, Bellingham, WA, 1993).
[8]H. J. Song, M. J. Rack, K. Abugharbieh, S. Y. Lee, V. Khan, D. K. Ferry, and D. R. Allee, J. Vac. Sci. Technol. B **12**, 3720 (1994).
[9]H. Sugimura, T. Uchida, N. Kitamura, and H. Masuhara, Appl. Phys. Lett. **63**, 1288 (1993).
[10]N. Kramer, H. Birk, J. Jorritsma, and C. Schonenberger, Appl. Phys. Lett. **66**, 1325 (1995).
[11]S. C. Minne, Ph. Flueckiger, H. T. Soh, and C. F. Quate, J. Vac. Sci. Technol. B **13**, 1380 (1995).
[12]A. Majumdar, P. I. Oden, J. P. Carrejo, L. A. Nagahara, J. J. Graham, and J. Alexander, Appl. Phys. Lett. **61**, 2293 (1992).
[13]H. Sugimura and N. Nakagiri, Jpn. J. Appl. Phys. **34**, 698 (1995).
[14]S. C. Minne, H. T. Soh, Ph. Flueckiger, and C. F. Quate, Appl. Phys. Lett. **6**, 703 (1995).
[15]P. M. Campbell, E. S. Snow, and J. P. McMarr, Appl. Phys. Lett. **66**, 1388 (1995).
[16]K. Matsumoto, M. Ishii, K. Segawa, Y. Oka, B. J. Vartanian, and J. S. Harris, Extended Abstracts of 1994 International Conference of Solid State Device and Materials, Yokohama, Japan, 1994 (unpublished), p. 46.
[17]E. S. Snow and P. M. Campbell, Science **270**, 1639 (1995).
[18]S. R. Manalis, S. C. Minne, and C. F. Quate, Appl. Phys. Lett. **68**, 871 (1996).
[19]S. R. Manalis, S. C. Minne, A. Atalar, and C. F. Quate, Rev. Sci. Instrum. (to be published).
[20]S. C. Minne, S. R. Manalis, and C. F. Quate, Appl. Phys. Lett. **67**, 3918 (1995).
[21]M. Tortonese, Yamada, R. C. Barrett, and C. F. Quate, *Proceedings of Transducers '91*, IEEE Publ. No. 91 CH2817-5 (IEEE, New York, 1991), p. 448.
[22]10^{20} cm^{-3} under the ZnO, 5×10^{18} cm^{-3} in the piezoresistor.
[23]Y. Kanada, IEEE Trans. Electron. Devices **ED-29**, 64 (1982).
[24]S. C. Minne, S. R. Manalis, A. Atalar, and C. F. Quate, Appl. Phys. Lett. **68**, 1427 (1996).

Monte Carlo simulation of inclined incidence of fast electrons to solids

Y. M. Gueorguiev[a] and G. M. Mladenov
Laboratory of Physical Problems of Electron-Beam Technologies, Institute of Electronics, Bulgarian Academy of Sciences, Tzarigradsko shose 72, 1784 Sofia, Bulgaria

D. I. Ivanov
DIPI Group, B. Petkov Str. 65, Sofia, Bulgaria

(Received 27 October 1995; accepted 24 May 1996)

In the present work inclined incidence of accelerated electrons to solids is simulated using Monte Carlo technique. Spatial distributions of absorbed electron energy density in a 125 nm poly(methylmethacrylate) resist layer on bulk Si substrate are obtained for angles of incidence 30°, 45°, and 60° at two beam energies—25 and 50 keV—together with the energy and angular distributions of the backscattered electrons. The results show strong asymmetry of the exposure distributions. Their peaks are significantly lower, wider, and 40–100 nm shifted, and their shapes are different in comparison with those for normal incidence of electrons. The ratio between the maximum values of exposure distributions due to the forward scattered and the backscattered electrons decreases with decreasing angle of incidence. These peculiarities of exposure distributions may lead to enhanced proximity effects and cause deviation from required pattern shapes. Therefore, if inclined incidence of accelerated electrons to any surface occurs during electron beam lithography, it has to be taken into account. © *1996 American Vacuum Society.*

I. INTRODUCTION

Monte Carlo simulation (MCS) is the most widely used technique for modeling the electron scattering in solids and particularly of obtaining spatial distributions of the absorbed electron energy density in resists as applied to electron-beam lithography (EBL) due to its advantages over other (especially experimental) methods.

However, one of the advantages of MCS, namely the possibility of modeling the scattering of fast electrons impinging on targets of complex topography is, to the best of our knowledge, not properly studied and exploited so far in regard to EBL, but it may be of interest in some particular applications (e.g., fabrication of high-resolution concave Brag–Fresnel zone plate x-ray lenses, EBL of very large scale integrated VLSI devices with stepped or other kind of complex topography surfaces). This problem can be considered, in the first approximation, as the simulation of inclined incidence of accelerated electrons to solids. There are several articles concerned with the latter one,[1–4] however they consider mainly backscattering of electrons and particularly energy, angular, and spatial distributions of backscattered electrons. As far as we know, all the results on energy deposition of fast electrons reported hitherto are obtained for normal incidence of primary electrons.

Therefore, the aim of the present work is to simulate the inclined incidence of fast electrons to the target surface and to obtain spatial distributions of absorbed energy density in the resist, deposited on substrate of the material most widely used in microelectronics, namely silicon.

[a]Electronic mail: JGEORG@BGEARN.ACAD.BG

II. MONTE CARLO SIMULATION

Since the MCS of electron scattering in solids is discussed in detail elsewhere,[1,5–10] here only the main features of the present model will be described.

In our program we use the conventional Monte Carlo technique based on a single scattering model and continuous slowing down approximation assuming the screened Rutherford elastic scattering cross section with a screening parameter of Nigam *et al.*[11] and the Bethe energy loss equation. In order to prevent the failure of the latter at low electron energies the parabolic extrapolation of $(dE/dx)^{-1}$ derived by Rao-Sahib and Wittry[12] is introduced below a certain electron energy. In this way the incident electrons can be pursued until their energy decreases to 50 eV.

Another peculiarity of the energy loss calculation is that instead of the conventional two-point difference scheme we involve the three-point difference scheme as proposed by Valiev *et al.*[13] The application of the three-point difference scheme increases the accuracy of the energy loss calculation, especially at low beam energies as well as for heavy targets.[14]

To account for differences in both the scattering and the stopping properties of materials at the interfaces between different layers of the target, a procedure for recalculation of the free path length and the energy loss is included.[15]

III. RESULTS AND DISCUSSION

In this work calculations are performed on an IBM 4381 machine using a total of 30 000 electron trajectories for each simulation. The spatial distributions of absorbed energy density are obtained for a zero-width δ function at two different energies of primary electrons, namely 25 and 50 keV. The angle of incidence is 30°, 45°, or 60° with respect to the target surface.

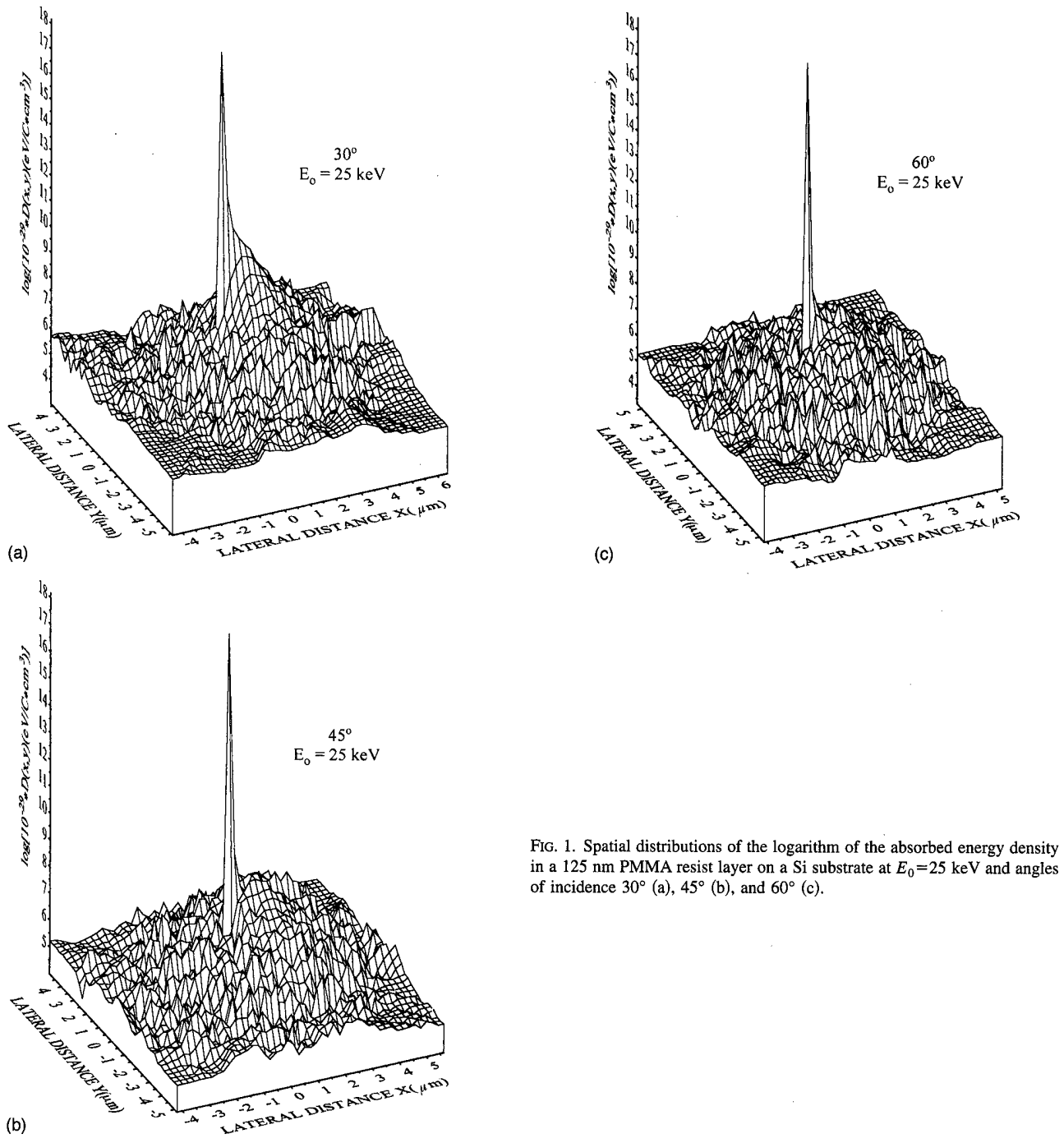


FIG. 1. Spatial distributions of the logarithm of the absorbed energy density in a 125 nm PMMA resist layer on a Si substrate at E_0=25 keV and angles of incidence 30° (a), 45° (b), and 60° (c).

In Figs. 1 and 2 the spatial distributions of the logarithm of absorbed energy density in 125 nm poly(methylmethacrylate) (PMMA) resist layer on bulk Si substrate at two beam energies: 25 and 50 keV, respectively, and angles of incidence 30° (a), 45° (b), and 60° (c) are shown, while in Fig. 3 the lateral distributions of absorbed energy density for the same conditions are presented in comparison with those for normal incidence. Since the resist layer is chosen to be relatively thin, the energy loss of electrons in the resist is averaged over the whole of its thickness. It is necessary to point out that the electron beam propagates in the plane $x0z$ ($y=0$ and azimuth angle $\varphi=0$) and, in the case of inclined incidence, is directed to the positive side of the x axis.

As was expected, in all these figures strong asymmetry of the distributions at inclined incidence of electrons is well seen. The exposure of the resist caused by backscattered electrons at $x>0$ is significantly greater compared with that at $x<0$ as well as with that of normal incidence of electrons, especially at lower beam energy (25 keV). In addition, the peaks of the distributions which are due to the exposure by the forward scattered electrons also differ from those for normal incidence. They are significantly wider (half widths at

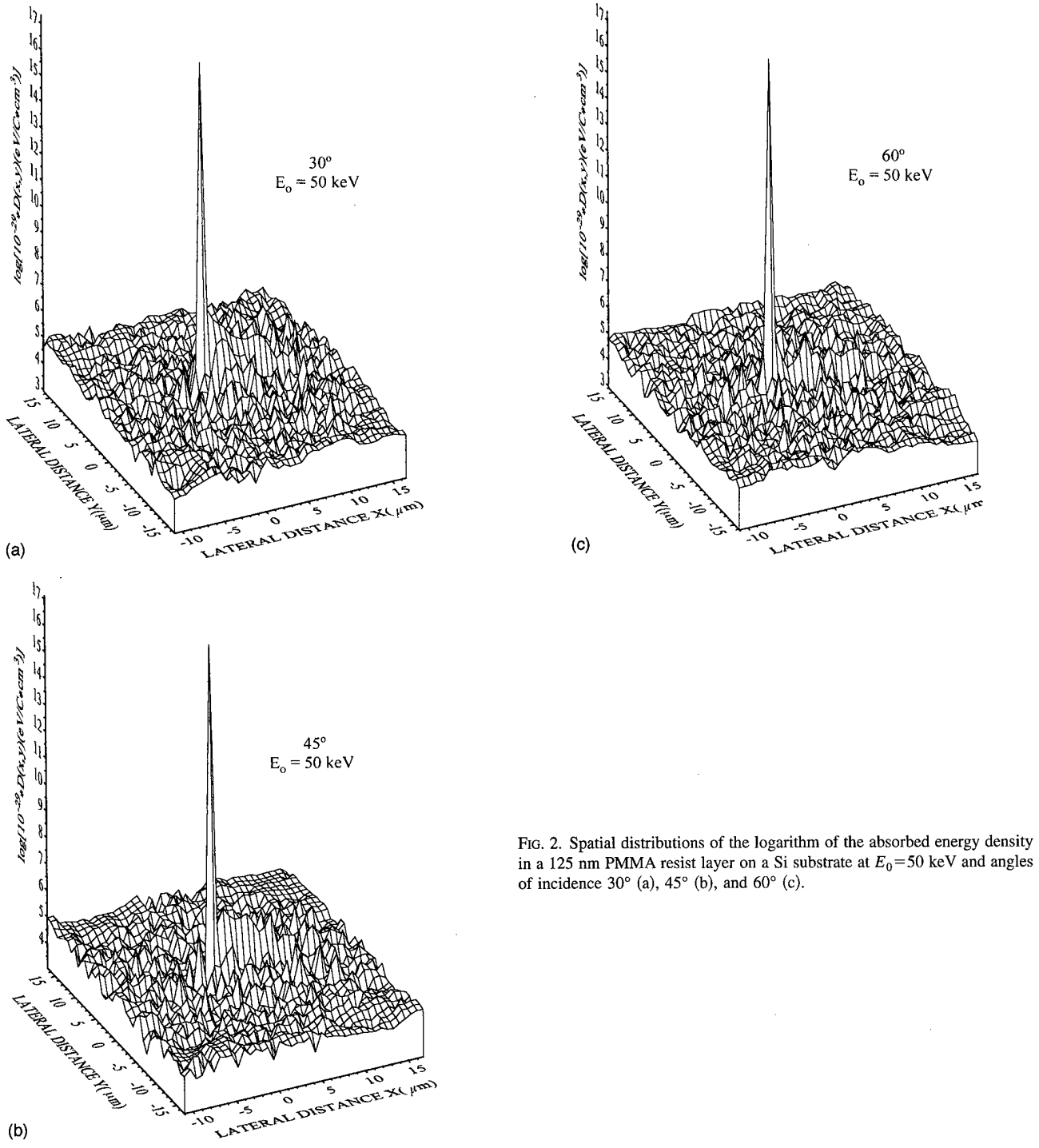


FIG. 2. Spatial distributions of the logarithm of the absorbed energy density in a 125 nm PMMA resist layer on a Si substrate at $E_0=50$ keV and angles of incidence 30° (a), 45° (b), and 60° (c).

the levels of turning of the peaks into backscattered distributions in the direction of $x>0$ are listed in Table I) and lower (the ratios of maximum values of the peaks at inclined incidence of electrons to those at normal incidence are also listed in Table I) and their maximum values are shifted from the incident point in the direction of $x>0$ (see Table I). Such distributions may lead to enhanced proximity effects not only at $x>0$ but, also surprisingly, at $x<0$, due to considerably lower ratios between exposures caused by forward scattered and backscattered electrons for all cases of inclined incidence.

Figure 4 compares the simulated energy distributions of backscattered electrons at $E_0=25$ keV and angles of incidence 30°, 45°, 60°, and 90° with that for bulk Al (whose atomic number $Z=13$ is close to that of Si, $Z=14$) obtained experimentally at $E_0=30$ keV, normal incidence, and takeoff angle 45° by Kulenkampff and Spyra.[16] Here dn/dQ is the number of backscattered electrons leaving the target with an

TABLE I. Calculated values of some parameters of the peaks of exposure distributions at inclined incidence of electrons.

Parameter	Half widths of peaks (nm)		Max_{incl}/Max_{normal}		Shift of maximum values of peaks (nm)	
E_0 (keV)	25	50	25	50	25	50
30°	600	600	0.041 2	0.054 9	90	100
45°	430	430	0.062 0	0.077 7	70	80
60°	350	350	0.121 5	0.148 7	40	50

energy in the interval $[Q_i(=E_i/E_0), Q_i+dQ]$, multiplied by the number of energy intervals in the range [0,1], and normalized by the number of incident electrons. Note that $\int f(Q)dQ > 1$ since dn/dQ is not normalized by the number of all backscattered electrons. This normalization is avoided because it makes the levels of all energy distributions almost equal and does not allow to show the increase in the number of backscattered electrons with the decrease in the angle of incidence.

In Fig. 4 it can be seen that the distribution for the target PMMA/Si at normal incidence of electrons is close to the experimental distribution for the Al target. With decreasing angle of incidence the level of distributions increases and their maximum values move to greater energies (see Table II).

In Fig. 5 the angular distributions of backscattered electrons at E_0=25 keV and angles of incidence 30°, 45°, 60°, and 90° are shown in comparison with sin θ. Here θ is the escape polar angle of backscattered electrons measured from the target plane and $dn/d\theta$ is the number of backscattered electrons leaving the target under an angle in the interval $[\theta_i, \theta_i+d\theta]$, multiplied by the number of angular intervals in the range [0,180°], and normalized by the number of incident electrons. Again, $\int f(\theta)d\theta > 1$ for the same reason as

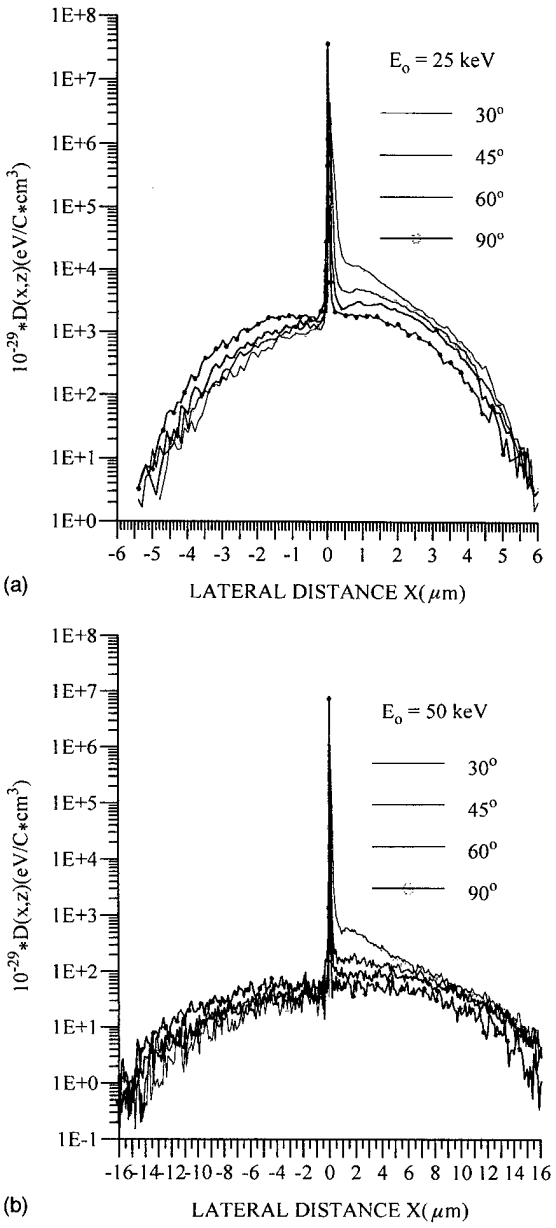


FIG. 3. Comparison between the lateral distributions of the absorbed energy density in a 125 nm PMMA resist layer on a Si substrate at angles of incidence 30°, 45°, 60°, and 90° and E_0=25 keV (a) or E_0=50 keV (b).

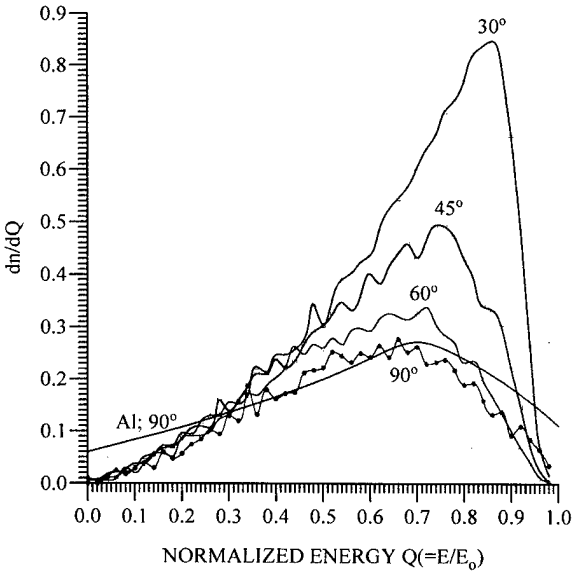


FIG. 4. Comparison of simulated energy distributions of backscattered electrons from a 125 nm PMMA/Si target at E_0=25 keV and angles of incidence 30°, 45°, 60°, and 90° with an experimental distribution for Al target at E_0=30 keV and normal incidence.

TABLE II. Calculated maximum values and their positions of the energy and angular distributions of backscattered electrons at $E_0=25$ keV.

Parameter	$(dn/dQ)_{max}$	Position of $(dn/dQ)_{max}$	$(dn/d\theta)_{max}$	Position of $(dn/d\theta)_{max}$
30°	0.847	0.86	0.720 3	48°
45°	0.495	0.74	0.458 1	66°
60°	0.339	0.72	0.342 2	78°
90°	0.277 5	0.66	0.240 6	90°

in the case of $f(Q)$. It is seen that the distribution at 90° can be considered as a Gaussian one rather than as being proportional to sin θ, but in our opinion it is a reasonable estimate of the actual angular distribution at normal incidence of electrons. The decrease in the angle of incidence leads to an increase in the level of distributions as well as to a shift of their maximum values to smaller angles (see Table II).

IV. CONCLUSIONS

In this article Monte Carlo simulation of scattering of fast electrons impinging at angles 30°, 45°, and 60° on the 125 nm PMMA/Si target surface is performed. The spatial distributions of absorbed energy density in the resist layer are obtained at two beam energies, namely 25 and 50 keV, together with the angular and energy distributions of backscattered electrons. Strong asymmetry of backscattered parts of exposure distributions is observed as well as considerable changes of forward scattered parts in comparison with those for normal incidence of electrons. These features of the distributions may cause a significant enhancement of proximity effects and have to be taken into account when inclined incidence of fast electrons to any surface occurs during EBL.

The results show that the decrease in the angle of incidence leads to an increase in the levels of energy and angular distributions of backscattered electrons as well as to the shift of maximum values of energy distributions to greater energies and of angular distributions to smaller angles.

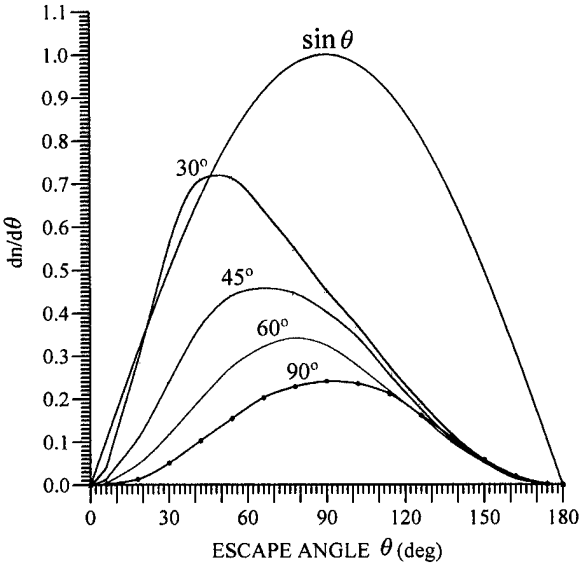


FIG. 5. Comparison of simulated angular distributions of backscattered electrons from a 125 nm PMMA/Si target at $E_0=25$ keV and angles of incidence 30°, 45°, 60°, and 90° with sin law.

ACKNOWLEDGMENT

This work was partially supported by the Bulgarian National Science Fund.

[1]K. Murata, T. Matsukawa, and R. Shimizu, Jpn. J. Appl. Phys. **10**, 678 (1971).
[2]K. Murata, J. Appl. Phys. **45**, 4110 (1974).
[3]S. Hasegawa, Y. Iida, and T. Hidaka, J. Vac. Sci. Technol. B **5**, 142 (1987).
[4]Z.-J. Ding, R. Shimizu, T. Sekine, and T. Sakai, Appl. Surf. Sci. **33/34**, 99 (1988).
[5]R. Shimizu and T. E. Everhart, Optik **36**, 59 (1972).
[6]D. F. Kyser and K. Murata, IBM, J. Res. Dev. **18**, 352 (1974).
[7]K. Murata, D. F. Kyser, and C. H. Ting, J. Appl. Phys. **52**, 4396 (1981).
[8]K. Murata, H. Kawata, K. Nagami, Y. Hirai, and Y. Mano, J. Vac. Sci. Technol. B **5**, 124 (1987).
[9]R. Shimizu and Z-J. Ding, Rep. Prog. Phys. 487 (1992).
[10]J. Georgiev, G. Mladenov, and D. Ivanov, Thin Solid Films, **251**, 67 (1994).
[11]B. P. Nigam, M. K. Saunderson, and T.-Y. Wu, Phys. Rev. **115**, 491 (1959).
[12]T. S. Rao-Sahib and D. B. Wittry, J. Appl. Phys. **45**, 5060 (1974).
[13]K. A. Valiev, A. N. Kirillov, B. N. Kovtun, T. N. Machviladze, and M. M. Mkrtchian, Proc. Inst. Gen. Phys. Acad. Sci. USSR **8**, 5 (1987) [in Russian].
[14]J. Georgiev, *Vacuum Electron and Ion Technologies*, Proceedings of the 8th International School, September 26–October 1, 1993, Varna, Bulgaria, edited by D. Karpuzov (NOVA, 1994), p. 406.
[15]J. M. Georgiev, Proceedings of the 4th International Conference on Electron Beam Technology, June 6–11, 1994, Varna, Bulgaria (Sofia, Bulgaria, 1994), p. 320.
[16]H. Kulenkampff and W. Spyra, Z. Phys. **137**, 416 (1954).

Generation mechanism of distortion aberration in a symmetric magnetic doublet for an electron beam projection system

Mamoru Nakasuji and Hiroyasu Shimizu
2nd Designing Department, Industrial Supplies and Equipment Division,
Ohi Plant, Nikon Corporation, 6-3 Nishi-Ohi 1-Chome, Shinagawaku, Tokyo 140, Japan

(Received 25 October 1995; accepted 1 May 1996)

Radial and azimuthal distortion aberrations are increasingly a function of the image side lens bore radius in the range from 1.25 to 5 times as large as the maximum image field radius. This phenomenon is inconsistent with our previous understanding. An assumption is made that these large distortions for the large bore radii come from the influence of the magnetic field of one on the other, thereby destroying the symmetry. This assumption is confirmed from the following simulation. When these distortions are calculated for the ideal case where the magnetic fields are calculated in the condition without the other lens, they are decreased to around $\frac{1}{10}$ of those for the case where the magnetic fields are calculated in the real condition with the lenses in proximity. When the object–image distance is 800 mm and the bore radii of lens 1 are 100 and 20 mm, the residual radial and azimuthal distortions are 1.5 and 0.7 nm, the beam blur is smaller than 45 nm for the beam semiangle from 0.05 to 0.7 mrad, where the main-field and subfield sizes in the image plane are 20 mm and 250 μm, respectively, the beam energy is 100 keV, and the space charge effects are neglected. © *1996 American Vacuum Society.*

I. INTRODUCTION

Today, worldwide, light optical exposure tools define the device pattern in the resist layer of silicon wafers using an *I*-line or deep ultraviolet optical stepper with wavelengths of 365 or 248 nm, respectively. However, there are several other lithography tools that use a much shorter wavelength. By using an x-ray, electron, or ion beam for the exposure instead of a light optical system, the resolution and depth of focus increase dramatically. In scanning-electron-beam lithography, for example, character projection lithography,[1] the size of the scan field is limited not only by the optical performance such as deflection distortion, blur, and nonvertical beam landing, but also by electronics such as the accuracy of the deflection amplifier and the resolution of the digital to analog converter (DAC) that defines the beam position on a Cartesian digital address grid, typically 2^{16}–2^{20} points along both the x and y axes. In projection-electron-beam lithography, the size of the main field is limited only by the optical performance such as field distortion, beam blur, and nonvertical beam landing. Projection systems are better suited to large field sizes which produce high-volume production and main-field stitching error-free systems.

For the lens system, the symmetric magnetic doublet,[1–3] and (PREVAIL) projection exposure with variable axis immersion lenses[4] have been studied. We obtained an aberration of less than 0.02 μm in distortion and 0.12 μm in beam resolution for a 27-mm-square field size through a computer simulation.[3] However, this large field image projection causes many problems; for example, it requires a high emittance electron gun. To resolve these problems, the large field is divided into many small subfields, and each subfield is projected with a dynamical aberration correction.[5–7] We also calculated the residual beam resolution and distortion (33 and 1.1 nm, respectively) with a dynamically compensated field aberration for an object–image distance of 1.5 m, and for 20 and 0.25 mm for the main-field and subfield sizes in the image plane, respectively.[8] For an object–image distance of 50 cm, residual aberration was calculated from the trajectory calculation after the correction of lens current, crossover, and beam positions.[9] The PREVAIL design is promising because it may permit large beam semi-angles if a large field is available with reasonable distortion. Unfortunately, its detailed calculation has not been reported yet.

For the lens bore radius it was stated that[10] the bore radius must be chosen sufficiently large so that aberrations higher than third order are not significant. The ratio of image field radius to image side bore radius was designated as 0.212. However, for a ratio of 0.667, good aberration characteristics were obtained,[11] where the aberrations were calculated from the trajectory calculation so that all the higher-order aberrations could be calculated. Therefore, we have questioned previous statements and studied aberrations as a function of the bore radius to ascertain whether the large lens bore radii are adequate for the reducing image projection system.

An optical system for the electron-beam projection system for a 4G-bit d-RAM pattern is designed and the aberrations are calculated.

II. LENS MODELS AND CALCULATION

Figure 1 shows a magnetic lens model for the symmetric magnetic doublet. The object to image distance and pole gap for lens 1 are kept to 800 and 560 mm, respectively. The crossover position at Z of 2140 mm is important position where is partitioned as 4:1 from the object to image. Wide gap lenses are used in which the fields are not concentrated to relatively small volumes because such wide gap lens are necessary to obtain small third-order aberration coefficients.[3] The object side and crossover side bore radii for lens 1 are

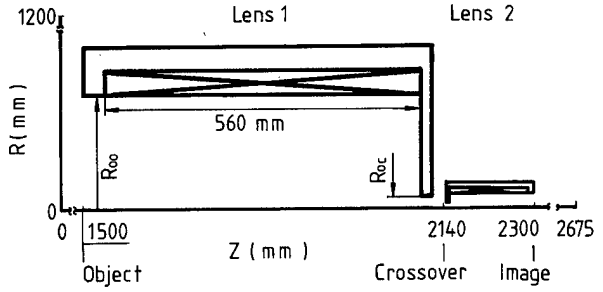


FIG. 1. Lens model and lens parameters for the symmetric magnetic doublet. The lens parameters for lens 2 are a quarter of those for lens 1.

R_{oo} and R_{oc} mm. The reduction factor is $\frac{1}{4}$ and all the lens parameters for lens 2 are a $\frac{1}{4}$ of those for lens 1. R_{oo} is varied from 50 to 200 mm and R_{oc} is varied from $\frac{40}{3}$=13.33 to 30 mm. For these models, the magnetic vector potential for each mesh point is calculated using the finite element method.[12] The total node numbers are 130 (Z direction)×40 (R direction). Third-order elements are used to obtain accurate magnetic vector potentials, so those for 390×120 points are calculated. The boundary condition is defined as the magnetic vector potential when the optical axis is zero. Over the area of $0<R<1200$ mm for $0<Z<2120$ mm, $0<R<77\,520-36Z$ for $2120<Z<2145$, and $0<R<300$ mm for $2145<Z<2675$, the computation using the finite element method was extended. The magnetic vector potentials between the nodes are interpolated with a third-order spline function. With the use of these magnetic vector potentials, the electron trajectories are calculated.

Figure 2 shows the calculated magnetic field line, the curved lines represent lines $A(R,Z)$=const. Since the excitation current is opposite to that between lens 1 and lens 2,

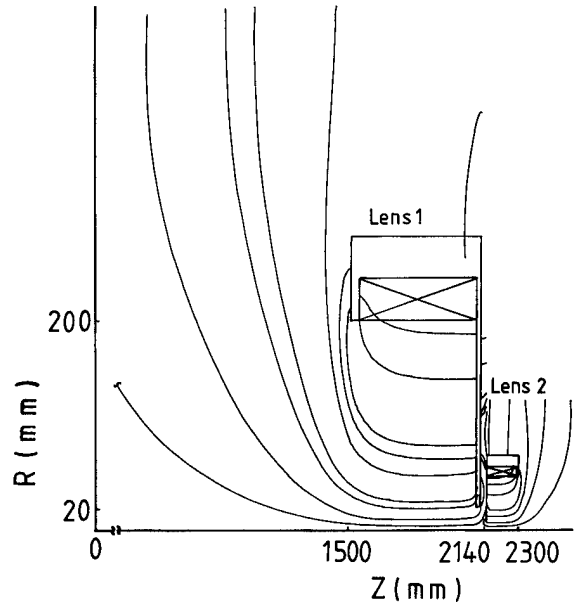


FIG. 2. Magnetic flux for the real lens system, where the object and crossover side bore radii R_{oo} and R_{oc} are 200 and 20 mm, respectively.

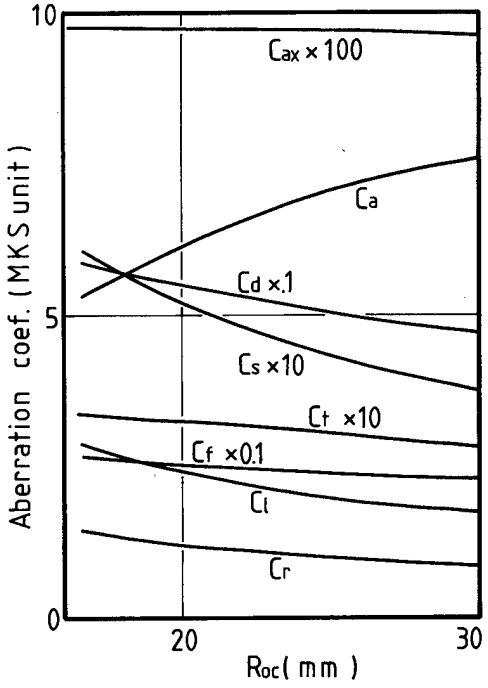


FIG. 3. Third-order aberration coefficients calculated from the axial magnetic field distribution and its first derivative as a function of the crossover side bore radius R_{oc} of lens 1, where the object side bore radius R_{oo} is 100 mm. Notations C_{ax}, C_a, C_d, C_s, C_t, C_f, C_l and C_r are axial chromatic (m/rad), astigmatism [(rad m)$^{-1}$], distortion (m^{-2}), spherical (m rad^{-3}), transverse chromatic (dimensionless), field curvature [(rad m)$^{-1}$], coma length (rad^{-2}), and coma radius (rad^{-2}), respectively.

the magnetic flux from one lens is repulsed by the magnetic flux from the other lens. These features will be discussed later.

The axial magnetic field distribution and its first derivative are also calculated. Using these values, third-order aberration coefficients are calculated when the crossover side bore radius R_{oc} and the object side bore radius R_{oo} are varied. The results are shown in Figs. 3 and 4, where C_{ax}, C_a, C_d, C_s, C_t, C_f, C_l, and C_r are axial chromatic, astigmatism, distortion, spherical, transverse chromatic, field curvature, coma length, and coma radius, respectively.

As seen in Fig. 3, only the astigmatism aberration coefficient is an increasing function of the crossover side bore radius R_{oc}, and all the other coefficients decrease slowly when R_{oc} increases. As seen in Fig. 4, the astigmatism and spherical aberration coefficients are decreasing functions of the object side bore radius R_{oo}, and the other coefficients are increasing functions of R_{oo}.

The field curvature, astigmatism, transverse chromatic, and distortion aberrations are calculated from the beam trajectories. The calculation procedures are as follows:[9] First, the lens excitation parameter k is determined. For the trial lens excitation parameter k_1, the principal ray trajectory is determined. Two electrons are emitted from the crossover to the object with two emission angles, θ_1 and θ_2, between the optical axis, and the trajectories are calculated. From the radial coordinates of those two trajectories at the object, the emission angle θ_{40} is calculated so that the radial coordinate

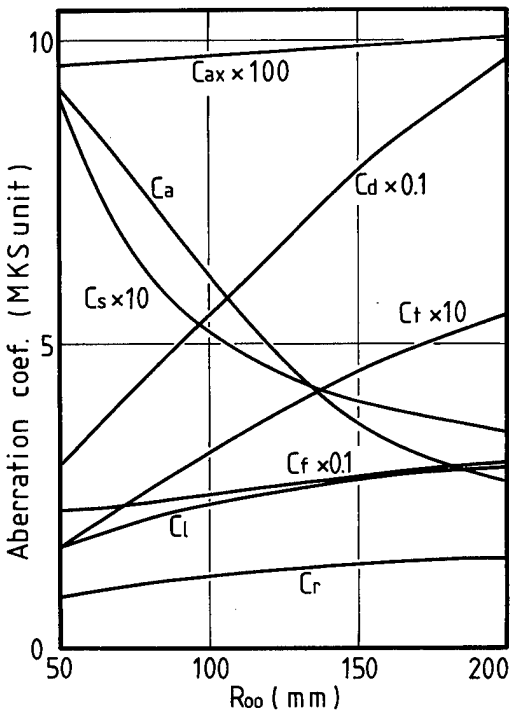


FIG. 4. Third-order aberration coefficients as a function of the object side bore radius R_{oo} of lens 1, where the crossover side bore radius R_{oc} is 20 mm. The notations are the same as in Fig. 3.

at the object R_o is 40 mm by interpolation or extrapolation. The trajectory with this θ_{40} is calculated from the crossover to the object. At the object, trajectory conditions are calculated by the method described in Ref. 9. These conditions contain X and Y coordinates X_0 and Y_0, and radial and azimuthal angles θ and ϕ, respectively. To obtain in-focus condition, four trajectories with the angular conditions $(\theta, \phi+0.1°)$, $(\theta, \phi-0.1°)$, $(\theta+0.1°, \phi)$, and $(\theta-0.1°, \phi)$ are calculated from the object to image. From these four trajectories, a focus position Z_1 for the lens parameter k_1 is calculated. For the next trial excitation parameter k_2, the same procedure is followed and a focus position Z_2 for the lens parameter k_2 is obtained. From these two focus positions, Z_1 and Z_2, for the lens parameters, k_1 and k_2, the in-focus lens excitation parameter k_0 is calculated by interpolation or extrapolation.

Second, for the various radial positions in the subfield, trajectories conditions are determined. For the lens excitation parameter k_0, the emission angle θ_{40} is calculated. Nineteen trajectories with $\theta_{40}\pm0.1n°$ $(n=0,1,...,9)$ are calculated from the crossover to the object. For each trajectory, the trajectory conditions at the object are calculated by the method mentioned above.

When four trajectories diverging radial and azimuthal direction $\pm0.1°$, respectively, per each 19 trajectory condition are calculated, the out-of-focus length ΔF and astigmatism difference ΔA are obtained.

When, for the principal ray trajectory conditions, 17 trajectories with beam energy of 100 keV, 100k±1 eV 100k±2 eV, 100k±5 eV, 100k±10 eV,... and 100k±200 eV are calculated, the transverse chromatic aberration may be obtained.

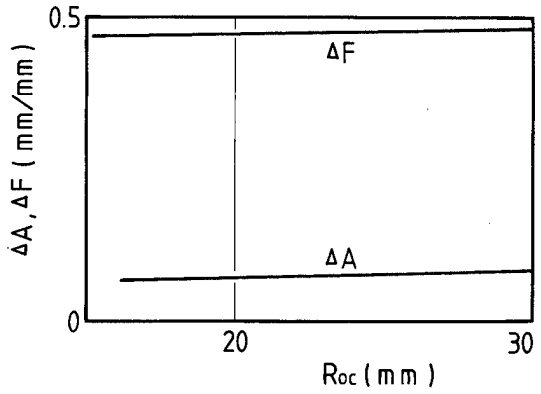


FIG. 5. Out-of-focus length ΔF and astigmatism difference ΔA as a function of R_{oc}, where R_{oo} is 100 mm.

For the 19 trajectory conditions, their trajectories are calculated from the object to image. From the radial and Y coordinates for the object and image R_o, R_i, Y_o (=0) and Y_i $(=\phi_d)$, the radial distortion $R_o/4-R_i$ and the azimuthal distortion ϕ_d are calculated.

Fig. 5 shows the astigmatism difference ΔA and the out-of-focus length ΔF as a function of the crossover side bore radius R_{oc}. Like the field curvature and astigmatism calculated by the third-order aberration coefficient shown in Fig. 3, they are almost independent of the crossover side bore radius, when the object side bore radius is 100 mm. Figure 6 shows the ΔF and ΔA as a function of the object side bore radius R_{oo}, where the solid line curve is the case where the crossover side bore radius R_{oc} is 20 mm, and the broken line curve is the case where $R_{oc}=R_{oo}/5$. As seen, the out-of-focus length ΔF is almost independent of R_{oo}, and the astigmatism length is a decreasing function of R_{oo}.

Figure 7 shows the radial and azimuthal transverse chromatic aberrations for the energy width of 1 eV for ΔR_c and $\Delta\phi_c$ as a function of the crossover side bore radius R_{oc}, where these aberrations are calculated for the trajectories emitted from the radial position of 40 mm at the object. As

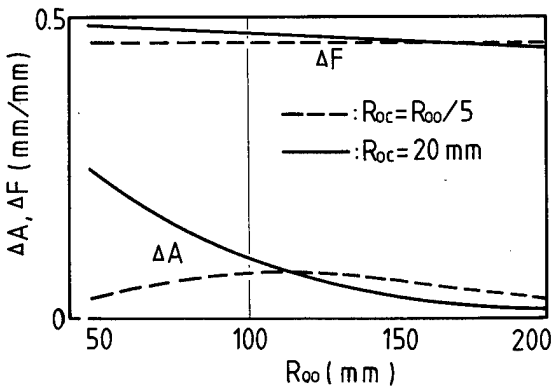


FIG. 6. ΔF and ΔA as a function of R_{oo}. The solid line curves are the case where R_{oc} is 20 mm, and the broken line curves are the case where R_{oc} is $R_{oo}/5$.

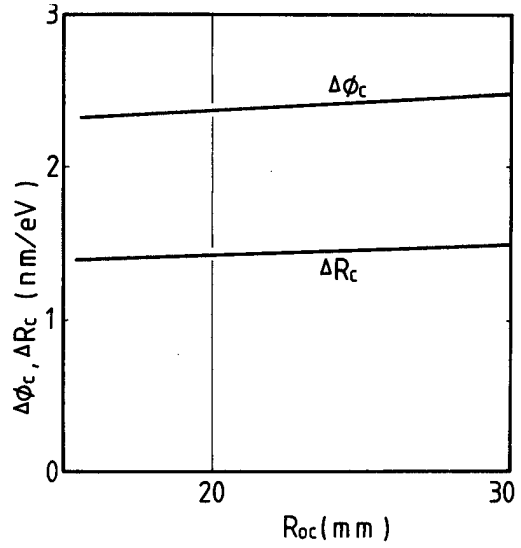


FIG. 7. Radial and azimuthal transverse chromatic aberrations per 1 eV for ΔR_c and $\Delta\phi_c$, as a function of R_{oc}, where R_{oo} is 100 mm.

seen, ΔR_c and $\Delta\phi_c$ increase slowly when the R_{oc} is increased. Figure 8 shows the beam landing angle θ_L with respect to the normal at the image. The radial and azimuthal transverse chromatic aberrations ΔR_c and $\Delta\phi_c$ per 1 eV are shown as a function of the object side bore radius R_{oo}. The radial and azimuthal transverse chromatic aberrations are a decreasing function of the object side bore radius R_{oo}. The beam landing angle θ_L increases from 0.4° to 2.5°, when R_{oo} increases from 50 to 200 mm.

The radial and azimuthal distortion $R_o/4-R_i$ and ϕ_d as a function of image radius are shown in the lower part of Fig. 9 ("real" case) where the bore radii R_{oo} and R_{oc} are 200 and 20 mm, respectively. The "ideal" case is discussed below in Sec. III. As seen, both distortions are linear function of R_i,

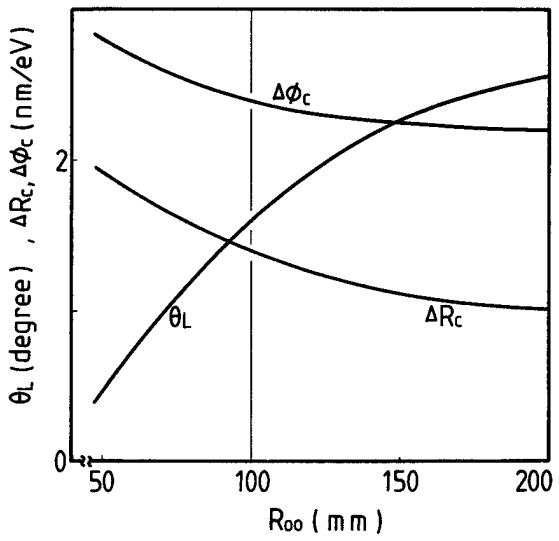


FIG. 8. Beam landing angle from the normal at the image θ_L degree, radial and azimuthal transverse chromatic aberrations ΔR_c and $\Delta\phi_c$, as a function of R_{oo}, where R_{oc} is 20 mm.

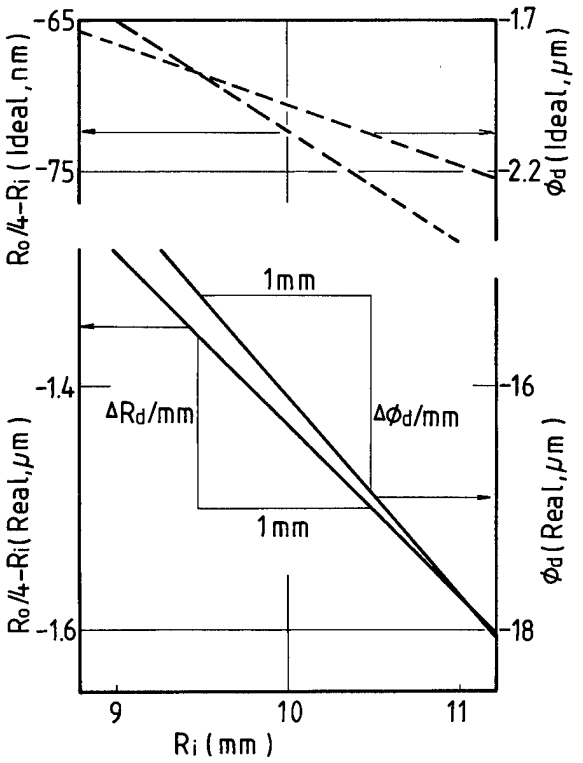


FIG. 9. Radial and azimuthal distortions $R_o/4-R_i$ and ϕ_d as a function of the image radius R_i. Bottom and top are for the real and ideal cases. Radial and azimuthal distortions per 1 mm image radius ΔR_d and $\Delta\phi_d$ are also defined.

so the distortion values per mm of the image radius ΔR_d (nm/mm) and $\Delta\phi_d$ (nm/mm) may be defined. When the radial distortion is corrected once for each subfield, residual

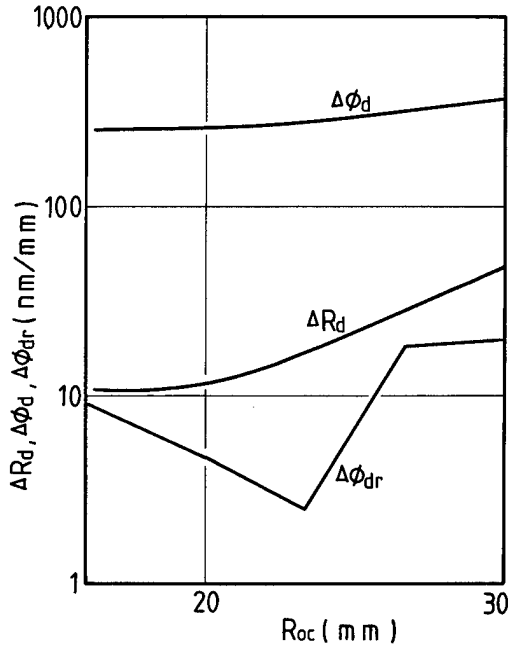


FIG. 10. Radial, azimuthal, and residual azimuthal distortions ΔR_d, $\Delta\phi_d$, and $\Delta\phi_{dr}$, respectively, as a function of the bore radius R_{oc}, where R_{oo} is 100 mm.

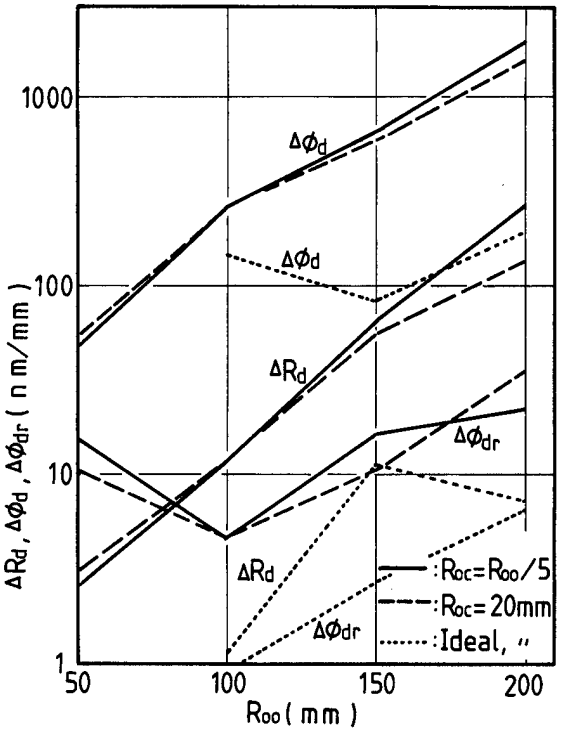


FIG. 11. ΔR_d, $\Delta\phi_d$, and $\Delta\phi_{dr}$ as a function of the bore radius R_{oo}, where the broken line curves are the case where R_{oc} is 20 mm, and the solid line curves are the case where R_{oc} is $R_{oo}/5$.

radial distortion is calculated as $\Delta R_d R_{\text{sub}}/2$, where R_{sub} is radial subfield size. When the azimuthal distortion is corrected to the main field center by rotation stage of the mask or wafer, residual distortion per 1 mm of the image radius is $\Delta\phi_d - \Delta\phi_{d0}$, where $\Delta\phi_{d0}$ is the azimuthal distortion per 1 mm of the image radius for the center of the main field.

Figure 10 shows the radial distortion per 1 mm of image radius ΔR_d, the azimuthal distortion per 1 mm of image radius $\Delta\phi_d$, and the residual azimuthal distortion per 1 mm of the image radius $\Delta\phi_{dr}$ as a function of the crossover side

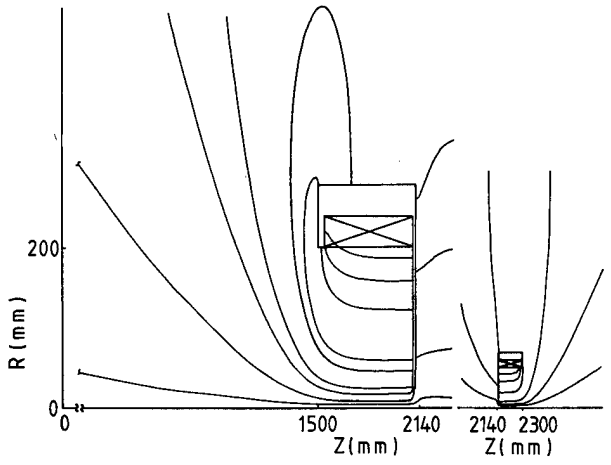


FIG. 12. Magnetic flux for the ideal lens system, where the lens parameters are same as in Fig. 2.

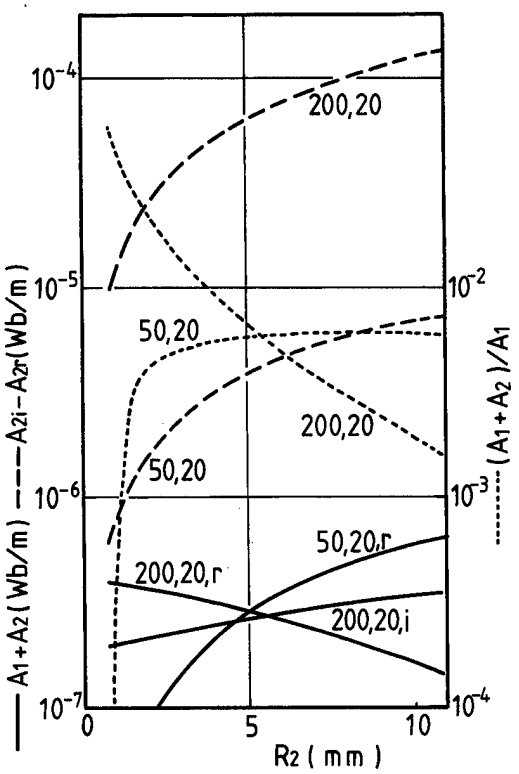


FIG. 13. Magnetic vector potential differences as a function of the radius R_2. Solid line curves is A_1+A_2, the broken line curves are $A_{2i}-A_{2r}$, and the dotted line curves are normalized values $(A_1+A_2)/A_1$, where subscripts 1 and 2 are for lenses 1 and 2, r and i are for the real and ideal cases, and 50 or 200 is R_{oo} and 20 is R_{oc}.

bore radius R_{oc}, where R_{oo} is 100 mm. As seen in Fig. 10, $\Delta\phi_d$ increases slowly when R_{oc} increases. The radial distortion per 1 mm of the image radius ΔR_d increases rapidly when R_{oc} increases. Though the azimuthal distortion $\Delta\phi_d$ is large, the residual azimuthal distortion $\Delta\phi_{dr}$ is smaller than 20 nm/mm.

Figure 11 shows ΔR_d, $\Delta\phi_d$, and $\Delta\phi_{dr}$ as a function of R_{oo}, where solid lines are the case where R_{oc} is $R_{oo}/5$ and the broken lines are the case where R_{oc} is 20 mm. As seen, the radial and azimuthal distortions increase rapidly when R_{oo} increases from 50 to 200 mm. As these calculated distortions include all-order aberration coefficients, these results are inconsistent with our previous understanding that the bore radius must be chosen sufficiently large that aberrations higher than third order are not significant. A sufficient large bore radius is harmful if wide gap lenses are used. The maximum residual azimuthal distortion is 40 nm/mm. When the subfield size is 250 μm square, residual azimuthal distortion is only 5 nm, and there is no problem.

III. DISCUSSION

A. Mechanism for generation of the distortions

To study the mechanism of generation of the radial and azimuthal distortions, the ideal case where the magnetic fields of lenses 1 and 2 are calculated in the absence of one another is studied. When Z is smaller or larger than the co-

ordinate of the crossover, the magnetic vector potentials for lens 1 or lens 2, respectively, are selected for the trajectory calculation. Figure 12 shows the magnetic flux profile for the ideal case, where the lens parameters are same as in Fig. 2. In Fig. 12, the magnetic flux formed by lenses 1 and 2 is distributed in the region where Z is smaller or larger than 2140 mm, respectively. In contrast, in Fig. 2, the magnetic flux formed by lens 2 is not distributed in the region where Z is smaller than 2140 mm, and that formed by lens 1 is not distributed in the region where Z is larger than 2140 mm, only where R is smaller than 100 mm.

From the comparison between the real and the ideal cases, it is expected that the magnetic vector potential in the region where the electron passes through are distorted in the real case. The magnetic vector potential A_1, as a function of radius at the center of lens 1 for Z of 1800 mm and A_2 as a function of radius at the lens center of lens 2 for Z of 2225 mm are studied, that is $A_1=A$ (1800 mm, R_1), $A_2=A$ (2225 mm, R_2), and $R_2=R_1/4$, where the lens parameters (R_{oo},R_{oc}) are (200, 20) and (50, 20). The vector potential A_1 for R_1 and A_2 for $R_1/4$ is of opposite sign and nearly equal in absolute value. Three types of the magnetic vector potential are calculated as a function of radius and are shown in Fig. 13. Solid line curves are A_1+A_2, broken line curves are $A_{2i}-A_{2r}$, and dotted line curves are normalized values $(A_1+A_2)/A_1$ for the real case. As seen in the solid curves, for the large bore radius and real case, A_1+A_2 is a decreasing function of radius R_2, and for the large bore radius and the ideal case and the small bore radius and real case, A_1+A_2 is an increasing function of radius R_2. As seen from the broken line curves, the difference between the real case and the ideal case for the large bore radius is ten times as large as that for the small bore radius. This means that for the large bore radius the magnetic vector potential is influenced more than that for the smaller bore radius. As seen in the dotted line curve for the small bore radius, the normalized magnetic vector potential difference $(A_1+A_2)/A_1$ is almost independent of R_2, however, $(A_1+A_2)/A_1$ for the large bore radius depends on R_2. This means that for the trajectories whose radial position is different, the smaller bore radius lens system gives nearly equal lens effect, however, the large bore radii lens system gives different lens effect for those trajectories. As a result, the distortions for the small bore radii lens system are small and those for the large bore radii lens system are large. Consequently it may be concluded that the generation mechanism of the large radial and azimuthal distortions for the large bore radii R_{oo} and $R_{oo}/4$ is that the magnetic fields of two lenses are influenced by each other, even if the crossover side bore radii R_{oc} and R_{oc} are small enough.

B. Design example

Requirements for the optics of the electron-beam projection system for the 4G-bit d-RAM pattern are assumed as follows. The main field and subfield sizes, distortion aberration, beam blur, and beam energy are 20 mm, 0.25 mm square, 20 nm, 45 nm, and 100 keV, respectively. Figure 14

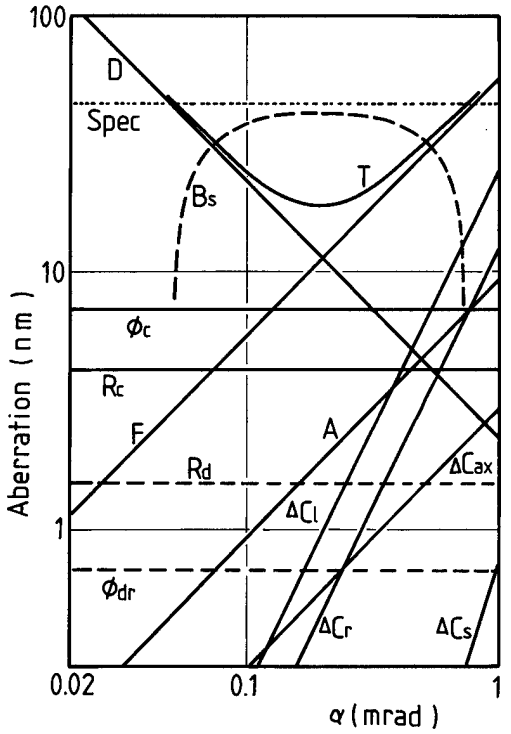


FIG. 14. Aberrations as a function of the beam half-angle α, T is the total beam blur, and B_s is the beam blur budget for the space charge effect.

shows aberrations as a function of the beam half angle, where the object–image distance, bore radii R_{oo} and R_{oc} are 800, 100, and 20 mm, respectively. They are calculated from the following equations: $D=0.61\lambda/\alpha$, $\Delta C_s=C_s\alpha^3$, $\Delta C_l=C_lR_m\alpha^2$, $\Delta C_r=C_rR_m\alpha^2$, and $\Delta C_{ax}=C_{ax}\alpha\Delta U/U$, where R_m is the maximum image radius and α is the beam semiangle, and where the spherical, axial chromatic, coma length, and coma radius aberration coefficients C_s, C_{ax}, C_l, and C_r, respectively, are calculated from Fig. 3. The field curvature and astigmatism aberrations F and A are calculated as $F=(0.25/2)\Delta F\alpha$ and $A=(0.25/2)\Delta A\alpha$, where ΔF and ΔA are shown in Fig. 5 at R_{oc} of 20 mm. The radial and azimuthal transverse chromatic aberrations are calculated as $\Delta U\Delta R_c$ and $\Delta U\Delta\phi_c$, where the energy width ΔU is 3 eV and ΔR_c and $\Delta\phi_c$ are shown in Fig. 7. The residual radial and azimuthal distortions are calculated from Fig. 10 as 1.5 and 0.7 nm, respectively. Total beam blur T is shown in the solid line curve. Since the beam blur specification is 45 nm, the beam blur budget for the space charge effect B_s may be shown as in the broken line curve. For the beam half angle from 0.05 to 0.7 mrad, the beam blur is smaller than 45 nm when the space charge effect is neglected. Though the astigmatism and the transverse chromatic aberrations are the decreasing function of the lens bore radii, these aberrations are not dominant in the total beam blur, as seen in Fig. 14. Therefore, the beam blur is not improved in the case of large bore radii.

IV. CONCLUSIONS

Radial and azimuthal distortion aberrations are an increasing function of the image side lens bore radius in the range

from 1.25 to 5 times as large as the maximum image field radius. This phenomenon is inconsistent with our previous understanding. An assumption is made that these large distortions for the large bore radii come from the influence of the magnetic field of one on the other, thereby destroying the symmetry. This assumption is confirmed from the following simulation. When these distortions are calculated for the ideal case where the magnetic fields are calculated in the condition without the other lens, they are decreased to around $\frac{1}{10}$ of those for the case where the magnetic fields are calculated in the real condition with the lenses in proximity. When the object–image distance is 800 mm and the bore radii of lens 1 are 100 and 20 mm, the residual radial and azimuthal distortions are 1.5 and 0.7 nm and the beam blur is smaller than 45 nm for the beam semiangle from 0.05 to 0.7 mrad, where the main-field and subfield sizes in the image plane are 20 mm and 250 μm, respectively, beam energy is 100 keV, and the space charge effects are neglected.

ACKNOWLEDGMENTS

The authors would like to thank T. Kuramoto, S. Yabumoto, Y. Miyazaki, and S. Suzuki of Nikon Corporation and Dr. M. Sogard of Nikon Research Center of America for their useful discussion through this study.

[1]H. C. Pfeiffer, IEEE Trans. Electron Devices **ED-26**, 486 (1979).
[2]M. B. Heritage, J. Vac. Sci. Technol. **12**, 2346 (1994).
[3]M. Nakasuji, S. Suzuki, and H. Shimizu, Rev. Sci. Instrum. **64**, 446 (1993).
[4]H. C. Pfeiffer and W. Stickel, Microelectron. Eng. **27**, 143 (1995).
[5]T. Asai, S. Ito, T. Eto, and M. Migitaka, Jpn J. Appl. Phys. **19**, Suppl. 19-1, 47 (1980).
[6]S. D. Berger, J. M. Gibson, R. M. Camarda, R. C. Farrow, H. A. Huggins, J. S. Kraus, and J. M. Liddle, J. Vac. Sci. Technol. B **9**, 2996 (1991).
[7]H. W. Koops, *Microelectronics Engineering 88* (Elsevier, North-Holland, New York, 1989), p. 217.
[8]M. Nakasuji and H. Shimizu, Jpn J. Appl. Phys. **33**, 1195 (1994).
[9]M. Nakasuji and H. Shimizu, Jpn J. Appl. Phys. **33**, 5086 (1994).
[10]M. B. Heritage, J. Vac. Sci. Technol. **12**, 1135 (1975).
[11]M. Nakasuji and H. Shimizu, Jpn J. Appl. Phys. **34**, 3746 (1995).
[12]P. F. Petric and V. D. Beck, Microelectron. Eng. **13**, 161 (1991).

Reticle fabrication by high acceleration voltage electron beam: Representative figure method for proximity effect correction [VI]

Takayuki Abe[a)]
ULSI Research Center, R&D Center, Toshiba Corporation, 1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki, Kanagawa 210, Japan

(Received 19 October 1995; accepted 1 May 1996)

Proximity effect correction for reticle making by the dose correction method is discussed. A new algorithm for calculating the optimum dose is proposed, which is based on the dose formula method and the representative figure method. Its main feature is that dose evaluation points are fixed at individual small regions whose size is sufficiently small compared with the backscattering range. The calculation speed of the computer is evaluated which is sufficient to suppress the correction time to less than the writing time. The required calculation speed is 500 MIPS×4 CPU at most for a minimum feature size greater than 0.2 μm on the reticle. This result suggests that the real-time proximity effect correction is possible for making reticles. Furthermore, when the algorithm is applied to x-ray mask fabrication and direct writing process, the calculation time can be suppressed to less than the writing time. © *1996 American Vacuum Society.*

I. INTRODUCTION

Electron beam (EB) writing systems have been powerful manufacturing tools for making reticles. Pattern dimension on wafer and the magnification rate of the reticle becomes smaller with succeeding LSI generations. Correspondingly, the minimum feature size on the reticle is also getting smaller, and higher accuracy is requested for reticle fabrication. The minimum feature size on the reticle will be 0.8 μ–0.5 μm and the permitted dimension error will be less than ±15 to ±10 nm for making 1 and 4 Gbit DRAM generations.

To obtain such a high accuracy, both the beam resolution and the forward scattering range of electrons in the resist must be small, of the order of 30 nm. Accordingly, an EB system with high acceleration voltage[1–3] will be required, even for reticle fabrication.[4,5] Reticle writing by high acceleration EB, however, is subject to many problems, for example, the proximity effect[6–13] and resist heating.[14,15]

In this article, proximity effect correction by the dose correction method is discussed for reticle fabrication. Dose correction method has an advantage to Ghost[8] and Phantom,[16] regarding high speed writing, because subsidiary exposure with a blurred beam is not needed. However, the time for calculating an optimum dose is crucially long. Objectives of this article are (1) to propose a new algorithm for calculating optimum dose, (2) to show that the time for calculating optimum dose can be suppressed to less than the writing time, when the new algorithm is applied to reticle fabrication, and (3) to x-ray mask fabrication and direct writing process. The new algorithm is based on the dose formula method[12,13] and the representative figure method.[16–19] In a modification of the previous algorithm,[18] the dose evaluation points are fixed at the center of the individual small regions whose size is sufficiently small compared with the backscattering range σ_b.

In this article, the new algorithm is shown to reduce the calculation time to about 1/30, compared with the previous algorithm, without deteriorating the total correction accuracy. The calculation speed of a computer is evaluated which is sufficient to suppress the correction time to less than the writing time. The required calculation speed is 500 MIPS×4 CPU at most for a minimum feature size greater than 0.2 μm on reticle. The result suggests that the real-time proximity effect correction is possible for reticle fabrication.

The contents of this article are as follows. In Sec. II, the assumptions and conditions of this article are explained. In Sec. III, the data amount for proximity effect correction is evaluated and the model system for real-time correction is explained. In Sec. IV, the new algorithm for calculating optimum dose is explained. The accuracy of the method is evaluated in Sec. V. In Sec. VI, the software system for evaluating the algorithm is explained. In Sec. VII, the calculation speed is evaluated. In Sec. VIII, the required calculation speed of computer is evaluated, which is sufficient to suppress the correaction time to less than the writing time. This article is summarized in Sec. IX. In Appendices A and B, the applications of the new algorithm to both x-ray mask fabrication and direct writing process are discussed. In Appendix C, the evaluation of the edge error is explained in detail.

II. ASSUMPTIONS, CONDITIONS, AND SOME NOTES

In this article, the double Gaussian approximation[7] for deposited energy is used. It is assumed that the high voltage EB system is used and the forward scattering range is 0. Then the deposited energy $E(x)$ at the position x is described as

$$E(x)=\int_{\text{pattern}} \frac{1}{1+\eta_E/2}\left\{\delta(x-x')+\frac{\eta_E}{\pi\sigma_b^2}\times \exp[-(x-x')^2/\sigma_b^2]\right\}dx', \quad (1)$$

a)Electronic mail: 000092030102F@tg-mail.t

where the η_E is the ratio between the deposited energy caused by backscattering electron and that by the forward scattering electron. In this article, the acceleration voltage of electron is set at 50 kV, and the backscattering range σ_b is set at 10 μm. Though, when the forward scattering range can be approximated by zero, our discussion can be used for any acceleration voltage by scaling all types of dimension by σ_b. The reason is that any quantity (e.g., the correction accuracy) depends on acceleration voltage only through σ_b.

In many articles, some assumptions are used for the *ideal* correction of the proximity effect. Case A: Ideal correction is to make the deposited energy for the entire exposed region uniform.[12] Case B: Ideal correction is to make the deposited energy at the edge of all figures uniform.[15] In this article, case A is used for simplicity. Though, the results of this article, such as the calculation speed, do not change, even when the other assumption is used.

The magnification ratio of reticle is assumed to be ×4. When length or dimension is described as 2 μm (w:0.5 μm) in this article, this means 2 μm on reticle and 0.5 μm on wafer. When dimension is described as 2 μm without any comment, this means 2 μm on reticle.

III. DATA AMOUNT AND REAL-TIME CORRECTION

In order to realize highly accurate proximity effect correction, figures must be cut into small figures whose maximum size is 2 μm or so. It is to suppress the edge error.[16] The cut small figures are called elements. The data amounts for elements are evaluated as follows.

Gate array of 130K gates was used to evaluate the worst case. The (black/white) inverted pattern of the wiring layer was selected for the evaluation because the layer has no array structure and is the most complicated layer. Chip size is 6×6 cm (w:1.5×1.5 cm) and the minimum feature size in the layer is 6 μm (w:1.5 μm).

For the systematic evaluation of data amount, it is assumed that the progress of device technology by one generation reduces the design rule to 1/2 and does not change the chip size. Then, the following relation holds for the data amount:

$$Q(\lambda)=(\mu/\lambda)^2*q(\lambda), \tag{2}$$

where $Q(\lambda)$ is the data amount for the future device with the minimum feature size of λ. The character $q(\lambda)$ is that for the pattern to which the present device is shrunk to have the minimum feature size of λ. The character μ is the minimum feature size of the present device without shrinkage; i.e., μ=6 μm (w:1.5 μm). The factor $(\mu/\lambda)^2$ is to make chip size invariant. Relation (2) holds not only for the data amount but also for the number of shots and the calculation time.

The data amounts of elements are shown in Fig. 1. Data size for an element was assumed to be 10 bytes. Figure 1 shows that the element division causes the increase of data amount to 10–20 G bytes. Then, it takes 5 h or so to store such a large volume of data to hard disk from a storage media such as a magneto-optical disk. This value is crucially large.

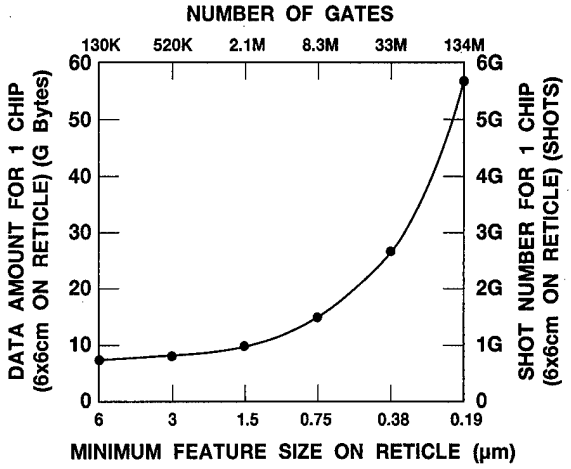


FIG. 1. Data amount of elements and shot number. Gate array was used for the evaluation. Chip size is 6×6 cm on reticle. Both element size and the maximum size of shot are 2×2 μm.

There are various possible solutions to this problem. One solution is (pseudo) real-time correction in an EB system. In order to realize real-time correction, the time both for obtaining the representative figures and for evaluating the optimum dose must be less than the writing time. In the following sections, we discuss how to satisfy this condition. Even if the real-time correction is not applied, the satisfaction of this condition is effective for quick turnaround time to make reticles.

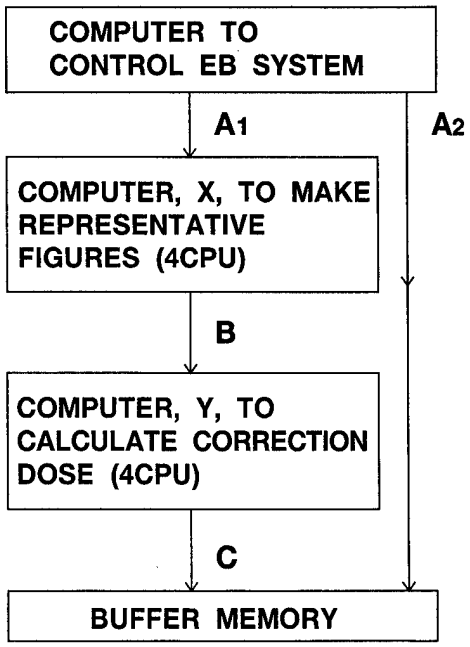


FIG. 2. Model system for real-time proximity effect correction. *A*1, *A*2, *B* and *C* are high speed bus, such as direct memory access (DMA) bus. Original figures are transferred through pass *A*1 and *A*2. Computer X prepares the representative figures Representative rectangles are transferred through pass *B*. Computer Y calculates the optimum correction dose. Calculated optimum dose is transferred through pass *C*.

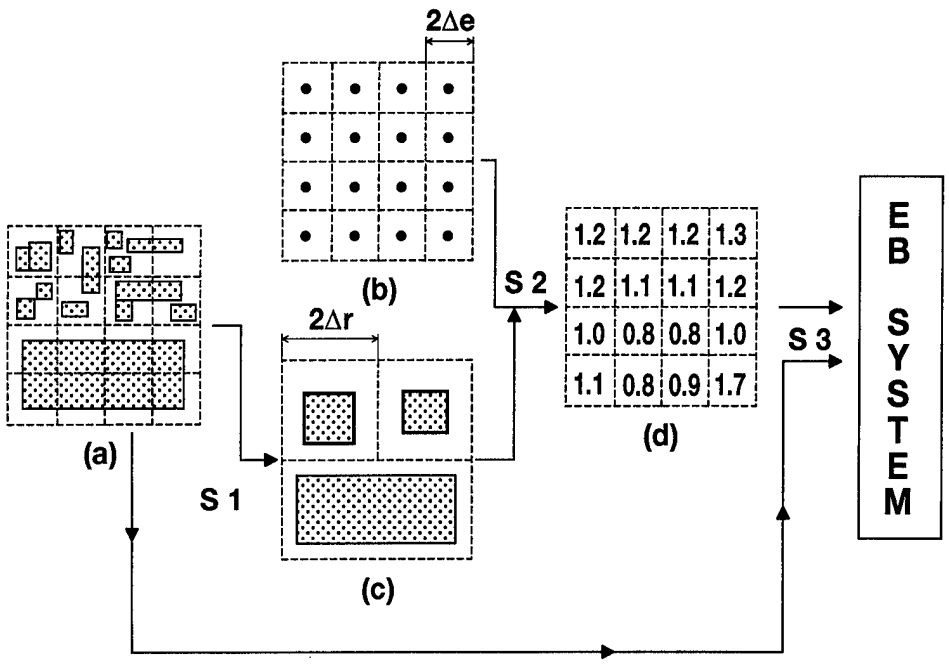


FIG. 3. The method for correction calculation (a) original figures and (b) elements, dots show the center of small region at which the correction dose is evaluated. (c) Reference rectangles (representative figures) and (d) calculated optimum dose. S1: preparation of representative rectangles (reference rectangles). S2: correction calculation using the dose formula method. S3: data transfer to EB system.

Figure 2 shows a simple model system for real-time correction, which is referred in the following discussions. The calculation process is carried out by the unit of stripe, which is the region exposed by one stage continuous motion. Pipeline architecture is assumed; that is when a stripe is processed in computer Y, the next stripe is processed in computer X. The data transfer time at the pass $A1$ and $A2$ is needed even if the proximity effect correction is not carried out. Therefore, the time is not considered in this article. The data transfer time at the pass B and C is discussed in the following section.

IV. NEW CORRECTION ALGORITHM

Figure 3 shows the new algorithm. The improvement from the previous algorithm[18] is that the dose evaluation point is fixed at the center in individual small regions for elements.

Step 1: Representative rectangles are made from the original LSI patterns in individual regions S_r whose size $(2\Delta_r \times 2\Delta_r\ \mu\text{m})$ is comparable with the backscattering range σ_b. This is used as a reference rectangle in step 2. In the case of the system in Fig. 2, Step 1 is carried out by computer X.

Step 2: The optimum dose is evaluated for the center (x_i, y_j) in the individual small regions S_e whose size $(2\Delta_e \times 2\Delta_e\ \mu\text{m})$ is sufficiently smaller than the backscattering range. Any suitable formula for the optimum dose is used by referring to the representative rectangles obtained in step 1. The following shows the example where Pabcovich's formula is used.

The formula of the optimum dose $D(x_i, y_j)$ at the position (x_i, y_j)[12] can be written as

$$D(x_i, y_j) = c[1/1 + kU(x_i, y_j)], \tag{3}$$

$$U(x_i, y_j) = \sum_m \phi_1(x_i, y_j, m) * \phi_2(x_i, y_j, m), \tag{4}$$

$$\phi_1(x_i, y_j, m) = \text{erf}[(r_m - x_i)/\sigma_b] - \text{erf}[(l_m - x_i)/\sigma_b], \tag{5}$$

$$\phi_2(x_i, y_j, m) = \text{erf}[(u_m - y_j)/\sigma_b] - \text{erf}[(l_m - y_j)/\sigma_b], \tag{6}$$

where c is a constant and k is an optimization parameter for correction. The values of the x coordinate at the right side, and the left side of the reference rectangle m, are denoted by r_m and l_m, respectively. The values of the y coordinate at the upper side and the lower side of the reference rectangle m are denoted by u_m and l_m, respectively. The character $\phi_1(\phi_2)$ corresponds to the correlation between a dose evaluation point and one reference rectangle for $x(y)$ coordinate. The summation Σ is carried out for the reference rectangles m a part of which or whole of which are located within a threshold distance, L_{th}, from the evaluation points (x_i, y_j). For example, the threshold distance L_{th} is set at $2\sigma_b$. The calculation is carried out by referring to a table of error functions for the reference rectangles.

Step 3: EB system exposes all shots in a small region S_e with the optimum dose assigned for the small region S_e.

This procedure reduces the total calculation time—compared with the previous algorithm—because of two reasons. The first reason is that the process for obtaining elements is abbreviated. Only the process for obtaining

reference rectangles is needed. The calculation time for the latter process is sufficiently shorter than that for the former process (1/2–1/5). The second reason is that the procedure reduces the calculation time for correction itself (step 2) for the following reason:

It is found from Eqs. (4) and (5) that the relations

$$\phi_1(x_i,y_1,m)=\phi_1(x_i,y_2,m)=\phi_1(x_i,y_3,m)=\cdots, \quad (7)$$

$$\phi_2(x_1,y_j,m)=\phi_2(x_2,y_j,m)=\phi_2(x_3,y_j,m)=\cdots \quad (8)$$

hold for arbitrary i, j, and m. This is caused by fixing the dose evaluation point at the center of individual small regions S_e. Because of relations (7) and (8), the number of references to the error function table is reduced. For example, suppose that the correction is carried out by a unit of a certain type of area A such as a subfield, and $N_e \times N_e$ is the number of small regions S_e in the area A. Then the number of references to the error function table for one reference rectangle is $2 \times N_e$. On the other hand, when the evaluation point is not fixed as the previous algorithm, relations (7) and (8) do not hold,[18] and the number of references to the error function table is $2 \times N_e \times N_e$. The fix of evaluation points reduces the number of references to the error function table to $(1/N_e)$. Thus, the high speed calculation is realized.

The number of calculations for correction itself (step 2) by this procedure depends on only the region sizes $2\Delta_e$ and $2\Delta_r$ (i.e., the backscattering range of electron σ_b) and does not depend on the number of original LSI patterns. Therefore, the calculation time of correction itself is independent of the design rule and device generation. This is the advantage of the representative figure method.[10]

V. CORRECTION ACCURACY

A new type of error is caused by fixing the dose evaluation point at the center. The error is called F error in this article.

Figure 4 shows the schematic explanation of both the worst case of the edge error[16] and the F error. It is found from Fig. 4 that the worst edge error exceeds the F error. This means that the new algorithm does not deteriorate the correction accuracy, compared with the previous algorithm in Ref. 18.

The worst error of the deposited energy is evaluated and is shown in the Fig. 5. The worst error appears at the edge of the large pattern whose size is sufficiently larger than the backscattering range σ_b. The error depends on the normalized small region size $(2\Delta_e/\sigma_b)$ for element, when the region size for reference rectangle is fixed. The error in the Fig. 5 involves all errors; that is, the edge error, the error of Pavcobich's formula and the intrinsic error of the representative figure method.

VI. SOFTWARE SYSTEMS

In order to evaluate the calculation speed, software system P3C (PPPC; preprocessor for proximity effect correction)[20] is used, and the new software system STEALTH (system for proximity effect correction for advanced lithography and reticle) is developed.

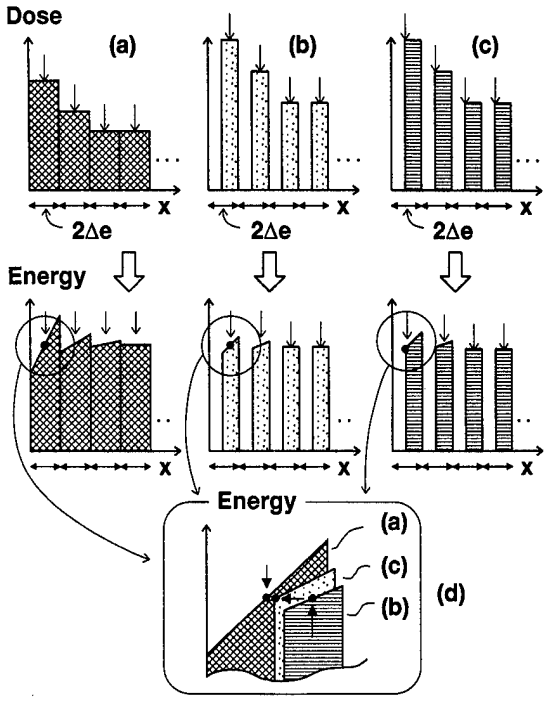


FIG. 4. Edge error and F error. Small arrows show the points at which the optimum dose is evaluated. (a) Shows the worst edge error, (b) shows the case that the evaluation point is at the center of gravity of pattern in a small region, and (c) shows the F error where the evaluation point is fixed at the center of small region. Deposited energy at the evaluation point is almost ideal for (a), (b), and (c) in that only the intrinsic error of the representative figure method and the error of correction algorithm appear at the point. The inclination of the deposited energy becomes maximum in case (a) because the effect of backscattering electron is maximum. Thus, the deposited energy deviates in a small region as shown in (d). The worst edge error exceeds the F error.

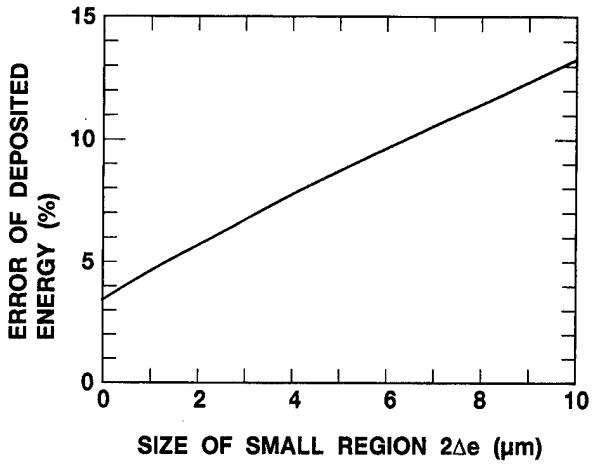


FIG. 5. Error dependence on the small region size for element $2\Delta_e$. Correction error of deposited energy was evaluated by numerical calculation. The region size for reference rectangle $2\Delta_r$ was fixed at 10×10 μm. The worst error appears at the edge of the large figure (whose size is sufficiently larger than the backscattering range σ_b).

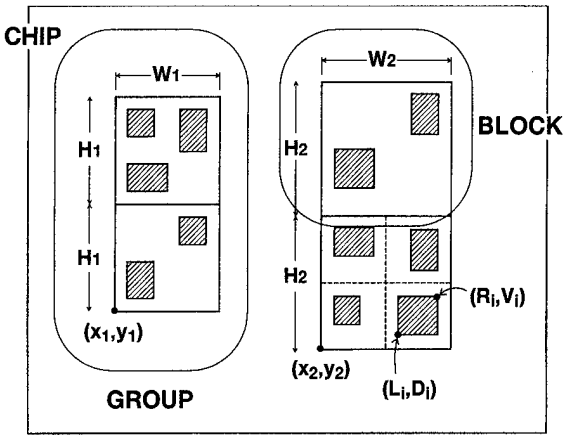


FIG. 6. Data format for STEALTH (1). Layout in a chip is shown. Shaded rectangles are the representative rectangles, which are referred to for correction calculation. Data consists of group, block, and representative figures. Corresponding data format is shown in Fig. 7.

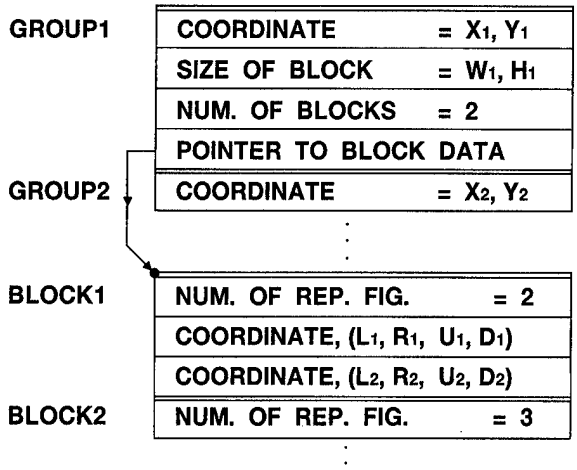


FIG. 7. Data format for STEALTH (2). Representative figures in Fig. 6 are described according to this format.

A. P3C (obtaining representative rectangles)

Details of the system are explained in Ref. 20. Only the rough sketch of the system is given in the following.

P3C carries out (step 1) in Sec. III and makes the representative figures from the original LSI pattern. The P3C system carries out the calculation in parallel using 4 CPUs. Data format for EB writing system EX-8D is used for input and output data. The output from a data conversion system is used as input data for the P3C system.

The data format is different from the format used in STEALTH which is explained in the next subsection. However, the difference is negligible for evaluating the calculation time.

B. STEALTH

The STEALTH software system carries out the correction calculation in (step 2) described in Sec. III.

The system does not use hard disk at the application (software) level. The STEALTH system accesses the input data in memory and stores the calculated optimum dose to the system memory. This is (1) to eliminate the disk access time and (2) for the application to the real-time system, such as the system shown in Fig. 2.

The data format for input data is schematically explained in Figs. 6 and 7. The data consists of the block in which some representative rectangles exist. The format is determined to reduce data handling time.

The STEALTH system calculates the optimum dose by the unit of the block which is called target block. The main sequence is as follows:

(1) (P-1) search nearby blocks which touch or are within the threshold distance L_{th} from the edge of the target block. Figure 8 shows the searched blocks and the target block.
(2) (P-2) obtain the correlations ϕ_1 and ϕ_2 by referring to the reference rectangles and using the error function table.
(3) (P-3) evaluate the value of U [calculate Eq. (4)].
(4) (P-4) calculate the optimum dose [Eq. (3)].

The STEALTH system is not applied to the arrayed pattern, because the array structure can be easily expanded by P3C without a burden. P3C can make one representative figure for several arrayed patterns in a small region S_r, and the obtained representative figure can be used several times for expansion. Such a process requires relatively little time.

When the STEALTH system is used in the model system in Fig. 2, attention must be paid to the data transfer time at pass B and C. It is assumed that (1) the data transfer speed is 20

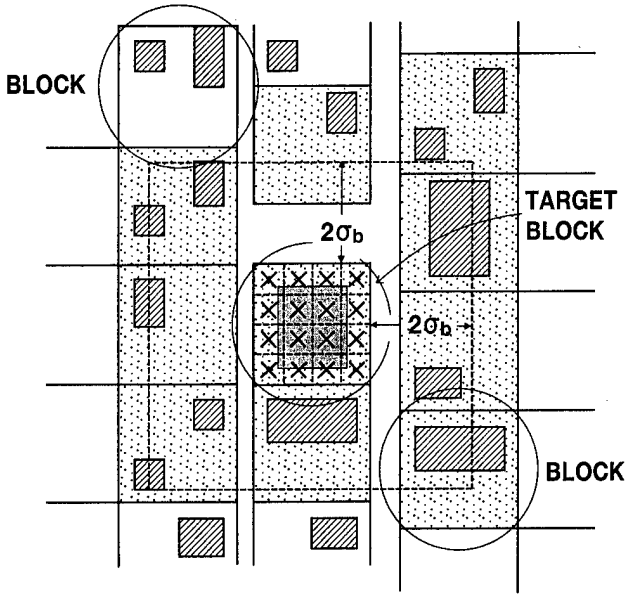


FIG. 8. Blocks used for correction. Position marked as (×) in a target block shows the center of small region at which the optimum dose is evaluated. Shaded rectangles are the reference rectangles (representative figures). Dotted rectangles show the blocks which touch or are within $2\sigma_b$ from the edge of the target block. Reference rectangles in the dotted blocks (including a target block) are referred to for correction calculation.

Mbytes/s, (2) the small region sizes $2\Delta_e$ and $2\Delta_r$ are 2 and 10 μm, respectively, and (3) the data is transferred by the unit of stripe whose size is 0.1×14 cm. Regarding the pass *B* (input of STEALTH), the data amount for one stripe is at most 22 Mbytes or so, when the data amount for one representative figure is 16 bytes. The corresponding data transfer time at path *B* is 0.8 s. Regarding the pass *C* (output of STEALTH), the data amount for one stripe is 35 Mbytes, when the data amount for one optimum dose is 1 byte. The corresponding data transfer time is about 1.8 s. The time for writing one reticle will be over 1000 s for the writing area of 200 cm^2–14×14 cm, as will be shown in Sec. VIII. Thus, the time for writing one stripe is over 7 s. The estimated values for data transfer time are sufficiently small, compared with the time for writing one stripe. The data transfer time does not limit the throughput of the EB system.

VII. CALCULATION TIME

A. P3C (step 1)

An engineering workstation (EWS) with 4 CPUs was used. The processing speed of each CPU is 52 MIPS. (Because superscalar architecture is adopted for the microprocessor, effective MIPS value depends on a program. The effective value is supposed to be larger than 52 MIPS for P3C. The MIPS value should be considered as a guideline. The value of SPECint 92 is 40 for the processor.) The P3C system carries out the calculation in parallel using 4 CPUs.

When the device pattern has array structure, the calculation time becomes short. For example, the calculation time for 64 M DRAM is less than 1 min for even the most complicated layer of the device. Thus, the gate array described in Sec. II is used for the worst case evaluation. Relation (2) is used for systematic evaluation, and the evaluated calculation time is converted to the condition that the size of writing area is 200 cm^2 on reticle.

The results are shown in Fig. 9. The calculation time is about 1/4–1/70 of the data conversion time. The acceleration ratio by 4 CPU multiprocessing is about 3.7 (370%), in the case of a 130 k gate array.

B. STEALTH (step 2)

Only 1 CPU of the computer described in the previous subsection was used. The small region size $2\Delta_r$ for reference rectangle is fixed at 10 μm. Block size was set at 20×20 μm, which corresponds to $2\sigma_b$. The threshold distance L_{th} is also set at $2\sigma_b$. Corresponding cutoff error is 0.25% at most, and is negligibly small.

Proximity effect correction was carried out for the region of 1×1 cm. The evaluation was carried out under the condition that the representative figures exist in all small regions S_r. The number of representative figures is 1×10^6.

The obtained (elapsed) calculation time is shown in Fig. 10. The calculation time is of the order of 1 min. The calculation speed per MIPS by STEALTH is 30 times faster than that by the previous system SPIDER II,[18] when the small region size $2\Delta_e$ is 2 μm.

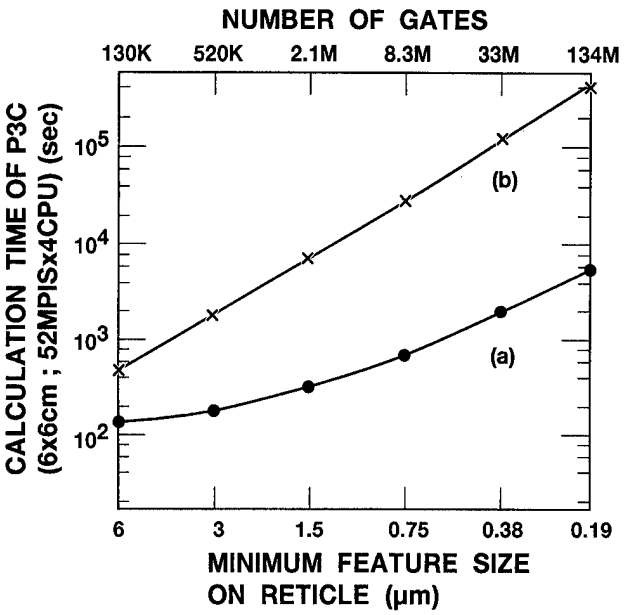


FIG. 9. Calculation time by P3C. (a) Time for obtaining representative rectangles by P3C. (b) Data conversion time (reference). Parallel processing using 4 CPU was adopted. Calculation speed of 1 CPU is 52 MIPS. Gate array was evaluated. Chip size was 6×6 cm on reticle. Small region size for reference rectangles $2\Delta_r$ was set at 10 μm on reticle.

Figure 10 shows that almost all of the calculation time consists of the time for evaluating Eq. (4) (P-3 in Sec. VI). The time required for searching the nearby block is small. This shows the effectiveness of the data format described in Sec. VI. The time for evaluating Eqs. (5) and (6) (P-2 in Sec. VI) is also small, as expected in Sec. IV. This shows the

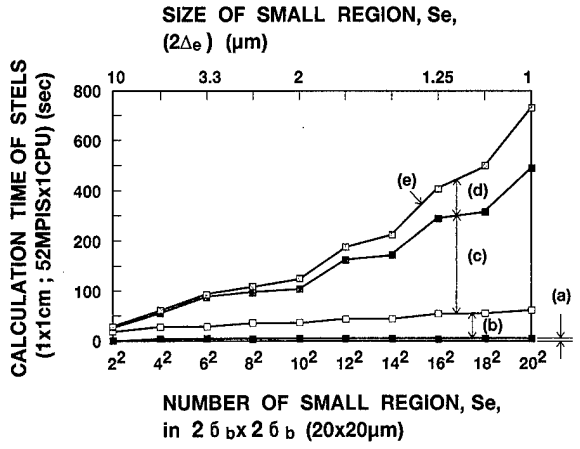


FIG. 10. Calculation time for correction itself (by STEALTH). 52 MIPS 1 CPU computer was used. Region of 1×1 cm on reticle was corrected. Small region size for reference rectangle $2\Delta_r$ was fixed at 10×10 μm. Calculation time depends on the small region size for element. (a) Time for searching nearby blocks (Sec. VI,P-1), (b) time for calculating the correlation ϕ (Sec. VI,P-2), (c) time for calculating Eq. (4) in Sec. IV (Sec. VI,P-3). (d) Time for calculating Eq. (3) in Sec. IV (Sec. VI,P-4). (e) Total calculation time by STEALTH.

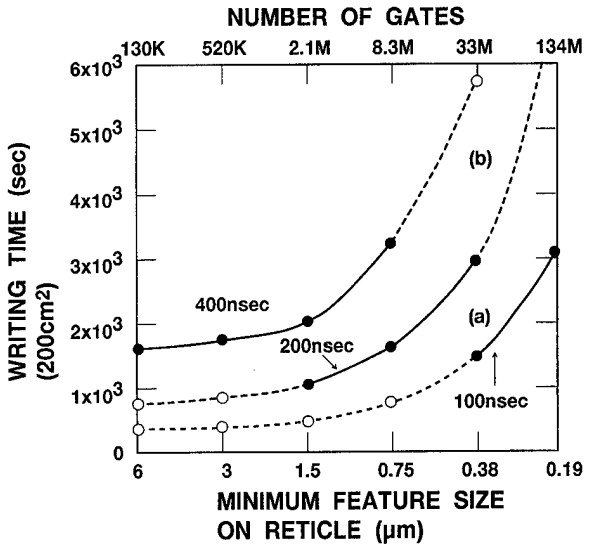


FIG. 11. Time for writing gate array on reticle. Writing area is 200 cm^2. Cycle time of EB system will decrease in accordance with the minimum feature size of device which will be written in the future. 100–400 ns shown in the figure is the cycle time of EB system. Black circles show the combination of the minimum feature size and the cycle time which is used or will be used in the future.

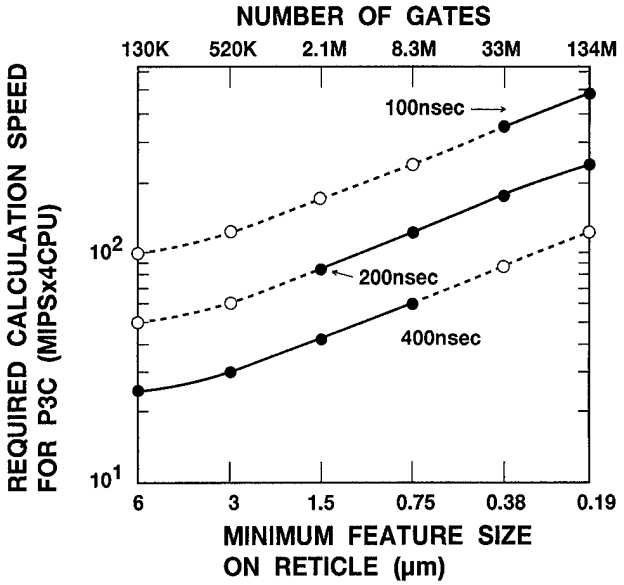


FIG. 12. Required calculation speed of computer for obtaining the representative figures by P3C. Parallel processing using 4 CPU was assumed. The figure shows the calculation speed of computer which can suppress the calculation time by P3C to less than the writing time. A small region size for reference rectangle was set at 10 μm. 100–400 ns shown in the figure is the cycle time of EB system. Black circles show the combination of the minimum feature size and the cycle time which is used or will be used in the future.

effectiveness of the new algorithm; that is, the fix of the evaluation point at the center of small region S_e.

VIII. DISCUSSION

In this section, the calculation speed of computer is evaluated, which is required to suppress the calculation time to less than the writing time.

A. Writing time

The shot number of the gate array described in the Sec. II was evaluated and is shown in Fig. 1. Equation (2) was used for the evaluation. The maximum shot size was assumed to be 2×2 μm.

The corresponding writing time is shown in Fig. 11. The writing area was set at 200 cm^2. The shot number in Fig. 1 was converted in proportion to the difference of the writing area. The writing time was evaluated as (shot number×cycle time), where cycle time means the sum of the time for one shot exposure and the preparation time from the end of one shot exposure to the start of the next shot exposure. The other times, such as the time for positioning a subfield, were ignored. The cycle time will decrease with the minimum feature size of device to be written by the EB system in the future. In this and the following subsections, three types of cycle time (400, 200, and 100 ns) were used for the evaluation. When the minimum feature size is 0.2 μm, the writing time becomes 3200 s–0.9 h.

B. P3C

In this and the next subsections, parallel processing using 4 CPU is assumed, and the calculation speed of the computer is evaluated which is subjected to suppress the calculation time to less than the writing time (in Fig. 11). It is assumed that the correction error of 6% is permissible regarding the deposited energy. Correspondingly, the sizes of small region S_e and S_r are fixed at 2 and 10 μm, respectively, from Fig. 5.

Figure 12 shows the evaluated calculation speed of computer, which is required for P3C to finish the calculation within the writing time. The estimation was carried out by comparing the writing time and the calculation time of P3C for a unit area.

The required calculation speed increases with each successive device generations for the following reason: The complexity of the device pattern increases with the device generation, and the corresponding calculation time by P3C increases. On the other hand, the number of shots (i.e., the writing time) does not increase much with each device generation.

When the minimum feature size is 3.0 μm (w:0.8 μm), and the cycle time is 400 ns, the corresponding writing time is 30 min. The required calculation speed of computer is 40 MIPS/CPU×4 CPU. Present EWS satisfies the condition. Highest calculation speed is required, when the minimum feature size is 0.2 μm (w:0.05 μm) and the cycle time is 100 ns. The corresponding writing time is 3200 s. The required calculation speed is 500 MIPS/CPU×4 CPU. Although the present EWS does not satisfy the condition, the required calculation speed is not a problem, because microprocessors of 1000 MIPS are expected to be produced when the minimum feature size is 0.2 μm (w:0.05 μm).

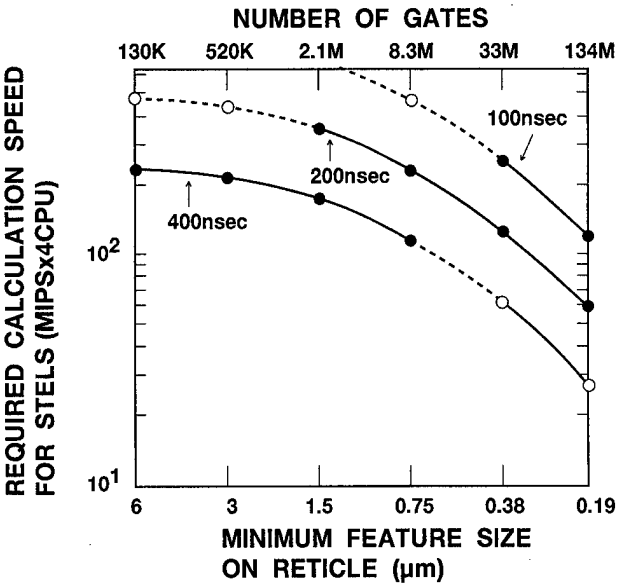


FIG. 13. Required calculation speed of computer for obtaining the correction dose by STEALTH. Parallel processing using 4 CPU was assumed. The figure shows the calculation speed of computer, which can suppress the calculation time for correction itself to less than the writing time. Small region sizes for reference rectangle $2\Delta_r$ and for element $2\Delta_e$ were set at 10 and 2 μm, respectively. 100–400 ns shown in the figure is the cycle time of EB system. Black circles show the combination of the minimum feature size and the cycle time which is used or will be used in the future.

C. STEALTH

The acceleration ratio by the 4 CPU parallel processing is assumed to be 3.6. The real ratio will be better than the value because STEALTH accesses high speed bus and does not access low speed hard disk.

Figure 13 shows the calculation speed which is required for STEALTH to complete the calculation time within the writing time. The estimation was carried out by comparing the writing time and the calculation time of STEALTH for a unit area.

The required calculation speed decreases with each successive device generation under the condition that the cycle time is independent of the device generation. The reason is (1) the calculation time for self-correction is independent of the device generation because of the representative figure method; on the other hand, (2) the number of shots (i.e., the writing time) increases with each successive device generation.

When the minimum feature size is 0.2 μm (w:0.05 μm) and the cycle time is 100 ns, the corresponding writing time is 3200 s. Then, the calculation speed of 120 MIPS/CPU×4 CPU is required. Present EWS already satisfies the condition. On the other hand, highest calculation speed of the computer is required when the minimum feature size is 1.5 μm (w:0.4 μm) and the cycle time is 200 ns. The corresponding writing time is 1100 s. The required calculation speed is 350 MIPS/CPU×4 CPU. This calculation speed can be effectively realized, for example, by using two computer systems of (200 MIPS/CPU×4 CPU/system×two systems).

Although the system shown in Fig. 2 was assumed in this article for simplicity, other types of systems will also realize real-time correction. For example, super multiprocessing computer and dedicated data processing hardware are other candidates for real-time correction.

IX. SUMMARY

The new algorithm for obtaining the optimum correction dose has been proposed and discussed for the application to reticle writing. The representative figure method and dose formula method are used for main algorithm. A characteristic feature of the algorithm is that the dose evaluation point is fixed at the center of small regions, S_e. This leads to the high speed correction calculation without deteriorating the correction accuracy. The STEALTH system has been developed which realizes the correction algorithm. The calculation time by STEALTH was 1/30 of that by the previous system.[18]

Under the assumption that parallel processing is adopted, the calculation speed of computer has been evaluated which can satisfy the condition that both (1) the time for obtaining the representative rectangles, and (2) the calculation time for correction itself are less than the writing time. The evaluation has been carried out by using gate array pattern until the minimum feature size reaches 0.2 μm on reticle (0.05 μm on wafer). The required calculation speed was 500 MIPS/CPU×4 CPUs at most.

This result suggests that (1) the calculation time for correction can be suppressed to less than writing time, and (2) the real-time proximity effect correction is possible for reticle fabrication. Furthermore, when the new algorithm is applied to x-ray mask fabrication and direct writing system, the calculation time can be suppressed to less than the writing time, as shown in Appendices A and B.

Note added in proof: Certain items discussed in this article are also discussed in Ref. 20. In the present article, however, these items are discussed in greater detail or are modified.

ACKNOWLEDGMENTS

It is mentioned in Sec. II that in view of the large amount of EB data, reading the data from strage media will present a crucial problem. The author would like to thank researchers working on EB system development at NTT Corporation, for pointing out this problem at a meeting between NTT and Toshiba. The author is also grateful to Mr. S. Yamasaki for help in developing the P3C software system and to Dr. T. Takigawa and Mr. I. Higashikawa for their encouragement.

APPENDIX A: SCALING AND APPLICATION TO X-RAY MASK FABRICATION

Highly accurate proximity effect correction is needed to make a 1:1 x-ray mask.[21–23] The new algorithm of this article can be applied to x-ray mask fabrication.

When the acceleration voltage is 50 kV, the back scattering range of the electron σ_b for the x-ray fabrication is about 1/4 of that for reticle fabrication. Suppose that all types of length are scaled to (1/4) as shown in Table I, corresponding

TABLE I. Scaling relation between x-ray mask fabrication and reticle fabrication. Acceleration voltage is 50 kV.

Scaling	Reticle fabrication → (1/4)	X-ray mask fabrication
Backscattering range σ_b	10 μm	2.5 μm
Maximum shot size	2 μm	0.5 μm
Small region size $2\Delta r$	10 μm	2.5 μm
Small region size $2\Delta e$	2 μm	0.5 μm
Minimum feature size on reticle (example)	0.75 μm	0.2 μm
(on wafer)	0.2 μm	
Writing region on mask (example)	200 cm^2	$200/4^2=12.5$ cm^2
(on wafer)	$200/4^2=12.5$ cm^2	

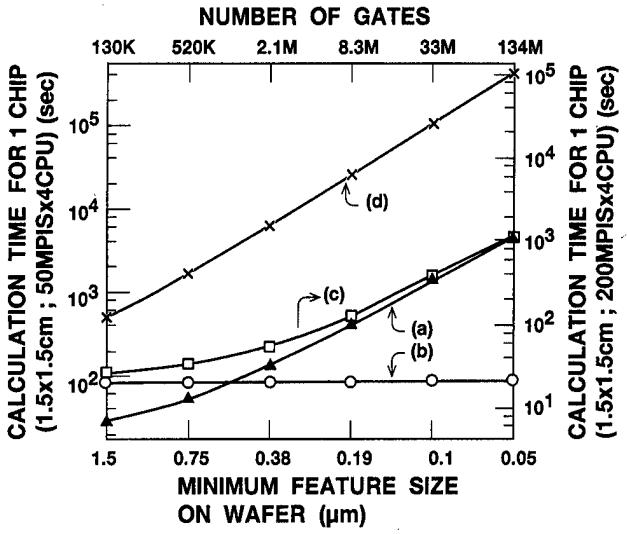


FIG. 14. Application to the direct writing process. Time for calculating optimum dose. (black/white) Inverted pattern of gate array was used for evaluation (wiring layer). Chip size is 1.5×1.5 cm. Small region sizes for reference rectangle and for element were set at 10 and 2 μm, respectively. (a) Time for obtaining reference rectangles; (b) calculation time for correction itself; (c) total time for correction (a)+(b); (d) data conversion time (reference).

to the difference of the backscattering range. The writing region of 12.5 cm^2 (~3.5×3.5 cm) is reasonable for the x-ray mask fabrication because 2–4 chips will be written in a mask. Table I shows that the maximum shot size for x-ray fabrication is (1/4) of that for the reticle fabrication. This condition is also reasonable because the optimum size of the maximum shot decreases with the minimum feature size of pattern.

Under the scaling conditions of Table I, the discussion and the results in this article can be directly applied to the x-ray mask fabrication. When the cycle time is 200 ns and the minimum feature size is 0.75 μm (~0.2 μm×4) on reticle, the writing time is 1600 s and the required calculation speed for STEALTH is 230 MIPS, for the reticle fabrication under the conditions of Table I. Correspondingly, when the cycle time is 200 ns and the minimum feature size is 0.2 μm, the writing time and the required calculation speed for STEALTH are 1600 s and 230 MIPS, respectively, as well as for x-ray mask fabrication.

This discussion suggests that real-time correction can be carried out for x-ray mask fabrication.

APPENDIX B: APPLICATION TO DIRECT WRITING PROCESS

In this Appendix, all types of length are measured on a wafer.

The algorithms mentioned in Sec. IV can be applied to direct writing process. When the data format in Ref. 17 is used, the amount of dose data for one chip is at most 30 Mbytes. Thus, the data amount is not problem and real-time correction is not needed in the case of direct writing process. However, the high speed correction calculation has a great advantage, because the correction calculation itself occupies 50%–99% of the total calculation time when the previous algorithm is used.[20]

The calculation time—using the new algorithm—is evaluated. The gate array described in Sec. III and Eq. (2) is used for the evaluation. The parallel processing with using 4 CPU is assumed. Two calculation speed of 1 CPU are assumed: 50 MIPS and 200 MIPS. The acceleration ratio by the parallel processing is set at 3.7 for the calculation of the correction itself.

The result of the evaluation is shown in Fig. 14. The new algorithm in this article completes the calculation for correction itself within 2 min or so. The value is negligibly small compared with the writing time. The calculation time by STEALTH is about (1/30) of that by the previous system.[18]

Even when the minimum feature size is 0.05 μm, the total calculation time is 1.5 and 0.4 h by using 50 MIPS×4 CPU and 200 MIPS×4 CPU computers, respectively. The time for writing 100 chips was evaluated as about 11 h.[20] The total time for correction is sufficiently small compared with the writing time.

APPENDIX C: EDGE ERROR

When the representative figure method is used, the whole pattern in a small region S_e is exposed by a uniform dose. This causes the error. The error is called *edge error* because the error becomes maximum at the edge of a large pattern.

In Ref. 16, the estimation of the edge error was given; however, the detail of the estimation was abbreviated. The detail is described in the following: Figs. 15(a) and 15(b) show the deposited energy with and without proximity effect correction, respectively. It is assumed that the profile of the deposited energy in the region (p) in Fig. 15(a) is the same as that in the region (q) in Fig. 15(b). Then the edge error $\pm\delta E$ is approximated by the error $\pm\delta E^*$.

In the following, the error $\pm\delta E^*$ at the edge of the lines and spaces (L/S) pattern is discussed. Pattern density is de-

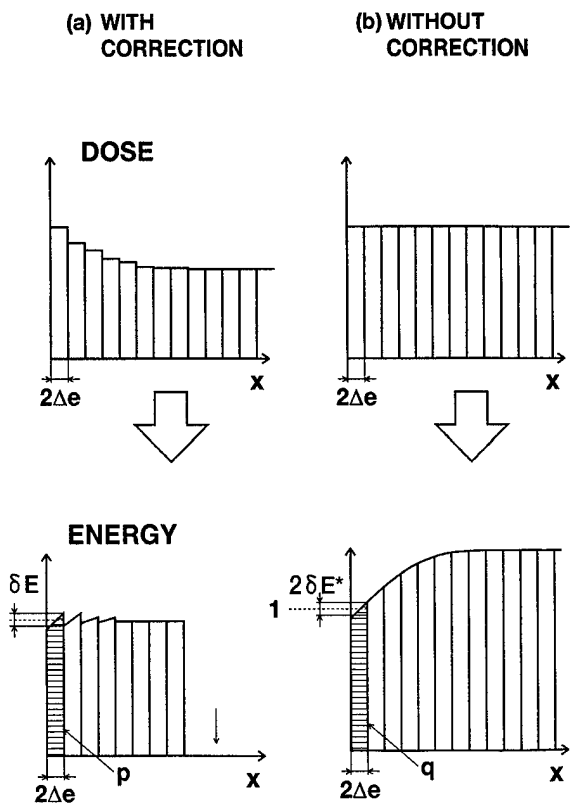


FIG. 15. Edge error. Finiteness of the small region size for element causes the edge error. Uniform dose is assigned for all figures in a small region; this causes the edge error. (a) and (b) show the case with and without proximity effect correction.

noted as p. When p is 1, the L/S pattern becomes a large figure and the error takes the largest value. The size of small region is described as $2\Delta_e$.

The deposited energy is described as

$$E(x)=\int_{\text{pattern}} \frac{1}{1+\eta_E/2}\left\{\delta(x-x')+\frac{\eta_E}{\sqrt{\pi}\sigma_b} \times\exp[-(x-x')^2/\sigma_b^2]\right\}dx'. \quad \text{(C1)}$$

Assuming that the linewidth and space width $\ll\sigma_b$, and the edge of the L/S pattern is at $x=0$, Eq. (C1) is modified as

$$E(x)=\int_0^{\infty} \frac{p}{1+\eta_E/2}\left\{\delta(x-x')+\frac{\eta_E}{\sqrt{\pi}\sigma_b} \times\exp[-(x-x')^2/\sigma_b^2]\right\}dx'. \quad \text{(C2)}$$

Setting $u=x-x'$, Eq. (C2) is modified to

$$E(x)=\frac{-p}{1+\eta_E/2}\int_x^{-\infty}\left[\delta(u)+\frac{\eta_E}{\sqrt{\pi}\sigma_b}\exp(-u^2/\sigma_b^2)\right]du. \quad \text{(C3)}$$

The inclination Q of the deposited energy at the edge ($x=0$) is obtained by the second term of the Eq. (C3) (the backscattering term) as

$$q=dE(x)/dx|_{x=0}$$

$$=\frac{-p\eta_E}{\sqrt{\pi}\sigma_b(1+\eta_E/2)}\left|d\left\{\int_x^{-\infty}[\exp(-u^2/\sigma_b^2)]du\right\}/dx_{x=0}\right.$$

$$=p\eta_E/[\sqrt{\pi}\sigma_b(1+\eta_E/2)]. \quad \text{(C4)}$$

Therefore, the edge error $\pm\delta E$ is described as

$$\delta E\sim\delta E^*=q\Delta_e=p\Delta\eta_E/[\sqrt{\pi}\sigma_b(1+\eta_E/2)]. \quad \text{(C5)}$$

When $p=1$, the equation coincides with the estimation of Ref. 16 and gives the maximum value of the edge error.

The estimation of the edge error is useful not only for using the representative figure method but also for other aspects of proximity effect correction, as follows:

Case 1: When the character projection method[24–26] is used and the dose correction method is applied, the dose in one character becomes uniform. This causes an error,[21,27] and the error can be considered as a sort of the edge error. Thus, the estimation gives the guideline for the maximum character size.

Case 2: When the dose correction method is used and the representative figure method is not used, the correction dose is assigned to a small figure which is also called element. Dose for an element becomes different from that for the adjacent element. This causes the connection error between two shots. This error also can be considered as a sort of edge error, and can be evaluated by Eq. (C5). The equation gives the maximum size of element even when the representative figure method is not used.

Case 3: When Ghost[18] or Phantom[16,17,19] is used, alignment is needed for subsidiary exposure. Equation (C5) gives the rough guideline for the permissible alignment error.

Case 4: When the dose formula method is used, the table of error function must be made. One value of the error function is assigned to a finite range of distance, such as $(x/\sigma_b=0.5)-(x/\sigma_b=0.55)$. The finite grid size causes the calculation error. The error is also roughly estimated by Eq. (C5) by setting $p=1$.

In cases 1–3, Eq. (C5) gives the size (and alignment error) of 2 μm as a guideline to suppress the error less than 3.3%; when the acceleration voltage is 50 kV, the substrate is Si and η_E is 0.8. In case 4, the grid size of 0.1 μm is sufficient to suppress the error less to than 0.17%.

[1]T. R. Neil and C. J. Bull, Electron. Lett. **16**, 621 (1980).
[2]T. Takigawa, K. Kawabuchi, M. Yoshimi, and Y. Kato, Microelectron. Eng. **1**, 121 (1983).
[3]G. A. C. Jones, P. M. Sargent, T. S. Noris, and H. Ahmed, J. Vac. Sci. Technol. B **3**, 124 (1985).
[4]T. Matsuzaka, H. Ohta, N. Saitou, K. Kawasaki, T. Kohno, and M. Hoga, Jpn. J. Appl. Phys. B **12**, 6018 (1993).
[5]T. Abe, Proc. SPIE **2254**, 194 (1994).
[6]J. S. Greeneich and T. Van Duser, IEEE Trans. Electron Devices **21**, 286 (1974).
[7]T. H. P. Chang, J. Vac. Sci. Technol. **6**, 1271 (1975).

[8]G. Owen and P. Rissman, J. Appl. Phys. **54**, 3573 (1983).
[9]M. Parikh, J. Appl. Phys. **19**, 4371 (1979).
[10]M. Parikh, J. Appl. Phys. **19**, 4378 (1979).
[11]M. Parikh, J. Appl. Phys. **19**, 4383 (1979).
[12]J. M. Pavkovich, J. Vac. Sci. Technol. B **4**, 159 (1986).
[13]T. Abe and T. Takigawa, J. Appl. Phys. **65**, 4428 (1989).
[14]H. I. Ralf, G. Duggan, and R. J. Elliott, in *Proceedings of the 10th International Conference on Electron and Ion Beam Science and Technology* (Electrochemical Society, Pennington, NJ, 1092), p. 219.
[15]F. Murai, S. Okazaki, and N. Saito, J. Vac. Sci. Technol. B **5**, 105 (1987).
[16]T. Abe, S. Yamasaki, R. Yoshikawa, and T. Takigawa, Jpn. J. Appl. Phys. II B **3**, L528 (1991).
[17]T. Abe, S. Yamasaki, T. Yamaguchi, R. Yoshikawa, and T. Takigawa, Jpn. J. Appl. Phys. A **30**, 2965 (1991).
[18]T. Abe, S. Yamasaki, R. Yoshikawa, and T. Takigawa, J. Vac. Sci. Technol. B **9**, 3059 (1991).
[19]S. Yamasaki, T. Abe, R. Yoshikawa, M. Itoh, F. Shigemitsu, and T. Takigawa, Jpn. J. Appl. Phys. B **30**, 3103 (1991).
[20]T. Abe (submitted).
[21]T. Fujino, H. Maeda, K. Moriizumi, T. Kato, and N. Tshbouchi, Jpn. J. Appl. Phys. B **33**, 6946 (1994).
[22]Y. Kuriyama, S. Moriya, S. Uchiyama, and N. Shimazu, Jpn. J. Appl. Phys. B **33**, 6983 (1994).
[23]S. Aya, K. Moriizumi, T. Fujino, K. Kamiyama, H. Minami, K. Kise, H. Yabe, K. Marumoto, and Y. Matsui, Jpn. J. Appl. Phys. B **33**, 6976 (1994).
[24]H. C. Pfeiffer, IEEE Trans. Electron Devices **ED-26**, 663 (1979).
[25]Y. Nakayama, S. Okazaki, N. Saitou, and H. Wakabayashi, J. Vac. Sci. Technol. B **8**, 1836 (1990).
[26]T. Fujino, K. Moriizumi, K. Kamiyama, H. Kusunose, S. Takeuchi, H. Morimoto, and Y. Watakabe, *Extended Abstract of 37th Spring Meeting* (The Japan Society of Applied Physics, Tokyo, 1990), p. 447 (in Japanese).
[27]T. Tamura, K. Nakajima, and H. Nozue, Jpn. J. Appl. Phys. B **33**, 6953 (1994).

Surface tension, adhesion and wetting of materials for photolithographic process

J. Bauer and G. Drescher
Institute for Semiconductor Physics, D-15230 Frankfurt (Oder), Germany

M. Illig
OEG G.m.b.H., Ringstr. 1003, D-15236 Frankfurt (Oder), Germany

(Received 18 July 1995; accepted 13 May 1996)

Surface tension of substrates and layers used in semiconductor technology (e.g., Si, SiO_2, Si_3N_4, AlCu, and WTi) as well as surface tension of materials of photolithography as resists and developers were determined by contact angle measurements. Resists and developers could be divided into two groups with polar and nonpolar behavior, respectively. The pretreatment of the substrates (cleaning, dehydration, and exposure to humidity) influences the surface tension. It was found that surface modification by primers, for example, hexamethyldisilazane and trimethylsilydiethylamine provides stable and reproducible surface tension. To reduce defect density of resist structures, a general optimization approach of surface tension was derived and successfully applied, using Si and AlCu as substrates in a typical photolithographic process. The results show a good adhesion of resist will be achieved if the work of adhesion is greater than 5 dyn/cm. A suggestion is given of process window of the contact angle is given for certain combinations of types of resist and developer.
© *1996 American Vacuum Society.*

I. INTRODUCTION

Surface tension of photolithographic materials like resist and developer has a great potential for the photolithographic process itself. For instance, homogeneity of spin coatings, planarization, and adhesion depend directly on surface tension of the materials involved. In this article, the influence of surface tension of photoresists and developers on adhesion and wetting is investigated and a method to optimize the process is suggested. In particular, for reducing feature sizes <1 μm the optimization of adhesion of these small resist structures is a factor of importance to decrease defect density.

Generally the lithographic process requires a modification of substrate surfaces before resist coating to prevent small features from ablation during developing, water rinsing and following wet etch processes. Primers like hexamethyldisilazane (HMDS), trimethylsilyldiethylamine (TMSDEA) and others[1–4] reduce surface tension of materials used in semiconductor technology like Si, SiO_2, Si_3N_4, AlCu and WTi, etc (Fig. 1). The primer reacts chemically with water deposits held by high-energy substrate surfaces or more closely regarded with bounded OH–groups. Thus surface tension is reduced and so is the possibility of penetration of developer, water, or etchants between resist and the modified substrate surface.

II. THEORETICAL BASICS AND EXPERIMENTAL METHODS

A. Surface tension measurement of solids and fluids

Applying the drop method (Figs. 2 and 3), the nonpolar part of surface tension σ^d (caused by dispersion energy between molecules) and the nonpolar part of surface tension σ^p (caused by dipole interaction, induced dipole moments and hydrogen bonds) are measured at once (measurement device: SURFTENS from OEG GmbH). The basic additive approach

$$\sigma_{ij}=\sigma_{ij}^p+\sigma_{ij}^d \tag{1}$$

is based on publications of Fowkes,[5–8] Gardon,[9] and Hansen; Beerbower[10] and Meyer *et al.*,[11,12] and Wu.[13,14] The model of harmonic mean is used to approximate interaction of low energetic phases (symbols: see Fig. 2):[13,14]

$$W_{\mathrm{asl}}=(1+\cos\,\Theta)\sigma_{\mathrm{gl}}=4\left[\frac{\sigma_{\mathrm{gl}}^d\sigma_{\mathrm{gs}}^d}{\sigma_{\mathrm{gl}}^d+\sigma_{\mathrm{gs}}^d}+\frac{\sigma_{\mathrm{gl}}^p\sigma_{\mathrm{gs}}^p}{\sigma_{\mathrm{gl}}^p+\sigma_{\mathrm{gs}}^p}\right]. \tag{2}$$

Work of adhesion W_{asl} is defined as reversible work necessary to create an interface area of 1 cm^2 between two different chemical materials with the phases solid-fluid or to separate them. This work of adhesion is defined by Dupré equation as follows:

$$W_{\mathrm{asl}}=\sigma_{\mathrm{gl}}+\sigma_{\mathrm{gs}}-\sigma_{\mathrm{ls}}. \tag{3}$$

In the following, all surfaces are modified by organic compounds. These compounds strongly reduce the spreading pressure. That's why the latter is further neglected ($\pi_s=0$).[5,14,15] Using two testing liquids of different polarity, it is possible to determine the surface tension of the substrate by an analytical solution of Eq. (2).[14] In our experiments, water (16–17 MΩ cm): $\sigma_{\mathrm{gl}}^p=50.7$ dyn/cm, $\sigma_{\mathrm{gl}}^d=22.1$ dyn/cm and methylene iodide (CH_2I_2): $\sigma_{\mathrm{gl}}^p=6.7$ dyn/cm, $\sigma_{\mathrm{gl}}^d=44.1$ dyn/cm were used. If the surface tension of the used substrates is known (σ_{s1}, σ_{s2}), the surface tension of liquids can be determined by solving Eq. (2) using the Newton–Raphson method. For support, a start value calculation was carried out successfully [see Eq. (4a)]:

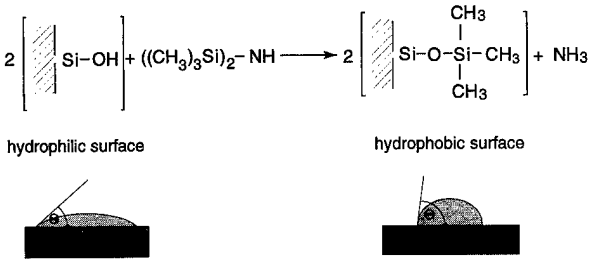


FIG. 1. Modification of high-energy hydrophilic surfaces by applying HMDS to improve resist adhesion.

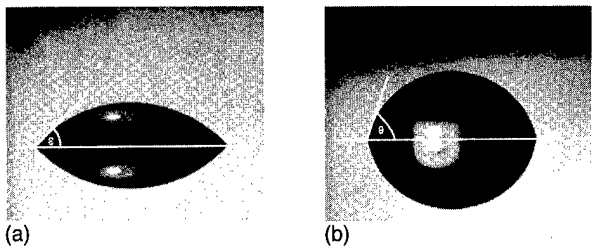


FIG. 3. Examples for contact angle of water θ_w after different surface modifications of Si with HMDS/5% TMDS (hotplate temperature T_p=120 °C) using only different prime times. (a) Prime time=2s, θ_w=40°; (b) prime time=20 s, θ_w=70°.

$$\sigma^d_{gl}(\text{start}) = \sigma^d_{s1}F(\Theta_1) + \sigma^d_{s2}F(\Theta_2),$$

$$\sigma^p_{gl}(\text{start}) = \sigma^p_{s1}F(\Theta_1) + \sigma^p_{s2}F(\Theta_2), \tag{4a}$$

$$F(\Theta_1) = \frac{1-(1+(\cos\Theta_i)/4)}{2(1+(\cos\Theta_i)/4)}. \tag{4b}$$

Measurement conditions were clean room <100, relative humidity 45%±2% at a room temperature of 21 °C. The drop volume was about 5–10 μl. After applying the testing liquid (<3 s) a picture was taken and contact angle could be calculated by the tangent method.

B. Valuation of adhesion

Figure 4 illustrates how liquid penetrates into the interface (resist/substrate) separating the resist. That occurs if the substrate has a greater affinity (small contact angle) for developer or water in the environment than for the resist. Adhesion of resist is determined by work of adhesion W_{a12} under air environment which is given by

$$W_{a12} = \sigma_1 + \sigma_2 - \sigma_{12}, \tag{5}$$

where σ_1 is the surface tension of the resist, σ_2 is the surface tension of the substrate, and σ_{12} is the interfacial tension between resist and substrate; and work of adhesion W_{l12} under environment of the liquid (developer or water)

$$W_{l12} = \sigma_{13} + \sigma_{23} - \sigma_{12}, \tag{6}$$

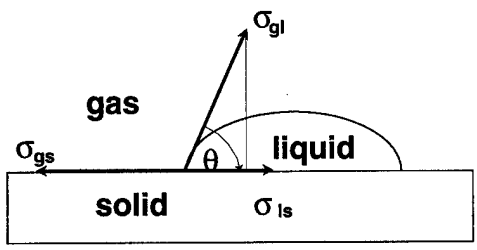


FIG. 2. Effect of surface tension using the method of drop contact angle: $\sigma_{gs}=\sigma_{ls}+\sigma_{gl}\cos\Theta$, $\sigma_{gs}=\sigma_s-\pi_s$, σ_{gs}: surface tension gas/solid, σ_{gl}: surface tension gas/liquid, σ_{ls}: interfacial tension liquid/solid, σ_s: surface tension solid under vacuum, and π_s: spreading pressure, e.g., reduction of surface tension in air or in a vapor, $\pi_s=\pi_e$ saturated condense pressure.

where σ_{13} is the interfacial tension between resist and liquid, and σ_{23} the interfacial tension between substrate and liquid. Calculation of interfacial tension may be carried out according to Eqs. (2) and (3) as follows:

$$\sigma_{ij} = \sigma_i + \sigma_j - 4\left[\frac{\sigma^d_i\sigma^d_j}{\sigma^d_i+\sigma^d_j} + \frac{\sigma^p_i\sigma^p_j}{\sigma^p_i+\sigma^p_j}\right]. \tag{7}$$

Also, the adhesion factor f_a is used for valuation of the quality of adhesion;[16]

$$f_a = W_{l12}/W_{a12}. \tag{8}$$

In the case of $f_a \leqslant 0$, no adhesion appears (spontaneous separation).

C. The value of wetting

Another important value in photolithographic processes is the wetting of photochemical materials. A uniform development velocity will assure adequate wetting of the entire resist surface with developer. In that instance, resist edge roughness and variations of the structures will be avoided and the reaction of the development process will take place, even in very small structures. A small wetting angle is crucial for a sufficient penetration depth in small structures. To remove residues completely, by the penetration of developer between the resist and the substrate, a certain wetting of the substrate surface with developer must be guaranteed. To avoid pinholes in the resist layer also, a good wetting of the substrate with resist has to be maintained during the coating process. The demands concerning the wetting angle are partially in contrast to the demands of the contact angle used for good adhesion. Several papers were published by Wu[13,14] investigating, by mathematical means, the optimization of wetting angle in order to produce polymer layers of high quality. In

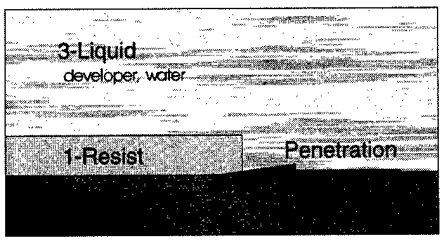


FIG. 4. Principle of resist separation in accordance to Yanazawa (Ref. 4).

TABLE I. Overview of kinds of resists and developers in respect to their surface tension based on experimental results (Tables II and III).

	Polar type (type A)		Nonpolar type (type B)	
	σ^p (dyn/cm)	σ^d (dyn/cm)	σ^p (dyn/cm)	σ^d (dyn/cm)
Resist (R)	20	30	10	30
Developer (D)	20	40	10	40

these papers, it was proven that wetting and adhesion are optimal if the polarity of the substrate and the wetting material are equal (polarity $x_i^p = \sigma_i^p/\sigma_i$). In practice the materials are more or less fixed, but modifications of the surface are a possibility to improve adhesion and wetting characteristics. For evaluation purposes, the wetting factor f_W due to Wu[14] is used and is reciprocal to defect density (unwetted area):

$$f_w = (1 - \lambda_{ls}/\sigma_{gs})^{-2}. \quad (9)$$

In the case of spontaneous wetting (contact angle $\Theta = 0$) $\lambda_{ls} \geq 0$ and $f_w \geq 1$ are valid. At contact angles $\Theta > 0$ defect density raises up because of $\lambda_{ls} < 0$ and $f_w < 1$. The wetting coefficient λ_{ls} is defined as the difference between Work of adhesion at substrate/liquid and work of cohesion of the liquid:[14,17]

$$\lambda_{ls} = W_{asl} - W_{cl} = W_{asl} - 2\sigma_{gl}, \quad (10)$$

where W_{cl} is the work of cohesion of the liquid. A further valuation of the wetting may be carried out regarding the contact angle itself. We assume that contact angle $<62°$ are suited for sufficient wetting.[18]

III. RESULTS

A. Surface tension of photochemical materials

The results of measurement of the surface tension of photochemical materials are shown in Tables I–III. I-line resists from HOECHST, SHIPLEY and ALLRESIST were investigated. The finding of surface tension of the liquids is based on several combinations of substrates with the investigated materials including their modification with adhesion promoter and other polymer layers, for example, polymethylmethacrylate (PMMA). The measurement precision of the surface tension of liquids is about 10%. In the resulting data the materials we used can be devided into two groups (Table I). We think the division can be generalized for all resists and developers. A decrease of surface tension of the developer (type B) could be better managed by adding special wetting agents (Table III, for instance, developer E6). Surface tension of the resist layers depends on softbake parameters and on material composition (resins, photoactive compounds, solutions and wetting agents). Spincoating was used to achieve a resist thickness of 1.5 μm and the softbake parameters were kept stable for all resists (hotplate: 60 s, 90 °C).

Many resists and developers can be devided in this sense. The measurements of Fadda *et al.*[17] show the obvious division of materials into two groups as in Table I also.

TABLE III. Surface tension of liquid photoresists and developers.

Liquids	Surface tension σ_{gs}^p (dyn/cm)	Surface tension σ_{gs}^d (dyn/cm)
Photoresists	1.5–2	33–34
MF501 (type B)	8	39
MFCD (type A)	20	42
AZ351 (1:5) (type A)	23	43
AZ726MIF (type A)	20	36
E6 (type B)	9	40
E6/without wetting agent (type A)	16	42

B. Surface tension of materials in semiconductor technology

Modified and pure substrate materials were used: (1) Si (2) SiO_2(TEOS)/phosphor doped SiO_2 (PSG)/Si_3N_4 layers, and (3) AlCu0.5(0.5%Cu) and WTi10(10%Ti) layers. An overview of the measured contact angles and the calculated surface tensions according to (2) is given in Tables IV–VI and Fig. 5. The measurement precision of the surface tension

TABLE II. Surface tension of photoresist layers (hotplate 90 °C, 60 s).

Resist type	Resist	Contact angle H_2O	Contact angle CH_2I_2	Surface tension σ_{gs}^p (dyn/cm)	Surface tension σ_{gs}^d (dyn/cm)
A	SPR511[a]	60°	38°	21	30
	SPR511[b]	63°	36°	19	30
B	AZ6612[a]	78°	39°	11	32
	AZ6612[b]	78°	43°	11	30
	AZ711B[a]	73°	32°	12	33
	AZ711B[b]	77°	31°	13	33
	AZ7510[a]	81°	43°	11	29
	AZ7510[b]	80°	38°	10	32
	ARP374[a]	79°	43°	11	30
	ARP374[b]	75°	48°	13	29

[a]Exposed.
[b]Nonexposed.

TABLE IV. Examples for priming with HMDS/5% TMDS.

Material	Contact angle H_2O	Contact angle CH_2I_2	Surface tension σ^p_{gs} (dyn/cm)	Surface tension σ^d_{gs} (dyn/cm)
Si	2°	34°	45	29
Si HMDS[a]	65°	52.5°	18	24
SiO_2	2°	33°	45	29
SiO_2 HMDS[a]	65°	59°	22	21
Si_3N_4	3.2°	33°	45	29
Si_3N_4/HMDS[a]	56°	39°	23	29

[a]Priming procedure: hotplate 20 s, 120 °C.

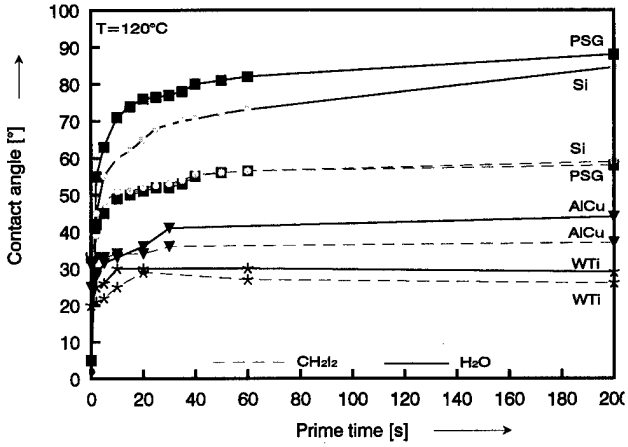


FIG. 5. Contact angle obtained with HMDS/5% TMSDEA for different priming time (hotplate 120 °C).

calculated by the measured contact angle is estimated to be about 5%. The measured values for Al (Fig. 5) show good correlation with those of Kaelble and Dynes.[19] They determined the surface tension of Al with $\sigma^p=38$ dyn/cm/$\sigma^d=28$ dyn/cm. Also, the decay of surface tension with increased time of exposure to humidity is documented in this article. Further values of surface tension of Al were published by Moreau[1] ($\sigma^p=38$ dyn/cm, $\sigma^d=25$ dyn/cm), Zissmann[20] (critical surface tension $\sigma_c=45$ dyn/cm) and Eley[21] ($\sigma=93$ dyn/cm).

Comparative values were found in publications concerning Si by Fowkes[5] ($\sigma=78$ dyn/cm) and Moreau[1] ($\sigma_c=78$ dyn/cm), concerning SiO_2 by Moreau[1] ($\sigma^p=35$ dyn/cm, $\sigma^d=15$ dyn/cm) and concerning Si_3N_4 by Yanazawa[4] ($\sigma=46.5$ dyn/cm).

Modification of surface properties were done with primer (HMDS/5% TMSDEA) in a gaseous atmosphere at a hotplate temperature of $T_p=120$ °C to reduce surface tension of high hydrophile materials. This was done with Si, SiO_2, and Si_3N_4 successfully, but one problem was a sufficient reduction of the surface tension of metals. For this it is necessary to use a different primer or to change the priming procedure. The differences between Si, SiO_2, and Si_3N_4 are small, for this reason, the investigations for the optimization of adhesion were carried out with Si as a representative material.

The pretreatment of the substrates (cleaning, dehydration and exposure to humidity) will greatly influence their surface tension (see Table V and VI). Contamination with water (dependent on relative ambient humidity) and organic materials on high-energy surfaces will occur during the time of exposure to air, which strongly reduces the surface tension of those substrates ($\pi_e \approx 316$ dyn/cm at SiO_2[8,14,15]).

Ambient oxygen and cleaning processes with oxidizing effect lead to thin oxide layers on substrate surfaces like on Si and on Al and thus to high-energy surfaces. We used several cleaning processes: SC1 cleaning (ammoniumhydroxide hydrogenperoxide mixture), SC2 cleaning (hydrocloricacid hydrogenperoxide mixture), HNO_3 cleaning and O_2-plasma treatment for Si (Table V), and for AlCu (Table VI).

AlCu surfaces were investigated by SIMS/AES after cleaning processes were completed. It could be shown that the change of contact angle and surface tension are influenced by the thickness of the oxide layer on the surface (Table VI).

By means of these spectra, a distinct lower thickness of the oxide layer on untreated AlCu (contact angle $\Theta_w=29°$, surface tension $\sigma^p=37$ dyn/cm, $\sigma^d=30$ dyn/cm) in respect to AlCu treated with strongly oxidizing medium HNO_3 (contact angled $\Theta_w=4°$, surface tension $\sigma^p=42$ dyn/cm, $\sigma^d=34$ dyn/cm) could be proved. The thickness of the Al–oxide layer d_{ox} of untreated AlCu was determined on the basis of the depth of escape of the Al-KLL Auger electrons (about 2.5 nm) and the existence of elementary Al-KLL signals to be about 1 nm$<d_{ox}<$2.5 nm. The thickness of the oxide layer of AlCu treated with HNO_3 is about 5–8 times thicker.

TABLE V. Contact angle and surface tension of Si for different treatments.

Pretreatment	Contact angle H_2O	Contact angle CH_2I_2	Surface tension σ^p (dyn/cm)	Surface tension σ^d (dyn/cm)	Surface tension σ (dyn/cm)
No pretreatment delivery quality	31°	41°	38	26,5	64,5
SC 1 cleaning	2°	32°	45	30	75
SC 1,2 cleaning	2°	33°	45	29	74
SC 1,2 cleaning exposure time=5 days	16°	38°	44	27	71
SC 1,2 cleaning O_2-plasma treatment	6°	35°	45	28	73

TABLE VI. Contact angle and surface tension of Al/Cu for different treatments.

Pretreatment	Contact angle		Surface tension		
	H_2O	CH_2I_2	σ^p (dyn/cm)	σ^d (dyn/cm)	σ (dyn/cm)
No pretreatment exposure time <10 min	29°	31°	37	30	67
1 day	32°	40°	31	30	61
5 days	73°	51°	15	25	40
HNO_3 cleaning DI-water	4°	20°	42	34	76
HNO_3 cleaning DI-water temp. 200 °C, 30 min	8°	24°	43	32	75
HNO_3 cleaning DI-water exposure time=5 days	28°	30°	37	30	68
HNO_3-cleaning DI-water O_2-plasma treatment	2.5°	34°	45	29	74

In accordance to the literature cited, it can be generalized that all high-energy surfaces (glass, quartz, metals, etc.) show, in practice, almost identical surface tension after short time of exposure to ambient humidity (i.e., some hours).[14,20] The polar part of those high-energy surfaces is about $\sigma^p \approx 40$ dyn/cm and the nonpolar component is about $\sigma^d \approx 30$ dyn/cm. But after increased time of exposure to ambient humidity, the surface tension is more and more reduced depending on cleaning processes used and ambient conditions.

The aim of special surface modification is to provide stable surface conditions for reproducible and nearly defect free process steps. To determine the surface tension the contact angle of water is well suited as shown in the following section.

C. Adhesion of resist in semiconductor technology

In Figs. 6 and 10, the measured values of the contact angle of water on modified surfaces of AlCu and Si are traced. The work of adhesion W_{a12} obviously reduces by increasing the process time of the HMDS coating. The real effect of improved adhesion—that depends on HMDS process time—can be estimated by means of the behavior of the resist layer during developing and rinsing with water [Figs. 6(b) and 6(c)].

The work of adhesion of resist/developer-combination RB/DA at $\Theta_w < 47°$ is negative and spontaneous separation occurs. In that experiment, a lifting of large resist areas was observed. For RB/DA combinations such surface conditions must be brought about by primer treatment to achieve water contact angles of $\Theta_w > 47°$.

Even slight HMDS treatment is sufficient for uncritical water rinsing after development [Fig. 6(b)]. That is caused by the high difference of polarity between water ($x_p = 2.3$) and resist ($x_p = 0.25,...,0.5$).

For experimental verification of adhesion of small resist structures the defect density of a test structure consisting of 1 μm×1 μm and 0.8 μm×0.8 μm dies was monitored (Figs. 7 and 8). Exposure of these structures was carried out by GCA 8000 *i*-line stepper (NA=0.35, coherent parameter $S=0.5$). Evaluation was done in the middle of the image area at an exposure dose of 1.8 H_0 (H_0 is the amount of exposure necessary to promote the total development of large areas of resist during the developing process). In our case, dip developing was used. Evaluation of the defect density was done by the light/dark field optical technique. Figure 9 shows examples of defect densities using SPR511 and AZ6612 resists. It can be generally concluded that for defect free structures in the submicrometer range work of adhesion of $W_{12} > 5$ dyn/cm is necessary (for developer and water treatment).

In Fig. 6, it shows the critical process of adhesion is most affected in the development process. For the resist/developer combinations RB/DA, RA/DA, and RB/DB, a surface modification for adjustment of a contact angle of water of $\Theta_w \geqslant 65°$ assures good adhesion quality of the resist structures even during development process step.

For combination of RA/DB, there were not observed any defects even at smaller contact angles but, in this case, the contact angle of water should not exceed 85° according to Fig. 6(c).

Generally, in cases combining resist with high polarity ($\sigma^p > 15$ dyn/cm) with developer of low polarity ($\sigma^p < 15$ dyn/cm), only a short primer process should be applied; whereas, the usage of resist with low polarity ($\sigma^p < 15$ dyn/cm) requires a longer primer process to adjust a contact angle of water of more than 65°.

These results are in accordance to the publication of Deckert and Peters[22] where they determined the best adhesion is obtained if substrate and resist have low values of σ^p.

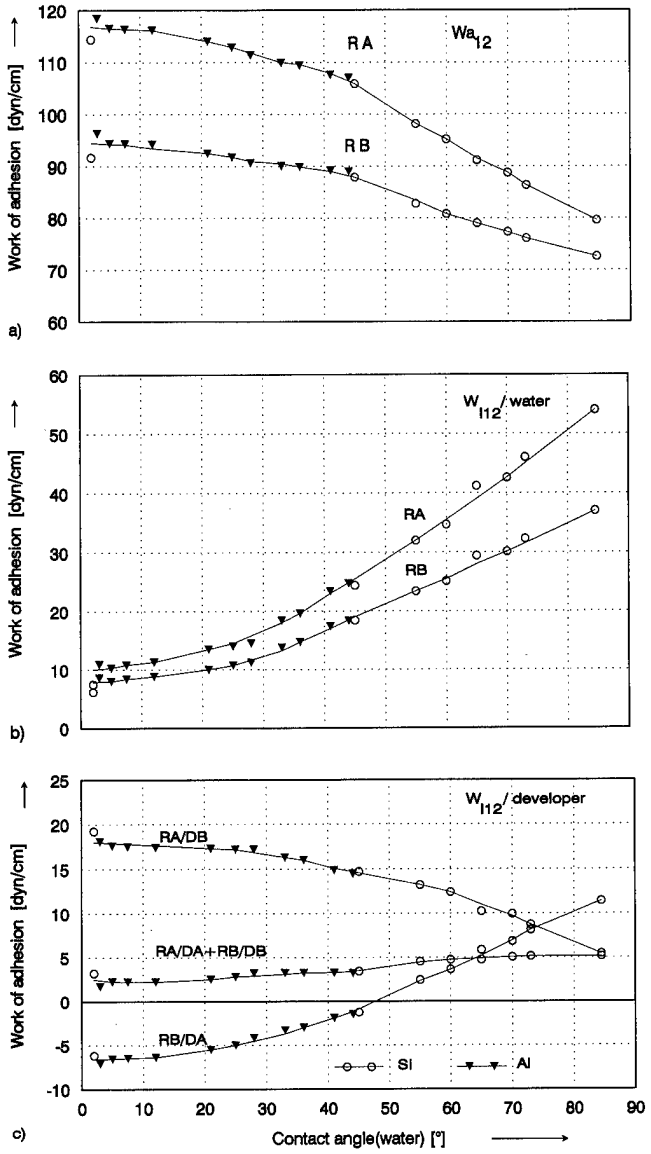


FIG. 6. Work of adhesion for resist as function of contact angle of modified Si and AlCu substrates (a) under air, (b) under water, and (c) under developer environment.

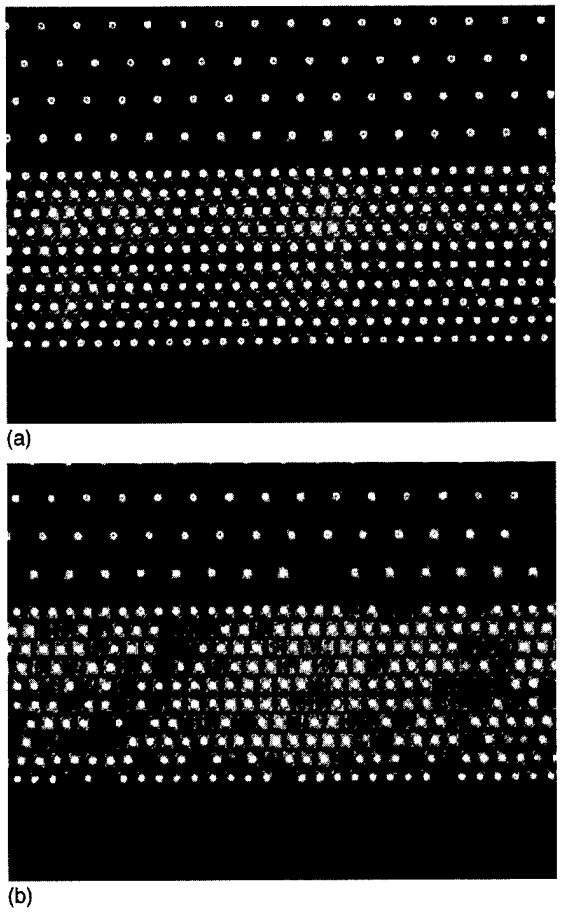


FIG. 8. Examples for measurement of defects of photo resist dies (0.8 μm $\times$0.8 mm) on AlCu with dark-field image: (a) photoresist/developer: SPR511/MFCD, substrate θ_w=25°, defect density=0%, (b) photoresist/developer: SPR511/MFCD, substrate θ_w=5°, defect density=30%.

In contrast to Deckert and Peters[22] recommending to combine substrates with high σ^p only with resist of low values of σ^p, we found, applying developer with low σ^p, substrates and resists with high σ^p could be combined successfully (combination RA/DB, in Figs. 6 and 10).

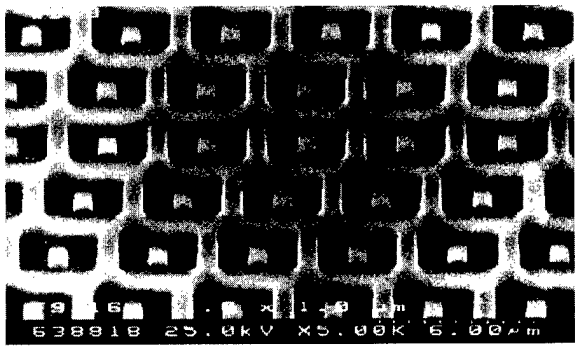


FIG. 7. Test structure for investigation of adhesion (SPR511 on AlCu, θ_w=25°).

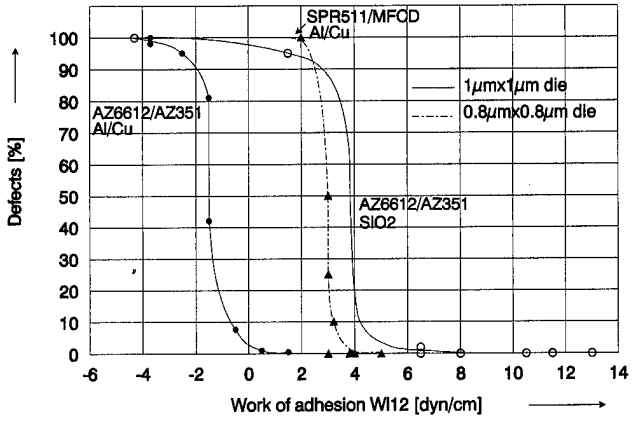


FIG. 9. Density of die defects depending on work of adhesion of resist on AlCu and on SiO_2 in developer solution (surface primer: HMDS/5% TMSDEA).

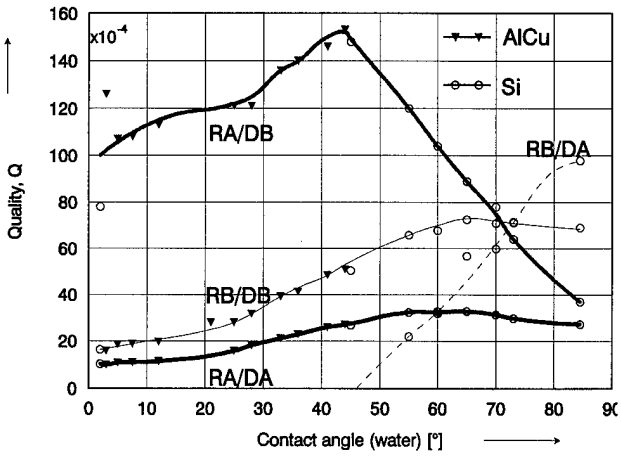


FIG. 10. Quality function [Eq. (11)] vs contact angle of water.

To apply resist, with low polarity at metal surfaces, it is necessary to use other primers (e.g., TMSDEA concentrated) for coating in a gaseous atmosphere or to coat those surfaces by spinning solutions of primer to adjust a contact angle of water of $\Theta_w > 50°$.

D. Wetting characteristics of resist and development

In Table VII, the measured contact angles and wetting factors for developer at Si, AlCu, and resist layers as well as for liquid resist on Si and AlCu are shown. To guarantee good wetting for any liquid a contact angle of that liquid (for example of developer and liquid resist) should be $\Theta < 62°$.[18] That corresponds to a wetting factor of $f_w > 0.30$.

Thus, problems concerning wetting should be only expected to occur using developer of type A (high surface tension $\sigma > 60$ dyn/cm) with substrates of very low surface tension ($\sigma < 35$ dyn/cm), e.g., after a long primer process (Table VII).

TABLE VIII. Limits for contact angle of water on materials of semiconductor technology modified with primer.

Resist/developer	(Example)	Contact angle (water) Θ_{min}	Θ_{max}	$\Theta_{optimal}$
RA/DB	(SPR511/MF501)	10°	75°	45°
RA/DA	(SPR511/MFCD)	60°	85°	70°
RB/DB	(ARP374/E6)	60°	85°	70°
RB/DA	(AZ6612/AZ351)	65°	85°	80°

E. Optimizing adhesion and wetting

For evaluation of all processed components the following quality function was defined:

$$Q = f_{ad} f_{aw} f_{wD1} f_{wD2} f_{wR}, \tag{11}$$

where f_{ad} is the adhesion factor of interface resist/substrate in developer, f_{aw} the adhesion factor of interface resist/substrate in water, f_{wD1} the wetting factor of developer on resist, f_{wD2} the wetting factor of developer on substrate, and f_{wR} the wetting factor of resist on substrate.

The results of this evaluation are traced in Fig. 10 (only positive values are depicted). On the basis of these results, the following limits of the contact angle of water on the substrates can be fixed and used for control purposes (Table VIII).

TABLE VII. Wetting parameters of developers (type A, B) and liquid photoresist on Si, AlCu and resist layers.

Substrate	H_2O Θ_W	Developer type A Θ_D	λ_{32}	f_{WD}	Developer type B Θ_D	λ_{32}	f_{WD}	Photoresist (liquid) Θ_R	λ_{R2}	f_{WR}
Si	2°	0°	2.6	1.07	3.6°	−0.1	1.0	19°	−2	0.95
(HMDS)	45°	26°	−6.1	0.82	23°	−4.1	0.87	24°	−3	0.89
	55°	41°	−15	0.59	37°	−10	0.68	37°	−7	0.76
	60°	46°	−18	0.51	41°	−12	0.62	40°	−8	0.72
	65°	51°	−22	0.43	44°	−14	0.56	42°	−8.5	0.69
	70°	54°	−25	0.38	48°	−16	0.50	43°	−9.5	0.65
	73°	57°	−27	0.33	50°	−18	0.47	44°	−10	0.62
	84.5°	64°	−43	0.24	55°	−21	0.38	46°	−11	0.58
AlCu	3°	0°	8.1	1.25	0°	6	1.18	0°	3.3	1.09
(HMDS)	7.5°	0°	5.5	1.16	0°	3.5	1.10	0°	1.3	1.04
	12°	0°	5.3	1.16	0°	3.4	1.10	0°	1.1	1.03
	25°	0°	1.5	1.04	0°	0.4	1.01	13°	−0.9	0.97
	33°	14°	−1.9	0.94	15°	−1.7	0.95	20°	−2.1	0.94
	41°	22°	−4.2	0.87	19°	−2.6	0.92	20°	−2.1	0.93
	44°	23°	−4.8	0.86	19°	−2.8	0.91	20°	−2.1	0.93
Resist type A	61°	36°	−11.4	0.66	25°	−4.7	0.83	...	...	...
Resist type B	81°	54°	−24.7	0.38	40°	−11.4	0.60	...	...	...

IV. SUMMARY

In accordance with experimental results (Figs. 8 and 9), it is possible to adjust optimal conditions for adhesion of resist on different materials by means of measurement of contact angle of water if surface tension of the applied resist and developer is known. It could be shown that a combination of resist with high σ^p and developer with low σ^p is a good prerequisite for adhesion and wetting. Whereas, using resist with high σ^p combined with developer with high σ^p, good adhesion can only be guaranteed if surface tension of the substrate is strongly reduced. We think that similar results can be achieved in etching processes combining resists with high σ^p with etchants with high σ^p, respectively. According to Fig. 9 for resist layers in liquids (developer, water, etchants), the work of adhesion should be >5 dyn/cm.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dieter Krüger and Ralf Kurps, Institute of Semiconductor Physics, Frankfurt (Oder), for the SIMS/AES investigations to analyze deposits before and after HMDS treatment. We also would like to thank Janett Guhr and John Dirks for their helpful discussions.

[1]W. M. Moreau, *Semiconductor Lithography* (Plenum, New York, 1988).
[2]R. Dammel, *Diazonaphthoquinone-based Resist* (SPIE Optical Engineering, Bellingham, WA, 1993), Vol. TT 11.
[3]Michielsen, V. B. Marriott, J. J. Ponjee, H. van der Wel, F. J. Touwslager, and J. A. H. M. Moonen, Microlectron. **11**, 475 (1990).
[4]H. Yanazawa, Colloids Surf. **9**, 133 (1984).
[5]F. W. Fowkes, in *Chemistry and Physics of Interfaces* (American Chemical Society, Washington, DC, 1971), pp. 153–167.
[6]F. W. Fowkes, Ind. Eng. Chem. **56**, 12–40 (1964).
[7]F. W. Fowkes, in *Wetting* (Society of Chemical Industry, London, 1967).
[8]F. W. Fowkes, J. Phys. Chem. **66**, 382 (1962); **67**, 2538 (1963).
[9]J. L. Gardon, J. Paint Technology **38**, 43 (1966).
[10]C. M. Hansen and A. Beerbower, in *Kirk-Othmer's Encyclopedia of Chemical Technology*, Suppl. 2nd ed. (Wiley, New York, 1971), pp. 889.
[11]E. F. Meyer and R. E. Wagner, J. Phys. Chem. **70**, 3162 (1966).
[12]E. F. Meyer, T. A. Renner, and K. S. Stec, Phys. Chem. **75**, 642 (1971).
[13]S. Wu, *Polymer Interface and Adhesion* (Marcel Dekker, New York, 1982).
[14]S. Wu, in *Recent Advances in Adhesion*, Proc. Am. Chem. Soc. Symp. (American Chemical Society, Washington, DC, 1971), pp. 45–64.
[15]N. Hackmann and A. C. Hall, J. Phys. Chem. **62**, 12 (1958).
[16]This factor is comparable to the factor introduced by Yanazawa (Ref. 4) derived from wetting tension, $f_{\text{wet}}=1-f_a$.
[17]E. Fadda, G. R. Amblard, A. P. Weill, and A. Prola, *Advances in Resist Technology and Processing XI*, Vol. 2195 (SPIE, Bellingham, WA, 1994), pp. 676–583.
[18]J. F. Padday, *Wetting, Spreading and Adhesion* (Academic, New York, 1978).
[19]E. G. Shafrin and W. A. Zisman, Am. Ceram. Soc. **50**, 478 (1967).
[20]W. A. Zisman, in *Adhesion Science and Technology*, edited by L. E. Lee (Plenum, New York, 1975), Vol. 9A, pp. 55–91.
[21]D. D. Eley, *Adhesion* (Oxford University Press, Oxford, 1961).
[22]Ch. A. Deckert and D. A. Peters, Thin Solid Films **68**, 417 (1980).

Polysilicon gate etching in high density plasmas. III. X-ray photoelectron spectroscopy investigation of sidewall passivation of silicon trenches using an oxide hard mask

F. H. Bell and O. Joubert[a)]
France Telécom, CNET, 38243 Meylan Cedex, France and LPCM UMR 110 IMN, 44072 Nantes Cedex 03, France

(Received 15 January 1996; accepted 3 June 1996)

The characteristics of poly-Si trench etching in high density plasma processes was studied by x-ray photoelectron spectroscopy. Poly-Si films on SiO_2-covered Si(100) substrates were masked with a 200-nm-thick oxide hard mask. The 200 mm wafers were then etched downstream using a helicon high density plasma source and a chlorine-based gas chemistry. After etching, samples were transferred under ultrahigh vacuum to a surface analysis chamber equipped with an x-ray photoelectron spectrometer. Regular arrays of trenches were used to determine the photoelectron signals originating from the tops, sidewalls, and bottoms of the features. A thin oxide film was found on the sides of the oxide masked poly-Si trenches. The origin of this film can be related to the sputtering and redeposition of oxide from the quartz tube of the helicon source located in the plasma generation region. A substantial amount of chlorine was present on the poly-Si sidewall of the features, whereas less chlorine was found on the oxide surfaces. The poly-Si sidewalls were covered by a small amount of oxygen. © *1996 American Vacuum Society.*

I. INTRODUCTION

High density plasma sources are increasingly being investigated to achieve high etch rates and selectivities and to improve control of etching profiles. In addition, ion energy control via wafer biasing is important to minimize substrate and gate electrode damage. Low pressure high density plasmas sources, such as the helicon source, can meet these requirements.[1–3]

In a recent publication, we demonstrated how to achieve anisotropic and selective etching of photoresist masked poly-Si over SiO_2 using a Cl_2/He gas chemistry. We developed a two-step etching process resulting in anisotropic etching profiles without any undesirable etching anomalies.[3] The chemical composition of the poly-Si trenches after etching and sample transfer under high vacuum by x-ray photoelectron spectroscopy (XPS) was also determined.[4] Regular arrays of lines and spaces were analyzed to determine the chemical composition on the tops, sidewalls, and bottoms of the features. A thin oxide film was found on the sidewalls of the photoresist and poly-Si, it was caused by the sputtering of the quartz tube located in the plasma generation region of the helicon source. A substantial amount of chlorine and carbon was formed on the sidewalls of the features during etching. Smaller amounts of carbon were found on the bottoms of the trenches. We concluded that the carbon rich film and the thin oxide film on the sidewalls might play an important role in passivating the sidewalls of the features.

In the present study, we investigated the etching characteristics when replacing the photoresist mask by an oxide hard mask. Furthermore, we characterized the chemical topography of the features by XPS and compared the results to those obtained using the photoresist masked wafers.

[a)]Electronic mail: olivier.joubert@cns.cnet.fr

II. EXPERIMENTAL PROCEDURE

The experimental system has been described in detail in recent publications.[3,4] The ultrahigh vacuum (UHV) cluster tool consists of four chambers. The load-lock chamber is located in a clean room Class 10 and provides the loading of up to twenty-five 200 mm wafers. The wafer can then be transferred via a transfer chamber into the etching reactor or into the analysis chamber. Wafer transfer between the chambers of the cluster tool is computer controlled by a robot arm located in the transfer chamber. The analysis chamber is equipped with an x-ray photoelectron spectrometer (Fisons Surface Systems ESCALAB 220*i*), a single electrostatic lens flood gun (VG model LEG 41), and a mass spectrometer (Micromass 386). A nonmonochromatized x-ray source consisting of a water cooled standard twin anode (Al $K\alpha$, 1487 eV/Mg $K\alpha$, 1253.6 eV), and a monochromatized Al $K\alpha$ are provided for XPS analysis; further details about the analysis system can be found elsewhere.[4] The helicon reactor made by Lucas Labs[5,6] consists of a plasma generation region in which a high density plasma helicon source is operated at 13.56 MHz,[7–10] and a plasma diffusion region. Two solenoid magnets are necessary to support the plasma in the plasma generation region. Two other solenoid magnets around the diffusion region maintain a high plasma density on the sample and improve the plasma uniformity. The helicon plasma source is operated at 2500 W forward power and 50 W reflected power at a frequency of 13.56 MHz. The sample can be rf biased using a 600 W power supply at 13.56 MHz. Helium backside cooling at 10 Torr pressure on the backside of the wafer is provided to maintain a constant wafer platen temperature of 25 °C during etching.

The oxide and silicon etch rates were measured using a home-made, real time *in situ* HeNe laser ellipsometer.[11] The ellipsometer is an automated rotating ellipsometer working

in the polarizer-compensator-sample-analyzer (PCSA) configuration. The changes in the ellipsometric parameters Ψ and Δ (corresponding to the change in phase shift and amplitude of a reflected elliptically polarized light) are related to the etching of poly-Si or oxide. By determining the ellipticity of the reflected light, information can be obtained on the optical constants Ψ and Δ of the analyzed materials and its overlayers. After data acquisition, a three- or two-layer simulation program can be used to calculate the poly-Si and oxide etch rates (assuming that the refractive indices of poly-Si and oxide are $3.866+i\times(-0.028)$ and 1.465, respectively).

The etching conditions used in this study were identical to those used for the poly-Si gate patterning described in a recent publication.[3] An etching recipe was used to simultaneously switch off the gas flows, source power, and bias power after etching, resulting in a more realistic "fingerprint" of the plasma treated surface.[4] Following etching, the wafer was transferred into the UHV analysis chamber within 30 s. Although the XPS data acquisition required up to several days, analysis of the same area at the beginning and end of acquisition showed that the surface kept under ultrahigh vacuum remained stable.

The monochromatized Al $K\alpha$ x-ray source was used for all experiments performed in this study. The source is provided with a source defined small spot toroidal monochromator. We used the 1000 μm spatial resolution mode resulting in an x-ray spot diameter of less than 1 mm.

The 200 mm wafers used in this work were poly-Si gate structures patterned with an oxide hard mask. Silicon (100) was coated with a 6-nm-thick oxide film. Poly-Si with a thickness of 300 nm was then deposited on the oxide. The wafers were patterned with a 200-nm-thick oxide mask. At the center of the wafer a 18 mm^2 unmasked area allows the *in situ* measurement of Ψ and Δ with the HeNe laser ellipsometer. The smallest feature dimensions were 0.25 μm isolated and nested lines.

The mask design used in this study allows characterization of the chemical topography of the tops, sidewalls, and bottoms of the features by XPS. Photoelectron signals from different portions of the features are separated using geometrical shadowing of photoelectrons and differential charging of the insulating and conducting surfaces. The patterned areas analyzed in this study consist of regular arrays of trenches, a blanket substrate, and an unpatterned mask material, with the size of the areas being at least 1.5 mm^2. This geometry allows the x-ray beam (1 mm in diameter) to be focused on one individual array, the resulting signal being the average of many identical lines and spaces. Variations in the line and space widths enable analysis of the desired portion of the features. The electron energy analyzer detects photoelectrons from a $\pm 6°$ cone, as was experimentally determined.[4] Before analyzing the features, one wafer was etched and the dimensions of the patterned features were determined by secondary electron microscopy. This metrology step provides an accurate measurement of the height of the mask and the width of the polysilicon lines after the etching process. Therefore the array that allows analysis of the desired portion of the feature can be selected.

The quantitative surface analysis by XPS is dependent on several aspects such as shadowing of the photoelectrons, the relative response of the analyzer across the acceptance cone, and the angular response of the electron analyzer.[4,12] However, in our configuration, these effects can be considered to be negligible; further details on mask design and analysis procedures can be found elsewhere.[4]

After etching, the wafer is transferred into the analysis chamber. The wafer is placed in the focal plane of the electron energy analyzer and x-ray sources using a stage elevator. The ability to rotate the wafer around the vertical axis and to move it in all horizontal directions allows the same array of lines and spaces to be analyzed. Lines and spaces can be arranged parallel or perpendicular to the axis of the electron analyzer. Consequently, the tops of the features and bottoms of the spaces (in the parallel mode) or tops and desired parts of the sidewalls—depending on the space between the lines—of the features (in the perpendicular mode) can be analyzed. The angle between the monochromatized x-ray source and the electron analyzer is 57°, and the angle between the x-ray source and vertical axis of the sample holder is 12°. A charge coupled device (CCD) video camera installed on the "hat" of the chamber is aligned with the focal point of the x-ray and electron energy analyzer and allows selection of the desired area of the sample to be analyzed.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Selective and anisotropic polysilicon trench etching using an oxide hard mask

We now study the poly-Si etching profiles obtained using an oxide hard mask. The etching conditions were those established in a recent publication:[3] a two-step etching process is performed to etch the poly-Si gate anisotropically and highly selectively between the poly-Si and the gate oxide. First, a high rf bias power of 240 W (etch rate: 250 nm/min) is applied to obtain an anisotropic etching profile; second, 40 nm before reaching the gate oxide film, the rf bias power is switched to 20 W (etch rate: 60 nm/min). The low bias power applied during the second etching step increases the poly-Si to oxide selectivity and suppresses the trenching effect. Identical process conditions to the second etching step are applied during the 50% overetch time; further details can be found.[3] Figure 1 shows scanning electron microscopy (SEM) micrographs of oxide masked, poly-Si nested, and isolated gates etched using the two-step etching process. The micrographs show that the features are etched anisotropically without any trenching, bowing, or other undesirable etching anomalies. As with the photoresist mask, trenching can be observed when etching the sample using high bias power conditions. The trenching phenomenon is suppressed during the second low bias power etching step (not shown here). The etch rates for poly-Si measured by *in situ* real time ellipsometry are found to be similar for both types of masks (240 nm/min during the first-step process and 60 nm/min during the second-step process). Indeed, the difference in

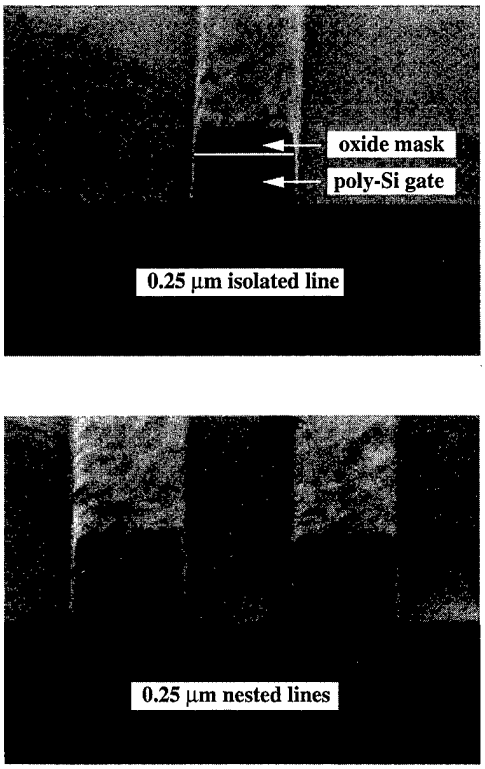


FIG. 1. Secondary electron micrograph of a poly-Si/SiO_2 0.25 μm isolated and nested line and space structure masked with an oxide hard mask.

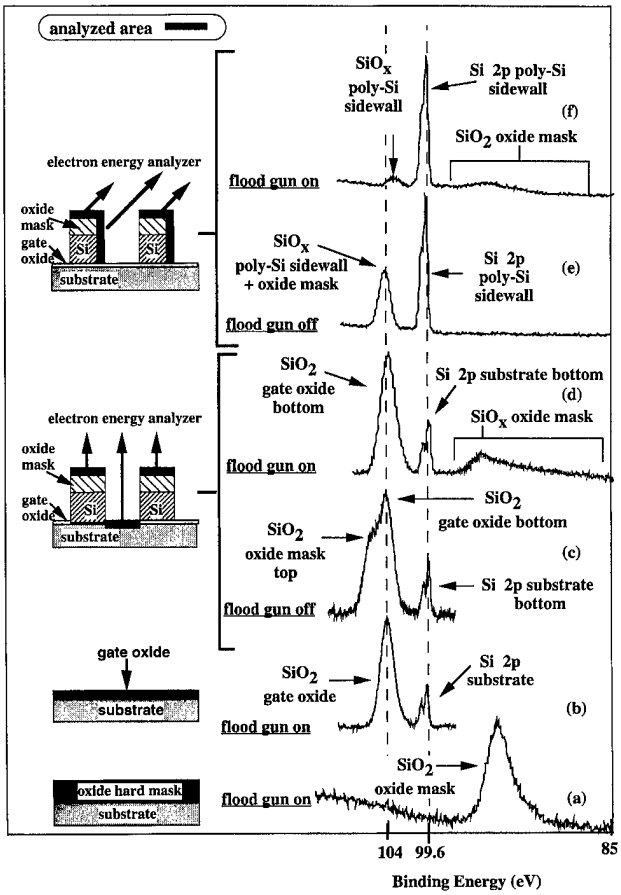


FIG. 2. Si 2p XPS spectra recorded on blanket oxide mask area (a), and blanket substrate (b). Spectra recorded with the 0.35 μm lines and spaces aligned parallel to the electron energy analyzer, and the charge neutralizer turned off (c) and turned on (d). The spectra in (e) and (f) are recorded with the 0.35 μm lines and spaces aligned perpendicular to the electron energy analyzer. Again, spectra were taken with the charge neutralizer turned off (e) and turned on (e).

terms of etching behavior between the photoresist and oxide masked samples is the gate oxide etch rate measured by *in situ* real time ellipsometry. The selectivity in an open area increases by a factor of >3 when replacing the photoresist mask by an oxide hard mask. These studies will be presented in detail in another article.

B. Chemical topography of the etched features determined using x-ray photoelectron spectroscopy

In a previous article, we investigated the chemical topography of the etched features when using a photoresist mask.[4] In this article, similar experiments were conducted using an oxide hard mask. The XPS spectra shown were recorded on wafers etched using the process conditions described in Sec. III. The mask material was a 200-nm-thick oxide hard mask. The thickness of the poly-Si was 300 nm and the gate oxide thickness 6 nm. The elements present on the etched surfaces were silicon, chlorine, and oxygen. No carbon was found on any of the analyzed surfaces. A charge neutralizer, providing low energy electrons, was used to charge surfaces in order to differentiate the XPS peaks resulting from insulating oxide mask surfaces and conducting silicon surfaces. In Sec. III B 1, we present a complete set of the Si 2p, Cl 2s, and O 1s spectra. In Sec. III B 2, we give a quantification of all elements present on the surfaces, followed by a comparison of the surface chemistry on photoresist and oxide hard mask features.

1. Determination of the chemical constituents on oxide masked poly-Si structures using XPS

Figure 2 shows a complete set of Si 2p spectra recorded on a blanket oxide mask [Fig. 2(a)], an unpatterned gate oxide [Fig. 2(b)], 0.35 μm line and space (L/S) features with the poly-Si lines aligned parallel to the electron energy analyzer [Figs. 2(c) and 2(d)], and on 0.35 μm L/S features with the lines perpendicular to the electron energy analyzer [Figs. 2(e) and 2(f)]. When the polysilicon lines are aligned parallel with respect to the electron energy analyzer, only photoelectrons from the tops and bottoms of the features contribute to the XPS signal. When the polysilicon lines are perpendicular to the electron energy analyzer, the analyzed areas are the tops and sidewalls of the features.

The broad peak observed in Fig. 2(a) can be attributed to the SiO_2 originating from the oxide hard mask. This peak is shifted from 105 to 92 eV due to the charging of the surface when the charge neutralizer is turned on. Not shown here is the spectrum with the charge neutralizer turned off. For the latter case, the oxide peak is located at 105 eV and no Si 2p peak at a binding energy of 99.6 eV is detected on the oxide.

Figure 2(b) shows the Si $2p$ peak recorded on an unpatterned gate oxide area. The peak at 104 eV can be attributed to the SiO_2 from the gate oxide, and the Si $2p$ peak at 99.6 eV originates from the underlying silicon substrate. Figure 2(c) shows the Si peak recorded in 0.35 μm L/S parallel to the electron energy analyzer with the charge neutralizer off. Since no silicon is present on top of the oxide mask [see Fig. 2(a)], the peak at 99.6 eV can unambigously be attributed to the Si $2p$ peak originating from the underlying silicon substrate. The spectrum shows two other peaks located at 104 and 105.2 eV. The SiO_2 peak at 104 eV results from the gate oxide film and the peak at 105.2 eV from the oxide hard mask (slightly shifted to a higher binding energy due to the charging of the insulating oxide mask). The spectrum is also recorded with the charge neutralizer on [Fig. 2(d)]. The SiO_2 resulting from the oxide mask is shifted to smaller binding energies (between 86 and 96 eV) whereas the Si $2p$ and SiO_2 from the substrate and the gate oxide remain unshifted. The integrated peak area of the shifted SiO_2 peak from the oxide mask is identical to the integrated area of the SiO_2 peak recorded with the flood gun off. Figure 2(e) shows the Si $2p$ spectra recorded in 0.35 μm L/S features perpendicular to the electron energy analyzer with the charge neutralizer turned off. Due to shadowing, the analyzed areas are the tops of the oxide mask and sidewalls of the features. The peak at 99.6 eV can be attributed to the Si $2p$ peak originating from the poly-Si sidewall. The spectrum also shows a SiO_2 peak (between 101 and 106 eV) coming from the sidewalls of the features and tops of the oxide mask. The spectrum is also recorded with the charge neutralizer on [Fig. 2(f)]. The SiO_2 peak is now separated into two components, whereas the Si $2p$ peak at 99.6 eV remains unshifted. One SiO_2 component is located at 103 eV, whereas the second SiO_2 component is shifted to smaller binding energies (between 87 and 97 eV). The sum of the integrated peak areas of the shifted and unshifted SiO_2 peaks is almost identical to the integrated area of the SiO_2 peak (within 5%) recorded with the flood gun off. The shifted SiO_2 peak can be attributed to the SiO_x film present on the sidewalls and tops of the insulating oxide mask, whereas the unshifted SiO_x peak results from the conducting poly-Si sidewall. Due to limitations in mask dimensions, no spectra could be recorded that only involved analysis of the tops and sidewalls of the oxide mask.

The Cl $2s$ spectra recorded on the same sample are displayed in Fig. 3. Only spectra obtained with the charge neutralizer turned on are shown. Figure 3(a) shows the Cl $2s$ peak recorded on the blanket oxide mask. The peak is shifted from 270 eV to a lower binding energy at 261 eV due to the charging of the insulating oxide mask. Figure 3(b) shows the Cl $2s$ spectrum recorded on an unpatterned gate oxide area. The Cl $2s$ peak is located at 270 eV indicating that the thin gate oxide on top of the silicon substrate behaves like a conducting surface. Cl $2s$ spectra recorded in 0.35 μm L/S features with the polysilicon lines parallel to the electron energy analyzer are shown in Fig. 3(c). The Cl $2s$ peak splits into two peaks at binding energies of 270 and 261 eV. The unshifted peak at 270 eV can be attributed to chlorine present

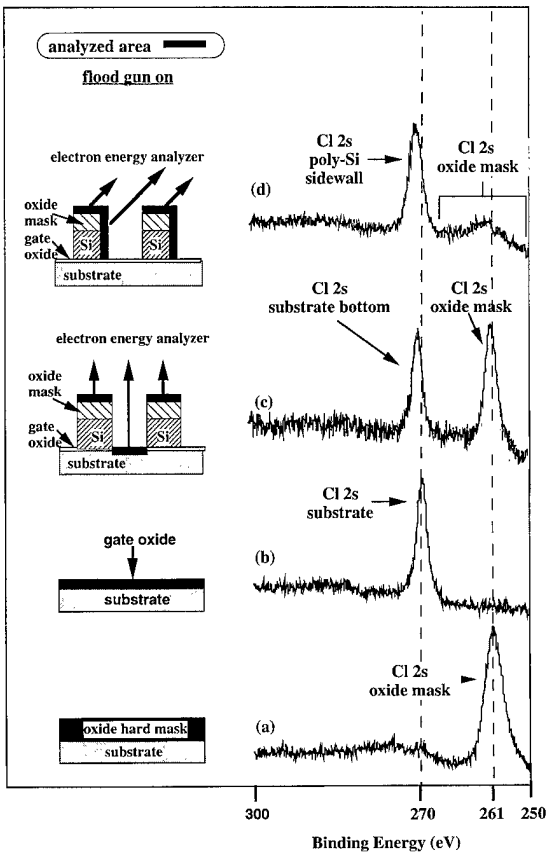


FIG. 3. Cl $2s$ XPS spestra recorded on blanket oxide mask area (a), blanket substrate (b), 0.35 μm lines and spaces aligned parallel (c) and perpendicular (d) to the electron energy analyzer. The charge neutralizer was turned on for the acquisition of all spectra.

on the bottom of the conducting gate oxide, whereas the second peak can be attributed to chlorine present on top of the surface of the insulating oxide mask. Figure 3(d) shows the Cl $2s$ peak recorded with the lines and spaces aligned perpendicularly to the electron energy analyzer (the analyzed areas are the tops and sidewalls of the feature). The unshifted Cl $2s$ peak at 270 eV can be attributed to chlorine present on the conducting polysilicon sidewall. The Cl $2s$ peak at 260 eV can be attributed to chlorine present on the tops and sidewalls of the oxide hard mask. Note that no carbon (at the binding energy of 284 eV) could be observed on any of the analyzed areas.

O $1s$ spectra are shown in Fig. 4. Similar to Si $2p$ and Cl $2s$, the O $1s$ peak can be differentiated when the charge neutralizer is turned on. Figure 4(a) shows the O $1s$ spectrum recorded on a blanket oxide mask. The O $1s$ peak shifts from 534 to 522 eV when the charge neutralizer is turned on. The O $1s$ spectrum recorded on the unpatterned gate oxide area is displayed in Fig. 4(b). The peak remains unshifted at a binding energy of 533 eV. When the features are analyzed with the lines and spaces parallel to the electron energy analyzer, the O $1s$ peak splits into two components. The unshifted peak at 533 eV can be attributed to oxygen resulting from the conducting bottom of the features (gate oxide) and the peak at 522 eV results from the insulating top of the

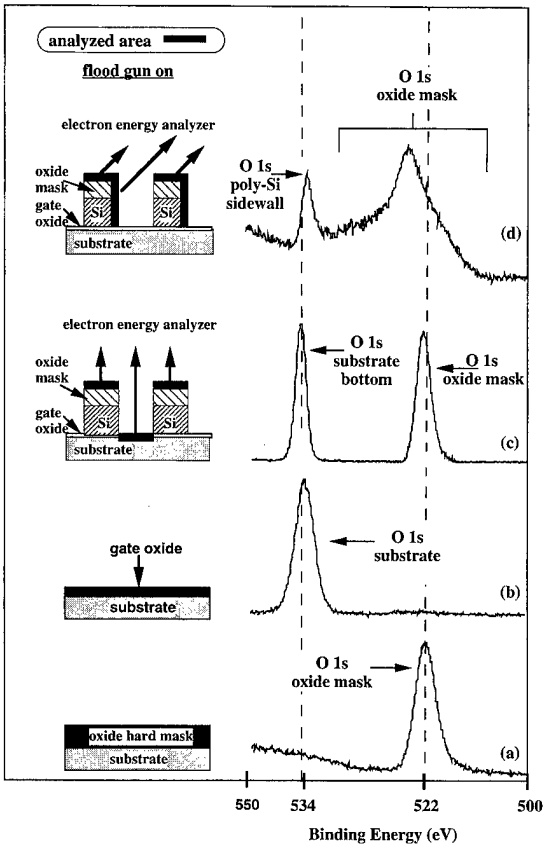


FIG. 4. O 1*s* XPS spectra recorded on blanket oxide mask area (a), blanket substrate (b), 0.35 μm lines and spaces aligned parallel (c) and perpendicular (d) to the electron energy analyzer. The charge neutralizer was turned on for all spectra.

oxide mask. The O 1*s* spectra displayed in Fig. 4(d) originates from the tops and sidewalls of the oxide mask and from the sidewalls of the polysilicon. The O 1*s* peak is split into two peaks. The unshifted peak at 533 eV originates from the conducting poly-Si sidewalls, whereas the broad peak centered at 524 eV can be attributed to oxygen from the tops and sides of the oxide mask.

2. Quantitative analysis of the chemical concentrations on the oxide masked features

We now present the results of the quantitative analysis of the chemical element concentrations on the features. The quantitative coverages of silicon, chlorine, and oxygen are derived from spectra recorded with the charge neutralizer turned on. The integrated peak intensities were divided by the Scofield cross sections (Si 2*p*: 0.82; Cl 2*s*: 1.69; O 1*s*: 2.93).[13] The angle between all the analyzed areas and the electron energy analyzer being 45°, direct comparisons can be drawn between the coverages of the different elements present on the tops, sidewalls, and bottoms of the features. The calculated concentrations of the elements present on the surfaces of the features are summarized in Table I.

a. Blanket substrate (gate oxide), unpatterned oxide mask, top of the oxide mask, bottom of the substrate, and oxide mask (sidewall and top). The calculated element concentrations on a blanket substrate (corresponding to the gate oxide); the unpatterned oxide mask, the top of the oxide mask, and the bottom of the features are rather similar. Slightly less SiO_2 is calculated on the blanket substrate and bottoms of the substrate. This observation is due to the underlying silicon substrate, which contributes to the XPS signal, therefore reducing the total amount of SiO_2. A smaller amount of chlorine is adsorbed on the substrate bottom. Since the chlorine coverage increases with ion energy bombardment,[4] the lower chlorine concentration might be due to a loss of chlorine ion energy when reaching the bottom of the feature. The element concentrations calculated from the tops and sides of the oxide mask show a lower total amount of chlorine and a higher amount of SiO_2 and oxygen. The chlorine concentration found is 30% less than on the blanket oxide mask. Assuming that the top of the oxide mask contributes to the chlorine signal by 16.1% and by taking into account the geometry of the features, the true chlorine coverage on the sides of the features can be estimated to be less than 9%. This result can be explained by the lack of ion bombardment on the sidewalls of the features, which reduces the adsorption of chlorine on the vertical surfaces.

TABLE I. Element concentrations calculated using quantitative XPS analysis after the standard etching process. The mask material was a 200-nm-thick oxide hard mask; the gate oxide thickness was 6 nm.

	After standard process			
Element concentration (%)	Si 2*p*	SiO_2	Cl 2*s*	O 1*s*
Blanket substrate	3.8	31.9	15.3	49.0
Blanket oxide mask		34.0	16.1	50.0
0.35 μm L/S poly-Si sidewall	51.3	8.1	32.5	7.9
0.35 μm L/S oxide mask top and sidewall	···	35.9	12.1	52.0
0.35 μm L/S oxide mask top	···	34.6	16.5	48.9
0.35 μm L/S substrate bottom	3.9	32.4	13.9	49.8

b. poly-Si sidewall. In addition to the expected high amount of silicon (0.513%), significant contributions of chlorine (0.325%), SiO_2 (0.081%), and oxygen (0.079%) are found on the sides of the poly-Si. The amount of chlorine is roughly double that of the concentration found on the oxide surfaces, indicating that the adsorption of chlorine on silicon surfaces is higher than on oxide surfaces.

3. Comparison of the chemical topography of oxide and photoresist masked poly-Si features

In our recent article,[4] we studied the chemical topography of photoresist masked poly-Si. We found that a thin oxide film is formed during etching on the sidewalls of the fea-

tures. We concluded that the source of the oxide film was not the redeposition of volatile etching products but the sputtering of the quartz tube located in the plasma generation region. We also observed a carbon rich film on the sidewalls of the features, possibly acting as an additional passivation layer on the sidewalls of the trenches.

We discuss here the chemical composition of the surfaces for both types of mask material, i.e., the sidewalls of the poly-Si, the bottom of the features, and unpatterned areas. We also observed a thin oxide film present on the sidewalls of the poly-Si trenches when using the oxide hard mask. The likely sources of this oxide are either the quartz tube located in the plasma generation region of the helicon source or the sputtering and redeposition of oxide species from the oxide hard mask. To determine the source of the oxide film, we calculated the ratio of the SiO_2 to Si $2p$ peak areas from the sidewalls of the poly-Si trenches. The ratio is 0.16 for the oxide mask and 0.17 when using a photoresist mask,[4] indicating that the thickness of the deposited oxide film is approximately the same for both types of masks. Since the process conditions were identical for both types of masks, we conclude that the same mechanism as that found for the photoresist masked sample is responsible for the oxide formation: the oxide is sputtered from the quartz tube during etching and redeposited on the sidewalls of the features. This result is rather surprising, since one could expect the sputtering of oxide species from the oxide hard mask to also contribute to the formation of the oxide passivation layer. Since this has not been experimentally observed, we suggest that the mask etching products are volatile and pumped away by the pumping system. On the other hand, the quartz tube in the plasma generation region of the helicon source is exposed to a higher plasma density than the wafer in the downstream region. In addition, the temperature of the quartz tube is much higher than that of the backside cooled wafer and the inner surface of the quartz tube is calculated to be about 1340 cm^2, 30 times higher than the oxide mask coverage on the wafer. Furthermore, an increase in roughness of the inner surface of the quartz tube was observed as a function of source time operation. As a consequence, the oxide surface of the quartz tube exposed to the plasma increases with time. This results in the deposition of an oxide film on a blanket silicon wafer (without biasing) at a rate of 15 nm/min, as measured by HeNe laser ellipsometry.

We found, on the other hand, a substantial amount of carbon on the sides of the poly-Si when using a photoresist mask, indicating that carbon from the photoresist is sputtered and redeposited on the sidewalls. The problem is to understand why carbon from the photoresist contributes to the passivation film on the sidewalls of the features. One plausible explanation is that the etch rate of the photoresist mask is approximately four to five times higher than the etch rate of the oxide mask. Consequently, more etching species from the mask material are present in the gas phase of the plasma. Assuming that the carbon-containing etching products are less volatile than the volatile oxide species from the oxide mask, the photoresist mask can contribute to the formation of a sidewall film. The results obtained using an oxide hard mask show perfect anisotropic profiles and no carbon present on the surfaces. We can therefore conclude that, by using a nonpolymerizing chlorine chemistry, a carbon rich layer may enhance anisotropic etching but is not necessary to obtain vertical sidewalls.

When using a photoresist mask, carbon is present on all surfaces and therefore contributes to the quantification of the element concentrations. In order to compare the chlorine coverage on the surfaces with both masks, the carbon contribution was excluded and the C $1s$, Cl $2s$, and Si $2p$ element concentrations were calculated again. Following this procedure, the chlorine coverage found on the sidewalls of the photoresist masked poly-Si increases from 21.4% (Ref. 4) to 30%, which is of the same order as the 32.5% calculated with the oxide mask. The amount of oxygen increases from 4.5% to 6.5% for the photoresist mask, which is slightly less than the calculated 7.9% for the oxide mask. In blanket areas, similar calculations show that the chlorine coverage is 16.7% for the photoresist mask and 15.3% for the oxide mask. Since the thickness of the gate oxide film was 4.5 nm for the photoresist masked sample and 6 nm for the oxide masked sample, no direct comparison of the oxygen contribution can be drawn.

In summary, comparison of the chemical topography of the photoresist and oxide masked samples shows that the surface modifications after etching are similar for both types of mask material. Approximately the same amount of silicon oxide, chlorine, and oxygen is present on the sides and bottoms of the features. The difference found is the carbon contamination on all surfaces when using the photoresist mask; the effect of the carbon contamination on the poly-Si to oxide selectivity will be discussed in a subsequent article.

IV. CONCLUSIONS

We have studied high density plasma etching processes of 0.25 μm polysilicon trenches. Polysilicon films on SiO_2-covered 200 mm silicon wafers were patterned using an oxide hard mask, and then etched downstream using a high density plasma helicon source. A two-step etching process developed for a photoresist masked sample is equally efficient at etching the features highly anisotropically with a high selectivity between the poly-Si and the gate oxide.

After etching, the 200 mm wafers were transferred (under high vacuum) to an ultrahigh vacuum analysis chamber equipped with an x-ray photoelectron spectrometer to determine the chemical composition of the surfaces. We found a thin oxide film on the sidewalls of the poly-Si which may passivate the sidewalls and therefore prevent lateral etching of the poly-Si gate. The thickness of this film found on the sidewalls was of the same order as the thickness obtained using a photoresist mask. The source of this film was attributed to the slight etching of the quartz tube in the plasma generation of the helicon source. Thus, the oxide mask material played no role in passivating the poly-Si trenches. The controlled erosion of the quartz tube might, however, be of significant importance for the long-term process stability. A

substantial amount of chlorine was found on the sidewalls of the poly-Si, whereas about half this amount was present on the oxide surfaces of the oxide mask, gate oxide bottom, and blanket gate oxide areas. A smaller amount of oxygen covered the sidewall of the poly-Si. The chlorine and oxygen coverages were comparable to those observed for the photoresist masked samples.

ACKNOWLEDGMENTS

One of the authors (F.H.B.) would like to thank Professor G. Turban, IMN at Nantes, for his support. Both authors would like to thank D. Henry for supporting the project. This work was performed in the context of GRESSI, the joint venture between France Télécom CNET and CEA LETI.

[1]I. Tepermeister, N. Blayo, F. P. Klemens, D. E. Ibbotson, R. A. Gottscho, J. T. C. Lee, and H. H. Sawin, J. Vac. Sci. Technol. B **12**, 2310 (1994).
[2]I. Tepermeister, D. E. Ibbotson, J. T. C. Lee, and H. H. Sawin, J. Vac. Sci. Technol. B **12**, 2322 (1994).
[3]F. H. Bell, O. Joubert, and L. Vallier, J. Vac. Sci. Technol. B **14**, 96 (1996).
[4]F. H. Bell, O. Joubert, and L. Vallier, J. Vac. Sci. Technol. B **14**, 1796 (1996).
[5]Lucas Labs Inc., 470C Lakeside Drive, Sunnyvale, CA.
[6]N. Benjamin, B. Chapman, and R. Boswell, SPIE, Santa Clara, CA, 1990 (unpublished).
[7]R. W. Boswell, Phys. Lett. A **33**, 457 (1970).
[8]R. W. Boswell, R. K. Porteous, A. Prytz, A. Bouchoule, and P. Ranson, Phys. Lett. A **91**, 163 (1982).
[9]R. W. Boswell and R. K. Porteous, Appl. Phys. Lett. **50**, 1130 (1987).
[10]A. J. Perry, D. Vender, and R. W. Boswell, J. Vac. Sci. Technol. B **9**, 310 (1991).
[11]A. Vareille, Y. Vuillod, and L. Thévenot, French Patent No. 88, 02784 4 March 1988.
[12]K. V. Guinn and V. M. Donnelly, J. Appl. Phys. **75**, 2227 (1994).
[13]J. H. Scofield, J. Electron. Spectrosc. **8**, 129 (1976).

Influence of patterning in silicon quantum well structures on photoluminescence

Hideo Namatsu,[a] Tomofumi Furuta, Masao Nagase, Kenji Kurihara, Kazumi Iwadate, Katsumi Murase, and Takahiro Makino
NTT LSI Laboratories, 3-1 Morinosoto Wakamiya, Atsugi-shi, Kanagawa 243-01, Japan

(Received 15 November 1995; accepted 28 May 1996)

This article reports on a study of the photoluminescence (PL) characteristics of $SiO_2/Si/SiO_2$ quantum well structures formed by the thermal oxidation of Si lines with rectangular cross sections and (111) side planes. The PL spectra have a peak at a wavelength of 760 nm regardless of Si width, indicating that the emission is due to the recombination of carriers at the Si/SiO_2 interface. The Si width required to obtain the maximum emission is about 5 nm, which is thicker than theoretically predicted. The reason for this is clarified by measuring the difference in the PL characteristics for patterned and unpatterned $SiO_2/Si/SiO_2$ film structures. In the film structures, the characteristics are found to vary by the patterning of the ultrathin Si layer, followed by oxidation; the Si thickness yielding the maximum intensity increases and a wide range of Si thickness contributes to the emission. This behavior is consistent with that of the oxidized Si lines and is possibly due to pattern deformation caused by oxidation. The stress accompanying the deformation is probably the reason for the increase in the PL intensity in thicker Si. Therefore, the PL characteristics for Si are strongly influenced by the patterning and subsequent oxidation. © *1996 American Vacuum Society.*

I. INTRODUCTION

The properties peculiar to nanometer-size silicon (Si) structures are expected to play an essential role in realizing new electronic and optoelectronic devices based on quantum effects. In particular, the optical properties of Si have been attracting a great deal of attention from the viewpoint of optoelectronics since strong visible photoluminescence (PL) from porous Si was observed.[1] A lot of work has so far, been carried out on either porous Si with 2–50 nm mean diameter fabricated by electrochemical anodization or chemical vapor deposited (CVD) films of nanocrystalline Si. In both cases, the PL emission was observed in various colors between red and blue because the maximum PL signal position varied in the energy range of 1.4–2.6 eV according to the difference in sample fabrication processes. On the basis of these results, various PL emission mechanisms, such as quantum confinement[1–4] and surface effects which include Si species covered by siloxene,[5] the Si–H complex,[6] and the SiO_2 matrix[7–12] have been suggested. Nonetheless, what the emission mechanism of PL from nanometer-size Si structures is remains unclear.

From the viewpoint of applying this strong PL to Si optoelectronics, the fabrication of well-controlled Si nanostructures is essential. For this purpose, nanolithography, especially with an electron beam, is effective because the patterns with the desired size can easily be fabricated with an ultrafine beam. In fact, nanometer-size resist patterns have been formed by electron-beam nanolithography.[13] By using such resist patterns as an etching mask, Si nanostructures were fabricated and PL measurements were tried on these structures.[14–17] However, Si nanostructures did not always show PL emission. For example, PL was not observed from 10-nm Si structures passivated with hydrofluoric acid.[18] This is due to the ambiguity of the relationship between the Si pattern size and PL, as well as the confusion surrounding the emission mechanism.

We have studied PL spectra from two-dimensional (2D) Si nanolines formed using electron-beam lithography, Si etching, and oxidation. In this article, we report on the relationship between the PL emission characteristics of the Si lines and the pattern size. The Si lines were fabricated by etching a (110) SIMOX (separation by implanted oxygen) wafer with a KOH solution. Owing to orientation-dependent etching, patterns with a rectangular cross section were formed.[14] The Si patterns were separated electrically from the substrate by the buried SiO_2 layer. After the width of the lines decreased due to oxidation, the Si pattern shapes and sizes were evaluated by TEM and PL spectra from the oxidized Si lines were measured. In addition, PL from patterned and unpatterned $SiO_2/Si/SiO_2$ film structures was measured to understand the characteristics of PL from the oxidized Si lines. Based on a comparison of these results, factors affecting PL from oxidized Si lines are also discussed.

II. EXPERIMENT

A 5–10 Ω cm (100) *p*-type SIMOX wafer with a 100-nm-thick top Si layer was used for nanoline fabrications. After washing the wafer surface with hydrofluoric acid, 50-nm-thick positype resist ZEP-520 (Nippon Zeon Co.) was spun onto it and exposed to the electron beam. Resist lines about 15 nm wide along ⟨112⟩ directions were formed with a period of 70 nm in a 100 μm×100 μm area after development. The Si surface exposed between the lines was oxidized by ECR oxygen plasma to a thickness of 2 nm[19] to form an etching mask. After removal of the resist film with acetone and ultrasonic agitation, the top Si layer of unoxidized regions was etched in a 20% KOH aqueous solution at 10 °C. Consequently, Si lines with rectangular cross sections

[a]Electronic mail: nam@nttica.ntt.jp

TABLE I. Si thickness after oxidation at 800 °C.

Sample	Oxidation time (min)	Remaining Si thickness (nm)[a] Middle	Top
a	0	14.6±1.8	14.6±1.8
b	30	11.6±2.4	12.7±1.8
c	80	6.5±1.4	10.1±1.5
d	94	4.6±2.0	9.1±1.5
e	125	1.0±1.4	7.7±1.3
f	170	0	7.1±1.4

[a]Si thickness obtained from TEM micrographs.

and (111) side planes were formed.[14,20] The width of the Si lines was decreased to the desired thickness by thermal oxidation in a dry oxygen ambient at 800 °C.

Thin-film structures were formed on (111) and (100) SIMOX wafers with a 7–14 nm-thick top Si layer. After resist patterning by lithography, the exposed Si layer was etched with chlorine-based gas plasma. Then film structures composed of 15-nm-thick thermal SiO_2, 0–7-nm-thick Si, and 400-nm-thick buried SiO_2 layers were obtained after thermal oxidation in a dry oxygen ambient at 800 °C.

PL measurements were carried out using 488-nm excitation from an Ar-ion laser and a monochrometer with a thermoelectrically cooled photomultiplier coupled to a GaAs cathode. A laser beam with a power of 4 mW was focused on a spot about 10 μm in diameter. The sample temperature was adjusted by controlling the amount of liquid He in the cryostat.

III. RESULTS AND DISCUSSION

A. Characterization of Si pattern shape after oxidation

Five samples about 15 nm in width and 100 nm in height were oxidized under the conditions shown in Table I. Figure 1 shows TEM micrographs of these samples before and after oxidation. The Si width decreases with increasing oxidation time. In addition, the Si lines become deformed as oxidation proceeds, as shown in Fig. 1. It is known that the oxidation of Si causes a volume expansion because SiO_2 is about two times larger than Si in volume per Si atom. The expansion tends to be suppressed at the corners of the pattern and therefore the stress generated there is greater than that in the middle region. The result is the pattern deformation shown in the figure; the oxidation rate decreases with increasing stress.[21] For instance, in Fig. 1(e), Si remains at the corners after oxidation, but the Si in the middle regions is completely oxidized. Even after a longer period of oxidation, the width of the Si at the corners remains about the same because of the large stress, as shown in Fig. 1(f). Such a large stress at the corners would affect the amount of stress in the middle region. It is estimated by simple simulation that the stress in the whole oxide on the Si line ranges from 10^9 to 10^{10} dyn/cm^2, but it is difficult to determine the stress exactly. The Si thickness in the middle and upper parts of the lines after oxidation is summarized in Table I.

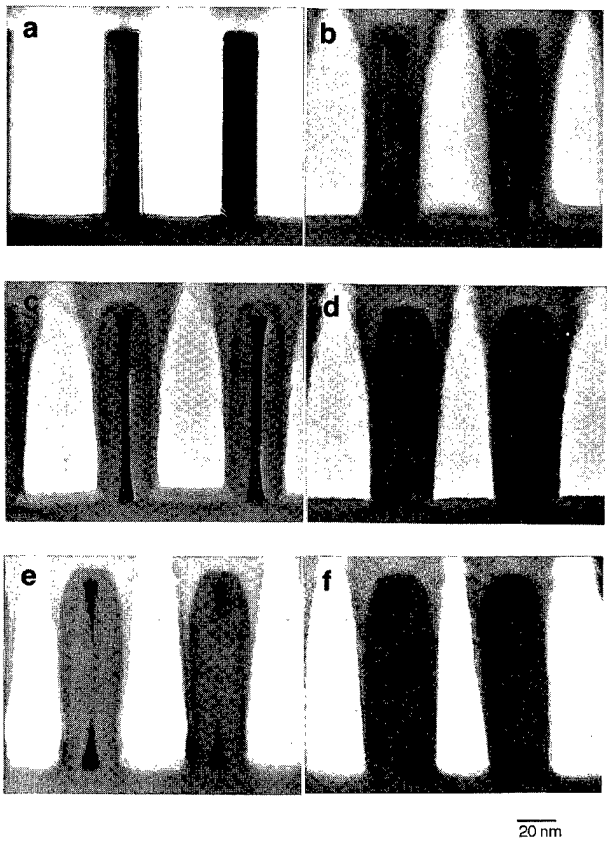


FIG. 1. TEM micrographs of the Si lines after (a) 0, (b) 30, (c) 80, (d) 94, (e) 125, and (f) 170 min of dry oxidation at 800 °C.

B. PL spectra from Si nanolines after oxidation

Figure 2 shows the PL spectra of the samples in Fig. 1. All oxidized Si lines clearly exhibit PL, whereas no PL was observed from sample (a), which was not oxidized. The PL intensity for samples (c) and (d), in which Si remains in the middle regions, is stronger than for samples (e) and (f), which have no Si in the middle. When the Si remaining is

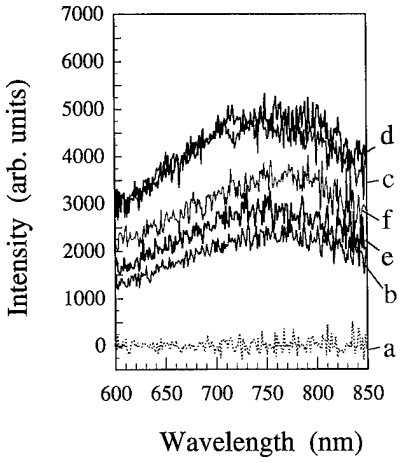


FIG. 2. PL spectra of the oxidized Si lines. These spectra are obtained by subtracting the background level from the original spectra. Sample temperature was maintained at 50 K.

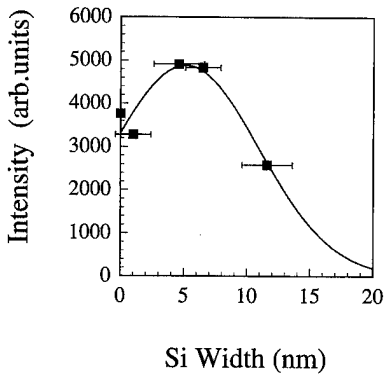


FIG. 3. The Si width dependence on the PL intensities at 760 nm. Sample temperature was maintained at 50 K. The Si widths were measured in the middle of the lines. The line is extrapolated under the assumption of a Gaussian distribution of the Si thickness.

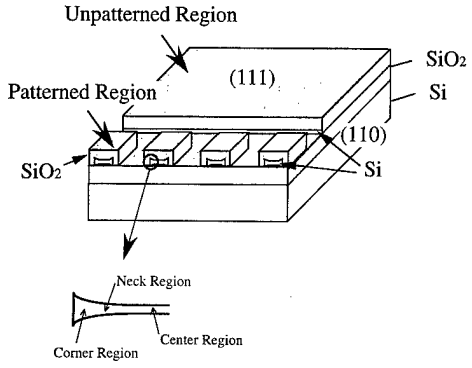


FIG. 5. The schematic of the $SiO_2/Si/SiO_2$ film structure used to examine the patterning effect. The patterns were fabricated by the etching of the top Si layer in (111) SIMOX substrates and subsequent oxidation. The Si thickness remaining was from 0 to 3 nm. The pattern size is 1 μm×0.75 μm.

still thick, like in sample (b), the intensity tends to decrease. On the other hand, the spectra for the oxidized Si lines have a peak at around a 760 nm wavelength (1.65 eV). This 1.65-eV PL is the same as the signal observed from oxidized porous Si[8,10,11] and a Si(100) 2D structure.[22,23] The spectra are broad and the peak position is independent of the Si width.

The dependence of the PL intensity at 760 nm on the Si width in the middle of the line is shown in Fig. 3. The intensity is maximum at a thickness of about 5 nm, which is thicker than estimated by the effective-mass approximation.[22] In this figure, the full width at half-maximum (FWHM) is considerably wide, indicating that a wide range of Si width contributes to emission. In addition, the emission is observed even when there is no Si left in the middle regions of the lines. This means the small amount of Si remaining in the corners is strongly related to the emission when the Si size in the corners becomes as small as the Si which contributes to the emission.

Figure 4 shows the temperature dependence of the maximum PL intensity. It decreases as the temperature increases, though the rate of the decrease is not so large from 20 to 80 K. The drop in PL intensity indicates that the ratio of nonradiative recombination processes, such as nonradiative carrier capture assisted by phonons on Si dangling bonds,[23] to radiative recombination becomes large with increasing temperature.

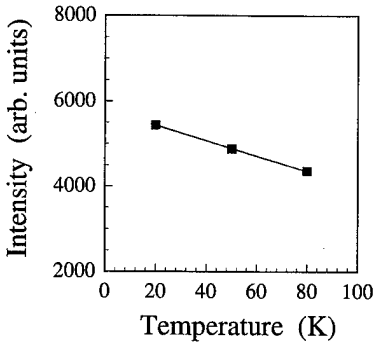


FIG. 4. PL maximum intensity as a function of sample temperature.

C. PL spectra from ultrathin Si film

In order to investigate the factors causing the interesting PL characteristics of the Si nanoline, we measured PL for patterned and unpatterned $SiO_2/Si/SiO_2$ film structures formed on (111) SIMOX substrates. The patterned $SiO_2/Si/SiO_2$ film structures were prepared by Si etching and subsequent oxidation so as to correspond to nanolines which were laid flat, as shown in Fig. 5. For these film structures, the Si-thickness effect on the PL intensities can be estimated exactly because the thickness variation of the Si layer is less than 0.5 nm. The Si film in the patterned structures is deformed just like the oxidized Si lines, although the degree of deformation is smaller than in the lines.

Figure 6 shows the PL spectra from unpatterned $SiO_2/Si/SiO_2$ film structures with several Si thicknesses in the pattern center region. The peak positions is near 770 nm.

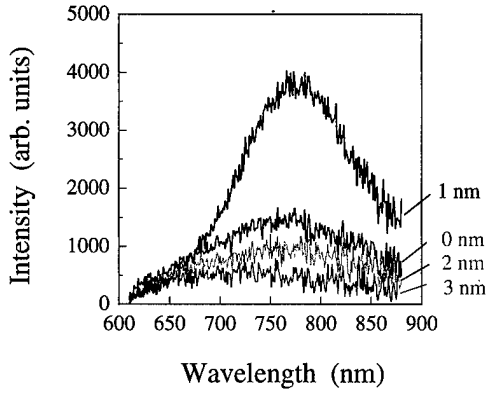


FIG. 6. PL spectra from the unpatterned regions of the $SiO_2/Si/SiO_2$ film structures as a function of Si thickness. Sample temperature was maintained at 50 K. The reason for the emission at the Si thickness of 0 nm is probably that Si grains remain even when the Si layer is expected to be completely oxidized and they contribute to emission.

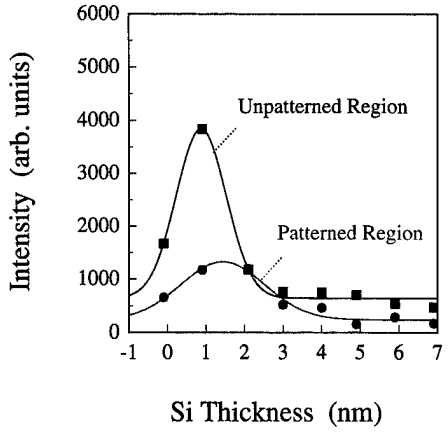


FIG. 7. The relationship between the maximum intensity and the Si thickness in the patterned and unpatterned $SiO_2/(111)Si/SiO_2$ film structures. The Si thickness was obtained by decreasing the Si thickness which corresponded to the oxidizing amount from an initial Si thickness. Sample temperature was maintained at 50 K.

In addition, strong PL emission is observed in the ~1-nm thick Si layer having (111) crystal plane as the top surface. The PL characteristics for the patterned structures are different from those for the unpatterned structures. The peak intensities for each Si thickness at 760–770 nm are shown in Fig. 7. The reason for the emission at the Si thickness of 0 nm is probably that Si grains remain even when the Si layer is expected to be completely oxidized and they contribute to emission. This is supported by TEM observation that a few Si particles remain in the sample indicated the negative thickness. As can be seen in Fig. 7, the maximum intensity in the patterned samples is lower than in the unpatterned ones. The proportion of the intensity drop is roughly consistent with that of the Si-area decrease caused by Si etching and oxidation. Other particular points of interest in Fig. 7 are that a wide range of Si thickness in the patterned samples contributes to the emission and the peak shifts to a somewhat larger Si thickness value than that of the unpatterned sample.

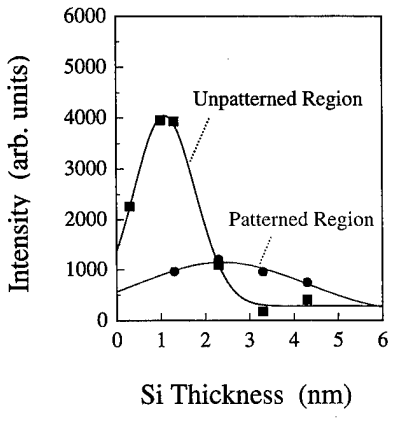


FIG. 8. The relationship between the maximum intensity and the Si thickness in the patterned and unpatterned $SiO_2/(100)Si/SiO_2$ film structures. Sample temperature was maintained at 50 K.

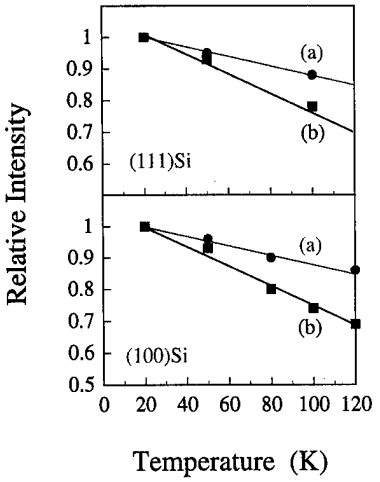


FIG. 9. PL intensity as a function of sample temperature in the (a) patterned and (b) unpatterned $SiO_2/Si/SiO_2$ film structure. The intensity at several temperatures is normalized by that at 20 K.

This behavior of the patterned sample is similar to that of the oxidized Si nanolines shown in Fig. 3, although the degree of the shift is different between the patterned film and the lines. Therefore, it is suggested that the strange behavior of PL from the oxidized lines is caused by the patterning and oxidation of the structures.

These PL phenomena due to the patterning of the ultrathin Si film are not special to the film structure formed in (111) Si. Figure 8 shows the relationship between the PL peak intensity, whose position is at the wavelength of 750 nm, and Si thickness in a (100) SIMOX wafer. It is interesting that the result from the (100) Si film is similar to the (111) Si film. It is presumed from Figs. 7 and 8 that there is not much difference in the excitation efficiency to the lowest excitation state in (100) and (111) Si, if the emission efficiency of (100) Si is similar to that of (111) Si. Anyway, whatever crystal planes the Si structure may have, to obtain maximum intensity the Si has to be quite thick, and a wide range of Si thickness contributes to the emission when the Si structure is patterned, though the intensity drops.

Figure 9 shows the temperature dependence of maximum PL intensities in the patterned and the unpatterned $SiO_2/Si/SiO_2$ layer structures. For both structures, the intensity decreases with increasing temperature in the range of 20–120 K. However, the proportion of the decrease for the patterned structure is smaller than that for the unpatterned structure. This is similar to the situation for the oxidized Si lines shown in Fig. 4. The cause is probably related to a small increment in the nonradiative process with increasing temperature. In the patterned samples, the probability of the phonon-related capture of carriers would not increase so much with increasing temperature because the lattice is under the strain due to the stress as discussed below.

D. Origin of PL from oxidized Si nanostructures

The fact that the PL peak position is at 760 nm regardless of the Si thickness, as shown in Fig. 2, indicates that the

emission does not come directly from the Si nanolines. This suggests that the mechanism of the PL emission from the oxidized Si lines is similar to that proposed for porous Si.[10] In that mechanism, radiative recombination occurs at the interface between Si and SiO_2: carriers become excited in the Si region where the band gap is widened due to quantum confinement effects and their energy eventually exceeds the energy separation of the radiative recombination centers at the Si/SiO_2 interface, i.e., 1.65 eV. The carriers then localize at the interface and the radiative recombination occurs there. Consequently, the peak position is independent of the Si width since emission is based on the constant band-gap energy of the interface regions. While this appears to be the primary mechanism involved in oxidized Si lines, it alone cannot explain the PL behavior in Fig. 3.

In the thin-film samples in which the Si layers were patterned, changing the number of divisions of the Si layer as a total Si area is the same, did not result in any clear change in the PL intensities. If the neck region between the corner and the center region of the pattern (see Fig. 5) were mainly related to the emission, the PL intensity would increase with increasing numbers of the divisions, i.e., the quantity of the neck regions. From these results, we conclude that the Si regions remaining abundantly in the center of the pattern chiefly contribute to the emission. This is expected to also be the case for the line pattern.

The reason that the Si yielding the maximum intensity in the patterned 2D–Si layers is thicker, as is seen in Figs. 3, 7, and 8, is thought to be as follows. As described in Sec. III A, the stress in the Si lines is not uniform. Such stress is expected to change the band structure, but exactly how it changes is difficult to estimate. For example, indirect band-gap energy is known to decrease by the simple application of hydrostatic pressure.[24] When tensile tangential stress is applied to a Si layer through oxidation, as in this experiment, the indirect band-gap energy is expected to increase.[22] On the other hand, it is expected that the number of dangling bonds acting as nonradiative centers increases at the Si/SiO_2 interface due to strain brought on by the large amounts of stress. Consequently, the number of carriers captured by the nonradiative centers increases, which could lower the emission efficiency at the interface of the patterned ultrathin (0–2 nm thick) Si region. As a result of the combination of the above phenomena, it is thought that Si thicker than predicted produces maximum intensity when the Si layer is patterned regardless of whether the silicon is in the form of lines or films. The reason the emissions come from Si regions with a broad range of thickness is probably due to the large variation in the stress generated in the lines or films.

IV. CONCLUSION

The Si lines with (111) side plane and rectangular cross sections were fabricated using (110) SIMOX substrates with electron-beam lithography and KOH etching. After the lines were oxidized at several thicknesses, PL was measured from the oxidized Si lines. Clear spectra with peak intensity at a wavelength of around 760 nm were observed and the emission was found to be due to the recombination of carriers at the Si/SiO_2 interface, though the detailed excitation processes are unknown. In addition, TEM images showing the Si thickness in conjunction with the PL intensity from the lines lead us to conclude that Si lines with a central thickness of 5 nm produce the maximum intensity. Furthermore, a wide range of the Si thickness contributes to the emission. The causes for this were clarified by discussing the difference in the PL characteristics for patterned and unpatterned $SiO_2/Si/SiO_2$ layer structures. First, it was found that the (111) Si layer in the unpatterned $SiO_2/Si/SiO_2$ layer structures has a maximum intensity at around a 1 nm thickness. With patterning, however, the Si thickness required to produce maximum intensity increases and a wide range of the Si thickness contributes to the emission, which is similar to the characteristics of the oxidized Si lines. Therefore, the changes in PL behavior are probably caused by the increased amount of stress produced by the Si patterning and subsequent oxidation process.

ACKNOWLEDGMENTS

The authors thank Dr. Seiji Horiguchi, Dr. Yasuo Takahashi, and Dr. Akira Fujiwara for helpful discussions and Dr. Katsutoshi Izumi for encouragement throughout this work.

[1]L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[2]H. Takagi, H. Ogawa, Y. Yamazaki, A. Ishizaki, and T. Nakagiri, Appl. Phys. Lett. **56**, 2379 (1990).
[3]R. Tsu, H. Shen, and M. Dutta, Appl. Phys. Lett. **60**, 112 (1992).
[4]T. Takagihara and T. Takeda, Phys. Rev. B **46**, 15578 (1992).
[5]M. S. Brandt, H. D. Fuchs, M. Stutzmann, J. Weber, and M. Cardona, Solid State Commun. **81**, 307 (1992).
[6]C. Tsai, K. H. Li, D. S. Kinosky, R. Z. Qian, T. C. Hsu, J. T .Irby, S. K. Banerjee, A. F. Tasch, C. Campbell, B. K. Hance, and J. M. White, Appl. Phys. Lett. **60**, 1700 (1992).
[7]R. Kumar, Y. Kitoh, and K. Hara, Appl. Phys. Lett. **63**, 3032 (1993).
[8]S. M. Prokes, Appl. Phys. Lett. **62**, 3244 (1993).
[9]H. Morisaki, H. Hashimoto, F. W. Ping, H. Nozawa, and H. Ono, Appl. Phys. Lett. **74**, 2977 (1993).
[10]Y. Kanemitsu, T. Ogawa, K. Shiraishi, and K. Takeda, Phys. Rev. B **48**, 4883 (1993).
[11]Y. Kanemitsu, H. Uto, Y. Matsumoto, T. Matsumoto, T. Futagi, and H. Mimura, Phys. Rev. B **48**, 2827 (1993).
[12]Y. Kanemitsu, T. Futagi, T. Matsumoto, and H. Mimura, Phys. Rev. B **49**, 14732 (1994).
[13]W. Chen and H. Ahmed, Appl. Phys. Lett. **62**, 1499 (1993).
[14]H. Namatsu, M. Nagase, K. Kurihara, K. Iwadate, T. Furuta, and K. Murase, J. Vac. Sci. Technol. B **13**, 1473 (1995).
[15]P. B. Fischer, K. Dai, E. Chen, and S. Y. Chou, J. Vac. Sci. Technol. B **11**, 2524 (1993).
[16]H. I. Liu, N. I. Maluf, R. F. W. Pease, D. K. Biegelsen, N. M. Johnson, and F. A. Ponse, J. Vac. Sci. Technol. B **10**, 2846 (1992).
[17]A. G. Nassiopoulos, S. Grigoroulos, E. Gogolides, and D. Papadimitriou, Appl. Phys. Lett. **66**, 1114 (1995).
[18]P. B. Fischer and S. Y. Chou, Appl. Phys. Lett. **62**, 1414 (1993).
[19]K. Kurihara, K. Iwadate, H. Namatsu, M. Nagase, and K. Murase, Microelectron Eng. **27**, 125 (1995).
[20]H. Namatsu, M. Nagase, K. Kurihara, K. Iwadate, and K. Murase, Microelectron. Eng. **27**, 71 (1995).
[21]H. I. Liu, D. K. Biegelsen, N. M. Johnson, F. A. Ponce, and R. F. W. Pease, J. Vac. Sci. Technol. B **11**, 2532 (1993).
[22]Y. Takahashi, T. Furuta, Y. Ono, T. Ishiyama, and M. Tabe, Jpn. J. Appl. Phys. **34**, 950 (1995).
[23]P. N. Saeta and A. C. Gallagher, J. Appl. Phys. **77**, 4639 (1995).
[24]D. L. Camphausen, G. A. N. Connel, and W. Paul, Phys. Rev. Lett. **26**, 184 (1971).

Computer simulations of porous silicon formation

Y. M. Weng and J. Y. Qiu[a)]
Materials Science Institute of Fudan University, Shanghai, People's Republic of China 200433

Y. H. Zhou
Electronic Engineering Department of Fudan University, Shanghai, People's Republic of China 200433

X. F. Zong
Materials Science Institute of Fudan University, Shanghai, People's Republic of China 20043

(Received 16 January 1996; accepted 3 June 1996)

A computer simulation has been carried out to study the dynamics of pore formation in porous silicon. Porous structures were generated by a simple three-dimensional Monte Carlo calculation, based on a modified percolation model for pore growth in lightly and heavily doped silicon substrates. The results are consistent with recent observations. As an approximate guide, the type formed morphology was shown to depend mainly on the doping level of the substrate. A strong dependence of the average porosity on the doping level of the substrate is also predicted. © *1996 American Vacuum Society.*

I. INTRODUCTION

The discovery of visible room-temperature photoluminescence from electrochemically etched highly porous silicon (PS)[1,2] has generated much interest in this material because of the possibility of an application to Si-based optoelectronic devices. A large number of papers on optical and structural properties of PS have been published during the last few years, but most basic questions concerning the origin of the photoluminescence (PL) and the morphology of PS film still remain open. No complete understanding of pore morphology exists, and it is likely that improvements in our understanding will come about through the application of high-resolution microscopy. Transmission electron microscopy and scanning tunneling microscopy have already proved essential in probing the porous structure. PS layers are inevitably fragile and this makes their evaluation more difficult, and necessitates the development of some novel approaches to specimen preparation. A correlation between the surface morphology of PS layers and their respective efficiency of the visible photoluminescence signal has been reported by Enachescu *et al.*[3] Several theoretical models are available, including the quantum model[4] and the diffusion limited model.[5] Very recently, a computer simulation of pore growth in n-type silicon has been developed by Erlebacher *et al.*[6]

It is well known that the significantly strong luminescence from PS is often observed on highly porous samples grown on lightly doped (nondegenerate) p-type substrates.[7] However, PS layers in lightly doped p-type silicon are more difficult to characterize by computer simulation due to the smaller length scales.[6] For lightly doped p-type silicon, a depletion is also set up but the field strength is relatively small in comparison to the conditions for pore formation in heavily doped (degenerate) silicon. Recent studies suggest that the structure from the PS layer obtained from a lightly doped p-type substrate is a network of interconnected nanocrystallites showing a disordered spongelike morphology rather than a series of parallel columns as from heavily doped n-type substrate.[8,9]

It is well known that the percolation model[5] can be described as random diffusion and flow of fluid in disordered porous materials. The approach is one of the most useful methods of understanding the properties of some disordered systems. It differs from the usual diffusion such as the diffusion limited model. The randomness in the percolation process is due to the disordered structure possessed by materials themselves. Recently, we have shown[10] that the Hansdorff dimension of PS surface is $D=1.88$, which is in good agreement with the numeral simulation for invasion percolation model. The pore formation in PS can be described by the dynamic process of invasion percolation of the HF solution in electrochemical anodization.

The purpose of this study is to perform a Monte Carlo calculation, based on a modified percolation model, to simulate the dynamics of pore formation on silicon substrate in electrochemical anodization, and to obtain a better understanding of the problem. The simulations presented here model the dissolution on a two- or three-dimensional grid representing a section of silicon. Disordered spongelike morphologies are formed through a simulated dissolution process by a modified percolation model. Many other groups have used computer simulations to model pore growth in silicon.[6,11] While in our model, an approximation for the electric field within the silicon bulk and its resultant influence on the dynamics of carrier transport in the vicinity of the pore front are considered. Previous models where the charge carriers follow random walks to the silicon/HF interface do not predict the spongelike feature of PS. We emphasize that our model helps to clarify many of not only the larger length scale but also the smaller length scale morphological characteristics of PS, although these preliminary results make no predictions about atomic scale features.

[a)]Electronic mail: jinglin@fudah.ihep.ac.c

II. MODEL FOR HOLE TRANSPORT DURING PORE GROWTH

A widely accepted mechanism of pore formation in anodized silicon implies the process of local dissolution of silicon during the anodization.[12] The anodic etching of silicon leads to the following reaction on the surface[13]

$$\mathrm{Si}+2\mathrm{HF}+(2-n)e^{+}=\mathrm{SiF_2}+2\mathrm{H}^{+}+ne^{-}, \quad (1)$$

where $n<2$, e^{+} and e^{-} represent hole and electron. Here it is assumed that holes take part.

SiF_2 is unstable

$$2\mathrm{SiF_2}=\mathrm{Si}^{(0)}+\mathrm{SiF_4}. \quad (2)$$

A very thin surface porous film (SPF) is formed by the deposition of elementary silicons $Si^{(0)}$ produced by disproportion reaction (2) on the surface of the lightly doped p-type substrate.[13] After this initial process, the surface porous film can be considered as a porous network connected by narrow throats, similar to a percolation. Under a small constant current electric field, the flow rate of the HF solution is constantly kept very low so that viscous force compared with capillary force can be neglected. The dynamic process of the random diffusion process of $Si^{(0)}$ and invasion percolation of HF solution in electrochemical anodization can be described by the model suggested in Ref. 14.

Our model is similar to the percolation model.[15] In the context of our simulation model, a 200×200×100 unit cubic lattice is taken into consideration. Each point in the cubic lattice represents a unit of silicon atoms.

A point can be one of several values, 0, 1, 4, or 8. If the value of a point is 0, it means that the point is an etched one. If the value of a point is 1, it represents that the point is an active one, 4 means that it is a surface point, and 8 means an inner point. In the anodization process, the anodic etching can only occur on the interface of HF acid and silicon. It is impossible for an inner atom to be etched, and the inner point represents just these kinds of silicon atoms. Also, the etching process occurs only at several places on the interface. If every place on the interface has the same probability to be etched, the porous morphology will not be received. The places where etching happens are active areas on the interface, and their counterparts in our model are active points. Other atoms on the interface which have little chance to be etched are represented by surface points in our model. The structure of a general simulation is shown in Fig. 1(a).

Our simulation was carried out at a 200×200×100 cubic lattice with a surface porous film as shown in Fig. 1(b). Initially, a very thin (thickness=1) SPF with a porosity of 0.1 was assumed. Under this initial condition, the surface could be regarded as a porous network, similar to percolation. On the SPF, a group of surface points were set to active points. They randomly dispersed on the top of a 200×200×100 cubic lattice and the top-most layer of the lattice is considered to be the silicon HF interface. The total number of the initial active points is named the seeds number, which represents the initial current value. Different porous silicon can be received under different current conditions. By changing the seeds number in our model, different morphologies of porous silicon can be produced.

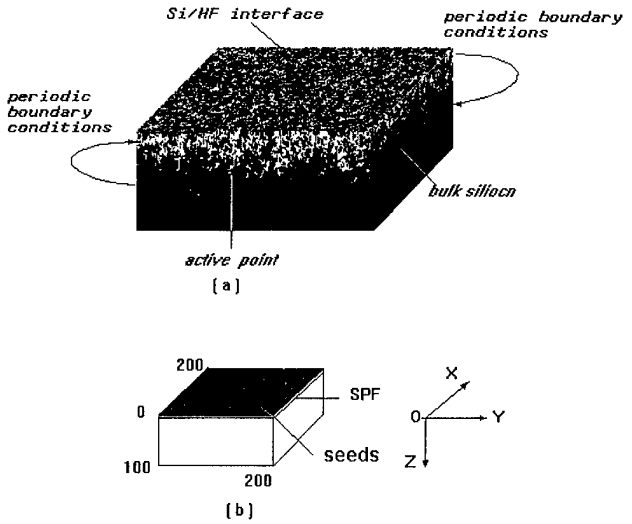


FIG. 1. (a) Simulation structure showing important features. (b) The 200 ×200×100 cubic lattice with a surface porous film, at which our simulation was carried out.

In the anodization simulation process, the cubic lattice was scanned several times. The scanning times represent the period of time of the anodization process. In one scanning, each point on the lattice was calculated. If the point currently scanned is an etched or inner point, no action will be taken. If it is a surface point, it has some probability to be changed into active point. This is one way to produce an active point. If the point currently scanned is an active one, it will be etched (its value will be set to zero) and adjacent inner points will be made surface ones. Also there is probability for one of the adjacent points to be changed into an active point, this is the other way to produce an active point. Because the porous silicon is formed under constant current condition, the total number of active points should be kept equal to the initial seeds number at any time.

Etching always occurs at the place where the static electric field strength is largest. It is known that the field strength is very large on the tip of a charged conductor. The HF acid can be regarded as a good conductor compared with the porous silicon whose resistivity is very high. On the tips of the pores, the field strength is larger than that at other places on the interface, thus etching always happens on tips. In our model, points next to an active point have more chances to be changed into active ones and etched at the next scan.

The process, in which an active point is etched and an adjacent points is changed into an active point, can be regarded as that of active point transfers from one place to another. With how much probability will the active point move downward or in other directions determines the morphology of the porous silicon. If the field strength along Z axis in Fig. 1(a) is not strong, it is a small length scale case.[6] In such a case, the bias of the holes to move toward the tips is not very strong, and its movement could be regarded to be

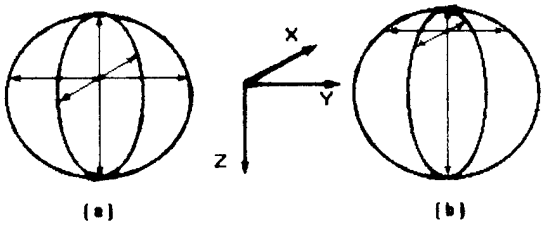


FIG. 2. The probability ball construction for the active points on a cubic lattice. Under little bias (a), the active point moves in Z direction with a little larger probability than in other directions. Under strong bias (b), the probability of active point to move in Z direction is much greater than that in other directions.

mainly random. Thus the holes will reach tips from different directions with almost equal probabilities. It can be regarded as that of the active points move in different directions with nearly equal probabilities. If the field strength along Z axis is very large, it is a large length scales case.[6] The holes are swept toward the pore tips by the electric field along Z axis, and the influence of random walk on holes is much weaker. The holes always reach the tips from below. It seems as if the active points always move downward along the Z axis. In a lightly doped silicon, the static electric fields near the tips in silicon are much weaker than those in heavily doped silicon. In lightly doped silicon, the probability of the active point to move downward is just a little greater than in other directions, as shown in Fig. 2(a). However in heavily doped silicon it is much greater [see Fig. 2(b)]. This difference causes the difference in morphologies.

III. RESULTS AND DISCUSSION

Here we describe the simulation results and discuss the effect of electric field on the resultant morphologies. Finally, we discuss the implications of this model for pore formation in silicon.

Figures 3–8 are pseudo-three-dimensional morphologies of PS. Because of the limitation of our computer, Figures 3–8 only show morphologies at 50×50 areas in one layer of the cubic lattice. The 50×50 area in figures showing surface morphologies are areas where "$Z=0$, $A<=x<=<=A+50$, $B<=y<=B+50$" in the cubic lattice. A and B are arbitrary numbers between 0 and 150. If the figures show cross-sectional view, the 50×50 areas are areas where "$x=A$, $0<=z<=50$, $B<=y<=B+50$" or "$y=A$, $0<=z<=50$, $B<=x<=B+50$" in the lattice. A is an arbitrary number between 0 and 200, and B is an arbitrary number between 0 and 150.

As results of simulation, Fig. 3 shows a surface morphology of PS grown on a lightly doped silicon (where a depletion layer is also set up but its width is relatively large, so the

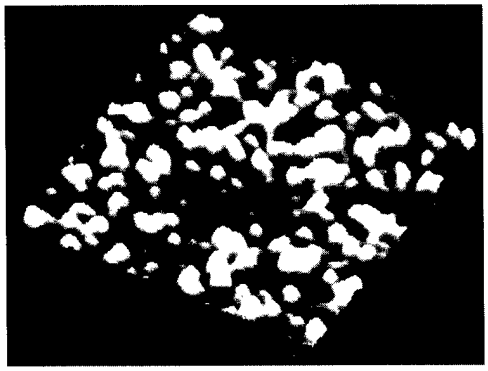


FIG. 3. Surface morphology (50×50 area) of PS with high porosity grown on a lightly doped silicon.

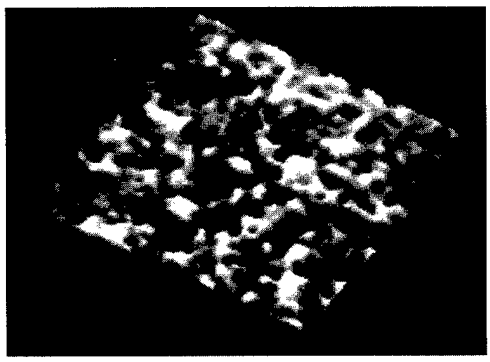


FIG. 4. Cross-sectional view (50×50 area) of the same case in Fig. 3.

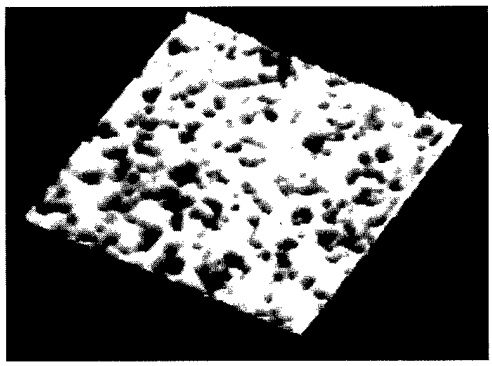


FIG. 5. Surface morphology (50×50 area) of PS with low porosity grown on a lightly doped silicon.

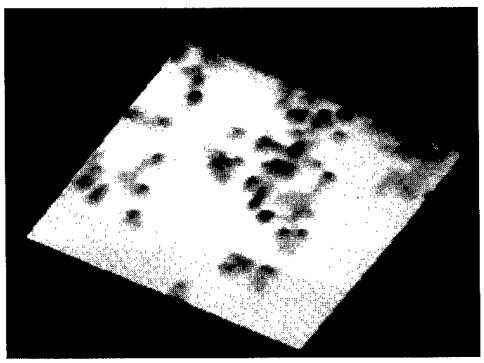


FIG. 6. Cross-sectional view (50×50 area) of the same case in Fig. 5.

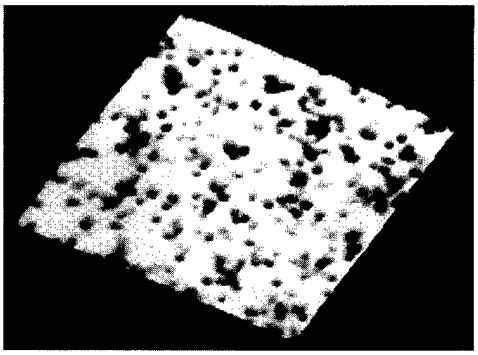


FIG. 7. Surface morphology (50×50 area) of PS with high porosity grown on a heavily doped silicon.

field strength is relatively small), which was formed with a large anodic current. Figure 4 shows the cross-sectional view of micrograph of the same case. We can find that the PS exhibits a spongelike morphology, the pores wind its way into silicon, and connect to each other. Also the silicon grains connect to each other. There is no tree or channellike feature. The porosity in a layer is about 80%. There is similarity in the morphologies of cross section and surface of PS. These results shown in Figs. 3 and 4 are in agreement with the experimental observations presented in Ref. 16. It must be emphasized that the intense luminescence is often observed on PS of high porosity, while the photoluminescence from PS of low porosity is very weak.

Figure 5 is the surface morphology of PS on a lightly doped substrate under a small anodic current. As a comparison, Fig. 6 shows the cross section of that PS. We can see the difference of morphologies between a PS under a small anodic current and that under a large current. First, the porosity of the surface is smaller. Second, the cross-sectional view of morphology is different with the surfaces. There is less connect among pores.

Figure 7 shows a surface morphology of PS grown on a heavily doped silicon (where a depletion layer is also set up but its width is relatively small, so the field strength is relatively large), which was formed with a large anodic current. Figure 8 is its cross-sectional view. There is clearly difference between different doping substrates. The surface is similar to that of the lightly doped one, but the feature of the cross-sectional view is different. The latter shows a channellike characterization.

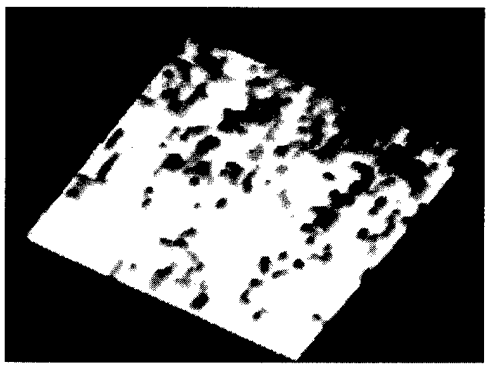


FIG. 8. Cross-sectional view (50×50 area) of the same case in Fig. 7.

Our results confirm that the effect of electric field biasing hole motion within the bulk of silicon is an important aspect of pore growth in silicon. An important new insight into PS emerges from these studies. In lightly doped silicon, either *p* or *n* type, the microstructure consisted of an apparently spongelike random array of fine hole markedly different from the heavily doped structure, which shows a channellike cross-sectional feature. In the case of pore propagation in heavily doped silicon, the local electric field distribution plays an important role in controlling the rate of interfacial charge transfer, which was different from the lightly doped silicon. Our results of dynamical simulation are in favorable agreement with those reported by Beale *et al.*[17]

The difference in two kinds of substrates is that the active points in heavily doped silicon move downwards much faster than those do in lightly doped silicon. This results in different morphologies in the two kinds of PS. The morphology is in accordance with the experimental results, therefore the model is usable. The simple model can show the main difference and explain some phenomena.

Otherwise, an obvious dependence of the porosity on the doping level of the substrate is also predicted. The porosity increases with the doping level.

IV. CONCLUSION

In summary, the principal results of this paper are as follows:

(1) A simple three-dimensional Monte Carlo simulation of PS formation has been developed.

(2) Our results confirm that the effect of electric field biasing hole motion within the bulk of silicon is an important aspect of pore growth in silicon. The model presented does not consider the details of the chemistry of the process, however, the formation of a porous nanostructure is shown to be determined by the semiconductor physics rather than by the chemistry.

Two types of morphologies of PS, spongelike and channellike structures, have been reported and details of their differing microstructure presented. The type formed was shown to depend only on the doping level of the substrate.

As an approximate guide, it appears that lightly doped silicon produces a fine network of pores whereas heavily doped material produces more of a channel structure.

(3) A strong dependence of the average porosity on the doping level of the substrate is predicted. The porosity increases with the doping level.

(4) The result clearly shows the presence of spongelike morphology in intense luminescent PS.

Our results confirm that the effect of electric field biasing hole motion within the bulk of silicon is an important aspect of pore growth in silicon. We believe that the present results provide a good guide to better understanding the structure and the properties of PS.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation.

[1]L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[2]J. S. Gardells, J. S. Rimmer, P. Dawson, B. Hanmiton, and R. A. Kublak, Appl. Phys. Lett. **59**, 2118 (1992).
[3]M. Enachescu, E. Hartmann, and F. Koch, Appl. Phys. Lett. **64**, 1365 (1994).
[4]R. Forrest and T. A. Witten, J. Phys. A **12**, 2109 (1979).
[5]R. L. Smith, S. F. Chuang, and S. D. Collins, J. Electron. Mater, **17**, 533 (1988).
[6]J. Erlebacher, K. Sieradzki, and P. C. Searson, J. Appl. Phys. **76**, 182 (1994).
[7]A. G. Cullis and L. T. Canham, Nature **353**, 335 (1991).
[8]I. Berbezier and A. Halimaoui, J. Appl. Phys. **74**, 5421 (1993).
[9]W. Yumin, Z. Xiangfu, Z. Renjing, C. Yi, Z. Wuning, and F. Zhineng, J. Infrared Millimeter Waves **12**, 164 (1993).
[10]W. Yumin, Z. Xiangfu, Y. Xingju, and C. Yufen, J. Fudan Univ. (Natural Sci.) **34**, 94 (1995).
[11]R. L. Smith and S. D. Collins, Phys. Rev. B **39**, 5409 (1989).
[12]V. Lehman and H. K. Foell, J. Electrochem. Soc. **137**, 659 (1990).
[13]T. Unagami, J. Electrochem. Soc. **127**, 476 (1980).
[14]D. Wihinson and J. Willemsen, J. Phys. A **16**, 3363 (1983).
[15]H. Shlomo and B. A. Daniel, Adv. Phys. **36**, 695 (1987).
[16]W. Yumin, Y. Xingju, C. Yufen, and Z. Xiangfu, Chin. Phys. Lett. **11**, 390 (1994).
[17]M. I. J. Beale, J. D. Benjamin, M. J. Uren, N. G. Chen, and A. G. Cullis, J. Cryst. Growth **73**, 622 (1985).

Comparison of advanced plasma sources for etching applications. V. Polysilicon etching rate, uniformity, profile control, and bulk plasma properties in a helical resonator plasma source

J. T. C. Lee,[a] N. Layadi, K. V. Guinn, H. L. Maynard,[b] F. P. Klemens, D. E. Ibbotson, and I. Tepermeister[c]
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

P. O. Egan and R. A. Richardson
Lawrence Livermore National Laboratory, Livermore, California 94550

(Received 8 February 1996; accepted 1 May 1996)

Etching of polysilicon features using a helical resonator plasma source is evaluated. Performance metrics consist of etching rate, etching rate uniformity, and profile control using HBr/O_2–He gas-phase chemistry. The effect of source power, rf-bias power, and reactor pressure on etching rate and uniformity is examined using a response surface experiment. Feature profile control is determined by examining nested and isolated lines and trenches using oxide mask/polysilicon/oxide structures. Good uniformity and vertical profiles are obtained at low reactor pressures, high source power, and rf-bias between 50 and 60 W. The operating point for best uniformity is at 3.5 mTorr, 3000 W source power, and 53 W rf-bias power. At this point, the etching rate is 3700 Å/min and the nonuniformity is less than 1.0%, over 125-mm-diam wafers. Radial profiles of electron temperature and ion density near the wafer surface are presented as a function of source power, rf-bias power, and reactor pressure. The ion density was found to be in the mid-10^{11} cm^{-3} range and electron temperatures were 5–7 eV. An increase in source power and reactor pressure results in an increase in ion density; however, the electron temperature shows a weaker dependence. Finally, these results are compared to those using helicon and multipole electron cyclotron resonance plasma sources evaluated in previous studies. We found that all three plasma sources provide high ion density at low pressures to meet performance demands for polysilicon etching; however, the helical resonator source offers somewhat higher etching rate and better bulk plasma uniformity. © *1996 American Vacuum Society.*

I. INTRODUCTION

The fabrication of ultra-large-scale integrated circuits requires increasingly tighter control of feature profiles in order to achieve high circuit densities.[1] Low pressure, high plasma density sources are becoming the accepted technology in many development laboratories and will be installed in significant numbers in production lines in the near future. Due to low pressure operation, these sources provide more anisotropic etching without sacrificing ion density and thus, etching rates. An important attribute of etch systems utilizing these sources is the ability to control the ion density and the ion energy independently through wafer biasing.[2] Many techniques are available for achieving high plasma densities at low pressures, including microwave excitation, such as in electron cyclotron resonance (ECR) sources,[3–6] and rf excitation, such as in helicon and helical resonator sources.[7,8] A quantitative evaluation and comparison of these plasma sources is necessary in order to make intelligent choices for a desired application.

The first two plasma sources we studied[9–12] were as follows, *(i) A multipole ECR, manufactured by Wavemat of Plymouth, Michigan*: The ECR plasma source consists of a 2.45 GHz microwave power supply and a microwave cavity surrounding a bell jar. High plasma densities are achieved by confining electrons in magnetic field cusps established by equally spaced permanent magnets surrounding the wafer plane. *(ii) An rf inductively coupled helicon, manufactured by Lucas Labs of Sunnyvale, California*: The helicon plasma source consists of an upper antenna section and a lower confinement section. Power at 13.56 MHz is coupled into the

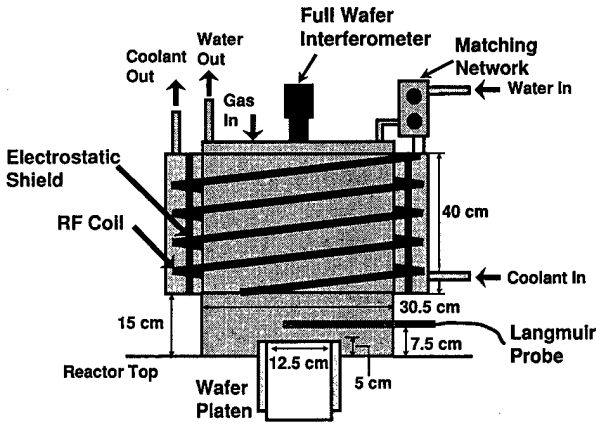


FIG. 1. Helical resonator plasma source.

[a]Electronic mail: tclee@bell-labs.com
[b]Present address: Lam Research Corporation, Fremont, CA 94538.
[c]Present address: Low Entropy Systems, Brighton, MA 02135.

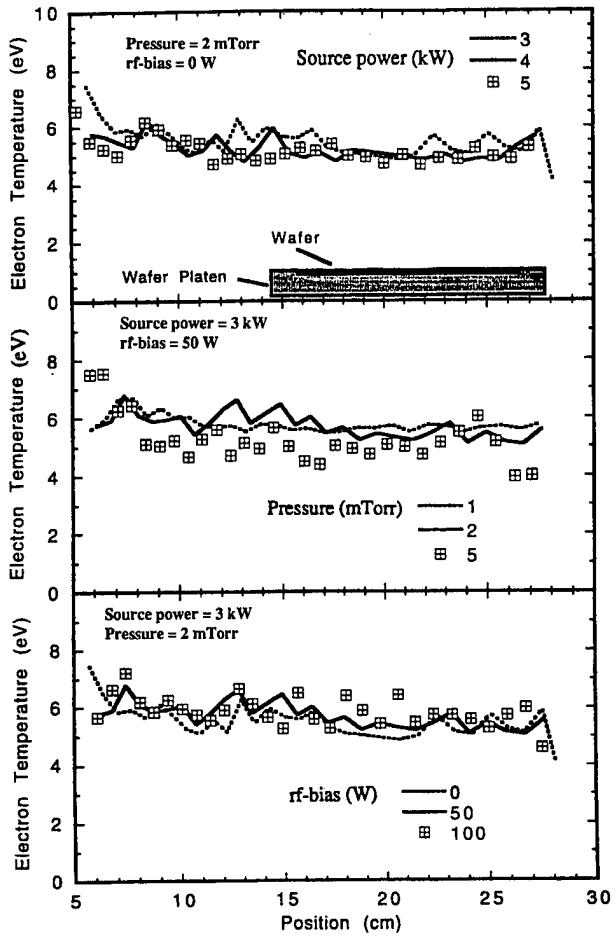


FIG. 2. Electron temperature (eV) in the helical resonator as a function of source power, reactor pressure, and rf-bias.

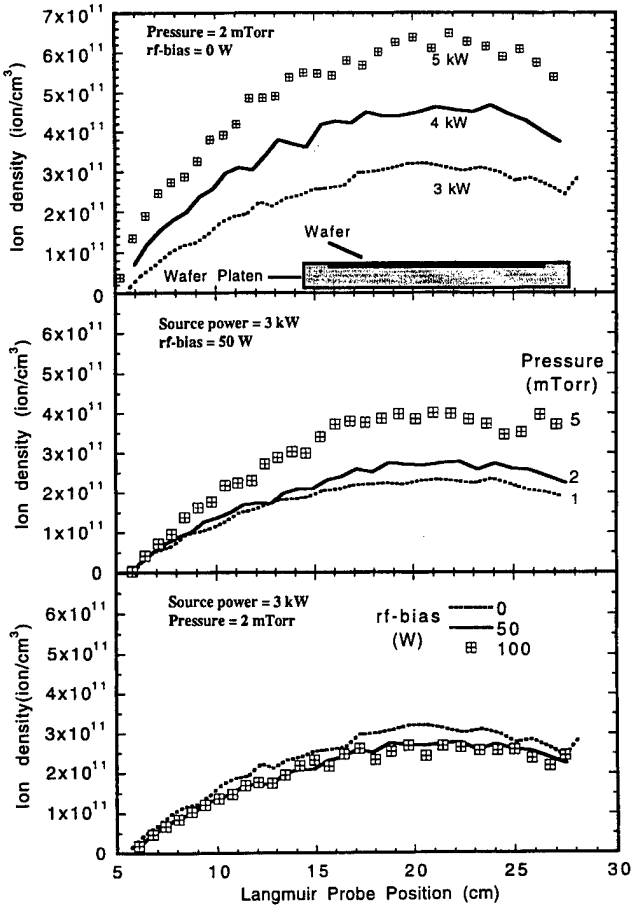


FIG. 3. Ion density (cm^{-3}) in the helical resonator as a function of source power, reactor pressure, and rf-bias.

plasma using a proprietary antenna design. Four solenoid magnets are used to control the magnetic field. The upper two magnets are necessary to support the helicon wave, while the bottom two magnets are used to improve uniformity and help maintain high plasma density at the wafer plane.

In our effort to compare commercially available, low pressure, high density plasma sources using identical reaction chambers to exclude platform dependent effects, we present here results obtained with the third source, a 12-in.-diam helical resonator manufactured by Prototech Research of Tempe, Arizona. We perform a statistical evaluation of polysilicon etching rates and uniformities as a function of source power, rf-bias power, and reactor pressure. Based on these results, operating conditions for optimum uniformity for the helical resonator source are determined. The effect of process parameters on polysilicon profiles is then evaluated and compared. We also present radial Langmuir probe measurements of the plasma properties above the wafer. These results are then compared with those obtained in our previous studies, using helicon and ECR plasma sources.

II. EXPERIMENTAL DETAILS

The setup of the helical resonator plasma source is shown in Fig. 1. The plasma source consists of a coil wrapped around a 12-in.-diam quartz tube. Radio frequency power up to 5000 W, at 13.56 MHz, is inductively coupled into the plasma through the coil. A Fararady shield inserted between the coil and the quartz tube reduces capacitive coupling that would result in sputtering of the quartz wall.

The 125-mm-diam wafer can be rf-biased (~13.56 MHz) through the wafer platen with a 600 W power supply. Helium backside cooling is used to remove heat from the wafer to the platen. Liquid nitrogen is circulated through the wafer platen to remove the heat, and the platen temperature can be controlled from −100 to +30 °C.

In order to characterize etching rates and uniformities, timed etches on 5000-Å-thick undoped, blanket polysilicon over 1000 Å of thermal oxide wafers were used. Etching was carried out for 15, 60, and 120 s and five-point thickness measurements (center point and four edge points) were taken before and after etching using a Nanometrics 216S scanning UV film thickness measurement instrument. All etching times were shorter than that required to clear the polysilicon film and reach the underlying oxide film. The initial polysilicon thickness is uniform to within ±1% of the average value obtained from the five-point measurement. Uniformity was determined as:

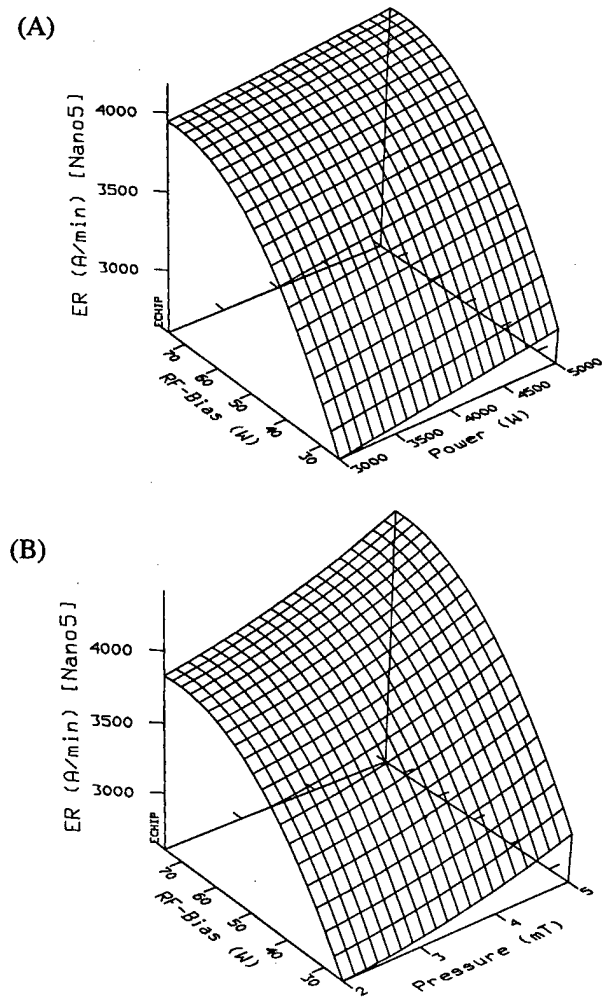


FIG. 4. Etching rates with the helical resonator, obtained from 5-point Nanometrics measurements: (A) as a function of applied source power and applied rf-bias power at 3.5 mTorr, (B) as a function of reactor pressure and applied rf-bias power at 4000 W applied source power.

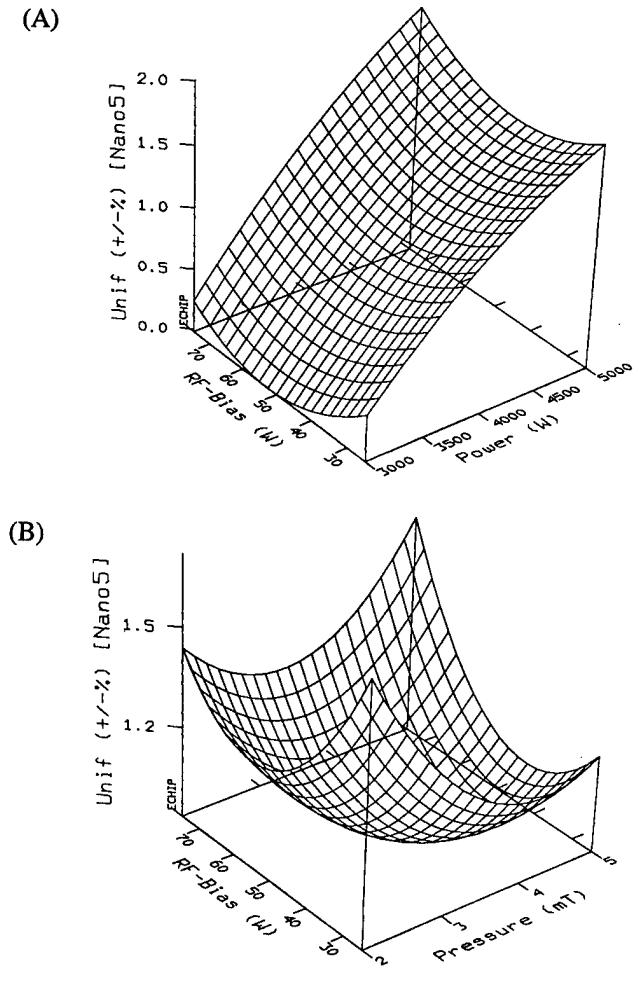


FIG. 5. Etching rate uniformity with the helical resonator, obtained from 5-point Nanometrics measurements: (A) as a function of applied source power and applied rf-bias power at 3.5 mTorr, (B) as a function of reactor pressure and applied rf-bias power at 4000 W applied source power.

$$\text{Uniformity } (\pm\%) = \frac{\text{ER}_{\max} - \text{ER}_{\min}}{\text{ER}_{\text{av}}} \frac{100}{2},$$

where $\text{ER}_{\min}$, $\text{ER}_{\max}$, and ER_{av} are, respectively, the minimum, maximum, and average etching rates.

Etching rates and uniformities were also measured by full wafer imaging (FWI) interferometry (Model No. 100-IS, made by Low Entropy Systems, Brighton, MA). The FWI (5- or 49-point) technique takes advantage of the plasma emission as the light source for determining *in situ* film thickness and etching uniformity over the full area of the wafer.[13]

The goal of any etching process development is to achieve the desired profiles while minimizing loss of process uniformity and etching rate. In order to evaluate profile control after etching, we etched polysilicon using hard masks since this will most likely be the mask of choice for 0.35 μm devices. The corresponding wafers consist of 2000 Å of boron phosphorus doped tetraethoxysilane (BPTEOS) hard mask over 5000 Å of polysilicon. The test pattern consists of isolated and nested lines and spaces ranging from 0.30 to 0.60 μm. A scanning electron microscope (Model No. DS-130C, made by International Scientific Instruments of Milpitas, CA) is used to evaluate etched profiles.

Bulk plasma properties (electron temperature and ion density) are determined by using a Langmuir probe (made by Plasma and Materials Technologies of Chatsworth, CA). The measurements were performed using argon discharges, due to the difficulty in maintaining probe cleanliness in HBr plasmas. More details concerning the probe data acquisition and analysis system are reported elsewhere,[10] where the same probe was used to characterize the ECR and the helicon plasma sources.

III. RESULTS AND DISCUSSION

A. Bulk plasma properties

The bulk plasma properties were spatially determined under the same range of conditions used to determine the etching rate and profile control. Langmuir probe measurements of electron temperature and electron density were performed as a function of the process parameters: source power, reac-

TABLE I. Etching conditions for response surface statistical experiment, mean etching rates, and etching rate uniformities obtained with the helical resonator from three techniques: (1) Nanometrics 5-point, (2) FWI 5-point, and (3) FWI 49-point measurements.

	Experimental conditions					Etching rate (Å/min)			Uniformity (±%)		
Run no.	Statist. position *XYZ*	Pressure (mTorr) *X*	Power (W) *Y*	rf-bias (W) *Z*	Etching time (s)	5-point Nanometrics	5-point FWI	49-point FWI	5-point Nanometrics	5-point FWI	49-point FWI
1	002	2.0	3000	75	90	3731	3239	3239	0.3	0.9	0.9
2	010	2.0	4000	25	120	2605	2302	2302	2.3	1.3	1.3
3	022	2.0	5000	75	75	3958	3392	3305	1.9	2.6	3.5
4	020	2.0	5000	25	120	2606	2278	2244	2.0	3.6	4.6
5	202	5.0	3000	75	90	4208	3762	3773	0.7	1.1	1.6
6	112	3.5	4000	75	90	4083	3661	3639	0.9	1.1	1.7
7	220	5.0	5000	25	120	3169	2834	2801	1.7	2.2	3.7
8	021	2.0	5000	50	120	3100	2808	2761	2.0	2.9	3.7
9[a]	022	2.0	5000	75	75	3862	3255	3183	2.7	3.8	4.5
10	101	3.5	3000	50	120	[b]	3178	3178	[b]	0.7	1.3
11	000	2.0	3000	25	120	2556	2178	2187	0.7	0.0	1.0
12	111	3.5	4000	50	90	3678	3228	3203	1.0	1.3	1.9
13	211	5.0	4000	50	90	4047	3596	3608	0.8	0.6	1.2
14[a]	202	5.0	3000	75	90	4239	3843	3820	0.7	1.1	1.6
15[a]	220	5.0	5000	25	120	3118	2873	2785	1.8	2.1	3.7
16	122	3.5	5000	75	75	4154	3586	3499	2.1	4.0	4.1
17	222	5.0	5000	75	75	4560	4023	3916	2.3	3.1	3.7
18[a]	000	2.0	3000	25	120	2580	2228	2231	0.5	0.9	1.9
19	200	5.0	3000	25	120	2708	2808	2808	0.4	0.7	1.5
20[a]	101	3.5	3000	50	90	3686	3178	3178	0.2	0.7	0.7

[a]Repeats.
[b]No data available due to clearing of polysilicon.

tor pressure, and rf-bias. Based on these results, the plasma source uniformity was determined. Figure 2 shows the radial dependence of the electron temperature (eV) as a function of source power, reactor pressure, and rf-bias power. The electron temperature is approximately 6.3 eV at a distance of 2.5 cm above the wafer surface. The electron temperature does not vary significantly with applied rf-bias power or reactor pressure. On the other hand, the ion density is strongly dependent on applied source power (Fig. 3). At the center of the wafer, the ion density increases from 2.5×10^{11} to 6×10^{11} cm^{-3} as the source power is increased from 3000 to 5000 W. The ion density varies from 2×10^{11} to 4×10^{11} cm^{-3} as the reactor pressure is increased from 1 to 5 mTorr. Unlike the applied source power and reactor pressure, rf-bias power only weakly affects ion density.

B. Etching rates and uniformities

The statistically designed experiment was based on the systematic variation of three main variables: reactor pressure (2–5 mTorr), plasma source power (3000–5000 W), and rf-bias power (25–75 W). Figure 4 shows the measured etching rates obtained from 5-point Nanometrics measurements (A) as a function of applied source power and applied rf-bias power at 3.5 mTorr, and (B) as a function of reactor pressure and applied rf-bias power at 4000 W applied source power. The rf-bias appears as the parameter having a stronger effect on the measured etching rate than the source power and reactor pressure. Based on the examined experimental conditions, etching rates in the range of 2600–3000 Å/min can be achieved at low reactor pressures, moderate applied rf-bias powers, and high source power.

Figure 5 displays the etching rate uniformities obtained from 5-point Nanometrics measurements (A) as a function of applied source power and applied rf-bias power at 3.5 mTorr, and (B) as a function of reactor pressure and applied rf-bias power at 4000 W applied source power. The source power shows larger effects on etching rate uniformity, as compared to the applied rf-bias power and reactor pressure, in our experimental conditions. The best uniformity is achieved at high applied source powers, low reactor pressures, and applied rf-bias powers between 50 and 75 W. At these conditions the uniformity is approximately ±0.2%.

Measurements of the etching rate and uniformity are also performed with FWI techniques (5 and 49 point). The corresponding results are summarized in Table I. The etching rate and etching rate uniformity values obtained from 5-point Nanometrics measurements are smaller than those obtained from 5-point and 49-point full wafer interferometry measurements. This difference may be caused by the assumed polysilicon refractive index between the two techniques. In addition, averaging on 49 points instead of 5 points explains the difference in the values obtained within the same (FWI) technique.

C. Profile control

Figure 6 shows the SEM micrographs of isolated lines (a1), nested lines (b1), and isolated trenches (c1) obtained

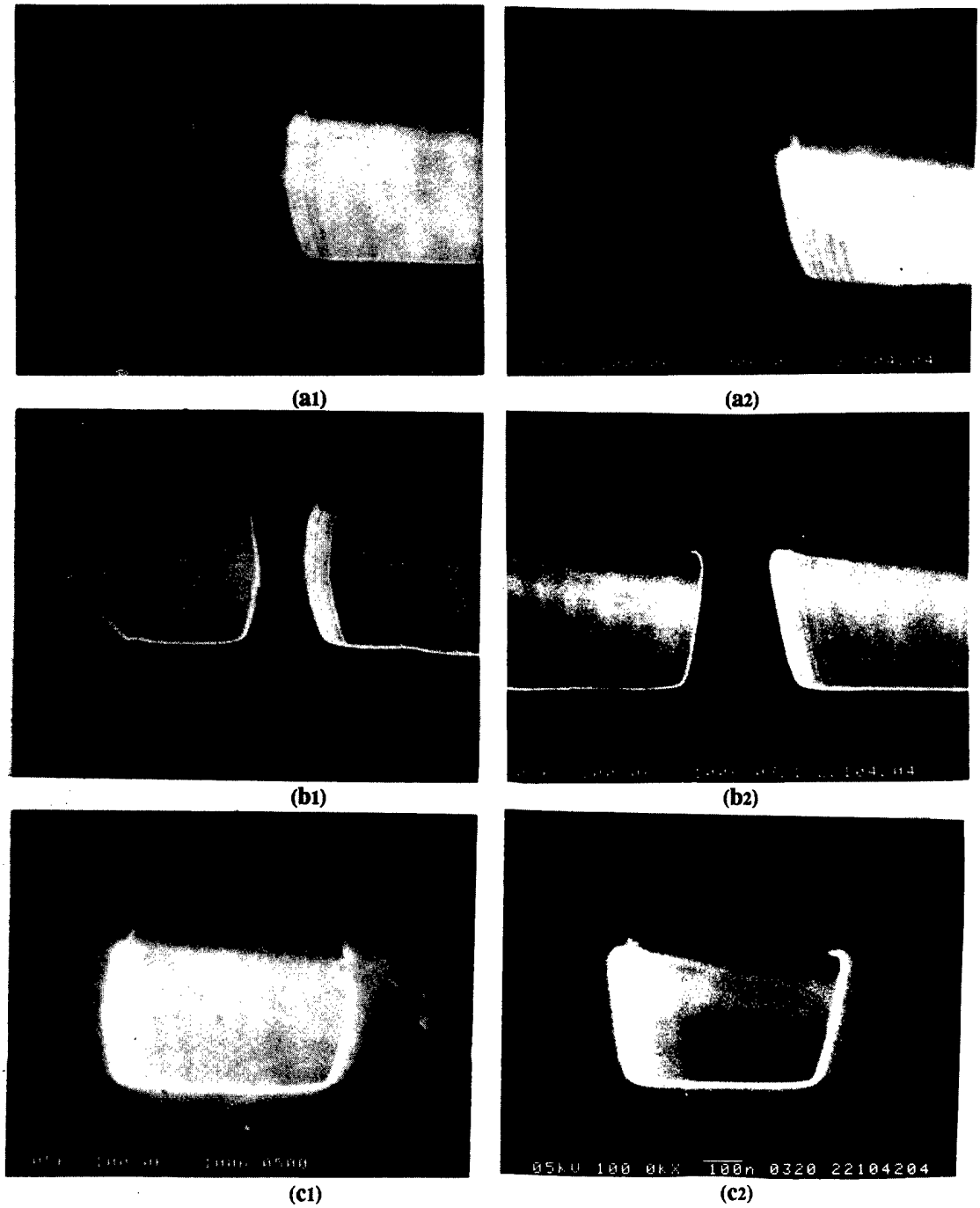


FIG. 6. SEM micrographs of isolated lines, nested lines, and isolated trenches obtained using pure HBr chemistry, at 2 mTorr, 50 W rf-bias, with the helical resonator (a1, b1, c1) at 3000 W source power, and with the helicon (a2, b2, c2) at 2500 W source power.

with the helical resonator at 2 mTorr, 50 W rf-bias, and 3000 W source power using 100 sccm of pure HBr. For comparison, we show in the same figure the corresponding SEM micrographs (a2, b2, c2) obtained with the helicon source, using the same chemistry, pressure, and rf-bias, but at 2500 W source power. For both sources, there is a slight undercutting of the mask but no microtrenching (i.e., the relatively faster etching rates observed near sidewalls). A decrease in the applied rf-bias power while holding both the reactor pressure and applied source power constant results in an increase in mask undercutting as indicated in Fig. 7. As the rf-bias decreases, two factors may contribute to the increase in undercutting. First, the decrease in rf-bias decreases the ion energy normal to the etching surface and therefore decreases the ratio of vertical to lateral etching rates. Second, the decrease in ion energy normal to the surface decreases the amount of oxide material sputtered from the mask. Oxide material deposited in the feature may act as a passivation layer. The sputtered oxide would be quickly removed from the bottom of the feature where significant ion bombardment

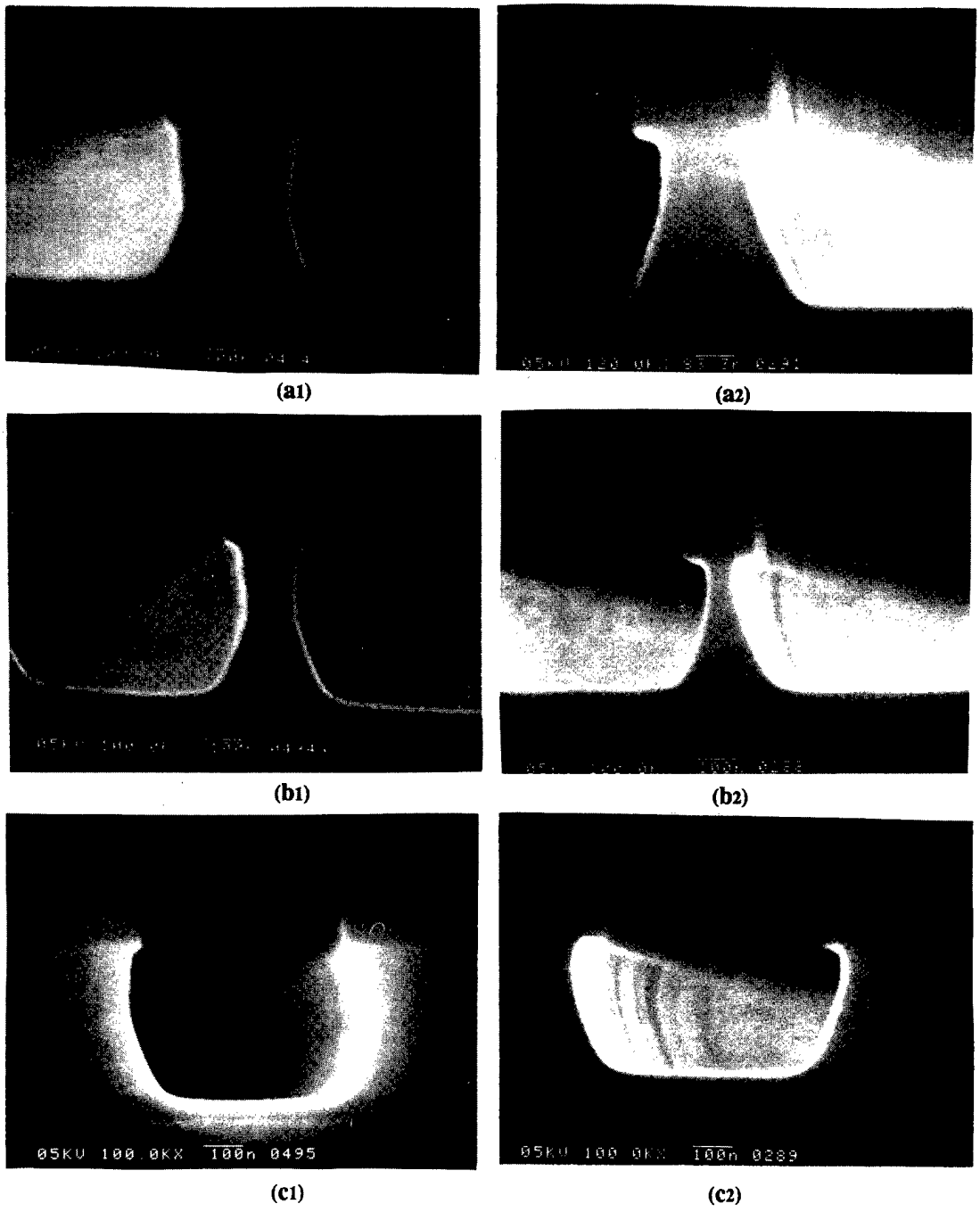


FIG. 7. SEM micrographs of isolated lines, nested lines, and isolated trenches obtained using pure HBr chemistry, at 2 mTorr, 25 W rf-bias, with the helical resonator (a1, b1, c1) at 3000 W source power, and with the helicon (a2, b2, c2) at 2500 W source power.

is present. However, oxide material deposited on the sidewall of the feature would greatly reduce the lateral etching rate. At lower applied rf-bias powers, the oxide material is no longer available to passivate the feature sidewalls.

An increase in reactor pressure from 2 to 5 mTorr at fixed rf-bias and source powers results in a very small increase in mask undercutting, for both the helicon and helical resonator sources (Fig. 8). The increase in mask undercutting with an increase in reactor pressure may be due to an increase in the reactive neutral concentration, i.e., a decrease in the ion-to-neutral ratio.

An optimized recipe for obtaining vertical profiles without mask undercutting was developed. The importance of acceptably high etching rates and high etching rate uniformity is maintained since a recipe resulting in vertical profiles without these attributes is clearly not useful. As discussed before, mask undercutting is most likely a result of the high lateral etching rate by free-radical species produced in the

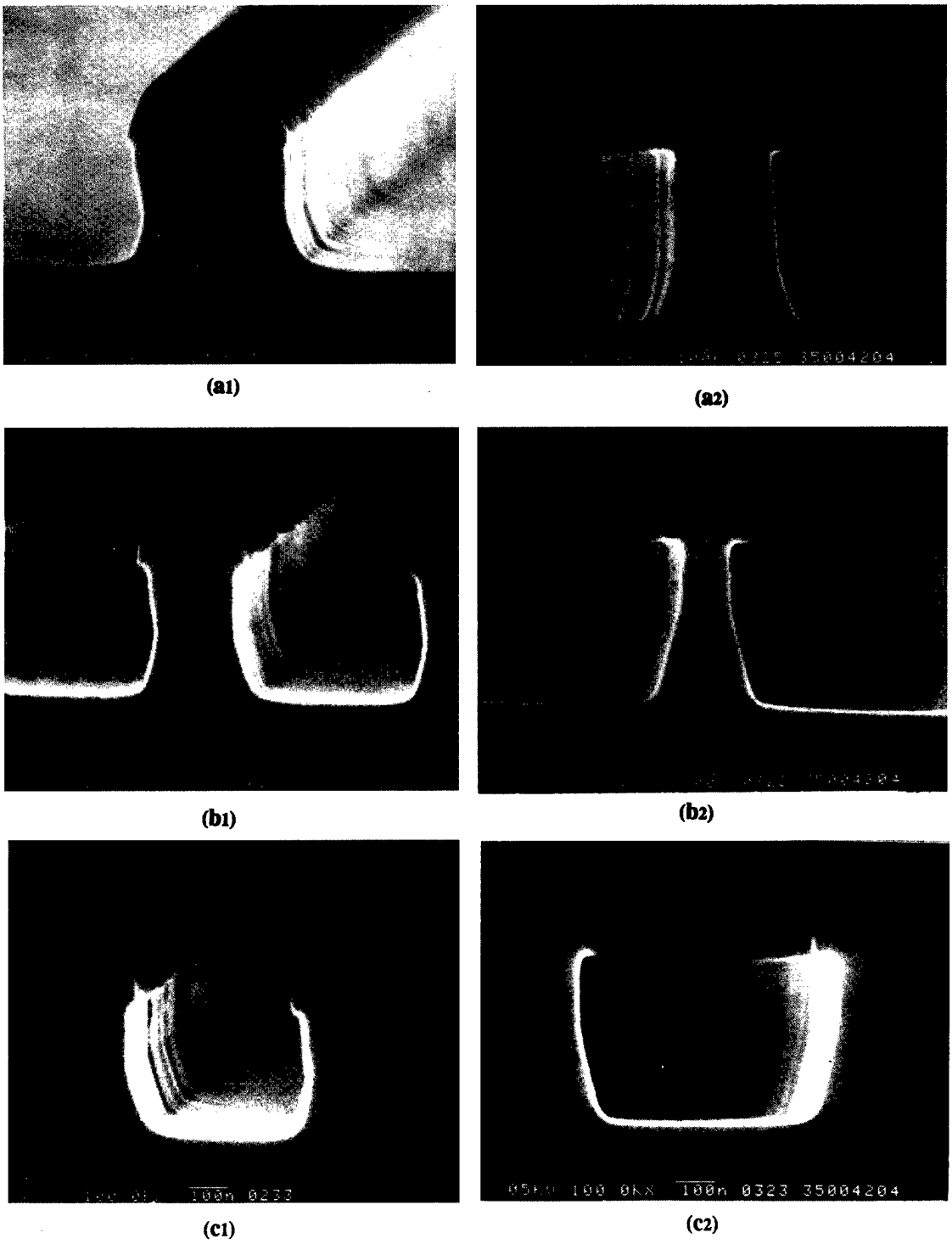


FIG. 8. SEM micrographs of isolated lines, nested lines, and isolated trenches obtained using pure HBr chemistry, at 5 mTorr, 50 W rf-bias, with the helical resonator (a1, b1, c1) at 3000 W source power, and with the helicon (a2, b2, c2) at 2500 W source power.

plasma. The presence of an oxide on the feature sidewall, either due to oxide mask sputtering or addition of oxygen to the 100 sccm HBr feed gas, decreases the lateral etching rate. To increase the oxide passivation of the sidewall in the optimum etching recipe, a small amount of oxygen was added to the feed gas. The optimum oxygen chemistry was 25 sccm of an O_2–He (20%–80%) mixture. Furthermore, to decrease the intrinsic etching rate of free-radical species, the wafer temperature was lowered to 0 °C. Figure 9 shows the vertical profiles obtained for the wafers etched in the helical resonator (a1, b1, c1) and the helicon (a2, b2, c2) plasma sources, using the optimized recipe (100 sccm HBr and 25 sccm O_2–He, 2 mTorr, 50 W rf-bias, 3000 W source power for the helical resonator and 2500 W for the helicon).

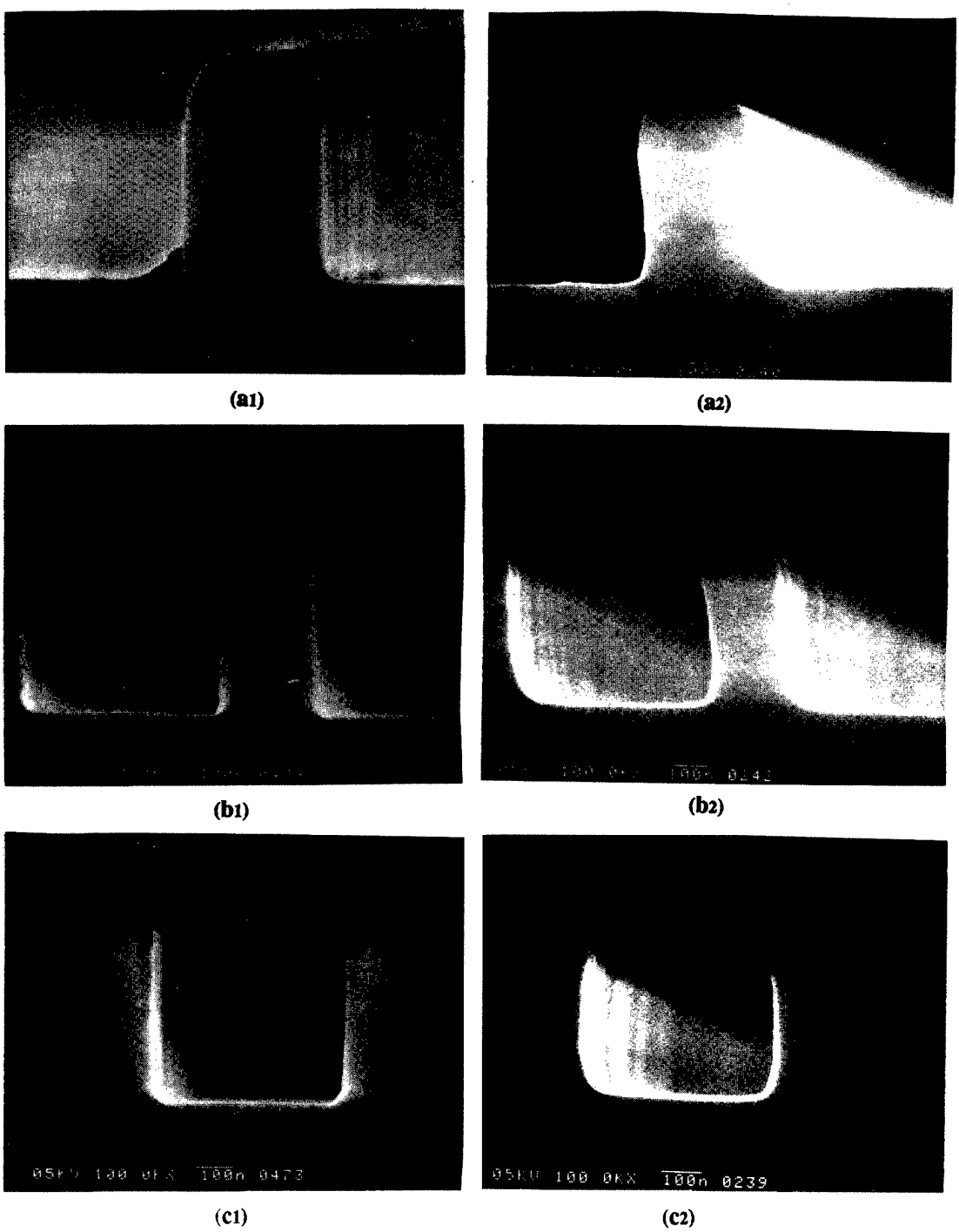


FIG. 9. SEM micrographs of isolated lines, nested lines, and isolated trenches obtained using the optimized recipe, a mixture of 100 sccm HBr and 25 sccm O_2–He, at 2 mTorr, 50 W rf-bias, with the helical resonator (a1, b1, c1) at 3000 W source power, and with the helicon (a2, b2, c2) at 2500 W source power.

IV. COMPARISON WITH OTHER SOURCES AND CONCLUSION

We have examined the etching performance of 0.35 μm polysilicon features in a helical resonator plasma source. We made similar studies for the helicon and electron cyclotron resonance sources previously.[9–12] In this section, we summarize the values of etching rates, uniformities, and bulk plasma parameters, in order to quantify the advantages and disadvantages among the three plasma sources.

(1) The three sources produce sufficient ion and reactive neutral densities to give etching rates in the 2000–3700 Å/min range. Also they provide high uniformity (less than $\pm 5\%$ across the 125-mm-diam wafers), as indicated in Table II. The helical resonator offers a higher etching rate and better uniformity, probably due to higher ion density induced by high excitation power and large plasma generation volume.

(2) The behavior of etched profiles is similar for all three sources as a function of rf-bias and reactor pressure. In all

TABLE II. Bulk plasma properties, mean etching rate, and etching rate uniformity for the electron cyclotron resonance (ECR), helicon (H), and helical resonator (HR) sources.

	Optimum etching rates and uniformities					Bulk plasma parameters		Plasma unif. (±%)		
Plasma sources	Source power (W)	Press. (mTorr)	rf bias (W)	Etch rate (Å/min)	Unif. (±%)	n_e (10^{11} cm^{-3})	T_e (eV)	From n_e	From T_e	From etching
ECR	1370	2.8	60	2300	1.1	0.5–3	5–9	3.6	4.1	4.1
H	2500	2.0	57	2400	2.6	2–5	5–7	2.7	3.3	3.3
HR	3000	3.5	53	3700	<1.0	2–6	5–7	<1.0	<1.0	<1.0

sources, vertical sidewalls are attained by adding oxygen to the feed gas and lowering the wafer platen temperature to 0 °C. The addition of oxygen results in an increase in passivation of feature sidewalls during the etching process, greatly decreasing mask undercutting. The decrease in wafer temperature further reduces mask undercutting by decreasing the lateral etching rate due to neutral species. The optimized profiles are the same for almost the same process recipe. Only the applied source power is unique for each source in the optimized recipes: 1000 W for the ECR, 2500 W for the helicon, and 3000 W for the helical resonator.

(3) Bulk plasma properties are similar for the polysilicon etching process window of 2–5 mTorr and 25–75 W rf-bias. As indicated in Table II, the ECR plasma density is slightly lower, consistent with observed lower etching rates. The best bulk plasma uniformity is obtained with the helical resonator source, due to a large plasma generation volume. However, the four solenoids in the helicon could in principle provide more flexibility in optimizing uniformity for different operating conditions; however, for these studies, no explicit optimization of the magnetic field in the helicon was performed.

(4) The similarity in the performance of the three examined sources indicates that additional factors other than etching rate, uniformity, and profile control may be key in deciding which source is best. These factors include plasma-induced damage, ease of source operation, and process reliability. Based on carrier lifetime, spectroscopic ellipsometry, and x-ray photoelectron spectroscopy measurements, these plasma sources produce minimal damage during silicon etching under optimum conditions.[12,14] Thus, manufacturability and ease of use will determine which source succeeds commercially. Ease of operation, plasma stability, and repeatable plasma ignition become relatively more important when deciding which source to use for a particular manufacturing application. For reproducibility and multistep process recipes automatic matching is necessary. The helicon and the helical resonator both have automatic matching and display high plasma stability which greatly simplifies the etching process optimization. The ECR source we used suffers from the lack of an automatic matching network and displays plasma instability—flickering and frequent mode changes—at high applied source power. This flickering is accompanied by large changes in reflected power. The lack of automatic matching for the ECR requires the use of dummy wafers to determine matching settings for each experimental condition, another disadvantage. Of course, one must consider that these advanced sources are continually being improved, and we expect technological problems such as tuning and mode stability in the ECR to be solved eventually.

In conclusion, we have examined three commercially available, low pressure, high density plasma sources using an etching platform consisting of identical reactor chambers to exclude platform dependent effects such as wafer platen design and reactor geometry. We studied the etching of 0.35 μm polysilicon gate structures for complementary metal–oxide–semiconductor devices. Etching rate, uniformity, profile control, and bulk plasma parameters were compared for the three sources. We found that the three plasma sources (ECR, Helicon, and Helical Resonator) can all meet sub-0.35-μm etching demands. The importance of manufacturability and ease of use will determine which source succeeds commercially.

ACKNOWLEDGMENTS

The authors thank V. M. Donnelly and R. A. Gottscho for helpful discussions.

[1]P. H. Singer, Semicond. Int. **5**, 50 (1993).
[2]M. A. Lieberman and R. A. Gottscho, *Physics of Thin Films: Advances in Research and Development*, edited by J. Vossen (Academic, New York, 1993).
[3]S. M. Gorbatkin, L. A. Berry, and J. Swyers, J. Vac. Sci. Technol. A **10**, 1295 (1992).
[4]D. X. Ma, T. A. Lin, and C. H. Chen, J. Vac. Sci. Technol. A **10**, 1217 (1992).
[5]P. K. Gadgil, D. Dane, and T. D. Mantei, J. Vac. Sci. Technol. A **10**, 1303 (1992).
[6]P. Mak, G. King, T. A. Grotjohn, and J. Asmussen, J. Vac. Sci. Technol. A **10**, 1281 (1992).
[7]A. J. Perry, D. Vender, and R. W. Boswell, J. Vac. Sci. Technol. B **9**, 310 (1991).
[8]D. E. Ibbotson, J. M. Cook, and C. P. Chang, in *Proceedings of the 12th Symposium on Dry Process* (IEE, Tokyo, 1990), p. 93.
[9]I. Tepermeister, N. Blayo, F. P. Klemens, D. E. Ibbotson, R. A. Gottscho, J. T. C. Lee, and H. H. Sawin, J. Vac. Sci. Technol. B **12**, 2310 (1994).
[10]I. Tepermeister, D. E. Ibbotson, J. T. C. Lee, and H. H. Sawin, J. Vac. Sci. Technol. B **12**, 2322 (1994).
[11]G. W. Gibson, H. H. Sawin, I. Tepermeister, D. E. Ibbotson, and J. T. C. Lee, J. Vac. Sci. Technol. B **12**, 2333 (1994).
[12]N. Blayo *et al.*, J. Vac. Sci. Technol. B **12**, 1340 (1994).
[13]T. J. Dalton, W. T. Conner, and H. H. Sawin, J. Electrochem. Soc. **141**, 1893 (1994).
[14]N. Layadi, V. M. Donnelly, and J. T. C. Lee (unpublished).

Surface morphologies for Br-etched Si(100)-2×1: Kinetics of pit growth and step retreat

F. J. Williams and C. M. Aldao
Institute of Materials Science and Technology (INTEMA), Universidad Nacional de Mar del Plata-CONICET, Juan B. Justo 4302, 7600 Mar del Plata, Argentina

J. H. Weaver
Department of Materials Science and Chemical Engineering, University of Minnesota, Minneapolis, Minnesota 55455

(Received 7 March 1996; accepted 24 May 1996)

Temperature-dependent surface morphologies of Si(100) resulting from Br etching at 700 and 800 K have been studied using scanning tunneling microscopy. Quantitative analysis of linear pit distributions allows us to interpret the mechanisms involved in surface pattern formation. We show that linear chains growing at 800 K on 450-Å-wide terraces are in quasi-steady-state conditions in that their lengths follow the most probable Flory–Schulz distribution. Results obtained under equivalent conditions at 700 K, however, demonstrate that this distribution is not reached. The difference reflects the rate of step retreat relative to the rate of pit growth. At 700 K, pits that develop on terraces are erased before their length distributions reach quasi-steady-state values even though the overall morphology does not change. The surface morphologies that are reached at long times can then be understood by considering the temporal evolution of linear chains on terraces where quasi-steady-state conditions are not fulfilled. © *1996 American Vacuum Society.*

I. INTRODUCTION

Halogen etching is widely used in the dry processing of semiconductors. It is then important to understand the mechanisms of etching so that they can be better controlled.[1,2] Scanning tunneling microscopy (STM) can be used to investigate adsorption and etching phenomena with atomic resolution, yielding insight into surface structures and the kinetics of etching. Chander *et al.*[3,4] have shown that etch pit growth on Si(100)-2×1 can be quantified, that the size distribution and shape anisotropy of the pits can be deduced, and that differences in activation energies for Si removal along and perpendicular to the dimer row direction can be determined. From analysis of STM images of etched surfaces, it has been demonstrated that the lengths of Cl-induced linear etch pits on Si(100)-2×1 follow the most probable Flory–Schultz size distribution[5] when quasi-steady-state conditions are met[6] and under conditions where the probability of linear growth remains constant irrespective of pit length.

For a very wide terrace, it is expected that the longer the time of exposure to an etchant, the greater the damage on the surface. However, it is observed in our experiments that terraces reach a steady-state etched morphology and that this morphology is temperature dependent. In this article, we report analysis of STM results for bromine etching of Si(100)-2×1 at 700 and 800 K. Images were obtained under conditions of halogen flux and fluence such that surface morphologies would reach steady state, i.e., they would not change with further exposure.[6,7] The resulting distributions of terrace pits were, however, quite different. Through analysis of the pit size distributions at 700 K, we demonstrate the consequences of having a finite dimension for terraces and the connection between step retreat and pit growth. Significantly, steps and their separation play a crucial role in surface layer etching and the resulting morphologies.

II. EXPERIMENT

The experiments were performed in an ultrahigh vacuum chamber equipped with a Park Scientific Instruments STM. The base pressure of the system was 5×10^{-11} Torr. Silicon wafers oriented within 0.2°–0.5° of (100) and miscut toward [110] were rinsed in ethanol prior to introduction into vacuum. The wafers were degassed at 600 °C for several hours and then flashed to 1200 °C for 1–2 min. This produces clean, well-ordered Si(100)-2×1.[8] Sample temperatures were monitored with an optical pyrometer. Electrochemically etched tungsten tips were cleaned using electron bombardment. Scan dimensions were calibrated using the Si(100)-2×1 lattice and single height steps. All micrographs were acquired in a constant current mode at room temperature. An electrochemical cell was used to provide molecular bromine for dosing.[7] The samples were ~3 cm from the mesh electrode during exposure. During cell operation, the chamber pressure remained below 1×10^{-10} Torr. Stable cell currents of 10 μA were used, releasing $\sim3.1\times10^{13}$ Br_2 mol/s. Bromine exposures were done under conditions of constant flux for a predetermined amount of time (160 and 320 s) so that exposures are quoted in units of mA s. The samples were cooled rapidly to room temperature immediately upon terminating the flux.

III. RESULTS AND DISCUSSIONS

Figure 1 shows two occupied-states images of Si(100)-2×1 obtained after exposure to 1.6 mA s Br_2 (flux 10 μA, time 160 s). They correspond to etching at 700 and 800 K,

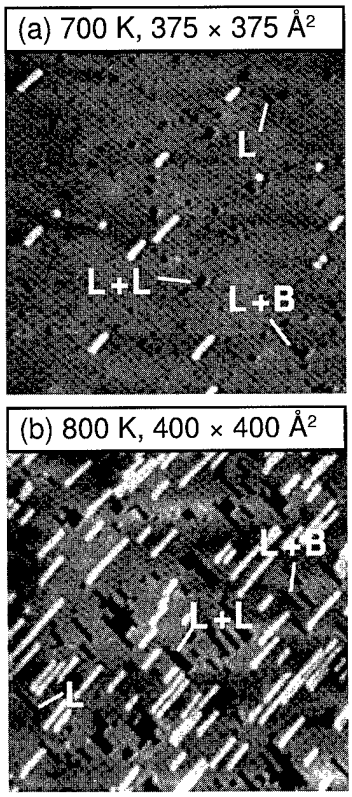


FIG. 1. Occupied-state STM image of Si(100)-2×1 after 1.6 mA s exposure to Br_2 at (a) 700 K and (b) 800 K, corresponding to 10^{16} atoms or about 16 monolayer equivalents of Br. Terrace etching produces one layer deep etch pits and Si regrowth islands. The pits can be characterized based on their shapes as being linear and one row wide, L, linear with a dimer vacancy on an adjacent row, $L+B$, and linear and two rows wide, $L+L$.

respectively. The 2×1 nature of the substrate is evident from dimer rows oriented along the image diagonals.[9] Bromine molecules adsorb dissociatively to produce Br atoms bonded to the dimer dangling bonds. While etching is minimal at low temperatures, Br can break dimer bonds at 700 and 800 K to form volatile species.[7] Etching involves Si removal from steps and from terraces. Si atoms along steps are more susceptible than their terrace counterparts since Si atoms are bound less tightly at steps than on pristine terraces and since steps act as accumulation sites for Br. Step etching then constitutes a "layer by layer" mechanism for Si removal. At the same time, steps also play an important role in the final terrace morphology since they erase pits that are growing sufficiently near the step.

Pits are initiated by removal of a single terrace dimer, creating a dimer vacancy. The desorption of a $SiBr_x$ unit is the limiting process since barriers for surface diffusion are much smaller than energies for desorption.[4] This dimer vacancy destabilizes adjacent dimers. Given sufficient time, the pit can grow laterally to produce features that are one atom layer deep. As a consequence of dimer breaking, some Si atoms are ejected onto the surface where they can form regrowth islands that are one layer high. Alternatively, a freed Si atom can be removed from the surface as a volatile species, contributing to the overall rate for dimer removal. In

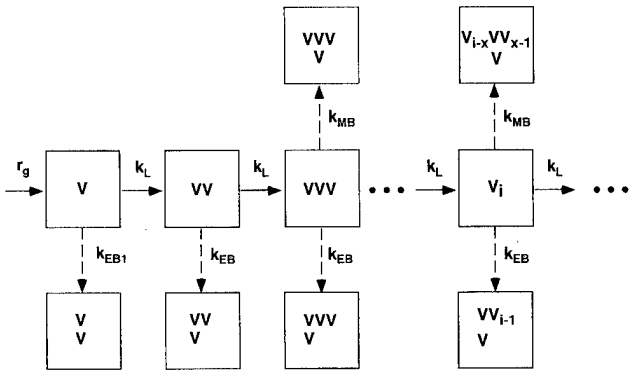


FIG. 2. Scheme depicting the growth of linear pits from the generation of single vacancies. Linear pits are generated by linear growth of shorter linear pits and are consumed by branching. Competition between different pathways produces the observed distribution.

principle, dimer vacancies can be mobile, as on the clean surface, but this mobility is not sufficient on halogen-rich surfaces to account for the observed morphologies.[7]

Figure 1 shows that the etch pits, whose formation is the main focus in this article, have different sizes and shapes. The smallest pits are the single dimer vacancies. Larger pits can be characterized as linear pits, **L**, that are one row wide, linear pits that have a dimer vacancy branch on an adjacent row, **L+B**, pits that are two rows wide, **L+L**, and so on. (They could be further characterized according to the location of the branch site.[10])

A. Kinetics of linear growth

Figure 2 depicts pit growth. A linear pit denoted VVV, for example, can grow by the removal of a dimer from either end of the pit (with a rate constant k_L). Alternatively, a linear pit can branch at either end (rate constant k_{EB}) or can branch at middle sites (rate constant k_{MB}). Note that a first neighbor dimer on a row adjacent to a single vacancy has an extra repulsive interaction due to the presence of an extra second neighbor compared to a dimer similar for longer linear pits. That is why the branching rate constant denoted as k_{EB1} is different from the general rate constant for branching at the end k_{EB} (in fact, $k_{EB1}>k_{EB}$).

After taking into account the statistical factors related to each event, the following equations can be written:

$$\frac{d[V_1]}{dt}=r_g-2(k_L+k_{EB1})[V_1], \tag{1}$$

$$\frac{d[V_i]}{dt}=2k_L[V_{i-1}]-2(k_L+2k_{EB}+(i-2)k_{MB})[V_i], \quad i=2,3,\ldots, \tag{2}$$

where $[V_i]$ is the surface density of linear vacancies and r_g is the generation rate of single vacancies. Since we are dealing with small amounts of dimer removal (a small fraction of a monolayer), r_g can be considered to be independent of

time. For simplicity, we consider that branching occurs only at pit ends. This has been shown to be the dominant branching process for linear pits of lengths $i<10$ (Ref. 10). These simplifications allow us to write analytical solutions to Eqs. (1) and (2), namely

$$[V_1]=[V_1]_{SS}\{1+(r-1)\exp[(f^{-1}-1)ct]\}, \tag{3}$$

$$[V_2]=p_L[V_1]_{SS}\{[1-\exp(-ct)]+(r-1)f \times[\exp[(f^{-1}-1)ct]-\exp(-ct)]\}, \tag{4}$$

$$[V_i]=p_L^{i-1}[V_1]_{SS}\left\{\begin{array}{c}1-\exp(-ct)\left[1+\sum_{k=1}^{i-2}[(r-1)f^{i-(k+1)}+1]\frac{(ct)^k}{k!}\right] \\ +(r-1)f^{i-1}(\exp[(f^{-1}-1)ct]-\exp(-ct))\end{array}\right\} i=3,4,\ldots, \tag{5}$$

where $[V_1]_{SS}$ is the density of single vacancies in quasi-steady state (term used here to denote the solution of the above equations for very long times). The term p_L represents the probability for linear growth, r takes into account the initial density of defects, and the two constants c and f are defined as

$$p_L=\frac{1}{1+\frac{2k_{EB}}{k_L}}, \tag{6}$$

$$r=\frac{[V_1]_{t=0}}{[V_1]_{SS}}, \tag{7}$$

$$c=2k_L+4k_{EB}, \tag{8}$$

$$f=\frac{k_L+2k_{EB}}{2k_{EB}-k_{EB1}}. \tag{9}$$

Since the surface morphology does not change with further exposure, the effects due to initial surface defects have disappeared as step etching has created fresh terraces. Thus, given sufficient time, step edges have moved to erase the original exposed surfaces, and the effects of possible initial defects will not appear in the mathematical expressions. If no defects are assumed on the freshly created terraces, we can consider the initial density of defects to be zero. Figure 3(a) shows the calculated evolution of linear vacancy densities as a function of time, ct, for $T=700$ K where f was determined using the interaction energy values from Ref. 10.

If the time required for the step to retreat across the width of the terrace is greater than the time needed to reach quasi-steady state, then Eq. (5) will reduce to $[V_i]_{SS}=[V_1]_{SS}p_L^{i-1}$ which is the most probable Flory–Schulz size distribution. In this case, the densities of linear pits reach their maximum values and a plot of size distribution establishes a straight line with slope equal to $\ln p_L$. On the other hand, if the time required for the step to retreat across the width of the terrace is lower than the time needed to reach quasi-steady state, then the linear pit densities do not reach their maximum values (according to the Flory–Schulz distribution). In this case, the slope in a plot of size distributions will present a steeper aspect, indicating that long pits are unlikely. Figure 3(b) shows the resulting size distribution as a function of $(i-1)$ for representative values of ct. For large enough ct, the distribution is a straight line. For smaller values of ct, it adopts the form of a smooth curve with a slight downward concavity. Accordingly, measurements that give size distributions can be used to assess the relative importance of step retreat.

B. Etching at 800 K

Figure 4(a) shows a plot of $\ln([V_i]/[V_1])$ vs $i-1$ for linear pits on Si(100)-2×1 after etching at 800 K. Linear regression

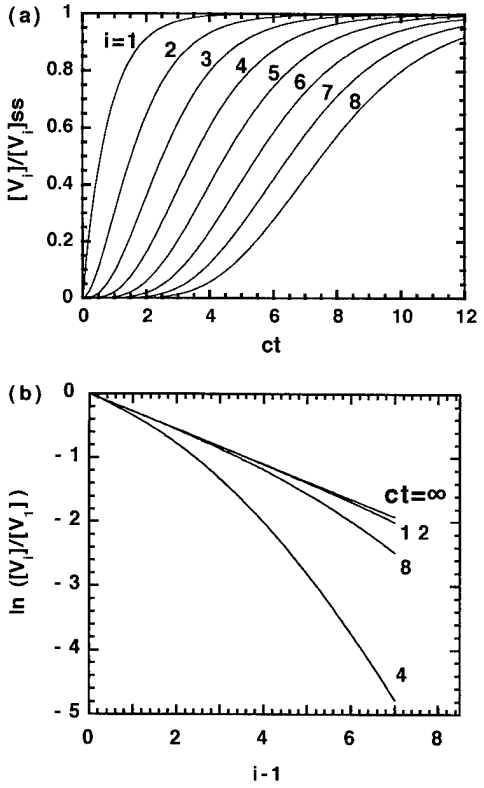


FIG. 3. (a) Plot showing the temporal evolution for linear pits of representative length i normalized to their final values. The term c is the sum of the rate constants for linear and branched growth, Eq. (8), and t is time. For insufficient values, $ct=4$, the linear pit distribution is far from the quasi-steady-state value, $ct=\infty$. (b) Plot of size distribution $\ln([V_i]/[V_1])$ vs $i-1$ for linear pits in terms of ct. By $ct=4$, only the areal distributions of very short pits are close to their final values. By $ct=12$, pits up to nine units in length are within a few percent of their final values. The distribution corresponding to $ct=\infty$ is the most probable Flory–Schulz size distribution for which quasi-steady-state conditions are met.

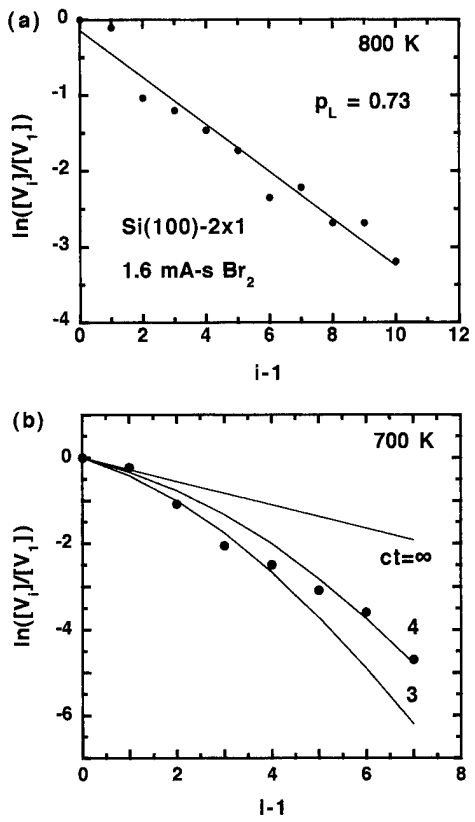


FIG. 4. (a) Plot of areal distribution of linear pits of size i, $[V_i]$, normalized to the areal density of a single vacancy, $[V_1]$, for Si(100)-2×1 after etching at 800 K with a Br_2 fluence of 1.6 mA s at a flux of 10 μA. The straight line indicates the fit obtained using the Flory–Schulz equation with the probability $p_L=0.73\pm0.05$. Linearity indicates that the probability of linear growth can be considered independent of pit length for $i<10$. (b) Plot of areal distribution of linear pits of size i after etching at 700 K [same fluence and flux as in Fig. 4(a)]. The curve for $ct=\infty$ shows the Flory–Schulz size distribution while those corresponding to $ct=3$ and 4 are distributions for which quasi-steady-state conditions are not fulfilled. The failure to reach quasi-steady state at 700 K reflects erasure of terrace pit distributions by retreating steps.

shows that most points fit a straight line. The slope under a 95% confidence interval gives $p_L=0.73\pm0.05$. Assuming, as in Refs. 3 and 4, that the rate constants can be related to activation energies via an Arrhenius relationship, then expressions of the form $k=\sigma\exp[-E/kT]$ can be written. The preexponentials should be equal since similar bond-breaking processes are involved in each etching event, namely the removal of a Si atom and the ejection onto the surface of the now-unpaired Si atom. Hence, $k_{EB}/k_L=\exp[-(E_{EB}-E_L)/kT]$. From the experimental value of p_L and Eq. (6), we find $E_{EB}-E_L=0.12\pm0.02$ eV. This analysis parallels that of Ref. 4 (etching of Si(100)-2×1 by Cl at 850 K) where the same difference in dimer removal energy, $E_{EB}-E_L$, was deduced. The good fit obtained with a straight line indicates that branching occurs preferentially at the ends of the pits, implying that k_{MB} is small, and that p_L is not greatly affected for $i<10$.

C. Etching at 700 K

In Fig. 4(b) we plot $\ln([V_i]/[V_1])$ vs $i-1$ for linear pits after etching at 700 K. The etching conditions, halogen flux, and exposure time were the same as those for etching at 800 K. Doubling the fluence to 3.2 mA s did not alter the pit distribution. It is apparent that etch pits and terrace chains are, in general, much shorter than those observed after etching at 800 K.

Analysis similar to that carried out above predicts a value of p_L of about 0.76 for etching experiments at 700 K. This assumes quasi-steady-state conditions, $ct=\infty$. As seen in Fig. 4(b), however, the experimental results indicate that the linear pit distribution is far from the quasi-steady-state condition. Indeed, a better fit is produced with ct of about 4. Thus, a given terrace is exposed to etching conditions for insufficient ct for it to reach quasi-steady state, the linear pits do not reach their final lengths, and we find behavior corresponding to $ct<\infty$.

We stress that the surface etched at 700 K is in a steady or unchanging-morphology state in the sense that exposures are sufficiently long that further etching does not alter the pit distributions. Although a particular feature will change its appearance, the overall morphologies are indistinguishable after reaching this steady state. However, the condition for (mathematical) quasi-steady state for the areal densities of linear pits cannot always be applied because many pits do not reach their final values.

IV. CONCLUSIONS

Step atoms are more susceptible to removal than their terrace counterparts. At low enough temperature, we expect that the ratio in etching rates for step and terrace atoms would be so high that step etching would be the only relevant mechanism. An increase in temperature has a more profound change for the rate of the higher energy process than for the lower energy process (as well as for the creation of the initial dimer vacancies). Accordingly, the ratio in their rates diminishes with temperature: the higher the temperature, the greater the relative amount of terrace etching. The time available for terrace etching between step erasures is then insufficient at 700 K, and linear pits do not reach their quasi-steady-state values. At 800 K, the largest analyzed pits can be considered to be in the quasi-steady-state condition. At 900 K, the rate of pit formation and growth is so high that pits grow into one another to form large structures that function like steps. Analysis of this situation is more complex.

Measurements at a fixed temperature as a function of terrace width should show the interplay of steps and terraces quite clearly. Increasing the width of a terrace etched at 700 K, for example, should give the pits a greater ct between erasure events and their population should increase toward quasi-steady state. For very narrow terraces, step retreat should always overwhelm pit formation. Indeed, studies of Si(100)-2×1 with a 4° miscut show very few terrace pits, although in that case the final state morphology is influenced by step–step interactions.[11]

ACKNOWLEDGMENTS

This work was supported by the U.S. Office of Naval Research and the National Council for Scientific and Technical Research (CONICET) of Argentina. The authors are pleased to acknowledge stimulating discussions with R. J. Pechman.

[1]R. J. Schutz, in *VLSI Technology*, edited by S. M. Sze (McGraw-Hill, New York, 1988), p. 184.
[2]H. F. Winters and J. W. Coburn, Surf. Sci. Rep. **14**, 161 (1992).
[3]M. Chander, D. A. Goetsch, C. M. Aldao, and J. H. Weaver, Phys. Rev. Lett. **74**, 2014 (1995).
[4]M. Chander, D. A. Goetsch, C. M. Aldao, and J. H. Weaver, Phys. Rev. B **52**, 8288 (1995).
[5]*Polymer Synthesis*, edited by P. Rempp and E. W. Merrill (Hüthig and Wepf, New York, 1986), Chaps. 2 and 3; P. J. Flory, Chem. Rev. **39**, 137 (1946).
[6]A general process can involve many subprocesses, each having a different time constant. Those that have reached their final state values are said to be in quasi-steady state since others may not have. Here, quasi-steady state implies that pits of length less than about 10 units have reached their final values; those that are extremely long, say 30 units, have not. As a step retreats across a terrace, it erases terrace pits. Hence, steps of long length never have sufficient time to reach final values. Processes that describe step retreat and pit growth are thermally activated with different energies so that the ratio in their rates will be dependent on temperature. The effect of pit erasure is then dependent on temperature and step width.
[7]D. Rioux, R. J. Pechman, M. Chander, and J. H. Weaver, Phys. Rev. B **50**, 4430 (1994).
[8]B. S. Swartzentruber, Y. W. Mo, M. B. Webb, and M. G. Lagally, J. Vac. Sci. Technol. A **7**, 2901 (1989).
[9]R. J. Hamers, R. M. Tromp, and J. E. Demuth, Phys. Rev. B **34**, 5343 (1986); R. J. Hamers, U. K. Kohler, and J. E. Demuth, J. Vac. Sci. Technol. A **8**, 195 (1990).
[10]J. R. Sánchez, C. M. Aldao, and J. H. Weaver, J. Vac. Sci. Technol. B **13**, 2230 (1995).
[11]Y. Gong, D. W. Owens, and J. H. Weaver, Phys. Rev. B (to be published).

Deep-etch silicon millimeter waveguide structure for electron acceleration*

T. L. Willke[a] and A. D. Feinerman[b]
Department of Electrical Engineering and Computer Science, The University of Illinois at Chicago, Chicago, Illinois 60607

(Received 4 January 1996; accepted 4 June 1996)

An ultrahigh vacuum compatible planar corrugated millimeter mm-waveguide structure (410-μm-deep) possessing bi-fold symmetry and a precision beam aperture (800 μm) has been fabricated using silicon processing technology, modeled with numerical analysis software, geometrically characterized, and compared to a similar waveguide fabricated using deep x-ray lithography (DXL) techniques. The waveguide was fabricated to operate at 60 GHz (λ=5 mm) with fields suitable for $2\pi/3$ phase advance operation. Multichip alignment technology was used to provide a semiclosed conducting surface with aperture-coupled periodic resonator cavities. A pair of Si/Pyrex composite metallized substrates patterned with corrugated geometries have been vertically stacked with 980-μm-diam Pyrex capillaries. Geometrical analysis of the muffin-tin waveguide was divided into two classifications: substrate feature error and die-to-die orientation error. Both types of error were characterized with the following results: feature accuracy was maintained to 0.1%–1.0% tolerances in all directions (5 μm or less in most cases) and die-to-die aperture distance agreed to within ~3% of theoretical calculation. Methods of improving these geometrical tolerances are suggested and critical issues are addressed. Electromagnetic testing of the mm waveguide has been investigated and a bead was fabricated for use in a bead-pull perturbation measurement of acceleration properties. The concluding section compares deep-etch silicon and DXL approaches for the fabrication of the "micro-linac." It is concluded that through further refinement of thermal and conductive properties that the silicon waveguide is a viable method of constructing a micro-linac mm waveguide, requiring less fabrication complexity, processing time, and capital equipment investment than DXL. © *1996 American Vacuum Society.*

I. INTRODUCTION

Electron acceleration at frequencies in the millimeter wavelength range requires conductive structures on the scale of several hundred microns with precise geometries. Gaussian random error analysis carried out by researchers at Argonne National Laboratory (ANL) indicates that approximately 10 μm(~0.28% error) or better of transverse dimensional accuracy is required to limit normalized electron beam emittance growth to <5%.[1,2] A general dimensional tolerance of 5 μm feature accuracy is necessary to maintain the structure's fabrication to within an order of the widely used 0.1% dimension accuracy rule applied to "macro" linear accelerators.

Henke designed a planarized double-sided muffin-tin geometry for use in mm-wave particle accelerator applications typically implemented with capacitive-plate loaded cylindrical waveguides.[3] Researchers at ANL have investigated 50 MeV 120 and 60 GHz linear acceleration systems utilizing Henke's planar geometry that are implementable with microfabrication techniques.[4] A 60 GHz structure has been developed at the University of Illinois' Microfabrication Applications Laboratory with adaptations of standard ultraviolet (UV) silicon fabrication methods (Fig. 1, Table I).

Each waveguide die consists of anodically bonded silicon and Pyrex wafers.[5] The Si ($\langle 110\rangle$ wafer orientation) is anisotropically etched to form the waveguide pattern. Concurrently, capillary seatings 600 μm wide are patterned to hold capillaries for stacking.[6] These seatings will provide a precision 800 μm iris opening between die when used in conjunction with 1-mm-diam cylindrical spacers. The $\langle 110\rangle$ silicon etch terminates on (111) planes to form ~70.54° parallelograms when viewed normal to the wafer surface.[7] An array of parallelograms is etched to form the basis for cavity structures separated by aperture flutes [Fig. 2(a)]. The nonorthogonal short sides of the parallelograms are removed with a wafer dicing saw, leaving central aperture flutes with a calibrated transverse width [Fig. 2(b)]. A wafer is processed in parallel to provide bulk material for cavity sidewalls. Two silicon strips with near-optically smooth side edges are placed adjacent to the aperture structures and anodically bonded in place establishing a corrugated pattern [Fig. 2(c)]. A uniform aluminum layer is applied to a minimal thickness of three rf skin depths (~1.2 μm at 60 GHz) for improved electrical conduction. Placing precision-drawn ($\pm$1-μm-diam accuracy) outer-diameter (o.d.) Pyrex capillaries into the lower die's anti-parallel seatings and forming a stack with a mirror image version of the die accurately controls die-to-die muffin-tin orientation.

The objective of this work is to demonstrate a bulk silicon micromachine based alternative to deep x-ray lithography (DXL) useful for fabricating mm-wave muffin-tin waveguide structures. The fabrication method is presented as well as an outline of an alternative x-ray fabrication process. Geometrical characterization data is examined for the 60 GHz silicon

*Published without authors corrections.

[a]Electronic mail: willke@cæ.wisc.edu
[b]Electronic mail: feinerma@eecs.uic.edu

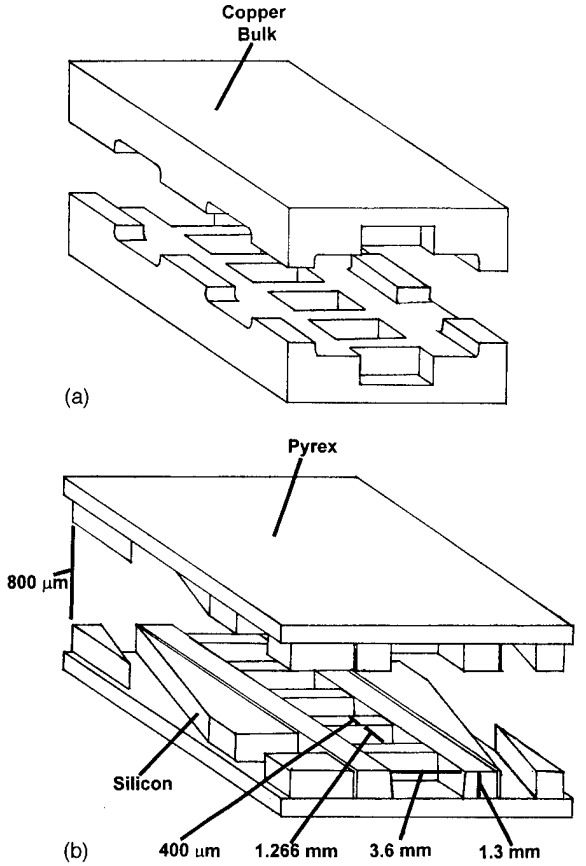


FIG. 1. Muffin-tin waveguide with beam aperture and interleaved side openings. (a) An ideal machined geometry; (b) the ⟨110⟩ silicon-fabricated equivalent.

muffin-tin structures and DXL-fabricated 240 GHz muffin-tin structures to determine each fabrication method's merits. Anisotropic etching of single-crystal silicon does permit bulk micromachining of the exact muffin-tin geometry and so the electrical impact of using an approximate geometry in the waveguide design is examined with the aid of numerical analysis.

II. SILICON WAVEGUIDE FABRICATION

Two identical but mirrored versions of the waveguide pattern were used for the top and bottom die. A 1.85 μm SiO_2 film was grown on double-polished 3-in.-diam (410-μm-thick) ⟨110⟩ silicon wafers. The design goal calls for 1.3-mm-deep structures but 410-μm-deep patterns were fabricated for proof-of-principle. The front-side SiO_2 layer was patterned with waveguide and capillary groove features while the back-side SiO_2 was completely removed in preparation for high-temperature anodic bonding (HTB) to Pyrex.

TABLE I. Geometry specifications for the muffin-tin waveguides (see also Fig. 1).

Specification	60 GHz silicon waveguide	240 GHz DXL waveguide
Free-space aperture gap	800 μm	300 μm
Muffin-tin cavity depth	410 μm (fab. target)	160 μm (fab. target)
Muffin-tin cavity width	3.6 mm	1.8 mm
Muffin-tin cavity length	1.266 mm	300 μm
Iris wall thickness	400 μm	100 μm

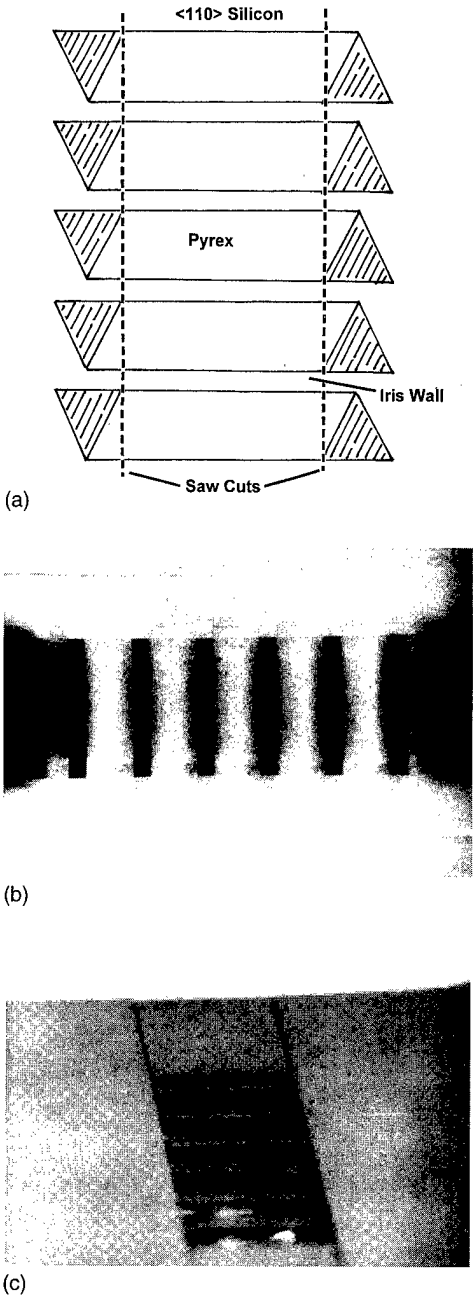


FIG. 2. Fabrication of right-angled muffin-tin corrugation pattern from ⟨110⟩ silicon. (a) A dicing saw determines the width of the vertically etched walls. The vertical walls become the irises in the muffin tin cavity and are 1.266 mm wide, 0.40 mm thick, and 0.41 mm high. (b) Top view of the iris walls bonded to a Pyrex wafer. The silicon is black and the Pyrex is gray. (c) Side strips have been bonded next to the iris walls to complete the muffin-tin cavities.

Figure 3 illustrates the setup for the HTB process. A double-polished 3-in.-diam 18 mil Pyrex wafer was placed on an unpolished Si wafer. The unpolished Si surface prevented accidental bonding between these two wafers while

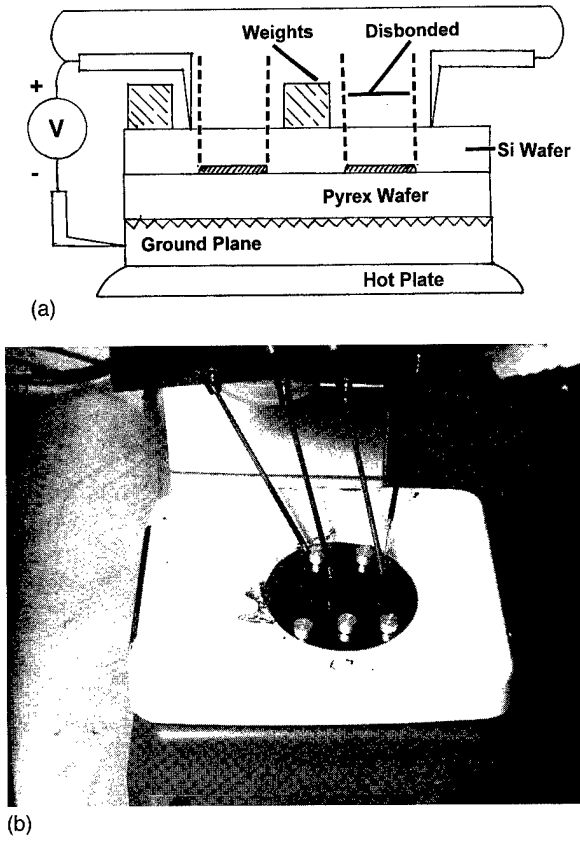


FIG. 3. High-temperature anodic bonding setup. A schematic of the bonding circuit is shown in (a) with a photograph of the setup in use (b).

providing a low resistance current return path. The backside of the patterned silicon wafer was pressed (with small weights) into intimate contact with the Pyrex wafer and bonded at ~50 °C by the application of a 500 V potential.

The anodic bond is attributed to the mobility of sodium ions in the glass. At an elevated temperature the ions in the glass drift to the Si ground plane, leaving behind a negatively charged depletion layer at the Si/Pyrex interface. The surfaces are pulled into tight contact and a permanent chemical bond takes place.[5] Intimate interface contact and similar thermal expansion coefficients of 3.25×10^{-6}/°C (0–300 °C) for Pyrex and 2.6×10^{-6}/°C for Si minimize shear stress during cool-down and postbond surface deformation that can introduce error into the structure's geometry.

After cool-down the bond was tested for strength (~350 psi) and the presence of defects. Following a brief BOE etch to remove unwanted oxide, a 40% by weight potassium hydroxide (KOH) (balance made of H_2O with isopropyl alcohol) 85 SDC silicon etch was performed with megasonic agitation. The etch was continued until the (111) plane edges terminated on the Pyrex plane and surface profilometry revealed no remaining Si in the recesses.

As illustrated in Fig. 4, the proper iris width is obtained by guiding a dicing saw in-line with two etched fiducial markings separated approximately by the wafer's diameter. The parallelogram fiducials' left and right obtuse angles define the iris width to photolithographic accuracy. A Micro

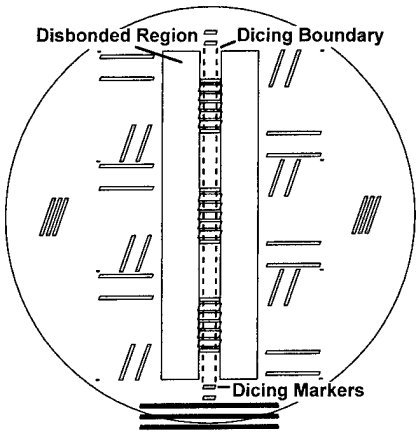


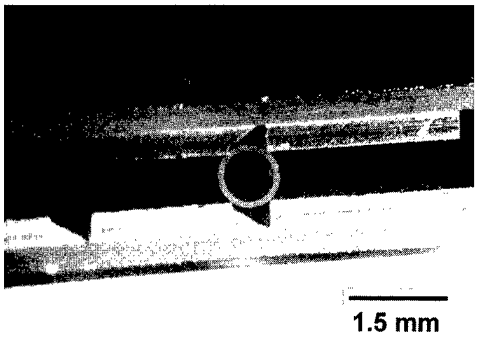
FIG. 4. AutoCad® mask drawing used to generate one of two waveguide pattern photomasks. The silicon is etched in the disbonded region to prevent it from bonding to the Pyrex wafer during the high temperature anodic bonding operation illustrated in Fig. 3(a). After bonding a dicing saw is used to free the unetched silicon in the disbonded regions. The dicing saw also removes the silicon to the left and right of the dicing markers. This step determines the width of the iris shown in Fig. 2(b).

Automation® dicing saw with a 0.9 mil blade traversed the structure 1 mil below the Pyrex plane in a rastering fashion that slightly overlapped neighboring tracks until the channel was ~1 mm wide (66 cuts) and then the procedure was repeated to define the second iris edge. Sidewall strips were constructed by patterning a third ⟨110⟩ wafer with a simple mask to define (111) plane edges prior to KOH etching. A pair of sidewall strips were held in contact with the central waveguide pattern as the second HTB process was performed; the structure was then one modular composite.

Aluminum was used as the conductor in the prototype structure. To achieve 1.38 μm of Al (1.38 μm is 3 rf skin depths with 15% overdeposition at 60 GHz) on the high-profile structures while minimizing deposition rate dependence on the angle of incidence, sputtering was used (300 Å min). After the sputter process a slight gap existed at the sidewall/iris interface that would affect the current conduction path. This issue is addressed in the discussion section. One composite of each image was stacked using glass capillaries (o.d.$=980\pm1$ μm to theoretically yield a ~775 μm aperture gap) cut to 0.6–0.8 cm lengths. Because of material availability, 980-μm-diam capillaries were used in place of 1 mm capillaries required for the design goal. The capillaries were seated into rectangular antiparallel grooves without bonding for the purpose of alignment evaluation (Fig. 5).

III. DXL WAVEGUIDE FABRICATION

Four waveguide sections were fabricated out of nickel (~160 μm thick) using a variation of LIGA DXL.[8] The structures were designed to be stacked using 484 μm o.d. capillaries. Cavity cell dimensions were 1.8 mm in width by 0.3 mm in length and each cell was separated by a 0.1 mm iris wall (Table I). A silicon wafer with SiO_2/Si_3N_4 thin films was used as a base substrate. A Ti/Cu/Ti thin film at 100 Å/100 Å/100 Å was deposited as an electroplating base.



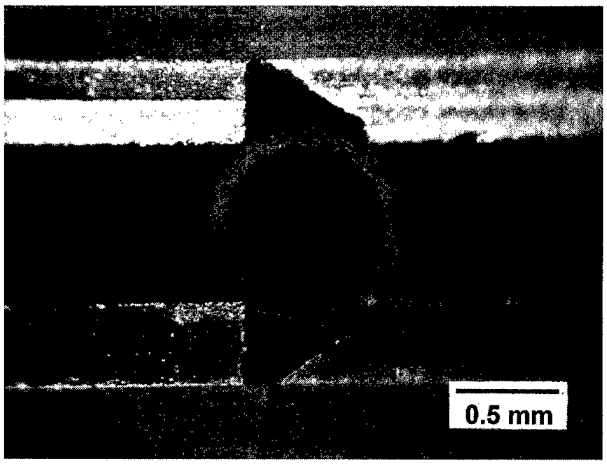


FIG. 5. Two views of a completed ⟨110⟩ silicon waveguide stack. 980 μm o.d. Pyrex capillaries were used to stack the die.

Poly(methyl methacrylate) (PMMA) was cast to a thickness of 200 μm and cured. A high-contrast x-ray projection mask was used at the 1 GeV Aladdin accelerator (Stoughton, WI) to expose negative features into the PMMA. Nickel was plated into the PMMA pattern to form one-half of the planar waveguide. The PMMA was removed to reveal the final structure (Fig. 6).

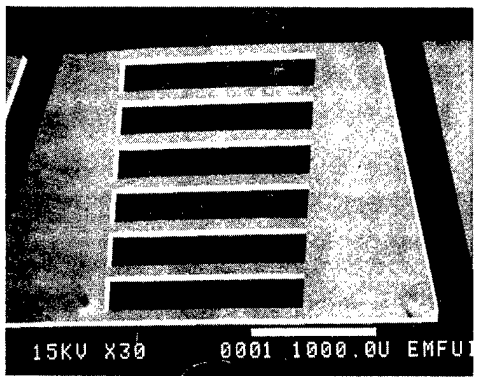


FIG. 6. SEM micrograph of a nickel DXL LIGA-fabricated muffin-tin structure. The plated nickel is approximately 160 μm thick.

TABLE II. MAFIA[9] simulation results for rf parameters.

Radio frequency parameter	No slope	25:1 slope	Parameter change
Frequency (GHz)	61.46	62.08	+1.0%
Q	2957	2799	−5.3%
r_0 (MΩ/m)	278.9	267.1	−4.2%
r_0/Q (kΩ/m)	93.1	95.4	+2.5%
α (1/m)	4.879	5.224	+7.1%

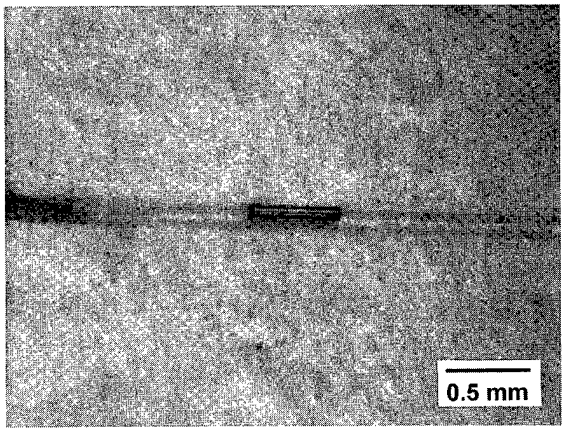
IV. MODELING AND CHARACTERIZATION OF THE WAVEGUIDES

Since anisotropic etches of silicon have finite selectivity, some amount of cavity sidewall slope was expected in the etched silicon waveguides. MAFIA[9] was used to test the electromagnetic effects of a sloped ⟨111⟩ iris wall. Cavity sidewall slope was considered to have less bearing on the electrical properties than the iris wall slope.

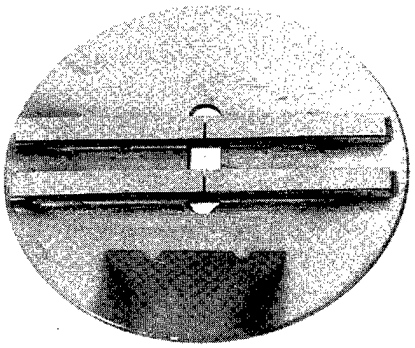
KOH ⟨110⟩:⟨111⟩ etch selectivity is typically between 100:1 and 200:1, though selectivity as high as 600:1 has been reported.[7] To magnify the effect of finite iris wall slope (m) on electrical performance degradation, a worst-case geometric model with $m=25$ was analyzed. Important rf parameters for the baseline structure, possessing infinite m, and for the $m=25$ structure are presented in Table II.

In the table, r_0 is the shunt impedance and α is the attenuation parameter. The shunt impedance is defined by $r_0=-E_{0z}^2/(dP/dz)$, where dP/dz is the rf power dissipated per unit length and E_{0z} is the amplitude of the axial electric field.[10] A high r_0 indicates efficient acceleration properties, since the E field contributes to acceleration while the dissipated power is wasted in the form of heat. The quality factor Q is a measure of the ability of a rf cavity to resonate. It is defined as the ratio of energy stored to energy lost, per radian of the rf wave. The r_0/Q value is an important measure of the effectiveness of an accelerating structure. Due to the highly variable ordinate projections of linear mesh densities in the MAFIA model, the percentages of parameter changes shown in Table II are not considered significant. It can be concluded from these data that the finite sidewall slope in the silicon structure will not significantly degrade the electrical performance of the waveguide. Variations in the impedance and resonant frequency due to the sloped walls can be compensated for by adjusting the mask geometry. The validity of the modeling computation may be verified by comparing the numerically computed r_0/Q value to empirical r_0/Q data gathered from field perturbation experiments.[10] This comparison was carried out on a 12 GHz model of the linear accelerator and is reported elsewhere.[11,12]

Due to the high operating frequencies of this mm-wave structure, the measurement of the electromagnetic fields using conventional perturbating objects was impossible. A microfabricated bead on a supporting insulator was processed to test the silicon waveguide. Commonly used perturbing structures include conductive spheres or ellipsoidal needles which are suspended in space by a relatively thin insulating string. These structures are typically used because they can be solved analytically and are easily fabricated. The micro-



(a)



(b)

FIG. 7. A microfabricated bead conductor (a) for use in electromagnetic field perturbation experiments and the shadow mask (b) used in its fabrication. Aluminum was sputtered onto an 80-μm-diam glass fiber through a mask opening approximately 550 μm in length to create the hollow conductive cylinder.

fabricated bead approximation to an ellipsoidal needle is an optical fiber coated with an aluminum film to form a hollow cylinder with a length of ~550 μm [Fig. 7(a)]. The bead was fabricated on 80-μm-diam-silica optical fiber by sputtering an Al film through a capillary shadow mask [Fig. 7(b)]. The sputter mask consists of two small-bore glass capillaries ground flat, mounted to a backing plate, and cut to provide ~550 μm precision gaps. Bead fibers were threaded into the large capillaries and metal was sputtered onto the regions exposed by the saw-defined slits. The mask allows two fibers to be sputtered simultaneously. A first generation bead was used in the electromagnetic field measurement on a 10× scale model of the 120 GHz version of the mm waveguide.[11] The results showed the expected shape of the magnetic field but the bare silica fibers were quite fragile. Also, the form factor for the bead depended on the bead's rotational angle due to the nonuniform thickness of the bead. The form factor is a scaling factor that is used to determine how the measured frequency perturbation relates to the field in the cavity for a particular bead geometry and material. Nylon thread provided a more rugged support for the microfabricated bead and subsequent bead experiments with 25–125 μm-diam nylon insulators gave accurate form factors. These beads were

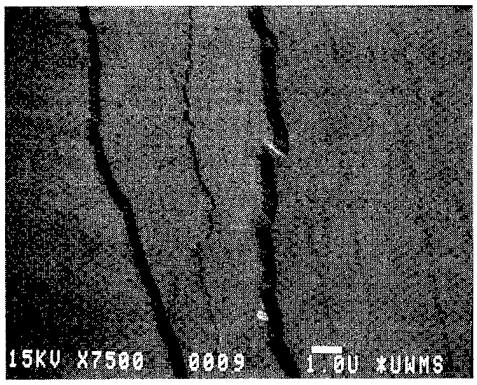


FIG. 8. SEM micrograph of the muffin-tin iris sidewall demonstrating boundary defects due to mask/crystal plane misalignment. The surface roughness between misalignment defects appears to be <0.2 μm.

able to accurately map out the electric field in the scale model.[12]

Physical characterization using scanning electron microscopy (SEM), optical and profilometry instrumentation was performed on eight silicon waveguide samples. Surface roughness requirements for the mm waveguide can be evaluated by comparing the electrical skin depth (δ) to the average surface roughness; a roughness on the order of or greater than δ would indicate that the conduction length that the current sees is greater than the physical length of the particular segment. Since δ=0.27 μm at 60 GHz for a copper conductor (the material of choice), a degraded conduction path would possess a visibly cloudy/dull surface finish. SEM and optical data suggest that all internal exposed surfaces possess a near-optical finish (≤2000 Å average surface roughness). A shingle step defect is present when the optical mask-to-crystal plane alignment is poor but the use of well placed rotational alignment marks can keep the spatial frequency of these steps to a minimum. Figure 8 illustrates this defect on the vertical iris wall of a silicon waveguide. Note that despite these shingle defects the surface roughness between the misalignment defects is <0.1–0.2 μm.

Cavity length accuracy (at the Pyrex/silicon interface) was measured with an inspection microscope and found to be ~±3 μm off the target dimension. Feature radii in the top plane were found to be <5 μm, though MAFIA analysis of the cavity structures has not proven that feature radii greater than 5 μm have a deleterious effect on the guide's acceleration properties. The feature height accuracy depends only on the accuracy with which the ⟨110⟩ wafers are sliced; high-precision custom thickness wafers are commercially available that offer ±1 μm thickness accuracy. Feature height runout >1 μm was not detected and sample bow (over a 2 cm length) was ~6 μm. Cavity width accuracy (limited by the dicing saw positioning and grinding accuracy) was found to be marginal, varying by ±10 μm with an additional error of up to +20 μm due to sidewall placement error.

SEM characterization of four DXL-fabricated samples suggested <0.1 μm of surface roughness with the exception of the top Ni plated surface, which had a 1–2 μm average surface roughness confirmed by profilometry data. Feature

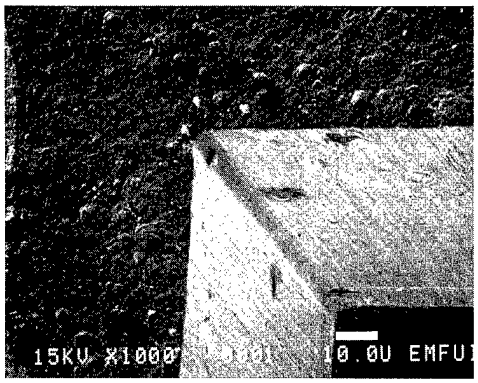


FIG. 9. High-magnification SEM micrograph of a muffin-tin cell corner in a LIGA-fabricated waveguide structure. The waveguide structure's sidewall roughness appears to be ~0.1 μm.

width was controlled to ~2 μm accuracy with a ~2 μm corner radius in the top plane (Fig. 9). The average feature radius in the profile crosssection was <0.25 μm. All features possessing widths ≳15 μm exhibited edge-bead plating effects. Surface profilometry revealed an average feature height of ~160 μm with a runout of 16 μm (adjusted to account for substrate bow).

The second source of geometrical error, that of top-to-bottom die orientation, has been evaluated for a silicon waveguide stack using optical microscopy. Wafer-to-wafer aperture distance was measured to be 750 ±8 μm (for a target aperture of 775 μm), representing a 3.2% dimensional error attributed to capillary seating sidewall slope (from ⟨111⟩ direction etching) and feature overetch of a few microns per die.

Die-to-die alignment was evaluated in two DXL waveguide stacks constructed with 484 ±1 μm o.d. capillaries and an adhesive apiezon wax (Fig. 10). SEM measurements indicated ±5 μm translational accuracy (in the beam aperture plane), which correlates with the calculated geometrical error due to the unfinished top Ni surface. Aperture size measurements were unreliable due to the large amount of nickel thickness variation in the DXL samples.

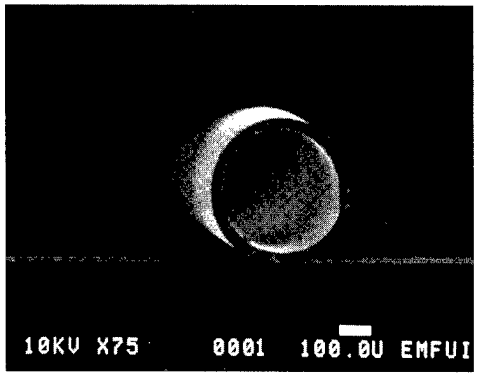


FIG. 10. SEM micrograph of a 240 GHz LIGA muffin-tin waveguide stack. 484 μm o.d. Pyrex capillaries were used to stack the two die.

V. DISCUSSION

The silicon processing method does not yield the high precision that DXL offers but it has the benefit of being inexpensive and relatively simple. The 120 GHz linear accelerator proposed would require 4.95 m of waveguide sections, or 140 three in. wafer substrates for 70 sections.[4] It can be assumed that the 60 GHz mm waveguide described in this paper would require at least the same number of 3 in. substrates. DXL manufacture of this many devices would be a very costly and time-consuming procedure, a tradeoff that probably will not justify the accuracy enhancement. Numerical analysis, to be verified by bead-pulling measurements, has demonstrated that iris wall slope does not significantly degrade the electric field characteristics. Other geometrical criteria described by error analysis have been met marginally for all topographical measurements.

Die-to-die iris separation (i.e., aperture height) and transverse iris width control, require improvement being unacceptable and marginal. The dicing saw step, which uses photolithographically defined markers to locate where the iris extents should terminate (i.e., determine iris width), introduces the largest error in the system. The 20 ±10 μm iris width variation (calculated over multiple wafer batch runs) corresponds to a 0.83% maximum feature width error, which exceeds the error limit suggested by electron beam analysis.

A second shortcoming of the iris design is lack of electrical continuity with the etched sidewall. A finite sidewall slope of 100:1 would yield a gap of ~13 μm in the full-scale waveguide; since sputtering would not bridge this opening a conductive filler, such as indium solder, must be considered. By adding a filler the surfaces may conduct freely but the corner geometry is not well defined. A better solution may be to use sidewalls defined by dicing saw cuts polished to a ~0.1 μm finish that would have an average slope >5000:1. This sidewall structure would meet the skin depth surface smoothness requirement and permit a sidewall-to-iris interface bridged by sputtered material. The gap would be ~1 μm and easily filled with a conductor that would not distort corner definition beyond ±5 μm.

Thermal load caused by rf power losses in the waveguide is an unaddressed issue in the silicon/Pyrex composite. The low 11 mW/cm °C thermal conductivity of Pyrex will not permit the 10 W/cm^2 °C of cooling required for thermal stability during operation of the accelerator.[13] Complicating proper cooling is a poor thermal conduction path beneath the sidewalls adjacent to the iris (Fig. 11). The 1 mil raster cuts which are a fundamental consequence of the silicon process presented form small free-space pockets in a ultrahigh vacuum environment thereby reducing thermal conductivity away from the central iris. Minimization of the Pyrex thickness by lapping and subsequently bonding the thinned Pyrex layer to silicon or diamond with microchannel cooling paths may provide an effective film conductivity coefficient of 10 W/cm^2 °C.

VI. CONCLUSION

Fabrication of a high aspect muffin-tin silicon waveguide has been developed as an alternative to a LIGA-fabricated

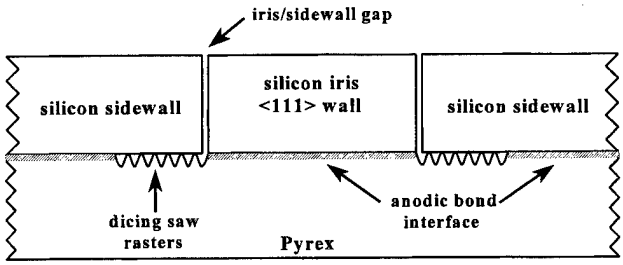


FIG. 11. Cut away transverse view of the muffin-tin structure. The hatched areas represent optically smooth interfaces. Anodic bonding is inhibited at the dicing saw rastered regions where the Si/Pyrex contact interface is minimal.

mm waveguide structure. The waveguide tolerances were to be held to ~5 μm for features and die-to-die alignment, with iris width controlled to ~10 μm accuracy. The LIGA structures were found to have excellent (better than 1–2 μm) feature accuracy in the feature plane but would require further mechanical finishing in order to meet total internal reflection and height tolerances.

It is believed that 2 μm feature accuracy (0.5% dimensional accuracy) and ~5–6 μm die-to-die alignment (0.7%–1.3% dimensional accuracy) is achievable in the ⟨110⟩ silicon waveguide through photomask/etch compensation and photolithographic process refinement. The current results are approximately a factor of 2 worse than this, which is considered a success for initial process development. Iris width control could be improved to ~±10 μm through controlled polishing, reducing the maximal transverse error to 0.28%.

To its benefit, the ⟨110⟩ waveguide is more cost effective in supplies and labor required for fabrication and can be produced with standard UV photolithography equipment, while the LIGA process requires time on a relatively expensive synchrotron radiation storage ring equipped with a coherent x-ray exposure station. It is estimated that the silicon waveguide requires ~1/5th of the finances on a per-waveguide section basis and has a fabrication cycle time ~1/4th that of the LIGA process.[14] The deep-etch silicon waveguide will require more refinement but is a viable alternative to the LIGA fabrication of a micro-linear accelerator waveguide.

ACKNOWLEDGMENTS

This research was supported by the Argonne contract "Miniature Linear Accelerators," 1993–96. The authors wish to thank Dr. Robert Kustom, Dr. Ali Nassiri, and Dr. Paul Matthews of Argonne National Laboratory for their invaluable contributions to the linear accelerator content of this article. They also wish to thank the University of Illinois' EMF for the resources used to produce the SEM micrographs herein.

[1]A. Nassiri, Internal Note, Argonne National Laboratory, Argonne, IL, 1994 (unpublished).
[2]R. L. Kustom, A. D. Feinerman, D. Grudzien, H. Henke, D. Horan, Y. W. Kang, A. M. Khounsary, P. J. Matthews, F. E. Mills, A. Nassiri, J. Song, and T. L. Willke, in Proceedings of the Linac 94 Conference, KEK, Tsukuba, Japan, 1994 (unpublished).
[3]H. Henke, TET Note No. 94/04, Technische Universitat Berlin, 1994 (unpublished).
[4]A. Nassiri, R. L. Kustom, F. E. Mills, Y. W. Kang, A. D. Feinerman, H. Henke, P. Matthews, J. Song, D. Grudzien, D. Horan, and T. Willke, IEEE, International Electron Device Meeting, Washington, DC, 1993 (unpublished), pp. 169–172.
[5]G. Wallis and D. I. Pomerantz, J. Appl. Phys. **40**, 3946 (1969).
[6]A. D. Feinerman, D. A. Crewe, D. C. Perng, S. E. Shoaf, and A. V. Crewe, J. Vac. Sci. Technol. A **10**, 611 (1992).
[7]W. R. Runyan and K. E. Bean, *Semiconductor Integrated Circuit Processing Technology* (Addison-Wesley, New York, 1990), pp. 242–293.
[8]H. Guckel, Nucl. Instrum. Methods Phys. Res. B **79**, 247 (1993).
[9]R. Klatt *et al.*, MAFIA, 1986 Linear Accelerator Conference Proceedings, Stanford, CA, 1986.
[10]F. Amman, in *Electron Linear Accelerators*, edited by P. M. Lapostolle *et al.* (Wiley, New York, 1970), Part B, pp. 39–173.
[11]P. Matthews, Y. Kang, T. Berenc, T. Willke, A. Feinerman, and R. Kustom, in Proceedings of the Fourth European Particle Accelerator Conference, London, 1994 (unpublished), pp. 1303–1305.
[12]P. J. Matthews, T. Berenc, F. Schoenfeld, R. L. Kustom, and A. D. Feinerman IEEE Trans. Microwave Theory Tech. (to be published).
[13]P. Matthews, A. Khounsary, and R. Kustom in American Physics Society Annual Meeting, Washington, DC, 1994 (unpublished).
[14]Microelectronics Center of North Carolina, MUMPS Project Brochure, 1994.

Minimized response time of optical emission and mass spectrometric signals for optimized endpoint detection

S. Thomas III, H. H. Chen, C. K. Hanish, J. W. Grizzle, and S. W. Pang[a)]
Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan 48109-2122

(Received 9 February 1996; accepted 29 May 1996)

The response times of optical emission spectroscopy (OES) and mass spectrometry (MS) have been measured for plasma etching of III–V heterostructures. For the Ga optical emission signal at 417.2 nm, a response time as fast as 0.2 s was obtained. The minimum response time of the $^{145}AsCl_2^+$ partial pressure, measured by MS, was found to be 0.9 s. The saturation times of the optical emission signal and the partial pressure signal have been shown to be related to the residence time of etch gases by varying the total gas flow rate and the chamber pressure. Decreasing residence time by reducing the pressure from 6 to 2 mTorr and maintaining a constant flow rate caused the saturation time of the Ga emission signal at 417.2 nm to decrease from 7 to 3 min. The $^{145}AsCl_2^+$ partial pressure signal saturated before the Ga emission signal. Endpoint detection for etching an AlInAs emitter and stopping on a GaInAs base of a heterojunction bipolar transistor was studied. Algorithms which monitor the change in Ga emission intensity have been developed to automatically stop the emitter etch with ~2 nm of the GaInAs base layer removed. Additionally, etching of GaInAs on an InP substrate was studied and the signal from OES detected the endpoint before the MS signal did due to the faster response time. © *1996 American Vacuum Society.*

I. INTRODUCTION

Dry etching is often used to fabricate structures of submicrometer dimension with vertical profiles. However, the selectivity between different materials is usually quite low. In these cases, large overetch times to clear a layer and stop without removing the next layer are impractical, and an endpoint detection scheme is necessary for reliable device fabrication. The ability to distinguish between separate layers has been shown using several techniques, including optical emission spectroscopy (OES),[1] mass spectrometry (MS),[2,3] laser reflectometry, and grazing angle optical emission interferometry.[4] AlGaAs layers as thin as 5 nm have been resolved in GaAs/AlGaAs heterostructures.[5] After detection and identification of the etch products, a scheme for endpoint detection is needed to stop the etch once the underlying layer has been reached. An optimized endpoint detection system should include sensors with the best response time and signal-to-noise ratio. This is essential for precisely stopping an etch when the thicknesses of the device layers are in the nanometer range. One such device is the GaInAs/AlInAs-based heterojunction bipolar transistor (HBT) with a GaInAs base layer as thin as 60 nm.[6] Using OES to monitor the Ga etch products, we have shown that a HBT emitter etch could be stopped with <10 nm of the GaInAs base layer removed.[7] In this article, the response time and saturation time of these *in situ* sensors are optimized to stop the emitter etch more precisely.

The response time of an *in situ* sensor, generally defined as the delay between changes in etch conditions and sensor signals, has been studied.[8] This parameter determines how quickly the control scheme can respond to changes in the plasma and stabilize the plasma conditions. It also sets a lower limit on how precisely an etch can be stopped once a layer has been completely etched. Another characteristic of sensor-based etch process monitoring which has been studied is saturation time. Once the change is detected, the saturation time represents the delay before the intensity of the monitored signal reaches steady state. The saturation time can be influenced by the residence time of gases in the plasma and chamber wall effects. Decreasing the saturation time increases the sensitivity of the *in situ* sensors to changes in etch rates relative to changes in residual etch product concentrations. The response time and saturation time have been studied for optical emission and mass spectrometric signals, and methods to reduce these are presented.

The effectiveness of OES for endpoint detection has been examined using scanning electron microscopy (SEM) for reactive ion etching of Si/SiGe heterostructures.[9] The SEM analysis showed that a significant amount of SiGe was etched before the Ge emission intensity increased. The response time of the Ge emission intensity was approximated at 3 s. In our study, reflectometry was used to evaluate the effectiveness of the endpoint detection scheme. By measuring the thickness of the GaInAs base layer before and after etching with reflectometry, the amount of base layer removed can be found. It has been demonstrated that modifications of the device material could improve *in situ* monitoring of dry etching in III–V compound heterostructures to an etch depth accuracy of ±8 nm.[10] However, if the response time can be reduced so that OES can be used to detect the endpoint with <8 nm of the base removed, then modifications of the device material are not necessary to improve etch depth accuracy

a)Electronic mail: pang@eecs.umich.edu

and to minimize the removal of the base layer.

II. EXPERIMENT

The etching system consisting of an electron cyclotron resonance (ECR) source and a rf-powered electrode has been described previously.[11] Microwave power at 2.45 GHz is coupled to the plasma using a copper input probe. Tuning of the plasma is obtained by adjusting the microwave power input probe length and the cavity dimension to minimize the reflected power. The power coupled to the stage can be controlled by fixing either the rf power from 0 to 500 W or the self-induced dc bias voltage ($|V_{dc}|$). The stage position can be adjusted so that the ECR source to sample distance can be varied from 8 to 27 cm. Gases are fed through mass flow controllers and the pressure is controlled by a throttled 1500 l/s turbomolecular pump. The volume of the chamber is 80 l.

One end of an optical fiber was placed at the quartz dome where the ECR condition was met to collect optical emission from the plasma. A Spex 500 scanning monochromator was used to detect emission intensity as a function of wavelength ranging from 300 to 900 nm. The resolution of the monochromator was 0.02 nm and the focal length was 0.5 m. The grating used was blazed for maximum efficiency at 630 nm. The signal was measured by a photomultiplier tube, then it was amplified, and sampled by a computer with LabVIEW running data acquisition and control software.

Mass spectrometric intensities were measured using a NGS partial pressure transducer (PPT). The mass spectrometer was designed to be portable which necessitated the positioning of the PPT 66 cm downstream from the plasma chamber. For a dedicated system, the PPT could be closer to the chamber, potentially decreasing response time. The electron energy of the ionizer was 70 eV and the base pressure of the PPT was kept $<5\times10^{-6}$ Torr by using a 0.6 mm orifice in conjunction with an 80 l/s turbomolecular pump. The mass spectrometer provides the ability to monitor two mass spectrometric signals simultaneously. The two signals can be updated as fast as 8 Hz, or one signal can be monitored alone at 10 Hz. All signals of interest including microwave power, rf power, pressure, gas flows, and signal intensities from the OES and MS systems can be collected and monitored in real time on the LabVIEW platform.

III. RESULTS AND DISCUSSION

It has previously been shown that when etching III–V materials such as GaAs and InP, optical emission from the group III elements, Ga and In, is readily detected using OES, while the partial pressures of the group V related etch products, $^{145}AsCl_2^+$ and $^{101}PCl_2^+$, are detected by MS.[7,12] The $^{145}AsCl_2^+$ and $^{101}PCl_2^+$ signals are proportional to the partial pressure of the chlorides formed in the ionizer of the PPT, and can be related to changes in the etch rates of GaAs and InP. Both OES and MS are able to monitor signals related to the etch gases. To make use of these signals for real-time control, it is important to measure the response of the sensors to changes in plasma conditions. Figure 1 shows the optical emission and mass spectrometric signals related to etch products of GaAs etched in a Cl_2/Ar plasma when the plasma is first ignited. The plasma conditions were 50 W microwave power, 125 W rf power, Cl_2/Ar flow at 4/16 sccm, chamber pressure of 2 mTorr, and 13 cm ECR source-to-sample distance. The microwave power was increased from 0 to 50 W and the plasma was ignited. It can be seen that the Ga emission intensity at 417.2 nm, which is sampled at 40 Hz, begins to increase as soon as the microwave power is enabled. By defining the response time as the time between when the microwave power rises and the time derivative of the Ga emission intensity is maximum, a response time of 0.2 s is obtained. There is a delay of 1 s before the $^{145}AsCl_2^+$ mass spectrometric signal increases from the background value shown prior to etch initiation. The cause of this delay could be due to two effects, either the integration time of the mass spectrometer or the transit time from the wafer surface to the detector. The integration time represents the desired rate of filtering and determines how frequently the signal is updated. For this study, the integration time was varied between 0.2 and 1 s. The transit time of the etch products is composed of two components, the residence time in the chamber and the time to travel from the gate valve to the ionizer of the mass spectrometer. Under typical etch conditions, the transit time was estimated to be 0.6 s.

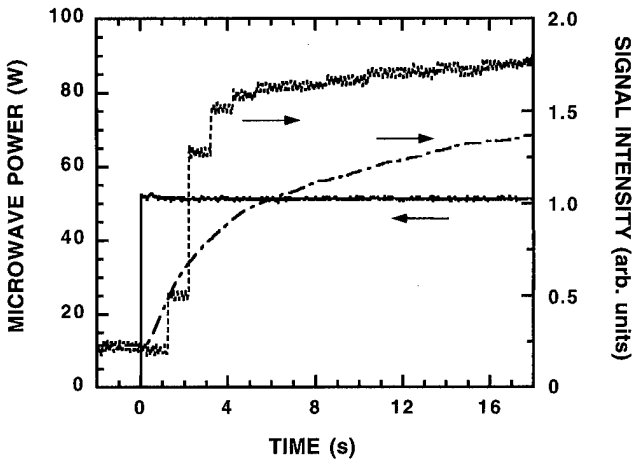


FIG. 1. Response of Ga optical emission at 417.2 nm (- – - –) and $^{145}AsCl_2^+$ partial pressure (- - - -) to ignition of plasma, as shown by increase in microwave power (———) from 0 to 50 W. The plasma conditions were 50 W microwave power, 125 W rf power, Cl_2/Ar flow at 4/16 sccm, chamber pressure of 2 mTorr, and the ECR source-to-sample distance was 13 cm.

Before sampling the Ga optical emission signal, the signal from the photomultiplier tube is amplified. The amplifier has a built-in low-pass filter to reduce high frequency noise. Figure 2 shows how the response time and signal-to-noise ratio vary with the filter cutoff frequency. In this case, the signal referred to is the time derivative of the Ga emission intensity. Increasing the cutoff frequency improves the time response of the signal by allowing the high frequency components to contribute to the signal. Thus, the minimum response time of 0.2 s occurred when the cutoff frequency was highest at 10 Hz. The disadvantage of increasing the cutoff frequency is that high frequency noise is no longer damped, leading to a

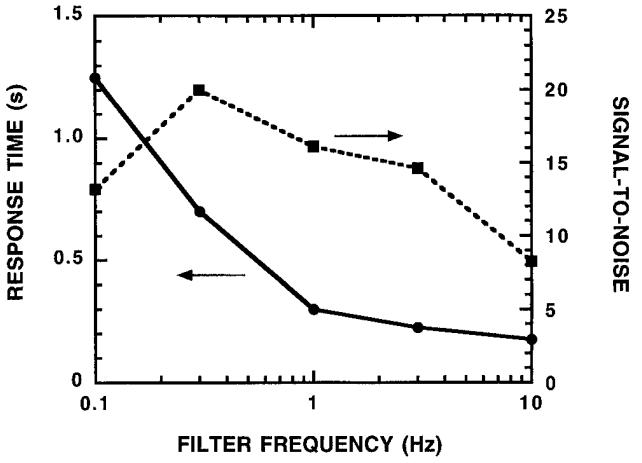


FIG. 2. Dependence of Ga optical emission response time on cutoff frequency of low-pass filter for the amplifier. Also shown is the signal-to-noise ratio of the time derivative of the Ga signal. The etch conditions were the same as in Fig. 1.

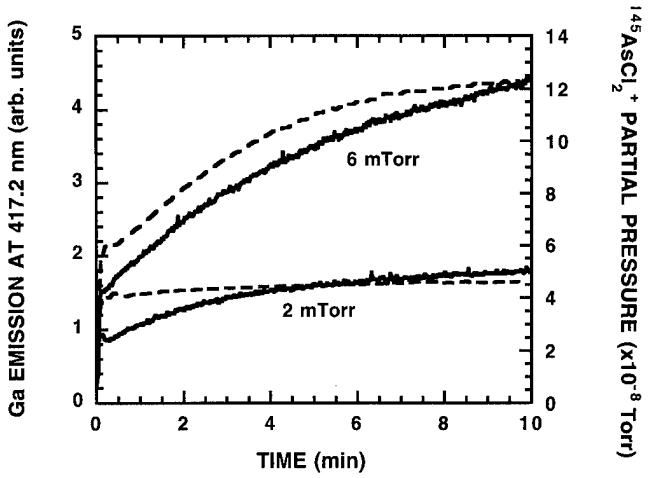


FIG. 4. Effects of pressure on the saturation time of etch product signals monitored by OES (———) and MS (- - - -). The etch conditions were 50 W microwave power, 100 W rf power, Cl_2/Ar flow at 3/27 sccm, and 13 cm source distance.

more noisy signal. This is shown in Fig. 2 by a decrease in the signal-to-noise of the Ga time derivative with increasing cutoff frequency. As the cutoff frequency decreases, the signal is also damped and the signal-to-noise decreases. The maximum value of the signal-to-noise was 20 and it occurred for a cutoff filter frequency of 0.3 Hz. The response time of the signal at this frequency was only 0.7 s. Therefore, for endpointing, a compromise is made between response time and sensitivity. At a filter frequency of 1 Hz, the signal-to-noise is 16 and the response time is 0.3 s. It is assumed that the response time will set a lower limit on the precision of the endpoint detection scheme.

Figure 3 shows how the response time of the $^{145}AsCl_2^+$ signal depends on the update frequency. For short integration times, the mass spectrometer responds quickly to changes in the $^{145}AsCl_2^+$ partial pressure and a minimum response time of 0.9 s is measured for an update frequency of 5 Hz. For processes that operate at higher pressure, the response time can be limited by transit times through differential pumping stages. Using *in situ* MS sampling of poly-Si deposited at 5 Torr, a response time limited by viscous flow was measured to be ~3 s.[8] The gas residence time is proportional to the gas pressure and inversely proportional to the gas flow rate. The residence time has been shown as a function of the gas flow at different gas pressures.[13] For our system, the residence time for a pressure of 2 mTorr and a total flow of 20 sccm is 0.6 s. This suggests that the minimum response time of 0.9 s is partially caused by the time needed to travel from the wafer surface to the mass spectrometer. The drawback to using a short integration time is the increased sensitivity to noise which is not averaged out effectively during the short integration time. This can be seen by the low signal-to-noise ratio of only 6.4 measured at an update frequency of 5 Hz. For longer integration times, the signal-to-noise is increased significantly by 10× to 64 at an update frequency of 1 Hz. At a frequency of 1 Hz, the response time of the mass spectrometer increased to 1.8 s. For endpoint detection, a frequency of 2 Hz is appropriate, since the response time is still limited by the transit time of etch products from the wafer surface to the sensor and a strong signal-to-noise of 28 is obtained.

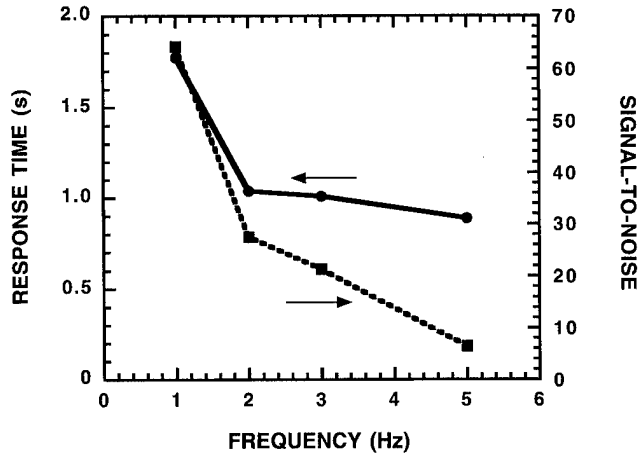


FIG. 3. Response time and signal-to-noise for the time derivative of the $^{145}AsCl_2^+$ partial pressure signal as a function of update frequency for the partial pressure measurement. The etch conditions were the same as in Fig. 1.

It should be noted from Fig. 1 that the signals do not saturate initially. This has been presented previously for the Ga intensity during etching of GaAs.[7] The saturation time was defined as the time required before the signal reached 80% of the steady state value. The saturation time of the Ga emission intensity at 417.2 nm decreased as the total gas flow rate increased. The saturation time for the optical emission and mass spectrometric signals are compared as shown in Fig. 4. The plasma is ignited at time 0 min and the signals begin to increase as the GaAs is etched. The $^{145}AsCl_2^+$ partial pressure behaves in a similar manner to the Ga optical emis-

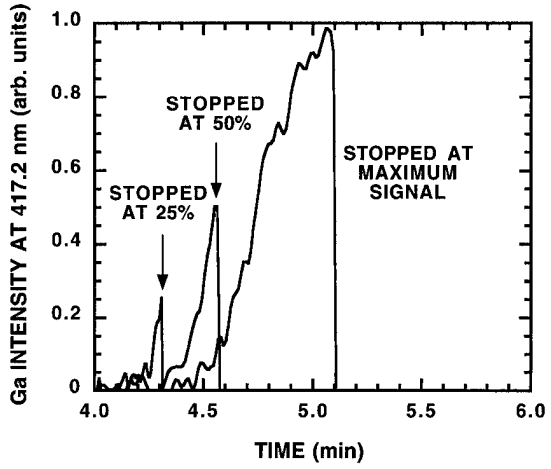


FIG. 5. Ga emission intensity at 417.2 nm as a function of etch time. The etches were stopped when the signal reached 25%, 50%, and 100% of the maximum expected signal. The etch conditions were 50 W microwave power, 100 W rf power, 2 mTorr, Cl_2/Ar at 3/27 sccm, and 13 cm source distance.

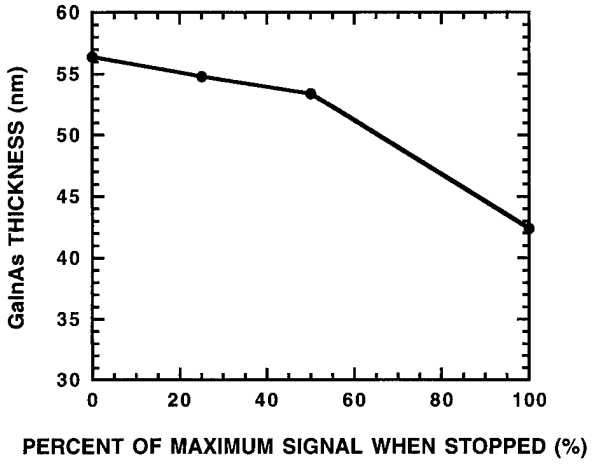


FIG. 6. Remaining thickness of GaInAs-base layer when stopping at three different Ga signal intensities. The etch conditions were the same as in Fig. 5.

sion at 417.2 nm. The $^{145}AsCl_2^+$ signal intensity at a pressure of 6 mTorr is 2.7× larger than that at 2 mTorr. However, it takes 4 min for the $^{145}AsCl_2^+$ signal at 6 mTorr to saturate, while it only takes 10 s to saturate at 2 mTorr. Also shown on Fig. 4 is the Ga emission intensity for the two different pressures. The saturation time of the Ga emission is higher at 6 mTorr than it is at 2 mTorr. At 2 mTorr, the Ga emission saturates in 3 min while at 6 mTorr it takes 7 min to reach 80% of the saturated value. The saturation times for the optical emission and mass spectrometric signals follow the same trend but MS provides lower saturation time. This could be due to differences between the two monitored species, or it could be related to the mass spectrometer being located downstream from the etching chamber while optical emission is taken directly from the plasma.

The endpoint detection algorithm for the emitter etch step of AlInAs/GaInAs-based HBTs samples the Ga optical emission signal at 417.2 nm, since this signal will rise once the GaInAs base layer begins etching.[7] For calibration, an AlInAs on GaInAs sample was etched to determine the maximum Ga emission intensity. This test structure consisted of 200 nm AlInAs on 60 nm GaInAs over an InP substrate. For the following etches, when the Ga intensity reached a specified percentage of the maximum value, a signal was sent to the etch system controller to disable the power. Figure 5 shows three traces of the Ga emission intensity as a function of time. For these three etches, the endpoints were defined as occurring when the Ga intensity reached 25%, 50%, and 100% of the maximum value. It can be seen that once the intensity reached the prescribed level, the plasma was turned off and the intensity dropped to zero immediately. The etch condition for all three was 50 W microwave power, 100 W rf power, 2 mTorr pressure, Cl_2/Ar at 3/27 sccm, and 13 cm source distance.

Prior to etching, the AlInAs and GaInAs layer thicknesses were measured using reflectometry. The AlInAs layer was found to be 209 nm and the GaInAs layer was 57 nm thick. After etching, the samples were remeasured to determine the remaining layer thicknesses. Figure 6 shows the remaining thickness of the GaInAs layer for the three different etches. The AlInAs layer had been completely removed in all three cases. When the etch was stopped after the Ga emission signal reached 25% of the maximum expected value, the remaining GaInAs thickness was 55 nm. Therefore, the etch had been stopped after only 2 nm of the base had been removed. For a 60 nm base layer, this represents only a 3% change in the base layer thickness due to overetching.

The rise in Ga emission intensity shown in Fig. 5 occurs at different times for each of the etches. The etch that was stopped when the intensity reached 25% of the maximum value, was the first to increase. The etch time to reach 25% was only 4 min 19 s. This can be compared to the slowest etch, in which the intensity did not reach 25% of the maximum until 4 min 38 s had elapsed. The difference in the AlInAs layer thicknesses for the two samples was measured to be <1 nm, such that the etch time difference is attributed to etch rate variation as opposed to thickness variation. The average etch rate for 10 samples was 47.1 ± 1.6 nm/min. The fastest etch rate was 48 nm/min while the slowest etch rate was 45 nm/min, a variation of only 6%. This is large enough that had the etching for the run with the fastest etch rate been carried out for the same time as the run with the slowest etch rate, an additional 15 nm of the GaInAs base layer would have been removed. Clearly, endpoint detection is required even when small run-to-run etch rate variations are anticipated.

In order to eliminate the need for a calibration step to determine the maximum Ga emission intensity, a more advanced endpoint algorithm was developed. This algorithm depends only on data from the current run and an estimated etch time. 1 min before the approximate endpoint, Ga emission intensity data are collected for 30 s. A least-squares fit through the optical data is used to model the optical trend of

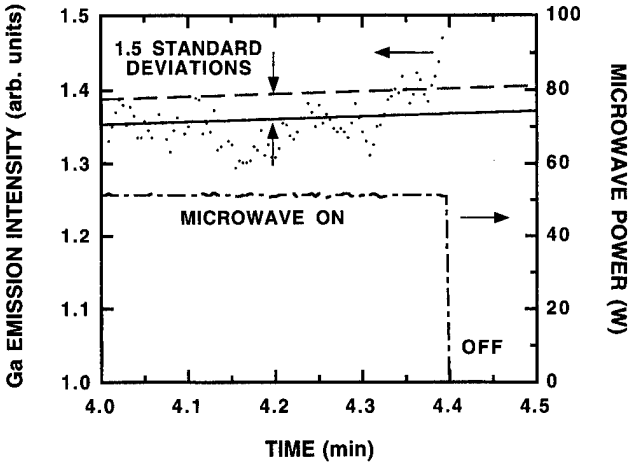


FIG. 7. Endpoint detection of an AlInAs/GaInAs structure using the changes in the Ga emission intensity as the GaInAs layer begins to etch. The model trend (———) and the upper bound (- - - -), 1.5 standard deviation above the trend, are shown with the microwave power, (- – - –). The etch conditions were the same as in Fig. 5.

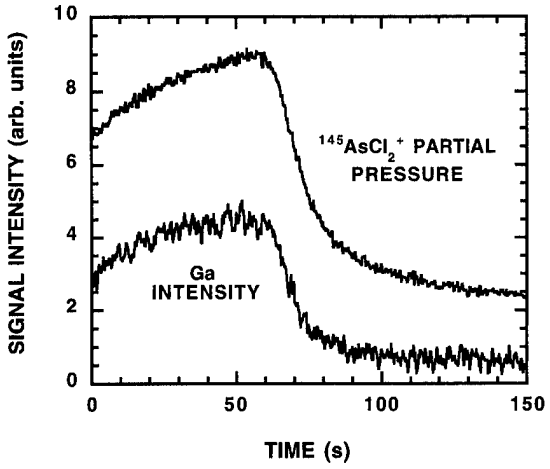


FIG. 8. Comparison of the Ga optical emission from OES and the $^{145}AsCl_2^+$ partial pressure from MS when endpointing GaInAs on InP substrate. The plasma conditions were 50 W microwave power, 100 W rf power, 1.5 mTorr chamber pressure, Cl_2/Ar flow at 3/27 sccm, and 13 cm source-to-sample distance.

the current run, and the standard deviation of the noise is calculated. At a sample rate of 4 Hz, the LabVIEW platform can make these calculations without missing any data collections. After modeling the trend, the algorithm looks for a rise in the Ga signal. The endpoint is determined when m consecutive data points are n standard deviations of noise above the calculated trend. Both m and n will be optimized to minimize the etching of the GaInAs base layer. LabVIEW sends a signal to the plasma system controller once the endpoint is reached and the etch is stopped automatically. Because this algorithm does not depend on data from other runs, it is robust to run-to-run changes in Ga emission intensity, noise level, and etch rate, as long as the etch rate does not increase so much that the etch ends more than 30 s before the expected time. Figure 7 shows this endpoint algorithm tested on an AlInAs/GaInAs/InP sample described previously. The endpoint was set when 4 consecutive data points of the Ga emission intensity were larger than 1.5 standard deviations of the noise. The approximate etch time was 4 min and the sample rate was 4 Hz. The endpoint was found at 4 min 23.5 s after etching only ~1 nm into the GaInAs layer. This is indicated by 4 consecutive points lying above the upper bound. The trend and standard deviation of noise are calculated using emission intensity data from 180 to 210 s. The microwave power, which was shut off at 4 min 24 s, is also shown.

Figure 8 shows the comparison of OES and MS signals when endpointing a GaInAs layer on an InP substrate. The plasma conditions were 50 W microwave power, 100 W rf power, 1.5 mTorr chamber pressure, Cl_2/Ar flow at 3/27 sccm, and 13 cm sample-to-source distance. As the GaInAs layer is removed and the InP substrate is exposed to the plasma, both the Ga emission and $^{145}AsCl_2^+$ signals are expected to decrease. The Ga optical emission intensity decreased at the same rate as the $^{145}AsCl_2^+$ partial pressure did. The time it took for the OES signal to drop 25% of the maximum is about the same as the MS signal. However, since the Ga optical emission responds faster than the $^{145}AsCl_2^+$ partial pressure does, one would expect the endpoint predicted by OES to occur before that predicted by MS. This has been verified by applying the endpoint algorithm described earlier. In this case, the endpoint is reached when 4 consecutive points lie below the lower bound, since the Ga emission and $^{145}AsCl_2^+$ partial pressure decrease once the InP substrate begins etching. Using the OES signal, the algorithm predicted that the endpoint occurred ~2 s before that predicted when using the MS signal. Therefore, OES is a more effective method for endpointing this structure.

IV. CONCLUSIONS

Signals from the etching system and the *in situ* sensors were collected simultaneously so that an accurate determination of the sensor response times could be made. The response time for OES varied from 0.2 to 1.3 s when different low-pass filter cutoff frequencies were used. A low-pass filter with a cutoff frequency of 1 Hz on the output of the photomultiplier tube provided a strong signal-to-noise and a short response time of 0.3 s. For short integration times, the response time of MS was limited to 0.9 s, the transit time of the etch products from the wafer surface. By increasing the integration time, the signal-to-noise improved at the expense of response time. The saturation time was reduced by decreasing the pressure and increasing the total flow rate. The $^{145}AsCl_2^+$ partial pressure measured by MS had a lower saturation time than did the Ga optical emission. For endpointing, once the sensor detected a change in the plasma, a signal was sent from LabVIEW to the etch system controller and the plasma was stopped. Etching of an AlInAs/GaInAs structure was stopped once Ga optical emission was detected, and reflectometry measurements showed that only ~2 nm of GaInAs was removed. OES is more effective for endpoint of

a GaInAs layer on an InP substrate since the endpoint predicted by OES occurred ~2 s before that predicted by MS.

ACKNOWLEDGMENTS

The assistance of L. Kamlet and F. Terry with the reflectometry measurements is greatly appreciated. This work is supported by the Advanced Research Projects Agency under Contract No. F33615-92-C5972.

[1]S. J. Pearton, F. Ren, C. R. Abernathy, and C. Constantine, Mater. Sci. Eng. B **23**, 36 (1994).
[2]D. L. Melville, J. G. Simmons, and D. A Thompson, J. Vac. Sci. Technol. B **11**, 2038 (1993).
[3]K. L. Seaward, N. J. Moll, D. J. Coulman, and W. F. Stickle, J. Appl. Phys. **61**, 2358 (1987).
[4]D. Angell and G. S. Oehrlein, Appl. Phys. Lett. **58**, 240 (1991).
[5]P. Collot, T. Diallo, and J. Canteloup, J. Vac. Sci. Technol. B **9**, 2497 (1991).
[6]U. K. Mishra, J. F. Jensen, D. B. Rensch, A. S. Brown, W. E. Stanchina, R. J. Trew, M. W. Pierce, and T. V. Kargodorian, IEEE Electron Device Lett. **EDL-10**, 467 (1989).
[7]S. Thomas III, K. K. Ko, and S. W. Pang, J. Vac. Sci. Technol. A **13**, 894 (1995).
[8]L. L. Tedder, G. W. Rubloff, I. Shareef, M. Anderle, D.-H. Kim, and G. N. Parsons, J. Vac. Sci. Technol. B **13**, 1924 (1995).
[9]H. H. Richter, A. Wolff, B. Tillack, and T. Skaloud, Mater. Sci. Eng. B **27**, 39 (1994).
[10]G. A. Vawter, J. F. Klem, and R. E. Leibenguth, J. Vac. Sci. Technol. A **12**, 1973 (1994).
[11]S. W. Pang, K. T. Sung, and K. K. Ko, J. Vac. Sci. Technol. B **10**, 1118 (1992).
[12]D. J. Kahaian, S. Thomas III, and S. W. Pang, J. Vac. Sci. Technol. B **13**, 253 (1995).
[13]K. Tsujimoto, T. Kumihashi, N. Kofuji, and S. Tachi, J. Vac. Sci. Technol. A **12**, 1209 (1994).

Reactive ion etching of sloped sidewalls for surface emitting structures using a shadow mask technique

B. Jacobs[a] and R. Zengerle
Deutsche Telekom, Research Center, 64295 Darmstadt, Germany

(Received 10 March 1995; accepted 11 April 1996)

A novel, very simple technique for the direct adjustment of the slope of facets in a reactive ion etching process is presented. The etching process is performed by screening the sample partially with an aluminum shadow mask, which is located inside the dark space of the plasma. The sidewall angle of etched lines, defined by a pattern mask on the sample, depends strongly on the distance between the shadow mask and the sample. The investigations are carried out with a methane-hydrogen plasma. The angle of inclination of the sidewalls can be varied in a controllable manner by more than 45°. This technique is well suited to fabricate 90° reflectors, which are necessary for three-dimensional optical interconnects with planar waveguide structures. To analyze the optical quality of the etched mirror plane, the beam transformation was investigated by reflectivity measurements using a single aluminum-evaporated mirror structure, which is illuminated by a focused laser beam at a wavelength of 633 nm. © *1996 American Vacuum Society.*

I. INTRODUCTION

Reactive ion etching (RIE) is a well established technique for the microfabrication of semiconductors. Its inherent high resolution and anisotropy permit a precise pattern transfer of features with geometrical sizes in the submicron region, which are often required in optoelectronic devices. In contrast to an ion beam etching (IBE) process, the main advantage of RIE is the simplicity of the etching process, which can easily be integrated in a cost efficient mass production line of semiconductor devices. However, serious problems arise for the application of RIE, if angled sidewall faces have to be generated for some special optical components like 45° mirrors for surface emitting lasers or optical diffraction gratings. An indirect influence of the sidewall profile can be achieved by controlled mask erosion or, with only a small variation range, by undercut etching.[1,2] A strongly asymmetric profile with different shapes of opposite sides of one structure can only be realized by angled RIE, where the substrate surface is tilted with respect to the incident ion flux, in a similar way as with the arrangement in an IBE process. Only a few papers concerning angled etching in a RIE chamber have been published up to now. All these investigations were carried out with special sample holders, where the substrate was mounted with a tilting angle on the cathode of the reactor chamber.[3–5] However the position of the sample leads to a disturbance of the electric field in the plasma sheath that causes an inhomogeneity of the ion flux across the surface of the tilted sample. The influence of such an inhomogeneity on the etching behavior has already been studied and different solutions have been proposed for the improvement of the homogeneity.

In this study we present another way to control the angle between ions and substrate by deflecting the ions during their path through the plasma sheath directly by a defined field distortion with a conductive mask. The range of influence depends on the special design of the mask. We have recently constructed a spot size transformer with a vertically tapered waveguide by using for the first time such a simple mask technique in a reactive ion etching process.[6] The variation of the etch rate along the waveguide, which leads to the vertical tapering over a distance of 500 μm, is caused by spreading the ion flux below the mask. Waveguide structures which are orientated perpendicular to the spreading direction show a strongly asymmetric sidewall profile, which is a clear indication for angled etching in the vicinity of the mask. Over a distance of several millimeters we observed only a weak dependence of the sidewall profile on the lateral position of the structure. Although in the present work, only a portion of the wafer was etched with asymmetric sidewalls, it may be possible in future work to design a mask with a periodically repeated design that allows processing of total wafers.

We present in detail the first results of angled etching of InP in a methane hydrogen-based plasma using a conductive shadow mask in the RIE process. The methane-hydrogen-based mixture is an efficient etchant for InP and related compounds, because it produces very smooth, mirrorlike surfaces and allows the fabrication of features with high aspect ratio due to a good selectivity between etch mask and substrate material.[7] We investigated the homogeneity of the etching behavior in the dependence on the sample position relative to the mask and studied the effect of the mask geometry on the shape of the etched facets. Finally, a deeply etched 45° mirror with a high aspect ratio, suitable for producing surface emitting planar waveguide structures, was fabricated and optically tested.

II. EXPERIMENT

Figure 1 shows the setup with the cross section of the applied mask for angled etching. The mask, which is formed from a 170 μm aluminum sheet, is directly fixed on the

[a]Electronic mail: jacobs@fz.telekom.de

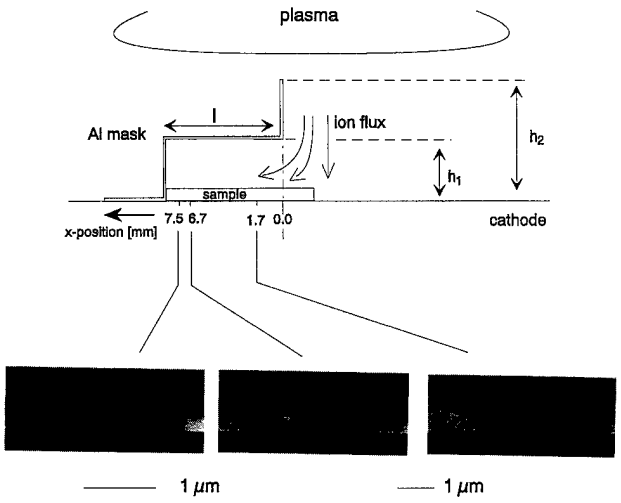


FIG. 1. Setup of the angled etching mode with a shadow mask and photographs of the profile of etched structures, placed below the mask in different positions.

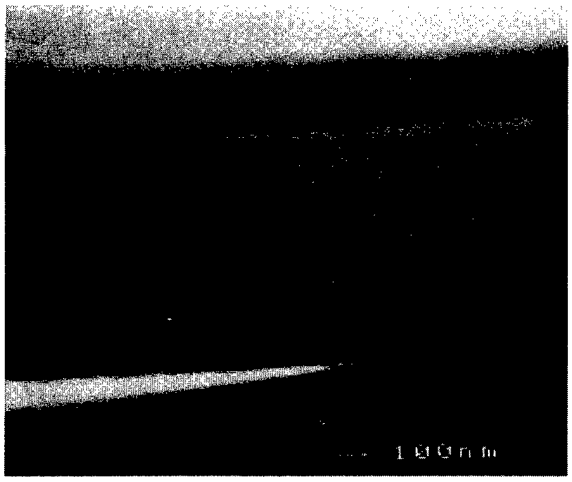


FIG. 2. View of the mirrorlike, sloped facet with tilting angle of 45°.

cathode. During etching the widely extended plasma above the mask provides a homogeneous ion density distribution at the beginning of the plasma sheath. The principle of ion deflection in the presence of the mask is illustrated by the different ion trajectories through the sheath. Far away from the mask the etching ions are accelerated in the plasma sheath directly to the negative biased cathode. Close to the mask, the positive ions are also attracted by the negatively biased mask, so that the inclination angle of the resulting ion flux on the cathode differs clearly from the vertical direction. To enhance the distortion of electrical field lines in the horizontal direction, a vertical sheet is added on the top of the mask.

We started our investigations with undoped InP substrates with a ⟨100⟩ surface. Line patterns with geometrical widths between 1 and 3 μm are defined on the samples by photolithography or electron beam lithography. We used two different mask materials: an organic resist AZ1512, and an evaporated Al_2O_3 film, structured by the lift-off technique. Both masks were 300 nm thick. The pattern transfer was performed by RIE etching with an electron cyclotron resonance (ECR) source at a process pressure of 10^{-2} mbar. We used a commercially available etching system, type Z401, from Leybold, which is described elsewhere in more detail.[8] The anode of a parallel plate reactor is replaced by an ECR plasma source and the distance from the cathode to the plasma chamber of the ECR source is around 20 cm. Around this ECR plasma chamber permanent magnets are arranged in a configuration that results in a toroidal magnetostatic field. The microwave power for the ECR discharge coupled into the plasma amounts to 200 W. Due to the additional ECR discharge the etch rate is increased to a sufficiently high value of 10 nm/min. The RF power was 40 W resulting in a bias voltage of 300 V. The plasma consists of a mixture of methane, hydrogen and argon (CH_4:H_2:Ar=8:50:1). Under these conditions the sheath thickness of the etch plasma, which seemed, as discussed later, to be an important parameter for angled etching, is roughly 2 cm.

III. RESULTS AND DISCUSSION

First we characterized the etching properties—etch rate, sidewall faces and surface morphology—in dependence on the position of the analyzed structure relatively to the mask. The lower part of Fig. 1 illustrates the characteristic shape of the etched sidewalls as observed at different positions on the sample. The line patterns are oriented perpendicular to the picture plane. In this direction the mask extending over 30 mm is long enough to ensure homogeneity of the ion flux. The mask is characterized by the height of the sidewall h_1 (9.4 mm), the length of the cover l (8.6 mm) and its total height h_2 (11 mm). The micrographs in the lower part of Fig. 1 show the significant influence of the shadow mask on the facetting of the structures in three different positions. Even in the range outside the covered region ($x< 0$), both sidewalls of the structures are strongly sloped as discussed later. Over a distance of 8 mm the wall angle on the left side opposite to the ion beam direction turns by more than 20°. The curvature of the sidewall remains nearly straight. So for all positions the realized left facets of the structures are acceptable for an application as a mirror plane. In particular, an angle of 45° is shown in the middle micrograph. Only on the top of the profile, a slightly convex shape at extremely tilted sidewalls may arise, which is illustrated on the left photo. In this case—close to the fixed end of the mask—the etch rate is significantly reduced. The right side of the line patterns directly facing the ion beam are not so well structured. There is a curved transition to the etched floor at the bottom of the facet and the tilting angle of this side depends only weakly on the lateral position.

Taking an appropriate masking material on the chip (Al_2O_3, the stripes being defined by electron beam lithography), the sidewalls of the etched structures themselves are very smooth as shown in Fig. 2. The roughness is well below

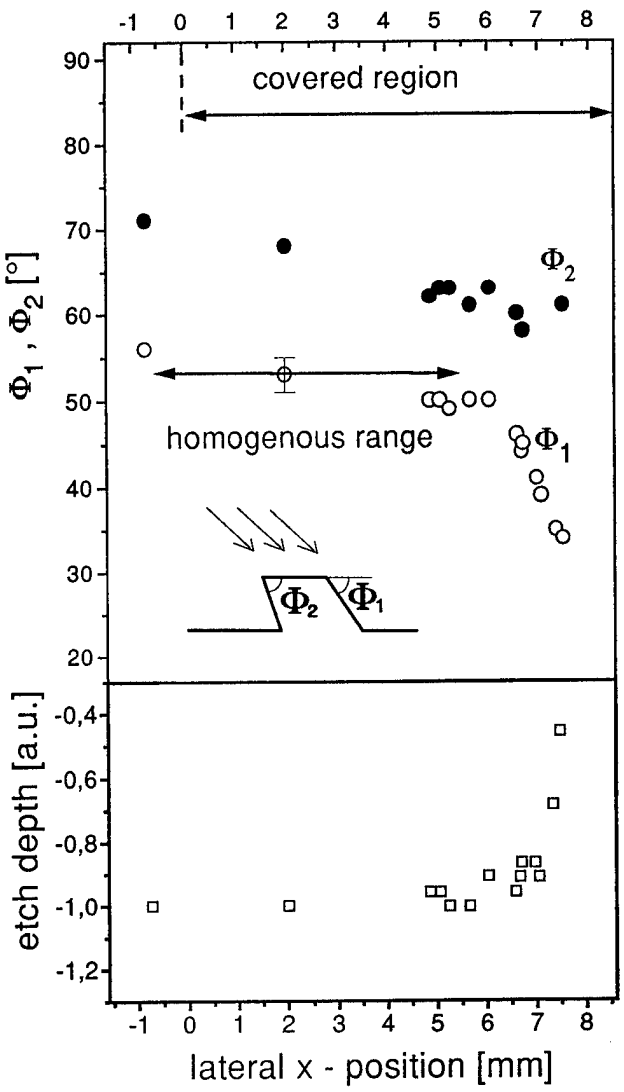


FIG. 3. Characterization of the profile of etched structures as a function of the lateral position.

50 nm and thus the mirrors are even applicable as reflectors for the visible light.

To quantify the etching behavior below the mask, etch depth and tilting angle Φ_1 and Φ_2 of both sidewall faces are plotted in Fig. 3 versus the lateral position of the structure. The definition of the angles is given inside the view graph of Fig. 3. The covered region begins in position $x=0$. From this point, the mask extends up to the right margin of the figure at $x=8.6$ mm. In the lower part of Fig. 3, the curve (open squares) represents the etch depth over this range of x. The depth remains constant along a distance of 6 mm. So in the first 6 mm under the mask no shadowing effect of ions is noticeable, which would lead to a reduction of the etch rate. Only along the last 2 mm up to the face of the aluminum mask, the etch rate drops. In the upper part of Fig. 3, the changes of the two wall angles Φ_1 and Φ_2 of the structure are shown. Φ_1 (open circles) corresponds to the wall, which is turned away from the ion direction. There is only a slight linear decay of Φ_1 with increasing x value around a value of 53°. The wall angle Φ_2 on the opposite side of the profile (filled circles) also decreases linearly; however, it is shifted towards an angle 15° higher. So the sidewalls are not parallel to each other. For patterns etched under the same conditions at normal incidence without using a shadow mask, both sidewalls of the resulting structures are symmetrical to the center of the line structure; however, they are also not parallel to each other. As defined in Fig. 3, this corresponds to a difference between the angles, $\Phi_2-\Phi_1>0$. This difference is slightly lower than that by using a shadow mask, as shown in an earlier publication.[9] Corresponding to the gradient of both curves a homogenous etching range of 5 mm length under the mask is achievable with an uncertainty of the facet angle of less than 2°. This is better than the technologically useful etching range, which results from the published properties of angled RIE with tilted sample holders.[5] In Fig. 3, the useful homogeneous range of 5 mm is marked by a horizontal line. The adjustment of the facet angle is mainly determined by the mask geometry.

We want to point out that a sidewall angle slightly different from 45° is still attractive for the fabrication of 90° deflecting mirrors as small beam corrections can be done externally by a lens array. In a three-dimensional architecture of optical interconnects with surface emitting lasers, a lens array between two optical chips usually serves to focus the single beam on a small area of the next stage. By asymmetric adjustment of the lens array, a compensation for a deviation from the ideal angle around 10° can be easily achieved. Consequently the angles achieved here for Φ_1 are in fact useful for deflecting mirrors.

Let us now consider the influence of the mask geometry (e.g., its total height) on the resulting pattern profiles. To illustrate the limit of an observable mask influence on the ion flux, first only vertical walls serve as "shadow masks." The total height of the wall was varied between 3 mm and 9 mm. In Fig. 4(a) the results of the sidewall face angles are shown versus the lateral positions of the structures. The straight, solid, horizontal line in the figure marks the wall angle of

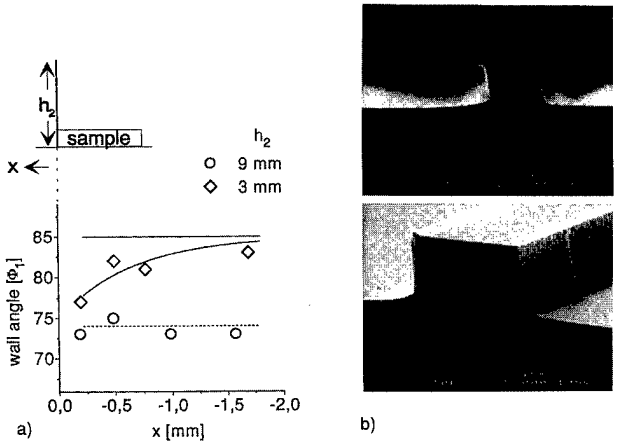


FIG. 4. (a) Influence of a vertical aluminum sheet on the facet angle in a RIE process. (b) Photographs of the corresponding etched structure at $x=-1.6$ mm (height of the shadow mask: 9 mm).

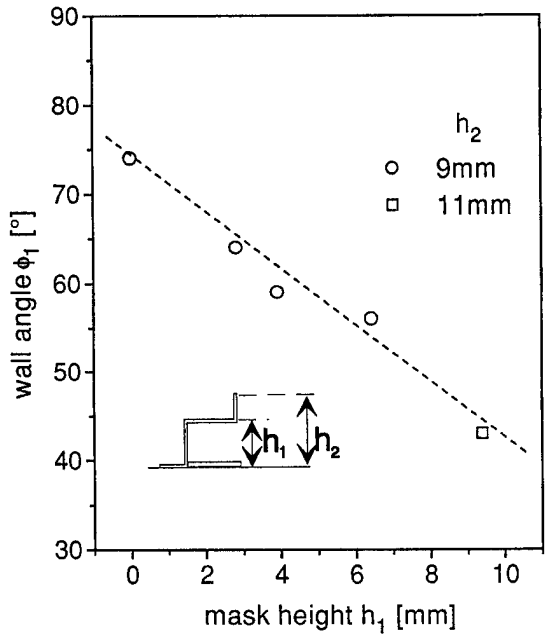


FIG. 5. Wall angle of etched structures as a function of the mask height.

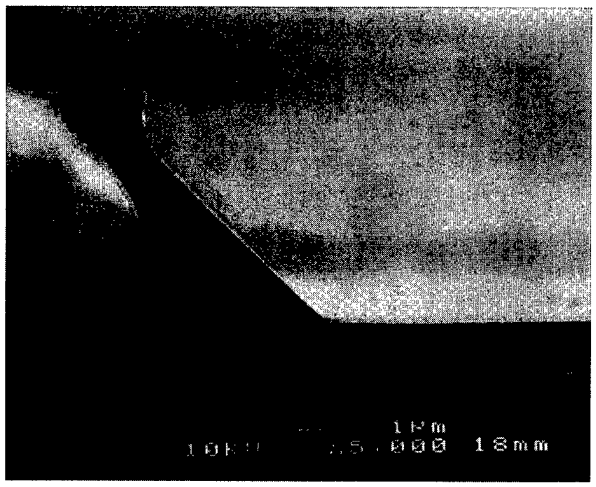


FIG. 6. Profile of a deep etched structure with a facet angle of 45°.

structures etched without shadow mask but using identical etching parameters as described above. For a height h_2=3 mm (represented by the circles) the asymmetric etching behavior due to the field influence of the conductive aluminum disappears already at a distance of 2 mm away from the wall. So the effect of this mask on the etching behavior is almost negligible. However, for h_2=9 mm, all structures have an identical slope of ≈74° over the same distance.

For the definition of the structures, which are discussed above [Fig. 4(a)], the organic resist AZ1512 was applied. The slightly inclined profile of such an etched structure as well as the surface quality of the sidewalls are shown in Fig. 4(b). As shown in the upper photo, very smooth, mirrorlike facets can be obtained also with an organic resist. Only on the top of the sidewall does a weak roughness occur. As can be seen in the lower photo, the erosion of the resist is strongly asymmetric. Especially the right side of the resist is removed by the sloped incident ions. So the etch selectivity achieved by using such a resist may not be sufficient for the required etch depth of several μm for mirrors.

Besides the total mask height h_2, the height of the mask cover, defined by h_1 in Fig. 1, mainly determines the slope of the etched structures. To illustrate the dependence between sidewall slope and mask geometry, the wall angles Φ_1 of structures, placed in comparable positions, are plotted versus the cover height h_1 in Fig. 5. The angle of the considered structures is related to the position where the etch rate is reduced to 80%, indicating that the ion flux starts to be shadowed noticeably by the aluminum cover. Up to this point the wall angle Φ_1 varies only weakly, as shown in Fig. 3. The viewgraph shows a linear decay of Φ_1 with increasing height from 74° to 43° for the mask 9 mm in height. The height of the plasma sheath is roughly 2 cm. So, in principle, the wall angle should be adjustable in a much larger range as shown; also below 40°, as long as the dimension of the mask does not exceed the plasma sheath. However, in our case the variation range of the mask was limited by the opening between transfer chamber and reactor of our equipment.

Considering the described height restriction to the applicable mask geometry, we finally want to demonstrate that the presented technique enables us to etch sloped facets with an ideal 45° angle for perpendicular deflection with a sufficient tolerance of 0.5 mm for the lateral adjustment, according to the assumed uncertainty of the angle as described above. Figure 6 shows the cross section of a deeply etched 90° deflecting mirror with high aspect ratio. An Al_2O_3 film was used as a pattern mask for this structure. The useful range for deflecting extends about 8 μm. To study the deflecting quality of the facet, the surface was evaporated with a 50 nm aluminum layer.

The optical quality of the mirror can be characterized by the distortion and the loss of a light beam after reflection. Therefore we measured the intensity distribution of the far field of a Gaussian beam, reflected at one of the deep etched micro mirrors. The measurement setup is shown in Fig. 7. We operate with visible light of a HeNe laser at 633 nm. The incident beam is focused by a microscope objective onto the mirror with a numerical aperture of 0.1. The width of the beam waist of 4 μm is much smaller than the useful deflecting range on the mirror of 8 μm, so that the edges of the mirror do not effect the shape of the reflected intensity distribution. To simulate the propagation of a beam, just leaving a waveguide, like in a surface emitting laser structure, the focus is pulled back from the mirror plane by a distance z_{focus} as shown in the enlarged part of Fig. 7. The incident beam is tilted against the sample around an angle of 15° to prevent a shading of the beam by the bottom of the sample. The far field is measured 110 mm away from the mirror by a scanning photodiode.

In Fig. 8 the far field intensity in direction parallel (w_h) and perpendicular (w_v) to the long side of the mirror, is plotted for a distance between focus and mirror z_{focus} of 10 μm, corresponding to the dimensions realized in surface

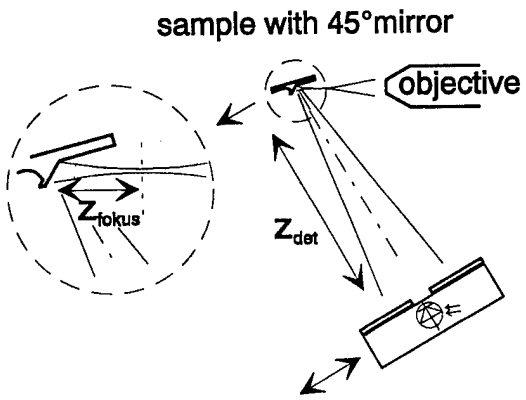


FIG. 7. Setup for the measurement of the far field intensities of reflected a Gaussian beam.

emitting structures.[10,11] In comparison to the intensity distribution, taken directly in the mirror plane by a video camera, as shown in the inset of Fig. 8, there is no significant distortion. The reflection loss relative to an undisturbed reference beam amounts to 0.6 dB. Considering the aluminum reflectivity of 90% at λ=0.63 μm, the additional loss, probably due to surface roughness, is 0.14 dB.

To characterize precisely the curvature of the mirror plane, causing a potential distortion of the beam, we have

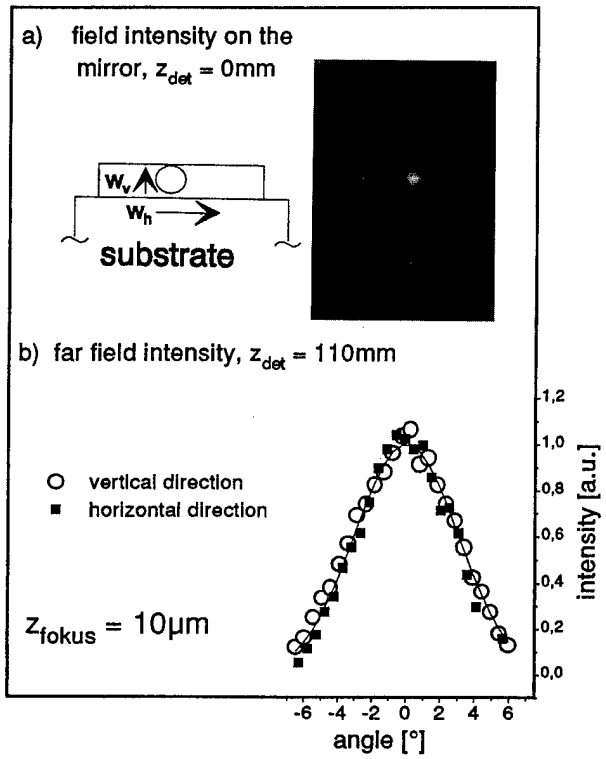


FIG. 8. Far field intensity of the reflected beam; inset: field intensity in the mirror plane.

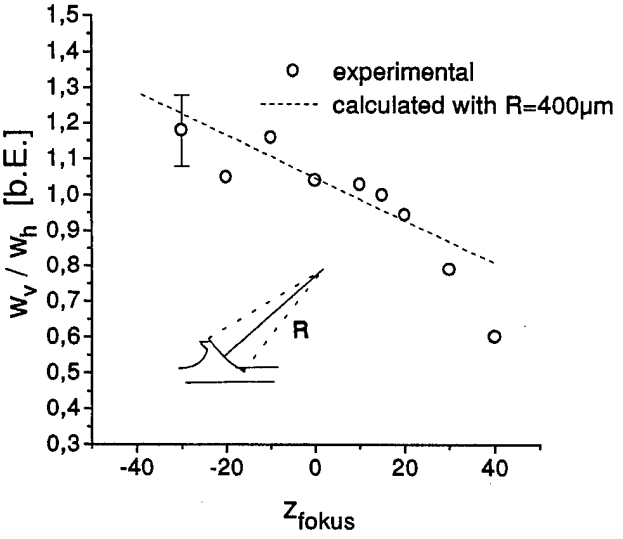


FIG. 9. Ratio of the beam diameter perpendicular and parallel to the long side of the micro mirror as a function of z_{focus}.

measured the $1/e$ widths w_v and w_h of the far field in both directions of the phase front in dependence on z_{focus}. The ratio w_v/w_h is plotted in Fig. 9 versus z_{focus}. By fitting the calculated relationship, derived from Gaussian beam optics, to the measured data, the radius of curvature R_m is determined to be 400 μm. Corresponding to this slight deviation from an ideal straight plane, no serious distortion of the reflected beam can be observed up to a distance z_{focus}=30 μm, when the beam diameter on the mirror becomes larger than the extension of the mirror plane.

IV. CONCLUSIONS

In summary, we have presented investigations of angled etching with a RIE process by deflecting the ions inside the plasma sheath directly with a conductive mask. We have shown that such a simple addition in the usual RIE equipment allows strongly sloped structures to be etched with wall angles as small as 40°. Besides this the methane-based etch chemistry provides mirrorlike sidewall faces. Deep etched structures with 45° angles show nearly a straight profile, which is suitable for application as a deflecting mirror. From optical measurements, performed with aluminum-evaporated structures, a minimum radius of curvature of the deflecting plane of 400 μm is determined. The study of the optical quality of such micro mirrors has shown that the angle dependent intensity distribution of the far field of the reflected beam is the quite identical to that of the incoming beam. The reflection loss amounts 0.6 dB. In principle, the wall angle of the facets can be adjusted by the height of the mask. The extension of the region with homogeneous etching conditions depends also on the total mask height. For a mask height of 9 mm the slope of etched structures varies with an uncertainty of 2° around an angle of 53° over a distance of 5 mm. This is better than the homogeneity obtained by angled RIE with a tilted sample holder.

[1]J. A. Bondur and R. G. Frieser, in *Plasma Processing*, edited by R. G. Frieser and C. J. Mogab (Electrochemical Society, Pennington, NJ, 1981), Proc. Vol. 81-1, p. 180.
[2]C. M. Horwitz, Appl Phys. Lett. **44**, 1041 (1984).
[3]L. A. Coldren and J. A. Rentschler, J. Vac. Sci. Technol. **19**, 225 (1981).
[4]T. Takamori, L. A. Coldren, and J. L. Merz, Appl. Phys. Lett. **53**, 2549 (1988).
[5]J. M. Kim, W. N. Carr, R. J. Zeto, and L. Poli, J. Electrochem. Soc. **139**, 1700 (1992).
[6]B. Jacobs and R. Zengerle, Electron. Lett. **31**, 794 (1995).
[7]C. Constantine, D. Johnson, S. J. Pearton, U. K. Chakrbarti, A. B. Emerson, W. S. Hobson, and A. P. Kinsella, J. Vac. Sci. Technol. B **8**, 596 (1990).
[8]K. H. Kretschmer, K. Matl, G. Lorenz, I. Kessler, and B. Dumbacher, Solid State Technol. **33**, 53 (1990).
[9]R. Zengerle, H. J. Brückner, B. Hübner, and W. Weiershausen, J. Vac. Sci. Technol. B **11**, 2641 (1993).
[10]J. H. Kim, R. J. Lang, and A. Larsson, Appl. Phys. Lett. **57**, 2048 (1990).
[11]G. A. Porkolab and E. D. Wolf, J. Vac. Sci. Technol. B **11**, 2552 (1993).

Comparison of the physical and electrical properties of electron cyclotron resonance and distributed electron cyclotron resonance SiO_2

M. Firon, M. C. Hugon, and B. Agius
Université de Paris Sud, Plateau du Moulon, BP 127, Orsay 91403, France

Y. Z. Hu, Y. Wang, and E. A. Irene
Department of Chemistry CB No. 3290, University of North Carolina, Chapel Hill, North Carolina 27599-3290

(Received 15 November 1995; accepted 10 May 1996)

A comparison was made of thin films of silicon dioxide deposited, at floating temperatures, using electron cyclotron resonance plasma enhanced chemical vapor deposition (ECR PECVD) and distributed electron cyclotron resonance plasma enhanced chemical vapor deposition (DECR PECVD). The refractive index, composition, and chemical bonding of the plasma oxides were determined by null and spectroscopic ellipsometry, nuclear reaction analysis, and Fourier transform infrared spectroscopy and were compared with thermal oxides. The damaged layer at the Si/SiO_2 interface resulting from ECR and DECR techniques was evaluated by spectroscopic ellipsometry. Finally, high frequency and quasi-static capacitance voltage characteristics and ramped current voltage measurements were performed to determine the electrical properties of the ECR and DECR deposited silicon oxide. Device quality SiO_2 thin films have been prepared using both deposition techniques: low interface state density [5×10^{10} $eV^{-1}\,cm^{-2}$ (ECR) or 2.5×10^{10} $eV^{-1}\,cm^{-2}$ (DECR)], and high critical field [5.2 MV/cm (ECR) or 6 MV/cm (DECR)] have been achieved.
© *1996 American Vacuum Society.*

I. INTRODUCTION

There exists a need for a low temperature process (<200 °C) for depositing high quality SiO_2 thin films for microelectronic applications[1] such as: passivation coatings, interlevel dielectrics, and gate dielectrics in metal–oxide–semiconductor field effect transistors (MOSFETs) and thin film transistors (TFTs).

Many attempts to develop an acceptable low temperature oxide deposition process have been made using both radio frequency plasma enhanced chemical vapor deposition (rf-PECVD)[2] and electron cyclotron resonance PECVD (ECR-PECVD).[3] The rf-PECVD technique incorporates considerable hydrogen (5–30 at. %)[4] and causes ion beam damage due to the high energy ions (100–1000 eV). The ECR technique uses a microwave frequency of 2.45 GHz to excite the plasma and the resonance condition is established by applying a 875 G magnetic field in the region of the plasma. The ECR-PECVD conditions enable the achievement of a high ion density plasma (10^{11} cm^{-3}) at lower pressures (0.05–0.5 Pa) than conventional rf-PECVD techniques.

This investigation considers the deposition of SiO_2 films by ECR-PECVD with both a classical reactor configuration and with the distributed system.[5–8] The goal of this research is to compare the physical and electrical properties of ECR and DECR deposited SiO_2 films. In a previous study,[9] we compared ECR and DECR oxidation mechanisms of Si and the resulting SiO_2 properties, and we reported that the ECR and DECR deposition techniques led to similar physical and chemical properties.

II. EXPERIMENTAL PROCEDURES

ECR-PECVD was performed in a custom built ECR plasma system.[10] The system consists of an independent vacuum processing chamber equipped with an ECR plasma source yielding 2.45 GHz microwaves, which are guided through a quartz window to the ECR plasma processing chamber. The vacuum chamber was evacuated to 10^{-5} Pa prior to processing using a turbomolecular pump. The DECR reactor, used for this study, was provided by Alcatel and has been described previously.[5–8] Briefly, the reactor consists of a 300 mm diam chamber, equipped with a load-lock; the residual pressure in both chambers is about 5×10^{-5} Pa. The microwave power is coupled to the plasma through eight antennas evenly distributed at the periphery of the deposition chamber.

Si samples used were commercially available [n-type (100) c-Si (resistivity: 1–2 Ω cm) for ECR and p-type (100) c-Si (resistivity: 8–10 Ω cm) for DECR]. Si samples were cleaned by a modified RCA procedure[11] with a final HF dip and de-ionized water rinse. After the cleaning procedures, the Si samples were blown dry using dry nitrogen and then loaded immediately into the load lock, and subsequently transferred to the main processing chamber. Ge and GaAs wafers were used for the determination of Si atomic composition.

For ECR experiments, diluted SiH_4 (3% SiH_4 in Ar) was used for the Si source. In all the work reported here concerning the ECR deposition, the microwave power was 300 W and the total gas flow rate was 30 sccm. For the DECR work, the total gas flow of undiluted SiH_4 and O_2 (20 sccm) and the microwave power (800 W) were also kept identical for all

films. We studied the effect of oxygen/silane flow ratios (R_O) (between 2 and 9) on physical and electrical properties of the deposited SiO_2 films, using the two microwave deposition techniques.

The ionic current for the oxygen plasma, as shown in Fig. 1(a) for ECR and Fig. 1(b) for DECR, was determined as a function of pressure. To obtain these characteristics, the substrate holder acted as a large planar Langmuir probe. Different behavior was observed: in the case of the ECR plasma the ion flux decreases as a function of the pressure whatever was the value of the microwave power, while for the case of the DECR plasma a maximum in the ion flux is found for a pressure in the range 0.1–0.6 Pa. In order to compare the properties of ECR and DECR plasma deposited films, the value of the ionic current should be the same. Since it is not possible to obtain ionic current characteristic curves for the O_2/SiH_4 mixtures (due to reaction), we use the oxygen plasma characteristics as an approximation. Thus, we have chosen a total pressure of 0.1 Pa and a microwave power of 300 W for the ECR method (2.35 mA/cm^2), and 800 W for the DECR method (2.1 mA/cm^2).

Film thicknesses and refractive indexes were measured at near half an ellipsometric period (80–100 nm) using null ellipsometry at 632.8 nm light for the ECR samples and 546.1 nm for the DECR samples. The refractive index was deduced using a single film model and assuming no absorption ($k=0$). The optical properties of the SiO_2 films have also been studied using spectroscopic ellipsometry (SE) operating between about 2 and 4 eV for ECR films and DECR films. The measured dielectric spectra in terms of ε_1 and ε_2 were modeled using the Bruggeman effective medium approximation (BEMA), which Aspnes and co-workers[12] and Collins and co-workers[13] have shown to be useful to characterize damage at the Si surface. To limit the number of parameters, we have assumed that the oxide film can be described by no more than two layers (interface layer and oxide layer) and that each layer was composed of no more than two of three possible components: a-Si, SiO_2, and voids, as we had done previously for the ECR-DECR oxidation study.[9] The layer thicknesses and the chemical compositions for each model were fitted using a linear regression analysis to provide the optimum agreement with experiment. The relative quality of the fit for each model is defined by an unbiased estimator δ. The δ formulas, δ_E for ECR samples, and δ_D for DECR samples, are

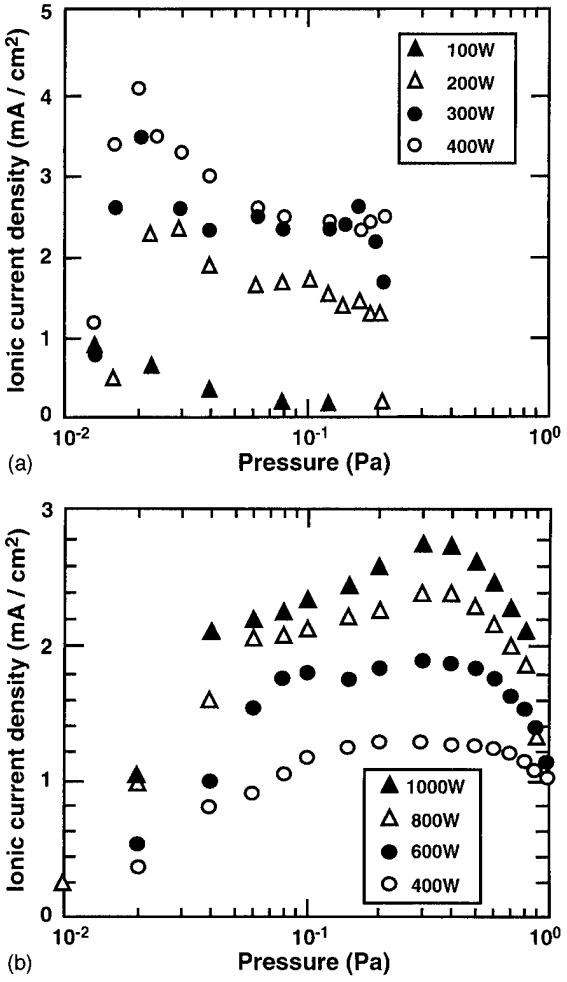


FIG. 1. Ionic current density as a function of the pressure for O_2 plasma for (a) the ECR technique and (b) the DECR technique.

$$\delta_E = \sqrt{\frac{1}{N-P-1}\sum_{i=1}^{N}\left[(\tan\psi_{\exp}-\tan\psi_{\mathrm{cal}})_i^2+(\cos\Delta_{\exp}-\cos\Delta_{\mathrm{cal}})_i^2\right]}, \tag{1a}$$

$$\delta_D = \sqrt{\frac{1}{2N}\left[\omega_\psi\sum_{i=1}^{N}(\psi_{\exp}-\psi_{\mathrm{cal}})_i^2+\omega_\Delta\sum_{i=1}^{N}(\Delta_{\exp}-\Delta_{\mathrm{cal}})_i^2\right]}, \tag{1b}$$

where N is the number of data points and P the number of unknown parameters. ω_Δ and ω_ψ are the weighting coefficients for Δ and ψ, respectively. Thus, the values for δ are only useful for comparison within a set of samples.

To confirm the ellipsometry results, the relative composition of Si and O was evaluated by nuclear reaction analysis (NRA) [^{28}Si(d,p), and ^{16}O(d,p), respectively],[14–16] while the hydrogen contamination was determined by the elastic recoil detection analysis technique (ERDA).[17]

Infrared measurements were performed using a Fourier transform infrared spectrophotometer operated with a mercury cadmium telluride (MCT) detector. Normal incidence

transmission was measured on the bare Si substrate, as background, and on the SiO_2 covered Si with a resolution of 2 cm^{-1}.

The etch rate of the films was examined using "*P*-etch" solution [HF(40%):HNO_3(65%):H_2O of 3:2:60] as reported by Pliskin.[18]

Al/SiO_2/Si MOS structures were prepared for $I-V$ and $C-V$ characterizations. Following film deposition, aluminium electrodes typically 200 nm thick with an area of 9 10^{-4} cm^2 for ECR films, and 500 nm thick with an area of 3 10^{-3} cm^2 for DECR films, were evaporated through a shadow mask. For the ECR films, the backside contact was made by abrading the wafer and applying a GaIn paste. For the DECR samples, Al was evaporated on the back side of the substrate. After metallization, the samples were subjected to postmetallization annealing in forming gas (10% H_2 in Ar for DECR samples and 10% H_2 in N_2 for ECR samples) at 450 °C for 30 min. The quasistatic $I-V$ characteristics were obtained with a HP 4140B picoammeter with the device biased in accumulation (voltage ramp of 50 mV/s). High frequency (1 MHz) and quasistatic (voltage ramp of 10 mV/s and 5 mV/s for ECR and DECR samples, respectively) $C-V$ curves were performed using an impedance analyzer and a picoammeter.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Oxide physical properties

The deposition rate (d_r), plotted as a function of the O_2/SiH_4 flow ratio, is shown in Figs. 2(a) for ECR and 2(b) for DECR. The same behavior is observed for the two deposition methods: d_r decreases when the O_2/SiH_4 flow ratio increases from 2 to 9. The depositions in the ECR and DECR-PECVD reactors involve chemical processes occurring in the gas phase and/or on the substrate surface. Under our experimental conditions of low operating pressure (0.1 Pa), the mean free path of reactive species is of the same order of magnitude as the dimensions of the reactor. This suggests that the chemical reactions between the silane and oxygen species occur on the substrate surface rather than in the plasma. The insets of Figs. 2(a) and 2(b) show a near linear relationship between d_r and the silane flow rate which indicates that the deposition is under a mass-transport control.

Figure 3 shows the refractive index versus O_2/SiH_4 flow ratio for (a) ECR-PECVD samples, and (b) DECR-PECVD samples. The dotted lines are from a thermal SiO_2 sample grown in dry O_2 at 1000 °C, which has a refractive index $n=1.465\pm0.002$. For both ECR and DECR deposition, the refractive index values are close to the thermal SiO_2. When the O_2/SiH_4 flow ratio is 9, the value of the refractive index is the closest to that of thermal SiO_2. As the O_2/SiH_4 flow ratio decreases, the refractive index increases indicating excess Si in the films.

For both the ECR and DECR depositions, we have found that the best fit optical model is a one layer model. Figure 4 shows the best optical models for ECR and DECR-PECVD samples deposited at different O_2/SiH_4 flow ratios. For

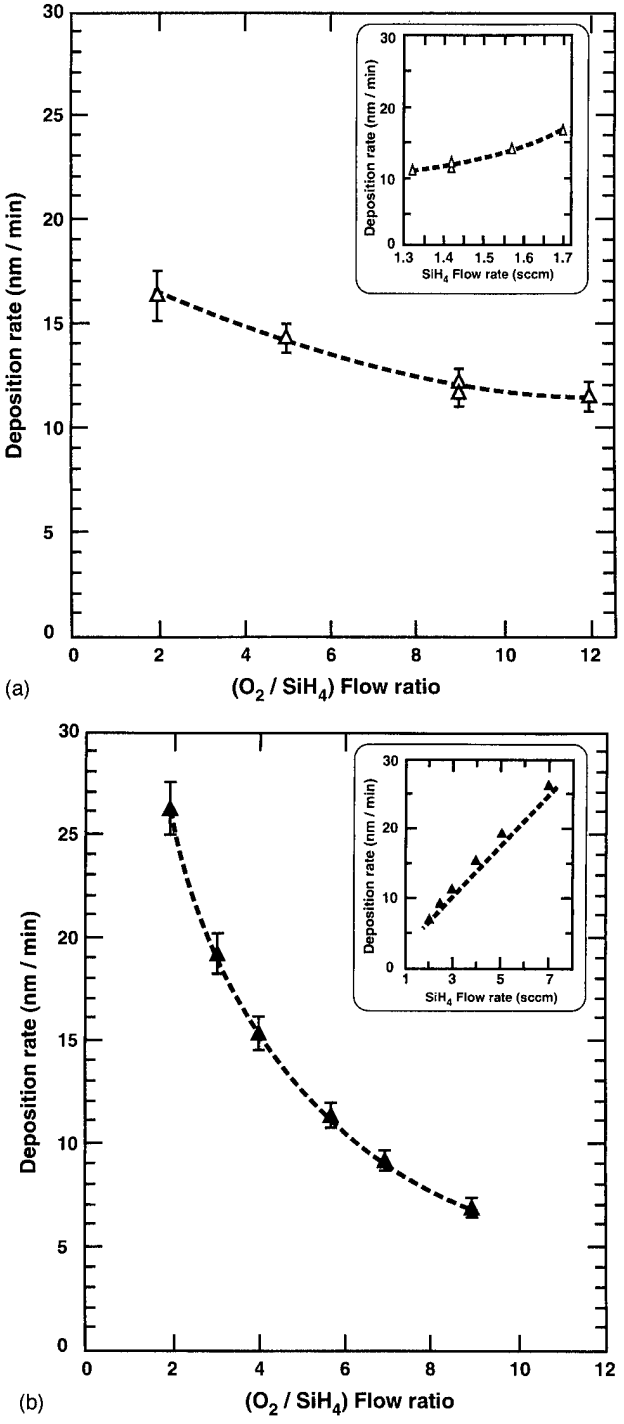


FIG. 2. Deposition rate as a function of the O_2/SiH_4 flow ratio and SiH_4 flow (inset) for (a) the ECR technique and (b) the DECR technique.

higher R_O values, SE indicates nearly 100% SiO_2 while for $R_O=2$ both DECR and ECR display measurable excess Si content. In addition SE indicates a small amount of excess Si in the DECR samples even at high R_O values. This is confirmed below with the Si analyses using NRA. Also, the volume percentage for *a*-Si decreases as O_2/SiH_4 ratio increases. This is in agreement with null ellipsometry results

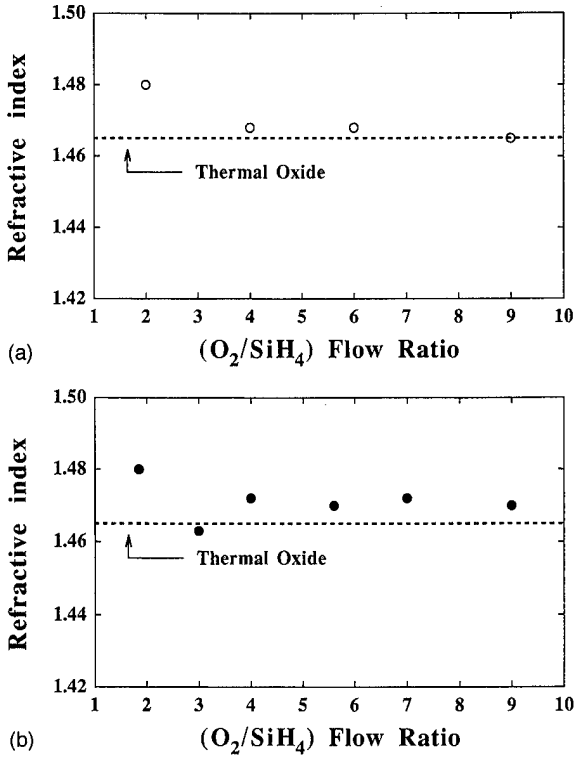


FIG. 3. Refractive index as a function of the O_2/SiH_4 flow ratio for (a) the ECR technique (632.8 nm) and (b) the DECR technique (546.1 nm).

which show that as the O_2/SiH_4 ratio decreases, the refractive index increases.

In order to confirm the film composition deduced from SE measurements, we have studied the composition by NRA and

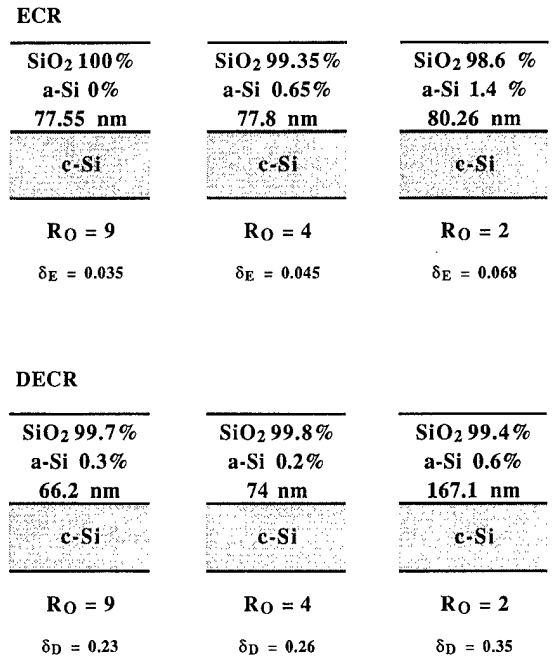


FIG. 4. Optical models calculated from SE data for different O_2/SiH_4 flow ratios for the ECR and DECR techniques.

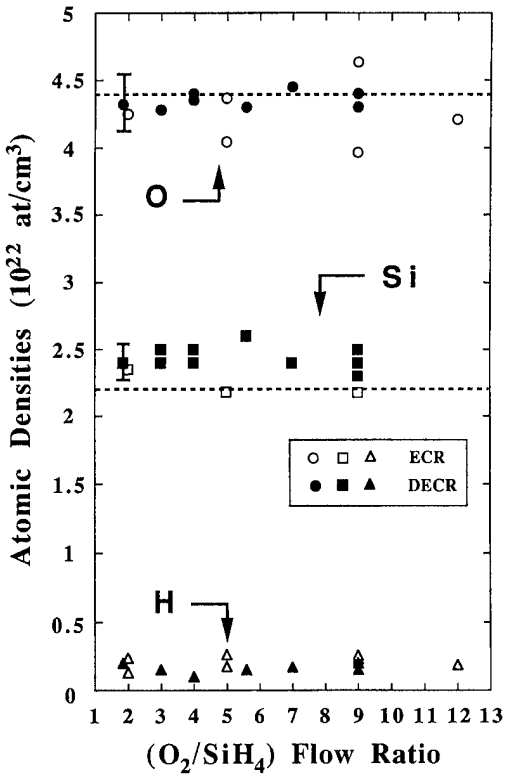


FIG. 5. Atomic densities of Si, O, and H in silicon oxide as a function of flow ratio for the ECR and DECR techniques (the dotted lines indicate atomic densities of Si and O in the thermal oxide).

ERDA. Films were prepared with a wide range of oxidant ratios and the results are illustrated in Fig. 5 where atomic densities of Si, O, and H in the films as a function of the gas phase composition are reported. These densities in at. cm^{-3} are deduced from NRA (accuracy: ±5 at. %) or ERDA (accuracy: ±5 at. %) and null ellipsometry measurements. NRA measurements are typically performed on GaAs or Ge substrates, and equivalent O content is found using Si substrates. It can be seen that O and Si atomic densities are close to those for thermal oxide (dotted lines). NRA confirms the null and spectroscopic ellipsometry which indicates that the DECR samples are slightly silicon rich (Fig. 6). In addition Figs. 5 and 6 indicate that the ECR stoichiometry is also less than 2 for O, but this is due to a slight oxygen deficiency. SE shows stoichiometry for high R_O samples. The formation of slightly silicon-rich DECR films could be due to an incomplete reaction, or the fact that oxygen is more volatile than silicon. From the two following experimental data: thicknesses of the films measured by ellipsometry, and silicon, oxygen, and hydrogen contents of the films measured by NRA or ERDA, we can deduce the mass density d which is shown in Fig. 6. The ECR and DECR film density is close to that for thermal oxide and displays no trend with composition.

Except for H, no other impurities were detected in any of the oxides. H is a frequent contaminant in the low temperature deposited SiO_2 films. The ERDA measurements gave a low hydrogen concentration (2 at. %) whatever the value of the O_2/SiH_4 ratio.

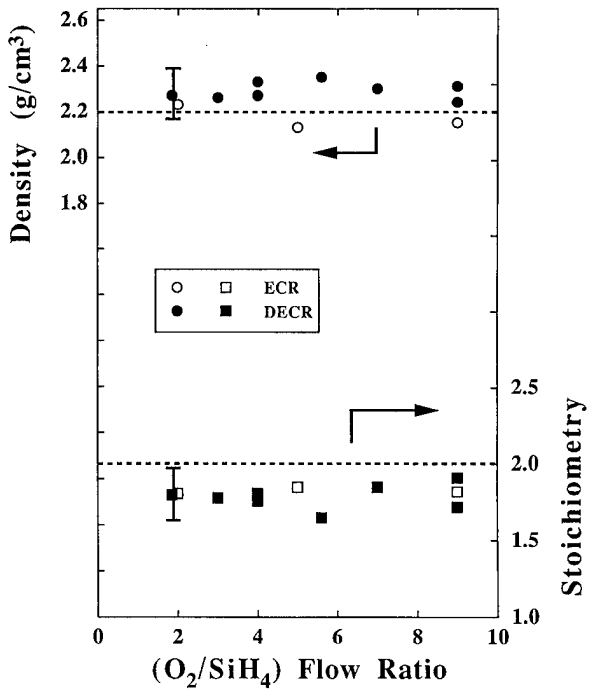


FIG. 6. Density and stoichiometry of ECR and DECR films as a function of gas flow ratio.

There are three characteristic infrared (IR) absorption bands for thermal SiO_2 near 450, 810, and 1075 cm^{-1} which correspond to rocking, bending and stretching vibration modes of Si–O–Si, respectively.[19] The Si–O stretching band position changes with the stoichiometry (x) of silicon oxide (SiO_x) deposited by plasma. When x is 1.9, the Si–O band frequency is about 1060 cm^{-1}, and it decreases as x does.[20] The position of this band is usually associated with the change of the oxygen bridge angle (θ). In force constant (k) models for vibrational properties, the frequency (ν) of Si–O–Si stretching mode can be expressed as[21,22]

$$\nu^2=(k/m_o)\sin^2(\theta/2), \tag{2}$$

where k is the nearest neighbor effective force constant, m_o is the mass of an oxygen atom, and θ is the O–Si–O bond angle which is assumed to be 144° for thermal SiO_2. When θ decreases, the associated vibrational frequency also decreases.[23]

The position of the Si–O stretching vibration band is shown as a function of O_2/SiH_4 ratio in Fig. 7, for both deposition processes. By way of comparison, the frequency of Si–O band observed for thermal SiO_2 is also reported. For both processes, the Si–O band frequency does not change with O_2/SiH_4 flow ratio. However, this band for DECR samples appears at a lower frequency than that observed for thermal and ECR SiO_2 spectra. This result is in good agreement with previous measurements (SE and NRA) in that the stoichiometry of SiO_2 deposited by DECR is 1.8 (Fig. 6). But ECR and DECR samples have the same stoichiometry. Thus,

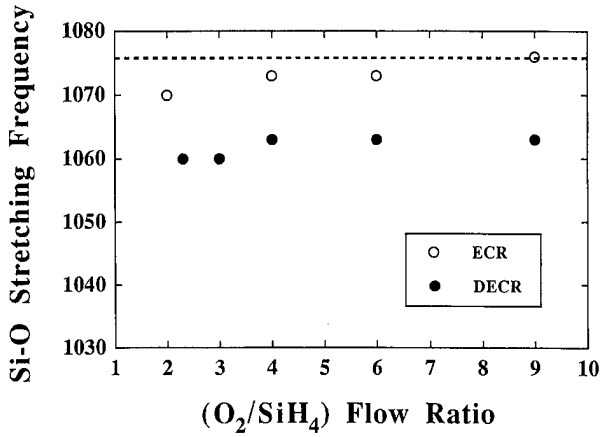


FIG. 7. Si–O–Si stretching frequency as a function of the R_O gas flow ratio for the ECR and DECR techniques. The dotted line indicates the value of thermal oxide Si–O–Si stretching frequency.

the difference between ECR and DECR data could be due to the difference in film density as was shown in Fig. 6.

Overall, our results show no trend of the physicochemical properties of the DECR and ECR films with the O_2/SiH_4 flow ratio. However, we do find a trend in the etch rate of DECR oxides in a HF solution with the gas phase ratios and with a comparatively smaller etch rate for ECR films (Fig. 8). The etch rate of DECR oxides decreases when O_2/SiH_4 flow ratio increases and it is lower for films deposited for $R_O=9$. The P etch rate is often associated with the deposition rate. Higher P etch rates are usually an indication of lower density for the same material. For $R_O<4$, the depo-

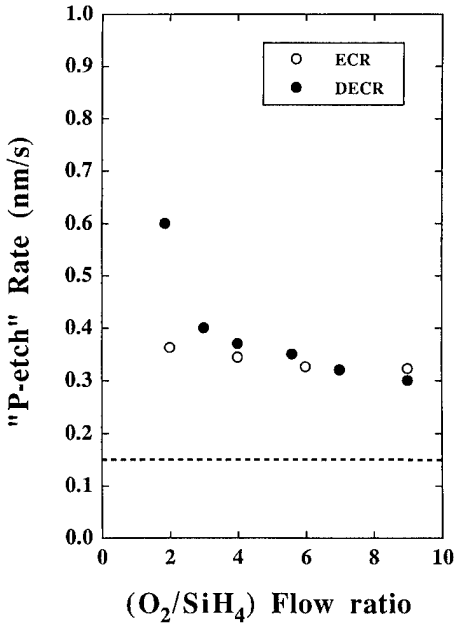


FIG. 8. P etch rate dependence on R_O flow ratio for the ECR and DECR techniques. The dotted line indicates the value of the thermal oxide P etch rate.

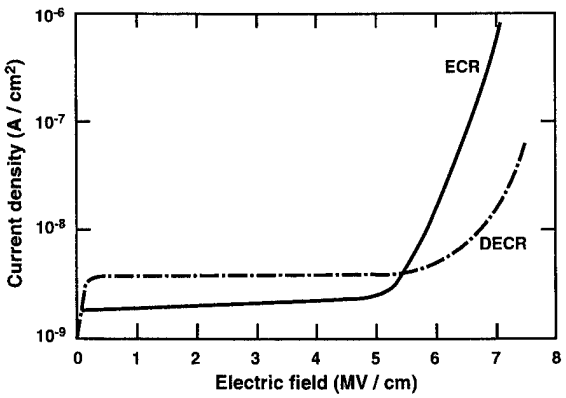


FIG. 9. $I-V$ characteristics of the ECR and DECR films deposited with a O_2/SiH_4 ratio of 9.

sition rate for DECR films is much higher than for ECR films. However from Fig. 6, we see that the DECR films display a larger density, but also a larger P etch rate (for $R_O<4$). Figure 5 shows that the lower x (SiO_x) for DECR is due to higher Si, but for ECR samples to lower O. Hence the point defect structure is decidedly different, and this may be implicated in the etch rate differences for our plasma prepared samples.

B. Oxide electrical properties

The electrical quality of the Si/SiO_2 interface and the SiO_2 film in metal–oxide–semiconductor (MOS) structures are traditionally defined by several measurements including the critical field (E_c), the resistivity (ρ), the interface state density (D_{it}) and the fixed oxide charges (Q_{ox}). E_c and ρ are deduced from quasistatic $I-V$ characteristics.[24] E_c is defined as the field required to pass 1 nA/cm^2 through the Al gate. High frequency and quasistatic $C-V$ curves are used to determine D_{it} and Q_{ox} from the Castagne and Vapaille method.[25]

Figure 9 shows typical $I-V$ characteristics for MOS structures where the SiO_2 was deposited with a O_2/SiH_4 ratio of 9. It is seen that for an electric field $E<5$ MV/cm for ECR films and $E<6$ MV/cm for DECR films, the characteristics are linear indicating ohmic behavior. These characteristics allow a determination of the resistivity (ρ) by

$$\rho=\frac{\Delta V}{\Delta I}\frac{A}{t}, \tag{3}$$

where A is the area of the MOS electrode and t is the thickness of SiO_2 film. Within the limit of accuracy in determining the values of ρ, this parameter which characterizes the bulk SiO_2 film, could be considered to be the same for the two techniques (1–2 10^{16} Ω cm). This result is in good agreement with physico-chemical properties, since we have also detected practically no difference between ECR and DECR films for high O_2/SiH_4 gas ratio films. The current injection from silicon substrate starts at an average field of 5–5.5 MV/cm for ECR films and $E=6-6.5$ MV/cm for DECR films. The critical field characterizes the beginning of

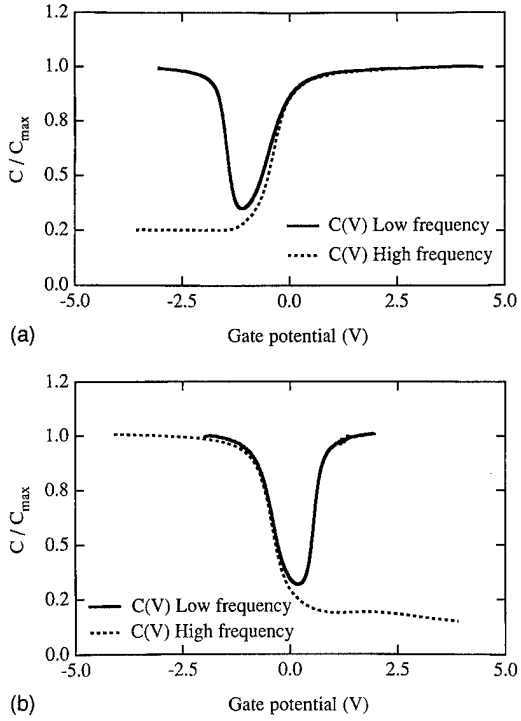


FIG. 10. High frequency (1 MHz) and quasistatic $C-V$ characteristics for O_2/SiH_4 flow ratio of 9 for (a) the ECR technique and (b) the DECR technique.

this injection: DECR films exhibit a slightly higher critical field (6 MV/cm) than the ECR films (5.2 MV/cm).

High frequency (1 MHz) $C-V$ characteristics of the MOS capacitors have been investigated to determine the fixed oxide charges. The flatband voltage as measured from these curves shows a shift towards positive values. By assuming that this shift is caused mainly by the work function difference and negative oxide fixed charges (Q_{ox}), we have estimated Q_{ox} to be about 7×10^{10} cm^{-2} for ECR films and -2×10^{10} cm^{-2} for DECR films for a gas flow ratio of 9.

Figure 10 shows the high frequency (1 MHz) and quasistatic $C-V$ curves for a O_2/SiH_4 flow ratio of 9. Further analysis[25] of these curves yields a midgap interface state density (D_{it}) of 5 and 2.5 10^{10} eV^{-1} cm^{-2} for ECR and DECR films, respectively. $2-3\times10^{10}$ eV^{-1} cm^{-2} is generally considered as the value found in production for Si/SiO_2 interfaces formed by the thermal oxidation of Si. Samples of SiO_2 prepared by other plasma deposition techniques including rf plasma[26] and remote rf plasma deposition[27] have also yielded good interfacial MOS electronic properties. Our electronic results for both ECR and DECR are comparable in terms of D_{it} and Q_{ox}.

In terms of electrical integrity, the previous results show that the ECR and DECR SiO_2 films are similar. These PECVD films present a critical field, a resistivity and an interface state density comparable to those of thermal oxides grown at much higher temperature (1000 °C).

IV. CONCLUSION

A comparison of ECR and DECR plasma CVD of SiO_2 processes using a variety of techniques shows that high qual-

ity deposited dielectric films are achievable with low interface damage and suitable electronic properties for MOS applications. No detectable plasma damage was found at the $Si-SiO_2$ interface. Nearly stoichiometric films were prepared at the higher oxidant ratios, although even these films were slightly off stoichiometry with $x<2$ for SiO_x. It appears that for DECR there was a slight Si excess, but for ECR there was a slight O deficiency when compared to thermal SiO_2. A high P etch rate was found which can rationalized based on the defect nature of the films rather than density. Overall the DECR-PECVD SiO_2 films were slightly electronically superior in terms of critical field and interface trap density, but this may be due to the better processing facilities for the DECR process.

ACKNOWLEDGMENTS

The travel portion of the research was funded by an NSF-CNRS cooperative research grant. The experimental research in the U.S. was funded in part by the National Science Foundation, NSF, and in France by le Centre National de la Recherche Scientifique, CNRS (GDR86).

[1]S. V. Nguyen, J. Vac. Sci. Technol. B **4**, 1159 (1986).
[2]J. Batey, E. Tierney, J. Stasiak, and T. N. Nguyen, Appl. Surf. Sci. **39**, 1 (1989).
[3]T. Hirao, K. Setsune, M. Kitagawa, T. Kamada, K. Wasa, K. Tsukamoto, and T. Izumi, Jpn. J. Appl. Phys. **27**, 30 (1988).
[4]T. V. Herak, T. T. Chau, D. J. Thomson, S. R. Mejia, D. A. Buchanan, and K. C. Kao, J. Appl. Phys. **65**, 2457 (1989).
[5]F. Plais, B. Agius, N. Proust, S. Cassette, G. Ravel, and M. Puech, Appl. Phys. Lett. **59**, 837 (1991).
[6]F. Plais, B. Agius, F. Abel, J. Siejka, M. Puech, G. Ravel, P. Alnot, and N. Proust, J. Electrochem. Soc. **139**, 1489 (1992).
[7]N. Jiang, M. C. Hugon, B. Agius, T. Kretz, F. Plais, D. Pribat, T. Carrière, and M. Puech, Jpn. J. Appl. Phys. A **31**, L1404 (1992).
[8]N. Jiang, B. Agius, M. C. Hugon, J. Olivier, and M. Puech, J. Appl. Phys. **76**, 1847 (1994).
[9]Y. Z. Hu, M. Li, Y. Wang, E. A. Irene, M. C. Hugon, F. Varnière, N. Jiang, M. Froment, and B. Agius, J. Vac. Sci. Technol. B **13**, 227 (1995).
[10]J. W. Andrews, Ph.D. thesis, University of North Carolina at Chapell Hill (1990).
[11]W. Kern and D. A. Puotinen, RCA Rev. **31**, 187 (1970).
[12]D. E. Aspnes, J. B. Theeten, and F. Hottier, Phys. Rev. B **20**, 3992 (1979).
[13]R. W. Collins, J. Vac. Sci. Technol. A **7**, 1378 (1989).
[14]G. Amsel, J. P. Nadai, E. D'Artemare, D. David, E. Girard, and J. Moulin, Nucl. Instrum. Methods **92**, 481 (1971).
[15]I. C. Vickridge, Nucl. Instrum. Methods B **34**, 470 (1988).
[16]T. Carrière, B. Agius, I. Vickridge, J. Siejka, and P. Alnot, J. Electrochem. Soc. **137**, 1582 (1990).
[17]V. Quillet, F. Abel, and M. Schott, Nucl. Instrum. Methods B **83**, 47 (1993).
[18]W. A. Pliskin and H. S. Lehman, J. Electrochem. Soc. **112**, 1013 (1965).
[19]F. L. Galeener and G. Lucovsky, Phys. Rev. Lett. **34**, 1414 (1976).
[20]P. G. Pai, S. S. Chao, Y. Takagi, and G. Lucovsky, J. Vac. Sci. Technol. A **4**, 689 (1986).
[21]P. N. Sen and M. F. Thorpe, Phys. Rev. B **15**, 4030 (1977).
[22]G. Lucovsky, Philos. Mag. B **39**, 513 (1979).
[23]G. Lucovsky, M. J. Mantini, J. K. Srivastava, and E. A. Irene, J. Vac. Sci. Technol. B **5**, 530 (1987).
[24]D. J. DiMaria, R. Ghez, and D. W. Dong, J. Appl. Phys. **51**, 4830 (1980).
[25]R. Castagne and A. Vapaillé, Surf. Sci. **28**, 157 (1971).
[26]J. Batey and E. Tierney, J. Appl. Phys. **60**, 3136 (1986).
[27]G. Lucovsky and D. V. Tsu, J. Vac. Sci. Technol. **85**, 2231 (1987).

Reactive-ion etching of WSi_x in CF_4+O_2 and the associated damage in GaAs

Yi-Jen Chan[a] and Chao-Shin Su
Department of Electrical Engineering, National Central University, Chungli, Taiwan 32054, Republic of China

Kuo-Tung Sung
Mosel Vitelic, Inc., Science-based Industrial Park, Hsinchu, Taiwan, Republic of China

(Received 2 January 1996; accepted 19 April 1996)

Assessments of WSi_x reactive-ion etching in terms of the different CF_4 to O_2 flow rate ratio were characterized. Based upon the evaluations from etching rates, side-wall profiles, surface roughness, and damages, we observed that the optimum etching condition was at a ratio of 10:1. The recovery of reactive-ion-etching-treated GaAs damaged layers through the thermal treatment was also investigated as a function of the annealing temperatures and duration times. These parameter evaluations were for the purpose of achieving a high performance GaAs metal–semiconductor field-effect transistor. © *1996 American Vacuum Society.*

I. INTRODUCTION

Reactive-ion etching (RIE) has been extensively applied to semiconductor device fabrication during the last decade. Compared with the other etching processes, a desired side-wall profile together with a reasonable etching rate and selectivity can be achieved by this technology after optimization, which is not always obtainable from either the wet chemical or the ion milling. The side-wall profile control is essential in maintaining a desired critical dimension, which is especially important for submicron devices. On the other hand, a higher etching rate is necessary for a larger throughput, while a better selectivity is able to avoid the overetch of the subsequent layer. In this study, we used a RIE process to etch away the WSi_x material, where WSi_x was the commonly used material for the gate terminal in self-aligned ion-implanted GaAs metal–semiconcdutor field-effect transistor (MESFET) fabrication. WSi_x films are highly resistive to the conventional wet-etching method, and the formation of WSi_x gates is generally realized by the RIE.[1–3]

Among the various etching chemistries for WSi_x films in the RIE, CF_4 plus O_2 is the widely used etchant gas for this material. In the first part of this study, we particularly focused on the optimization of etching conditions for WSi_x with the different CF_4 to O_2 flow ratio. These evaluations include the etching rate, side-wall profile, plasma damage, and surface roughness. As to the second part of this work, the removal of plasma damage was systematically investigated by the subsequent thermal treatment. This treatment demonstrated an effective way to remove the damage in GaAs active layers. Finally, WSi_x gate GaAs MESFETs were fabricated based on the optimum etching conditions and the following thermal treatment.

II. OPTIMUM CONDITION OF CF_4/O_2 FLOW RATE RATIO

The WSi_x material was cosputtered on the top of n-GaAs layers in an Ar-ambient chamber, and the x value is around 0.3. Due to a higher vapor pressure of WF_6 and SiF_4 at room temperature, fluorine-based enchants such as CF_4 or SF_6 are suitable for WSi_x etching. However, the addition of O_2 in the CF_4 plasma to a certain amount is accompanied by an increase of F radical density, resulting in an increase of WSi_x etching rate.[4] Figure 1 illustrates the WSi_x etching rate and the density of F radical as a function of different CF_4 to O_2 flow rate. The flow rate of CF_4 was 10 sccm, and this rate was fixed through the whole course of this work. F radicals were detected by optical emission spectroscopy in the RIE chamber at a wavelength of 703.7 nm. The etching rate was first increased by increasing the O_2 content, and reaching the highest value of 1600 Å/min at 4 sccm of O_2. By further increasing the O_2 amount, the density of F radical is diluted, and causes a decrease in the etching rate. F radicals at 703.7 nm followed almost the identical tendency as the etching rate curves, and can serve as a good signal to monitor the WSi_x etching process. The etching selectivity between WSi_x and GaAs was 28 at CF_4: O_2 flow rate ratio of 10:1. The etching rate of GaAs is insensitive to the O_2 concentration under such plasma condition, and the rate is around 20 Å/min.

As to the side-wall profile investigation, Figs. 2(a) and 2(b) show the scanning electron microscopy (SEM) cross section of WSi_x gates at an O_2 flow rate of 4 and 7 sccm, respectively, under a RF power of 30 W and a pressure of 50 m Torr. The etching time of each individual process was determined by the etching rate of WSi_x. Therefore, the etching time of O_2 flow rate of 4 sccm was the shortest.

The O_2 flow rate of 1 sccm demonstrated similar side-wall profiles as in the case of 7 sccm. The top Al metal was used as an etching mask to replace the photoresist, which was heavily attacked by the plasma bombardment. In the case of 4-sccm O_2 flow rate, which also showed the highest etching rate of WSi_x in Fig. 1, a large undercut profile with an angle of 28° was obtained. As to the examples of both 1 and 7 sccm, this undercut profile was reduced, and an ideal vertical side-wall profile can potentially be obtained. This change in side-wall profiles associated with O_2 concentrations is consistent with the results presented in Fig. 1; namely a higher F

[a]Electronic mail: yjchan@mbox.ncuee.ncu.edu.tw

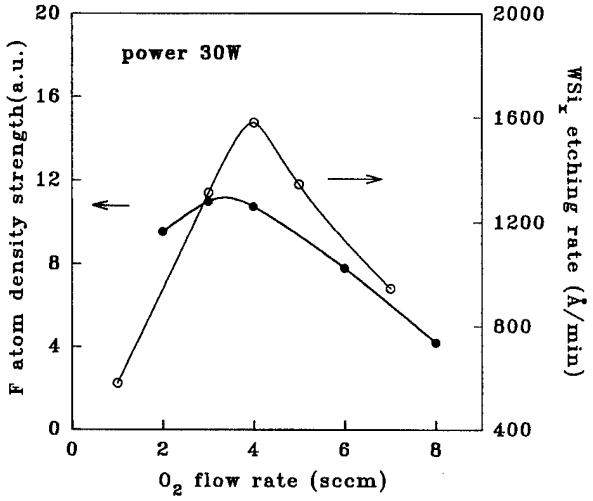


FIG. 1. WSi_x etching rate and associated F atom in RIE chamber as a function of O_2 flow rates.

radical density attacks WSi_x side walls more seriously resulting in a larger undercut angle. The dc bias was insensitive to the O_2 content in the RIE chamber, and the effect causing from pure ion bombardment can be ruled out. Therefore, either in a lower or a higher O_2 flow rate condition maintains a better WSi_x side wall.

Plasma damage caused in a RIE chamber is a potential problem, which will degrade the device performance. These damages may be correspondent to the structural deformation, stoichiometric or impurities.[5–7] The next characterization is associated with the plasma damage in the GaAs surface. We evaluated the sheet resistance (R_{sh}) in a GaAs active layer by means of a transmission-line method (TLM). The TLM patterns were defined by photolithographic techniques, and ohmic contacts were realized by evaporating the Ge–Au–Ni metal, followed by a 350 °C 1-min hot-plate annealing. Figure 3 shows the R_{sh} change in n-GaAs layers as a function of O_2 content after being exposed in the RIE chamber for 5 min. Each data point represents at least an average of five points, located in different areas. The original value of R_{sh} was 490 Ω/sq without the RIE treatment. The increase of R_{sh} indicates the enhancement of surface damage, resulting from the plasma bombardment in GaAs active layers.[8,9] This damage reached the highest value at O_2 flow rate of 4 sccm, and either increasing or decreasing the O_2 flow rate caused a decrease of surface damage. R_{sh} was 628, 756, and 500 Ω/sq corresponding to the O_2 flow rate of 1, 4, and 7 sccm, respectively. At 4-sccm O_2 flow rate, the WSi_x etching rate is the fastest and may induce a rougher surface morphology.[10,11] A rougher surface degrades the electron mobility, which makes a higher R_{sh}. Using the atomic force microscopy (AFM), GaAs surface etched under a 4 sccm O_2 did show a rougher morphology, and this issue will be further discussed in the next paragraph.

Based on the previous investigation, the O_2 content in a RIE chamber will affect the R_{sh} and may be related to the GaAs morphology. Therefore, the final part of this section is

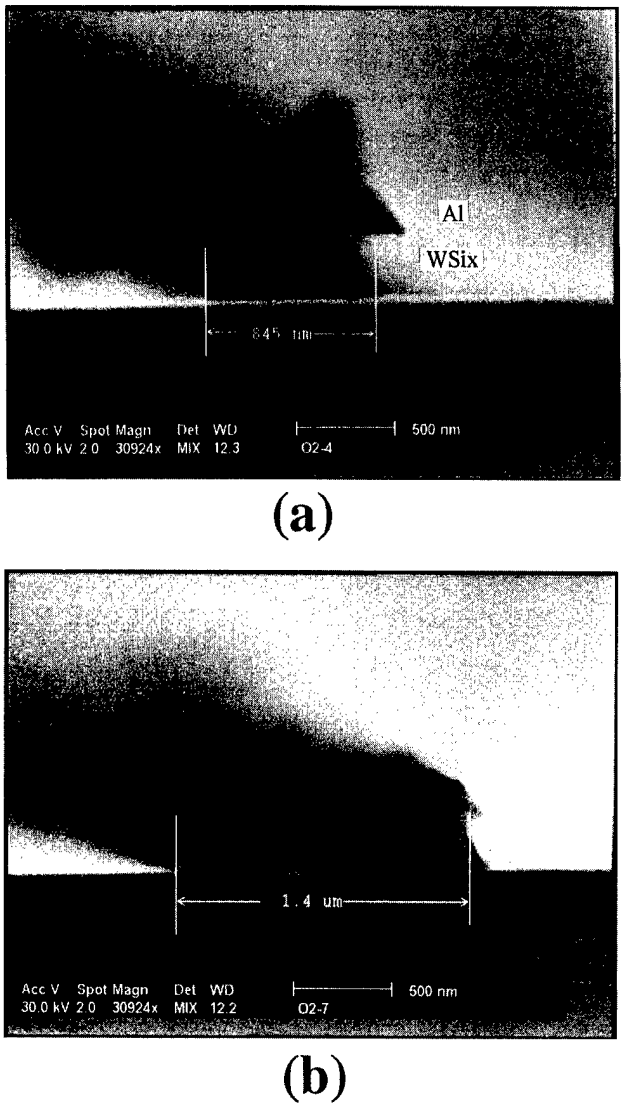


FIG. 2. SEM cross section of WSi_x gate profiles on GaAs substrate after RIE. (a) CF_4: 10 sccm, O_2: 4 sccm and (b) CF_4: 10 sccm, O_2: 7 sccm.

the investigation of surface roughness associated with the O_2 flow rate. Surface roughness was determined by an AFM. Figure 4 shows the average surface roughness of GaAs after being treated in a RIE chamber for 90 s at a RF power of 30 W. The original GaAs surface roughness without the RIE treatment was around 10 Å. Surface roughness reached the highest value of 34 Å at O_2 flow rate of 4 sccm, which corresponds to a higher R_{sh}. Either increasing or decreasing the O_2 content, as shown in this figure, resulted in an improvement of GaAs surface. Improving surface roughness after RTA may be associated with the evaporation of nonvolatile species, which were the side products after RIE. The other possibility is the rearrangement of surface atoms due to a higher surface mobility during the RTA treatment. Therefore, we conclude that there exists an optimum condition for the different O_2 content in a CF_4+O_2 gas mixture. This window could avoid the generation of too many F radicals, which causes a higher surface damage and a larger undercut

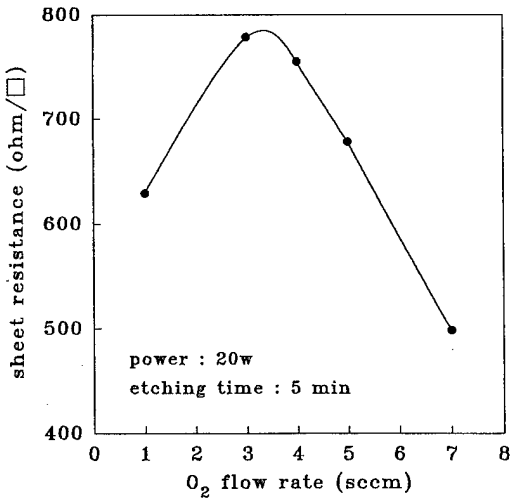


FIG. 3. Sheet resistance change of GaAs active layer as a function of O_2 flow rates after RIE treatment.

WSi_x profile. Either higher or lower of the O_2 content will substantially improve these undesired properties.

O_2 flow rates at both 1 and 7 sccm seem to be the optimum condition of etching away WSi_x materials. However, for a practical consideration, a higher O_2 content is not preferred since it increases the etching rate of photoresist, which was generally used as an etching mask. Figure 5 illustrates the photoresist etching rate as a function of different O_2 flow rates. The etching rate was enhanced monotonously by increasing the O_2 flow rate. An etching rate of 3300 Å/min was obtained at a flow rate of 7 sccm. Therefore, it is clear that under a higher O_2 flow rate condition, although it still can achieve reasonable etching properties, photoresist is not sufficient to protect the underlying materials. In consequence, we chose the O_2 flow rate of 1 sccm as the best etching parameter, and used this condition for the following damage recovery study.

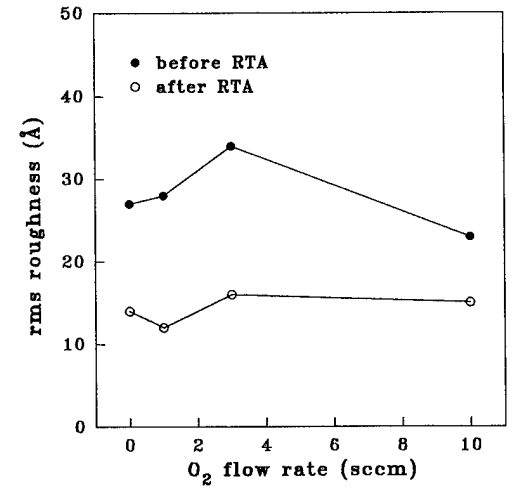


FIG. 4. Surface roughness evaluated by AFM of RIE treated GaAs before and after the RTA heat treatment.

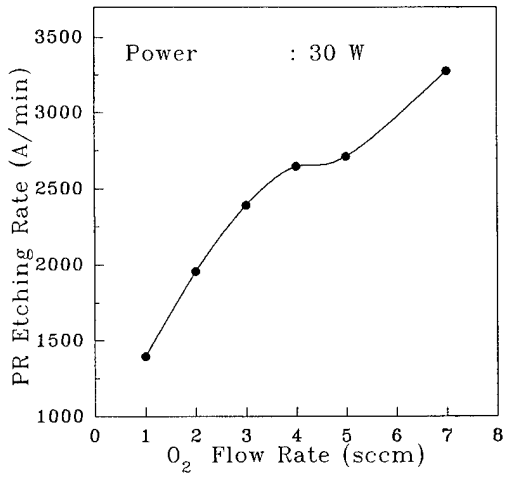


FIG. 5. The change of etching rate in photoresist vs O_2 flow rate in a CF_4+O_2 RIE chamber.

III. RECOVERY OF RIE PLASMA DAMAGE

Since the exposed GaAs active layer in a RIE chamber is unavoidably attacked by the plasma bombardment, in this section we focused on the technique to restore these damaged layers. A post-RIE thermal treatment is the typical technique to remove these damages,[12] and we used the rapid-thermal-annealing (RTA) to heal surface crystalline defects. The RTA provides a rapid, clean heat treatment, and the surface decomposition in GaAs can therefore be eliminated. The evaluation of damage and recovery of the GaAs surface was carried out by measuring the sheet resistance (R_{sh}) of active layers, which is determined by the TLM. Ohmic contacts were formed after the RTA recovery process, and the alloy condition was 350 °C, 1 min. Therefore, the recovery effect resulting from the ohmic alloy is limited.

A 1000 Å thick n-GaAs ($n=1\times10^{17}$ cm^{-3}) active layer was grown by a molecular-beam epitaxy (MBE) system on semi-insulating GaAs substrates. The original R_{sh} was 490 Ω/sq. These layers were subsequently placed in a CF_4+O_2 (10 sccm:1 sccm) RIE chamber to investigate the change in R_{sh} on n-GaAs layers. The R_{sh} increased significantly as a function of the exposed time and became saturated after a 2 min plasma treatment. The R_{sh} after a 10 min exposure became 940 Ω/sq under a dc bias of 740 V at 30 W. The higher value of R_{sh} indicates a reduction in mobility and sheet carrier density,[7] which corresponds to a degradation of the active layers.

As to the damage recovery evaluation, we used the sample treated in a 30-W, CF_4+O_2 plasma for 2 min, and the corresponding R_{sh} value was 900 Ω/sq RIE-treated GaAs wafers were surface unpassivated in the RTA chamber, and the Ar is the only gas used in the thermal cycle. R_{sh}, shown in Fig. 6(a), decreased systematically as a function of RTA temperatures. The R_{sh} became 590 Ω/sq after a 500 °C, 90-s treatment, which is quite closed to the as-grown sample (490 Ω/sq). However, for further increasing the RTA temperatures, the improvement of R_{sh} was limited. The time dependent RTA evaluation, shown in Fig. 6(b), demonstrated the exist-

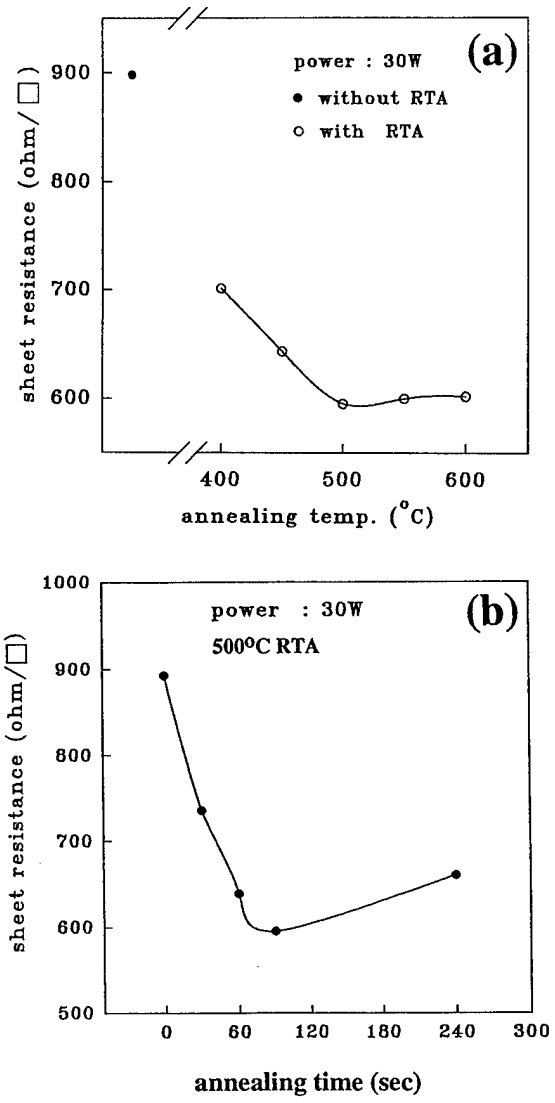


FIG. 6. The recovery of sheet resistance of n-GaAs associated with (a) RTA temperatures and (b) duration times.

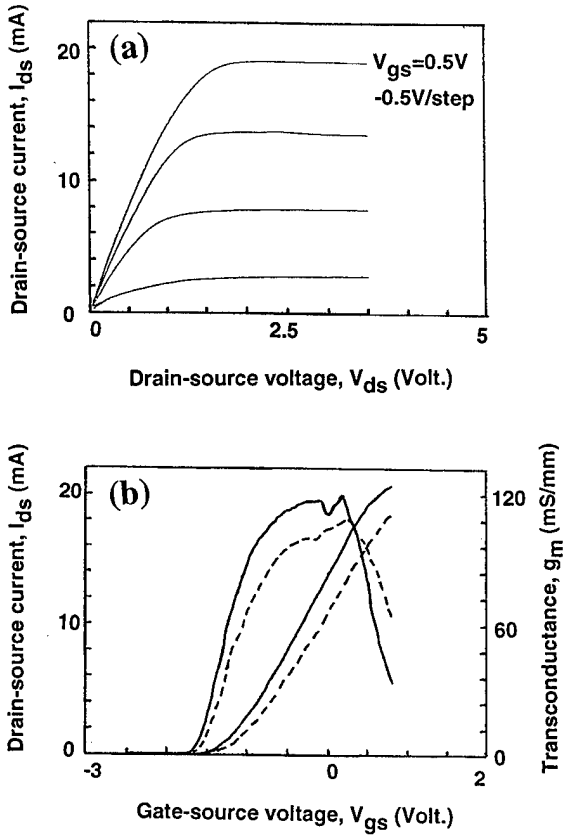


FIG. 7. (a) I_{ds}-V_{ds} output and (b) I_{ds}-g_m-V_{gs} transfer characteristics of RIE +RTA gate in GaAs MESFETs (gate length=1.2 μm). RIE gate GaAs MESFETs (dashed lines) are included in (b) for comparison.

ence of an optimum duration time for RTA. For a longer duration time, the surface decomposition may become severe, and cause an increase of R_{sh}. Therefore, we concluded that a 500 °C, 90-s RTA heat treatment is an optimum condition to remove the damages in GaAs active layers. Since Ar was the only gas used in the RTA chamber, conduction recovery resulting from the hydrogen passivation can, therefore, be ruled out. The effect of this thermal treatment can also be confirmed by the AFM investigation. GaAs surface roughness, shown in Fig. 4, became smaller after the RTA, which may be associated with the improvement of crystalline defects.

After fully characterizing the recovery of damaged n-GaAs layers, we applied this RTA process to the MESFET fabrication. The channel thickness of this MESFET was 1000-Å thick with a carrier density of 2×10^{17} cm^{-3}. 1.2-μm-long WSi_x gates were first defined by the RIE with a CF_4/O_2 flow ratio of 10:1, and the RF power was 30 W at 50 mTorr. This etching process was followed by a 500 °C, 90 s RTA heat treatment to recover the damaged layers. However, samples without being through the RTA treatment were also included to verify the effect of this recovery process. Both RIE+RTA and RIE treated MESFETs were finalized by a mesa etching and an ohmic formation process. The R_{sh} evaluated from the on-chip TLM test patterns revealed a value of 880 Ω/sq for the RIE-treated layer. As to the RIE +RTA treated layer, the R_{sh} dropped to 570 Ω/sq which is quite closed to its original value (550 Ω/sq).

Figure 7 demonstrates the $I-V$ characteristics of RIE +RTA treated GaAs MESFETs with a gate-length of 1.2 μm. The peak transconductance (g_m) was 120 mS/mm. As to the RIE treated GaAs MESFETs, attached in Fig. 7(b), this value was 110 mS/mm under the same bias condition. The associated threshold voltages (V_{th}) were −1.75 V for the RIE +RTA, and −1.56 V for the RIE gate, respectively. A lower V_{th} of RIE gate indicates a decrease of carrier density in the MESFET's channel. Electrons, under such circumstance, will be trapped inside the defect and cause a decrease of V_{th}. These trapped electrons will subsequently be released followed by a RTA treatment, which makes the V_{th} move to a negative voltage direction.

IV. CONCLUSIONS

RIE WSi_x etching has been optimized based on the different CF_4 and O_2 flow ratio. As far as the side-wall profiles and surface damages are concerned, we conclude that with a flow rate ratio of 10:1 is the best etching condition for WSi_x materials. Plasma damages caused in a RIE treatment can be restored by the subsequent RTA treatment. This thermal treatment recovers the electrical properties and also improves the surface roughness. Based on the optimum conditions of WSi_x etching and thermal process, GaAs MESFETs were fabricated and demonstrated the importance of these parameters.

ACKNOWLEDGMENTS

The authors acknowledge Professor J. I. Chyi for providing the MBE grown layers, and the financial support from the National Science Council, Republic of China (NSC 84-2215-E-008-003).

[1]F. R. White, C. W. Koburger, D. L. Harmon, and H. J. Geipel, J. Electrochem. Soc. **129**, 1330 (1982).
[2]R. L. Lee and F. L. Terry, Jr., J. Vac. Sci. Technol. B **9**, 2747 (1991).
[3]R. J. Shul, A. G. Baca, D. J. Rieger, and A. J. Horward, Electron. Lett. **31**, 317 (1995).
[4]M. E. Burba, E. Degenkolb, S. Henck, M. Tabasky, E. D. Jungbluth, and R. Wilson, J. Electrochem. Soc. **133**, 2113 (1986).
[5]H. F. Sahafi, G. F. Goldspink, A. A. Rezazadeh, A. P. Webb, and M. A. Carter, Electron. Lett. **28**, 2300 (1992).
[6]F. Ren, T. R. Fullowan, S. J. Pearton, J. R. Lothian, R. Esagui, C. R. Abernathy, and W. S. Hobson, J. Vac. Sci. Technol. A **11**, 1768 (1993).
[7]S. Agarwala, I. Adesida, C. Caneau, and R. Bhat, Appl. Phys. Lett. **64**, 2979 (1994).
[8]M. A. Foad, S. Thoms, and C. D. W. Wilkinson, J. Vac. Sci. Technol. B **11**, 20 (1993).
[9]K. K. Ko, S. W. Pang, T. Brock, M. W. Cole, and L. M. Casas, J. Vac. Sci. Technol. B **12**, 3382 (1994).
[10]R. Petri, P. Brault, O. Vatel, D. Henry, E. Andre, P. Dumas, and F. Salvan, J. Appl. Phys. **75**, 7498 (1994).
[11]K. T. Sung and S. W. Pang, Jpn. J. Appl. Phys. **33**, 7112 (1994).
[12]K. T. Sung and S. W. Pang, J. Vac. Sci. Technol. A **12**, 1346 (1994).

Characterization of electrical damage induced by CH_4/H_2 reactive ion etching of molecular beam epitaxial InAlAs

M. Achouche,[a] A. Clei, and J. C. Harmand
France Telecom, CNET Centre Paris B, 196 Av. H. Ravera, 92220 Bagneux, France

(Received 22 December 1995; accepted 19 April 1996)

In this article, we report investigations on the effects of methane/hydrogen (CH_4/H_2) reactive ion etching (RIE) of InGaAs/InAlAs/InP heterostructure materials for high electron mobility transistors, and especially on the electrical properties of the InAlAs layer after dry recess etching of the InGaAs cap layer. The ion etching induced damages in the barrier layer InAlAs are evaluated by diode current–voltage and capacitance–voltage measurements and deep-level transient spectroscopy (DLTS). The $I-V$ data indicate that RIE lowers the Schottky barrier height (ϕ_b) and increases the ideality factor. Using low pressure (10 mTorr) RIE processes, with various self-bias voltages, shows that the use of low ion energy is necessary to get good Schottky contacts. Rapid thermal annealing at 400 °C is shown to induce a limited decrease of the dry etching induced defects. However, full recovery of the electrical properties is not achieved. A significant improvement of the Schottky diode electrical characteristics and DLTS spectra is observed after wet etching a 60 Å thick InAlAs layer before metal deposition, indicating that the main damages are concentrated within a short distance from the surface. © *1996 American Vacuum Society.*

I. INTRODUCTION

The increasing importance of InP-based materials in the fabrication of microelectronic and optoelectronic integrated circuits requires the development of well-controlled processes for the fabrication of microstructure devices.

Extensive work on selective wet and dry etching processes concerning these materials has been reported[1,2] to allow the fabrication of devices such as high electron mobility transistors (HEMTs), heterostructure field effect transistors (HFETs) and heterojunction bipolar transistors (HBTs). For example, in HBT technology, etching of the emitter must be done very precisely and has to stop at the thin base region. For HFET, a controlled process, giving accurate reproducibility and etch selectivity of the InGaAs contact layer versus the InAlAs barrier layer, is necessary to insure low threshold voltage dispersion. Although conventional wet chemical etching does not induce damages in the layers, it does not offer the same capabilities as dry etching in terms of control of fine patterns, uniformity, anisotropy, and etching profiles.

The recent demand for extremely high performance InAlAs/InGaAs HEMTs[3] for microwave application or optical communications circuits, in the 1.3–1.5 μm wavelength range, has created a need for selective dry etching techniques. Reactive ion etching (RIE) with CH_4/H_2 has been developed and used by several groups for InP-based materials, but relatively few reports[4,5] have been published which demonstrate the application of this technique in transistor fabrication. Actually, RIE may induce surface damages caused by ion bombardment, stoichiometry changes due to preferential depletion of group V elements, redeposition of hydrocarbon polymer, and hydrogen passivation of donors. Such defects results in poor Schottky contacts, which induce excessive leakage current and low breakdown voltage that limit HEMT device performance. Such defects can also increase the low frequency noise of such devices.

In this article, we study the influence of the RIE of the InGaAs layer and exposure of InAlAs to the rf plasma on the electrical properties of TiAu–InAlAs Schottky contacts. Conventional $I-V$ and $C-V$ measurements were used to evaluate the electrical characteristics of the Schottky diode (barrier height, ideality factor, and leakage current) after etching the InGaAs cap layer. Deep traps induced by interaction of the plasma with the semiconductor surface were measured by deep-level transient spectroscopy (DLTS). Annealing studies on these diodes were carried out to remove damages. Spatial extension of the damaged layer was evaluated by wet etching thin layers at the InAlAs surface in order to recover good Schottky diode characteristics.

II. EXPERIMENT

The epitaxial structures used in this work were grown by molecular beam epitaxy (MBE) on n^+ InP substrates. They consist of a 0.25 μm thick n^+ InAlAs buffer layer, a 0.5 μm InAlAs barrier, and a 0.01 μm undoped InGaAs cap layer. Si was used as a dopant, and the carrier concentration was typically 2×10^{16} cm^{-3} in the InAlAs barrier layer.

Diodes were processed on these layers using a conventional technology. Alloyed AuGeNiAu was used as a back side ohmic contact. After etching the InGaAs cap layer with CH_4/H_2 using a SiN mask, the TiAu (50–300 nm) Schottky contact was deposited by e beam evaporation. Native oxides on the InAlAs surface were removed by a diluted HCl dip prior to metal evaporation. Schottky electrodes with a 2×10^{-3} cm^2 area were used for electrical characterization. Etching was performed in a planar diode RIE system (ALCATEL GIR 100) operating at a frequency of 13.56 MHz. As a result of a previous optimization of the etching parameters, a gas mixture of methane and hydrogen was used

[a]Electronic mail: mohand.achouche@bagneux.cnet.fr

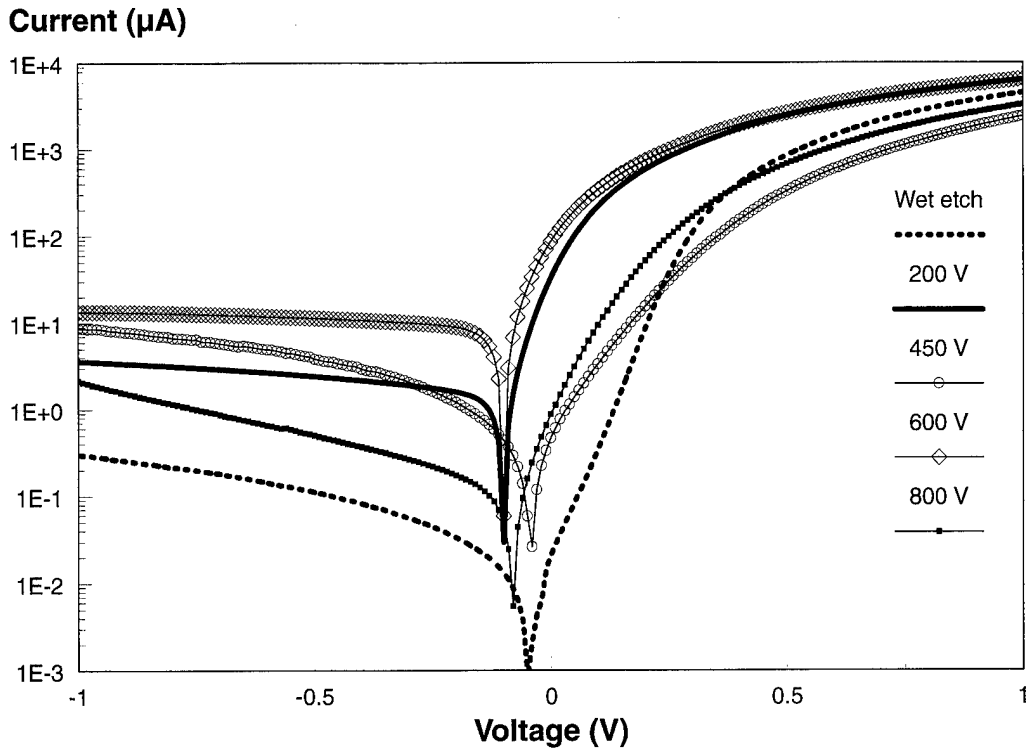


FIG. 1. Log I vs V characteristics of TiAu–InAlAs Schottky diodes at various self-bias voltages.

at a total pressure of 10 mTorr, with $1CH_4/6H_2$ ratio, and a total flow rate of 30 sccm. Several etching processes were compared with rf power applied to the cathode ranging from 20, 60, 100, to 140 W. The mean self-bias on the cathode was 200, 450, 600, and 800 V, respectively. Selective etching of InGaAs over InAlAs was obtained when the flux of the reactive hydrocarbon species was higher than the removal rate of the reacted InAlAs. The $(1CH_4/6H_2)$ mixture at a pressure of 10 mTorr results in low deposition rate of hydrocarbon polymers on the sample surface. Under these conditions, both InGaAs and InAlAs were etched for all rf powers. For example, the etch rate of InGaAs and InAlAs at 200 V is 15 and 3 nm/min, respectively, giving an etching selectivity of InGaAs over InAlAs of 5:1. Selectivity between InGaAs and InAlAs was observed to increase with plasma rf power. We used similar etch times (5 min) and gas residence times for these RIE processes. On some samples, rapid thermal annealing, in an Ar/H_2 atmosphere, during 20 s before Schottky metal deposition, was used to remove damages from the etched layer. The annealing temperature (400 °C) is typical of ohmic contact formation.

DLTS measurements were performed using a DL 8000 digital DLTS system from Bio-Rad. This system digitizes the capacitance or current transient. It then performs a numerical Fourier transformation, in which parameters are analyzed with the significant advantage that the deep-level characteristics can be obtained during measurements using a single temperature scan. The first sine coefficient spectrum $b1$ has been used in this study to characterize the transient evolution with temperature. For all measured samples, a reverse bias at -0.8 V, a filling pulse at -0.2 V, and a pulse width of 1 ms were used. Temperature was scanned in the 100–450 K range.

III. RESULTS AND DISCUSSION

A. Characterization of the RIE samples

1. I–V characteristics

Figure 1 shows a typical semilog plot of the $I-V$ characteristics of TiAu–InAlAs diodes as a function of self-bias voltage. The $I-V$ plot of a chemically etched sample is also shown for comparison.

These data are analyzed assuming a thermionic emission of electrons over the potential barrier. For a Schottky diode dominated by thermionic emission, the current versus voltage characteristics are given by

$$I=I_0\{\exp[q(V-IR_s)/nkT]-1\}, \tag{1}$$

where I_0 is the saturation current, n is the ideality factor, and R_s is the series resistance. Ideality factors are calculated under forward bias, from the slope of the linear region of log $I-V$ plots, and effective barrier heights ϕ_b from the extrapolated saturation current, using relations (2) and (3), respectively:

$$n=\frac{q}{kT}\frac{\partial V}{\partial(\ln I)} \tag{2}$$

and

TABLE I. Effective ϕ_b and real Schottky barrier height ϕ_{b0}, ideality factor n, series resistance R_s, and reverse current at -1 V calculated from $I-V$ data.

Type of processing	ϕ_b (eV) $(I-V)$	ϕ_{b0} (eV) $(I-V-T)$	n	R_s (Ω)	I_r (A)
Chemical etched	0.60	0.62	1.07	2.5	3×10^{-7}
RIE 200 V	0.50	0.48	1.13	3.12	6×10^{-6}
RIE 450 V	0.44	0.39	1.30	3.5	2×10^{-5}
RIE 600 V	0.42	0.39	1.40	3.9	5×10^{-5}
RIE 800 V	0.51	0.41	1.60	10	7×10^{-6}

$$\Phi_b=\frac{q}{kT}\ln(A^*T^2/I_0), \tag{3}$$

where A^* is the effective Richardson constant ($\sim$10 A cm^{-2} K^{-2} for InAlAs).

For the reference diode, transport characteristics dominated by thermionic emission ($n\leqslant1.10$) are observed; the effective barrier height calculated from I_0 is 0.60 eV. For the reactive ion etched samples, the $I-V$ characteristics could not be accurately fitted by a thermionic emission model. Comparing the $I-V$ characteristics, two observations can be made:

(i) The reverse as well as the extrapolated forward saturation current increase for reactive ion etched samples when compared to the wet etched samples. This effect may be related to a reduction of the effective Schottky barrier height.

(ii) Under forward bias, the current increases in the low voltage region, whereas the slope of the linear region of the log $I-V$ plots decreases in the high bias region.

Measured values of barrier height, ideality factor, series resistance and reverse current at -1 V calculated from $I-V$ data are summarized in Table I.

The decrease in barrier height after reactive ion etching may be related to near surface damages induced by ion bombardment during etching. This may be responsible for the observed current increase at low forward bias and also for the increase in the series resistance of the diodes. From Table I it appears that ϕ_b does not decrease monotonically with increasing self-bias: ϕ_b for 800 V is higher than for 450 and 600 V. Similar variations of Schottky contacts parameters have also been reported[6] for IBE and ion sputtering of Si and GaAs.

2. I–V–T characteristics

Analysis of the temperature dependence of the saturation current density was also conducted in order to evaluate the real barrier height from the Richardson plot $\ln(J_sT^{-2})$ vs T^{-1}.

a. *Forward bias.* Figure 2 shows the Richardson plot $\ln(J_sT^{-2})$ vs T^{-1} of the Schottky diodes. The corresponding barrier heights are reported in Table I. We obtain a "real" barrier height of 0.62 eV for the chemically etched sample, which compares well with the "effective" barrier height deduced by the thermionic emission model of the forward current. The values of the real barrier height for samples etched with 10 mTorr CH_4/H_2 for 200, 450, 600, and 800 V are

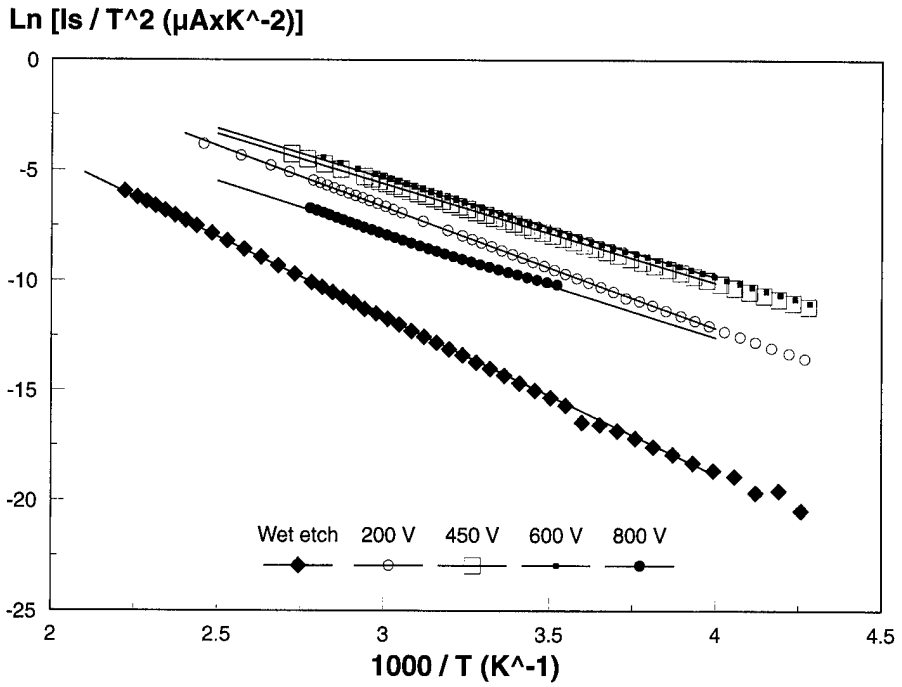


FIG. 2. Richardson plot ln (J_sT^{-2}) vs T^{-1} of the Schottky diodes.

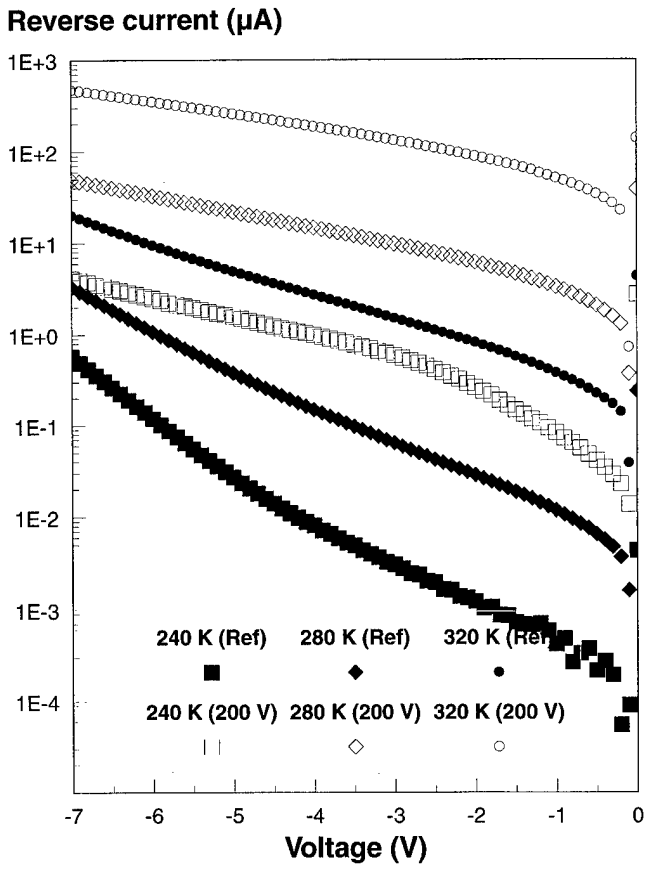


FIG. 3. Reverse $I-V-T$ plots for diodes processed with chemical etching and RIE at a self-bias of 200 V.

0.48, 0.39, 0.39, and 0.41 eV, respectively. From these results, it appears that the effective barrier height is higher than the real barrier height for the samples etched at 450, 600, and 800 V. The effective barrier height, as estimated from the saturation current of room temperature $I-V$ characteristics, includes the effects of the presence of a thin insulating layer between metal and semiconductor. Larger ideality factors (1.3, 1.4, and 1.6) are also observed on these Schottky diodes processed at higher ion energy ($\geq$450 V), resulting probably from the presence of this interfacial layer at the Schottky junction. Such interfacial layers may result from the formation of a heavily damaged region near the sample surfaces during RIE at high ion energy ($\geq$450 V) as will be discussed later.

The effects of CH_4/H_2 reactive ion etching on the electron conduction through the TiAu–InAlAs Schottky junction were analyzed using $I-V-T$ measurements under forward and reverse bias on both chemically etched samples and samples etched by RIE with a 200 V self-bias voltage.

From the Richardson plot of the saturation current density at forward bias, we deduce the values of the real Schottky barrier height as given in Table I (0.6 eV for the reference sample and 0.48 eV for the RIE sample at 200 V).

b. *Reverse bias.* Figure 3 shows the temperature dependence of the reverse current of these Schottky diodes. For the reference sample, the reverse current shows a sharp increase with increasing reverse bias, especially at low temperature. The plot of $\ln(I_R)$ vs $1/T$ shows two different slopes, giving activation energies of the reverse current ϕ_T of 0.61 and 0.29 eV for low and high reverse bias, respectively. The observed reverse $I-V$ characteristics at low bias can be explained by a thermionic emission model as reported by many authors,[7] and is similar to the barrier height deduced from the Richardson plot of the forward saturation current. For higher bias voltage, an activation energy around 0.29 eV is deduced from the $\ln(I_R)$ plot. This lower activation energy can be attributed to a defect-assisted tunneling mechanism which governs the current at high reverse bias. It should be noted that this value is close to the activation energy of a defect $E1\sim0.2$ eV, observed by DLTS.

For the RIE sample, an excess current flow is observed, and two slopes are found on the Richardson plot of the reverse current as shown in Fig. 4. At low bias, an activation energy of 0.5 eV is attributed to a generation current via plasma-generated defects (see Sec. IV). At larger bias, a tunneling current through defects located at an energy level of 0.45 eV is probably the main mechanism observed in RIE InAlAs. These results are in good agreement with the DLTS results discussed in Sec. III A 4.

3. C–V characteristics

Figure 5 displays the 1 MHz capacitance–voltage ($C-V$) characteristics of the devices. For the chemically etched sample, the $1/C^2$ vs V plot produces a straight line which slope can be used to calculate the donor concentration ($N_d\approx2\times10^{16}$ cm^{-3}), and the intercept with the voltage axis, the barrier height. We find a barrier height of 0.65 eV which is in good agreement with that calculated from $I-V$ and $I-V-T$ measurements. For the RIE sample, the diodes have a much reduced capacitance, and $1/C^2$ vs V is no longer linear, with dC^{-2}/dV decreasing at low bias. Due to the nonlinearity of such a plot, no useful information on the barrier height can be obtained.

The decrease of the measured capacitance after RIE is attributed to the additional series capacitance of the interfacial layer (polymer film and/or damaged layer). These variations of the reverse-bias capacitance of Schottky diodes after dry etching have already been reported in the case of GaAs.[6] It was concluded that ion bombardment creates donorlike damages near the metal/semiconductor interface: the presence of this interfacial layer and damage defects in the near surface region of the semiconductor induces a nonuniform charge density in the depletion region. The existence of such defects perturbs the parabolic conduction band and then affects the intercept of $1/C^2$ vs V plot.

4. DLTS measurements

The DLTS system used in this work is the same as that described by Weiss.[8] Assuming that the traps are fully occupied by the filling pulse, the time dependence of the space

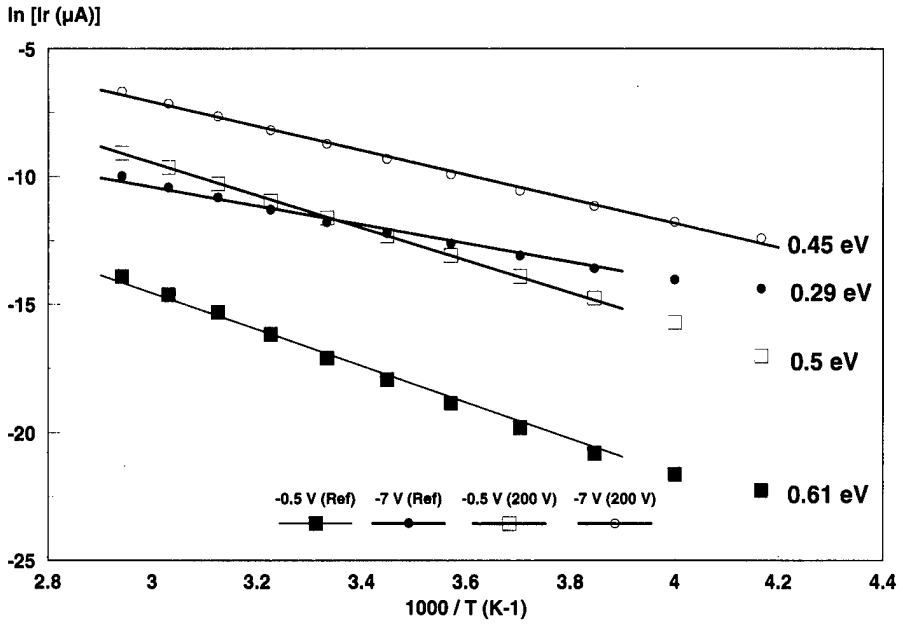


FIG. 4. Richardson plot of the chemically etched diode and RIE (200 V) etched diode.

charge capacitance $C(t)$ (the capacitance transient) is approximately exponential, for low trap concentrations ($N_t \ll N_s$):

$$C(t)=C_r-C_r\left(\frac{N_t}{2N_s}\right)\exp\left(\frac{-t}{\tau_{\text{emission}}}\right), \tag{4}$$

where C_r is space charge capacitance at reverse bias voltage, N_t is trap concentration, N_s is doping level, and τ_{emission} is emission time constant of electrons at the actual temperature of the diode. Extrapolation of the measured transient values to $t=0$ gives the value of $C(t=0)$:

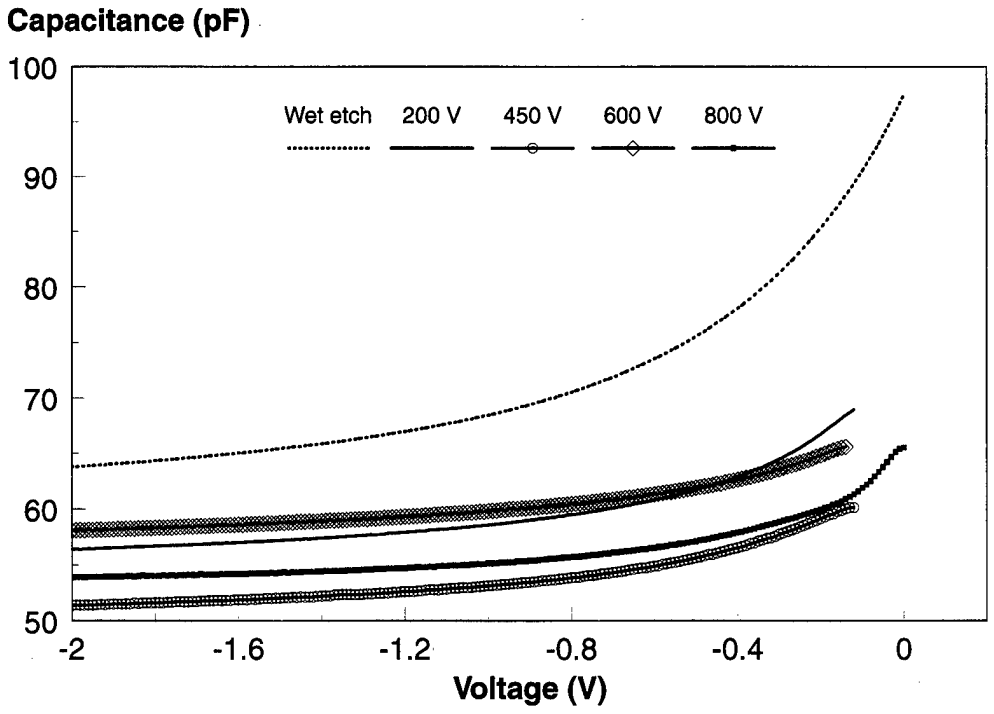


FIG. 5. C vs V (reverse bias) characteristics of TiAu–InAlAs Schottky diodes at various self-bias voltages.

$$\frac{C(t=0)}{C_r}=1-\frac{N_t}{2N_s}. \tag{5}$$

Thus all parameters in the above formula are known and the trap concentration N_t can be calculated. When a transient is recorded, we obtain one point on the Arrhenius plot from a Fourier analysis (at the measurement temperature).

The N_t value is obtained with bias values selected such that the depletion width is significantly larger than the transition region. Then, to make a more exact trap concentration calculation, an active volume has to be considered corresponding to the region extending from the position (Lr) where the Fermi level (Ef) crosses the trap level at (Et) during the filling pulse to the position (L_p) where the Fermi level crosses the trap level during the emission phase. Then the capacitance can be written as

$$C\approx C_r\frac{N_t}{2N_s}\frac{L_r^2-L_p^2}{W_r^2}, \tag{6}$$

where W_r is the depletion region width. This equation takes into account the width of the transition region under both pulse and reverse bias conditions.

For DLTS measurements, the Schottky diodes were reverse biased at -0.8 V, and the filling pulse voltage was -0.2 V to overcome the heavily damaged region caused by exposure of the InAlAs surface to the plasma. The present bias condition allows us to detect relevant traps located from 0.1 μm (depletion width at -0.2 V) to 0.19 μm (depletion width at -0.8 V) in depth from the InAlAs surface.

Figure 6(a) shows the typical DLTS spectra for an as grown InAlAs MBE sample. Three traps are evidenced, namely, E_1, E_2, and E_3, as previously reported.[9,10] Activation energies are close to 0.20, 0.42, and 0.62 eV for levels E_1, E_2, and E_3, respectively. Trap concentrations are 5.1×10^{14} cm^{-3} for E_1 and 7.0×10^{14} cm^{-3} for E_2. The defect concentration for E_3 is estimated to be 4.9×10^{15} cm^{-3}. Trap concentrations of our InAlAs MBE samples are in the same order of magnitude as those obtained by other groups.[9,10]

Figure 6(b) shows the DLTS spectra for the same sample after RIE. The trap activation energies, capture cross sections, and concentrations evaluated from an Arrhenius plot of the DLTS maxima are listed in Table II. These data show that RIE samples can be classified into two categories, depending on the self-bias. The first one is for samples etched at 200 and 800 V, whereas the second pertains to self-bias of 450 and 600 V. Reactive ion etching at 200 or 800 V results in devices having increased densities of defect E_2 (activation energy 0.42–0.50 eV), while the concentration of defect E_3 (0.6 eV) is found to decrease. DLTS spectra of samples etched at 450 and 600 V show a single maxima corresponding to defect E_2 (~0.45 eV). The trap concentration of this level is comparable to that obtained at 200 and 800 V. E_1 (0.20 eV) disappears for all etched samples.

DLTS spectra in Fig. 6(b) also show a shift of the peaks related to traps E_2 and E_3. This indicates a change in activation energies for levels E_2 and E_3 from those of the control sample. In Fig. 6(b), it can be observed that the peak amplitude of defects E_2 and E_3 does not increase with increasing self-bias. Furthermore, in all cases, the concentration of the traps at level E_2 is increased after reactive ion etching. However, the height of the DLTS peak related to the

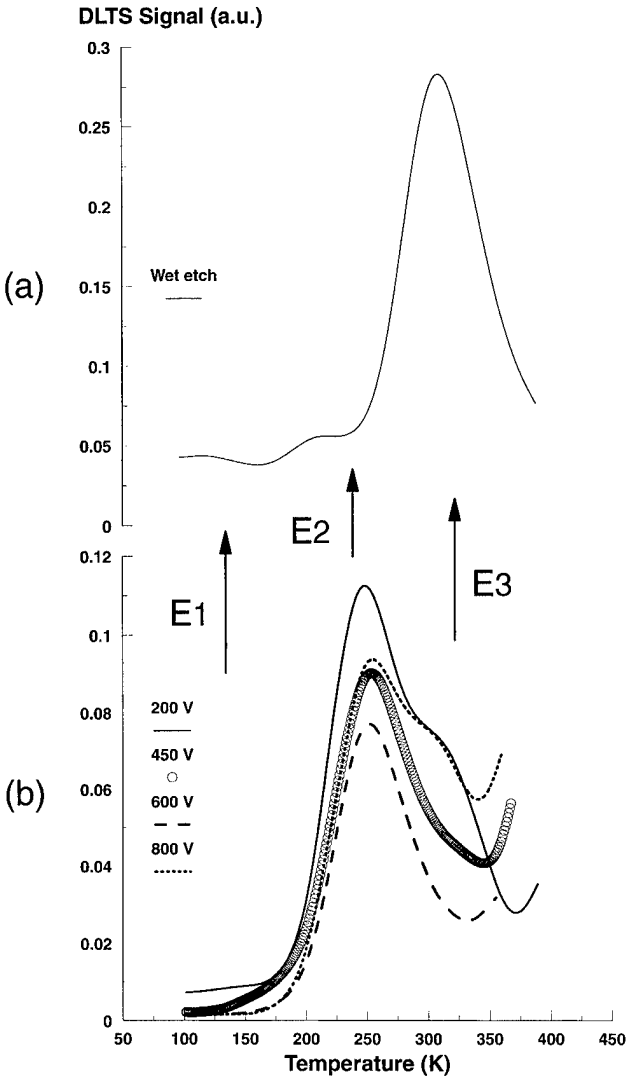


FIG. 6. DLTS spectra of typical as grown (a) MBE-InAlAs and (b) RIE-InAlAs at various self-bias voltages.

TABLE II. Activation energies E_t, capture cross sections σ_t, and trap densities N_t of the deep levels obtained from the Arrhenius plot data.

Type of processing	E_c-E_t (eV)	σ_t (cm^2)	N_t ($\times10^{15}$ cm^{-3})
	0.20	1.5×10^{-14}	0.515
Chemical etched	0.42	6×10^{-13}	0.701
	0.62	2×10^{-13}	4.86
RIE 200 V	0.45	5.34×10^{-14}	2.65
	0.69	6.43×10^{-12}	1.95
RIE 400 V	0.49	1.27×10^{-13}	2.0
RIE 600 V	0.45	2.93×10^{-14}	2.0
RIE 800 V	0.50	2.26×10^{-13}	2.06
	0.62	4.88×10^{-13}	1.28

TABLE III. Effective ϕ_b and real Schottky barrier height ϕ_{b0}, ideality factor n, and reverse current at -1 V for the annealed samples.

Type of processing	ϕ_b (eV) (I–V)	ϕ_{b0} (eV) (I–V–T)	n	I_r (A)
RIE 200 V+RTA10 s	0.54	0.46	1.10	2×10^{-6}
RIE 200 V+RTA20 s	0.57	0.48	1.10	1×10^{-6}
RIE 200 V+RTA30 s	0.54	0.40	1.07	5×10^{-6}
RIE 450 V+RTA20 s	0.46	0.39	1.15	1×10^{-4}
RIE 600 V+RTA20 s	0.49	0.38	1.40	2×10^{-4}
RIE 800 V+RTA20 s	0.43	0.30	2.00	1×10^{-3}

deep trap E_2 shows a saturation between 2 and 3×10^{15} cm^{-3}. This suggests that defects E_2 are not simple point defects, but, rather, complex defects. This is supported by the widening of the E_2 peak, also suggesting that it is related to complex defects.[15]

B. Annealing effects after RIE

1. I–V and I–V–T characteristics

We have investigated the effects of a thermal treatment in removing the damages created by RIE. Rapid thermal annealing was performed in Ar/H_2 atmosphere at 400 °C for 10, 20, and 30 s.

In Table III, we report the values of the real and effective barrier height, ideality factor, and reverse currents, at a -1 V bias for samples etched at 200 V and annealed for different times. An increase of the effective barrier height and a reduction of the leakage current at low reverse bias were observed after annealing. Temperature dependence of the saturation current density shows that the real barrier height did not change after rapid thermal annealing (RTA). These typical variations of the metal–insulator–semiconductor Schottky junction were $(\Phi_b)_{eff}>(\Phi_b)_{rea}$. Analysis of the reverse current of the annealed diode at different temperatures shows an increase of the tunneling current at high reverse bias. This is probably due to the formation of a native oxidelike insulating layer at the Schottky junction. At low reverse bias, a reduction of the leakage current is observed. These variations can be explained in terms of effective Schottky barrier enhancement and generation current reduction which is probably associated to the effectiveness of annealing in the reduction of the plasma-induced defects. However, full recovery was not achieved since there is only a limited decrease of the surface defects.

Rapid thermal annealing was also performed on diodes etched at 450, 600, and 800 V. The corresponding parameters are also reported in Table III. The results of the annealing on these diodes were different from those obtained with 200 V (low ion energy). Samples processed at 450 and 600 V (see Table III) did not show any significant change in effective and real barrier height. Moreover, the I–V characteristics of samples etched at 800 V degraded after annealing. An increase of the reverse current was observed after annealing of these diodes processed at higher ion energy ($\geq$450 V).

However, the values of the effective barrier height for all the etched and annealed samples were apparently higher by about 0.1 eV than the real barrier height determined by the Richardson plot. The enhancement of the effective Schottky barrier after annealing for the diodes etched at 200 V (low ion energy) can be attributed to the presence of a native interfacial oxidelike insulating layer. On the other hand, annealing of the samples etched at higher ion energy ($\geq$450 V) induces an incorporation of additional interface states which increases the reverse current by generation via plasma-induced defects, and also by tunneling through interface defects.

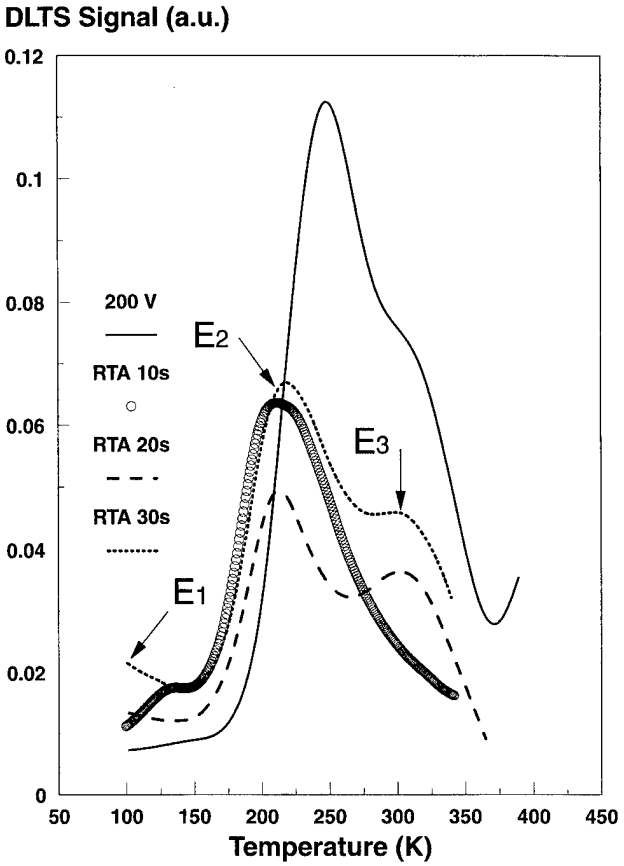


FIG. 7. DLTS spectra of InAlAs samples etched at 200 V and annealed for 10, 20, and 30 s at 400 °C.

2. DLTS measurements

Figure 7 shows a slight effect of thermal annealing on the reduction of the concentration of trap E_2 for the samples etched at 200 V. Trap concentration of level E_2 is estimated to be $(0.9–1)\times10^{15}$ cm^{-3} after annealing. Traps E_1 (0.2 eV) and E_3 (0.6 eV) reappear after RTA, with lower concentrations than for the reference sample. In Fig. 7, we can also observe the effectiveness of annealing in suppressing the shift in peak temperature for the deep trap E_2.

Figure 8 gives the measured DLTS signal for a sample etched at 200 V and annealed for 20 s. The activation energy and capture cross sections obtained from the Arrhenius plot data for each trap (E_1, E_2, and E_3) are used as parameters for a simulation of the theoretical DLTS signal (exponential capacitance transient). A broadening of the experimental

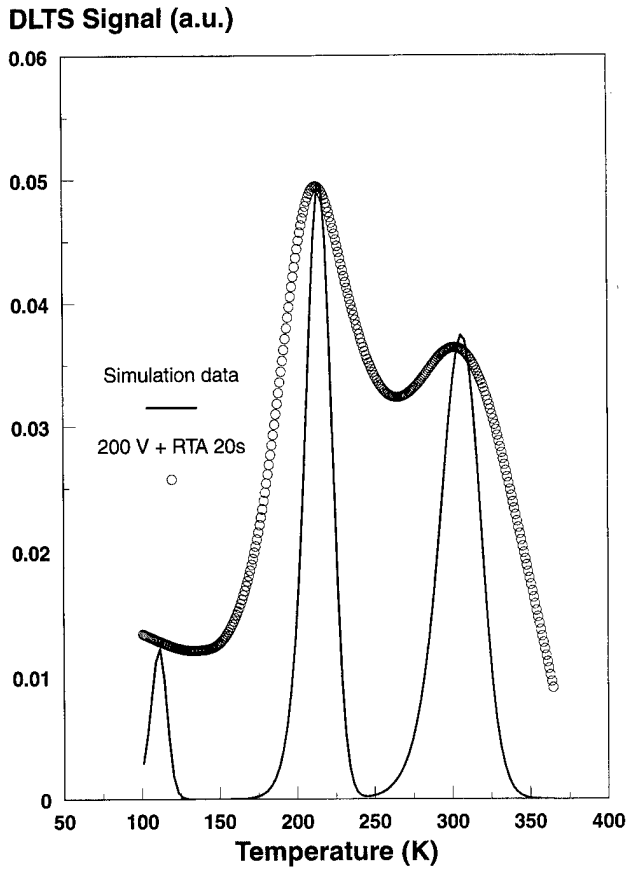


FIG. 8. Measured and simulated DLTS spectra of sample etched at 200 V and annealed for 20 s.

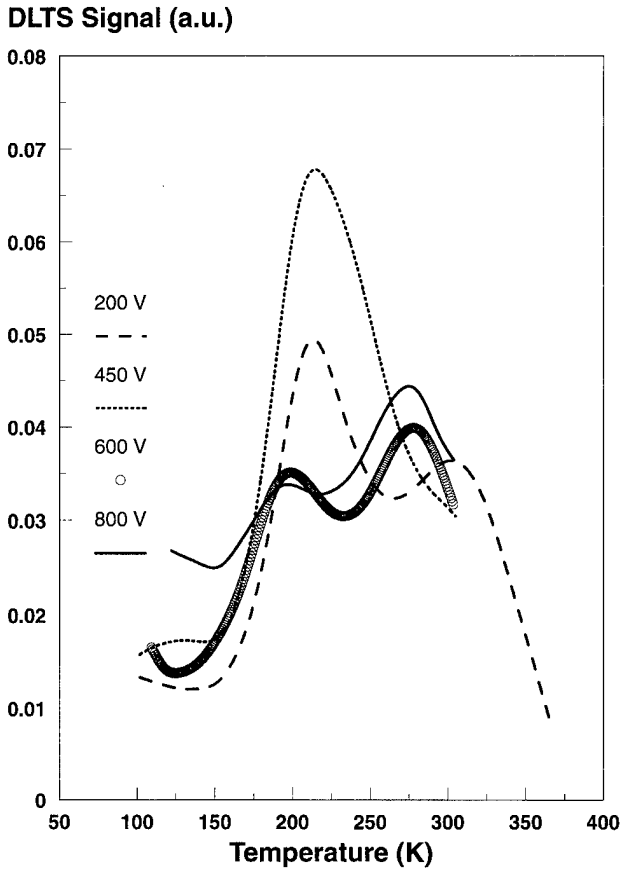


FIG. 9. DLTS spectra of InAlAs samples etched at various self-biases and annealed for 20 s.

peaks at 220 and 300 K is observed in Fig. 8, as compared to the simulated signals. Fourier analysis of the transient indicates that it is not a true exponential one, indicating an overlap of peaks corresponding to different activation energies. This effect was also observed by Omling *et al.* in GaAsP.[11] They attributed this behavior to an alloy broadening effect.

From these results, it appears that annealing samples etched at 200 V result in a small decrease of the trap concentration of level E_2, but the value remains larger than for the reference sample. On the other hand, after RTA, E_1 reappears at a low level (small DLTS peak height) and E_3 concentration remains constant, around 1×10^{15} cm^{-3}. However, the main effect on the DLTS spectra for samples etched at 200 V and annealed at 400 °C is a reduction of the broadening of the DLTS peaks at 220 and 300 K, suggesting a reduction of the energy distribution range of the deep level.

Figure 9 compares the effect of thermal annealing on DLTS signal versus temperature on the samples etched at various self-bias. The parameters of the observed traps are displayed in Table IV. All traps reappear after thermal annealing. However, as for the sample etched at 200 V, no well defined DLTS peaks related to deep traps E_2 and E_3 are observed after thermal annealing. The concentration of the deep-level E_2 after RTA is also higher (between 8×10^{14} and 1×10^{15} cm^{-3}).

C. Damage removal by wet chemical etching

1. I–V characteristics

The damaged layer thickness was investigated only for the samples etched at 200 V (low self-bias) by removing the damaged layer using wet chemical etching. Figure 10 shows

TABLE IV. Activation energies E_t, capture cross sections σ_t, and trap densities N_t of the deep levels after thermal annealing.

Type of processing	E_c-E_t (eV)	σ_t (cm^2)	N_t ($\times10^{15}$ cm^{-3})
RIE 200 V+RTA 10 s	0.26	3.57×10^{-12}	0.225
	0.42	2.36×10^{-13}	1.07
	0.20	3.68×10^{-13}	0.148
RIE 200 V+RTA 20 s	0.49	1×10^{-13}	0.927
	0.64	5.87×10^{-13}	0.907
	0.22	3.78×10^{-12}	0.300
RIE 200 V+RTA 30 s	0.47	1.68×10^{-12}	1.52
	0.61	2.38×10^{-13}	1.48
	0.24	1.51×10^{-14}	0.356
RIE 450 V+RTA 20 s	0.45	9.51×10^{-14}	1.08
	0.67	3.61×10^{-13}	1.38
	0.25	5.52×10^{-13}	0.227
RIE 600 V+RTA 20 s	0.50	1.97×10^{-12}	0.912
	0.61	1.37×10^{-14}	1.39
	0.24	1.95×10^{-13}	0.62
RIE 800 V+RTA 20 s	0.42	1.11×10^{-12}	0.812
	0.60	1.97×10^{-13}	2.17

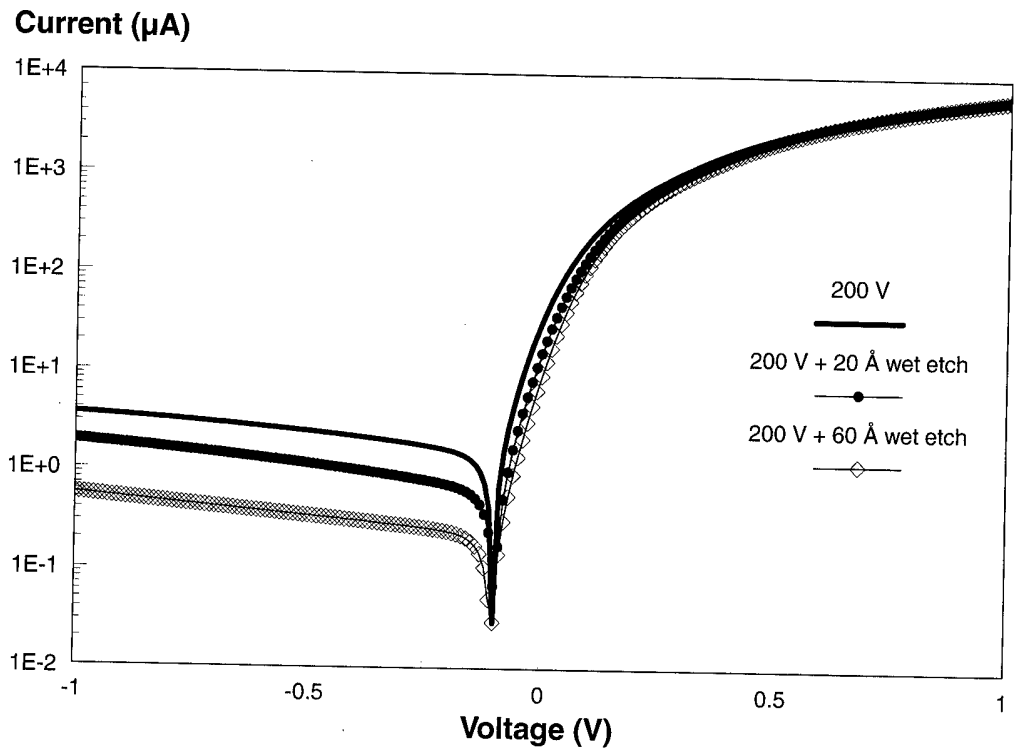


FIG. 10. Log I vs V characteristics of samples after RIE at 200 V and wet-chemical etching.

the changes in $I-V$ characteristics after CH_4/H_2 dry etching at 200 V and after the damaged surface is removed by step etching. The corresponding Schottky parameters are summarized in Table V, where it can be seen that a wet etch about 20 Å thick induces an increase of the barrier height from 0.50 to 0.54 eV, a decrease of the ideality factor (1.07) and a small decrease of leakage current (2×10^{-6} A at -1 V). On the other hand, a wet etch of 60 Å improves the electrical characteristics of the diodes and essentially the leakage current, which is about 6×10^{-7} A at -1 V. A wet etch of the InAlAs surface after RIE at 200 V also induces a decrease of the series resistance of the diodes down to 2.5 Ω. The current transport through the Schottky barrier, after a shallow wet etch of the InAlAs layer, can anew be described by a thermionic emission model ($n<1.1$). This result supports the presence of an interfacial layer near the semiconductor surface after dry etching, which is eliminated by wet chemical etching.

TABLE V. Effective Schottky barrier height ϕ_b, ideality factor, and reverse current at -1 V for the samples wet-chemically etched after the RIE treatment at 200 V.

Type of processing	ϕ_b (eV) ($I-V$)	n	R_s (Ω)	I_r (A)
RIE 200 V	0.50	1.13	3.12	6×10^{-6}
RIE 200 V+wet etch 20 Å	0.54	1.07	2.5	2×10^{-6}
RIE 200 V+wet etch 60 Å	0.57	1.05	2.5	6×10^{-7}

2. DLTS measurements

Figure 11 shows a set of DLTS spectra where the samples were chemically etched after RIE at 200 V. A reverse bias of -0.8 V and a filling pulse voltage up to 0 V was used in order to analyze the InAlAs layer. These data show that a chemical etching of 20 Å after RIE induces a decrease of the height of the DLTS peak related to the deep trap E_2, whereas the concentration of the trap E_3 becomes higher than that of E_2. This DLTS result confirms the association of increased density of E_2 traps after RIE to plasma-induced damage. Following a wet etching of 60 Å, we found a DLTS spectra similar to that obtained on the reference sample. The corresponding $I-V$ characteristics indicate the disappearance of residual damages. We can conclude that the thickness of the heavily damaged region for the RIE samples etched at 200 V, is about 60 Å.

Even though DLTS is unable to detect defect levels close to the surface, the variations of the dominating deep-level E_2 after step-by-step wet etching shows its dominance in the near surface region. This can be put together with observations made on GaAs after ion implantation, for which a strong increase of defect density toward the surface is also observed.[22] Thus, surface Fermi level pinning by the deep defect states should be taken into consideration.

IV. DISCUSSION

$I-V$ measurements performed on InAlAs Schottky diodes after RIE show a decrease in barrier height and an increase in

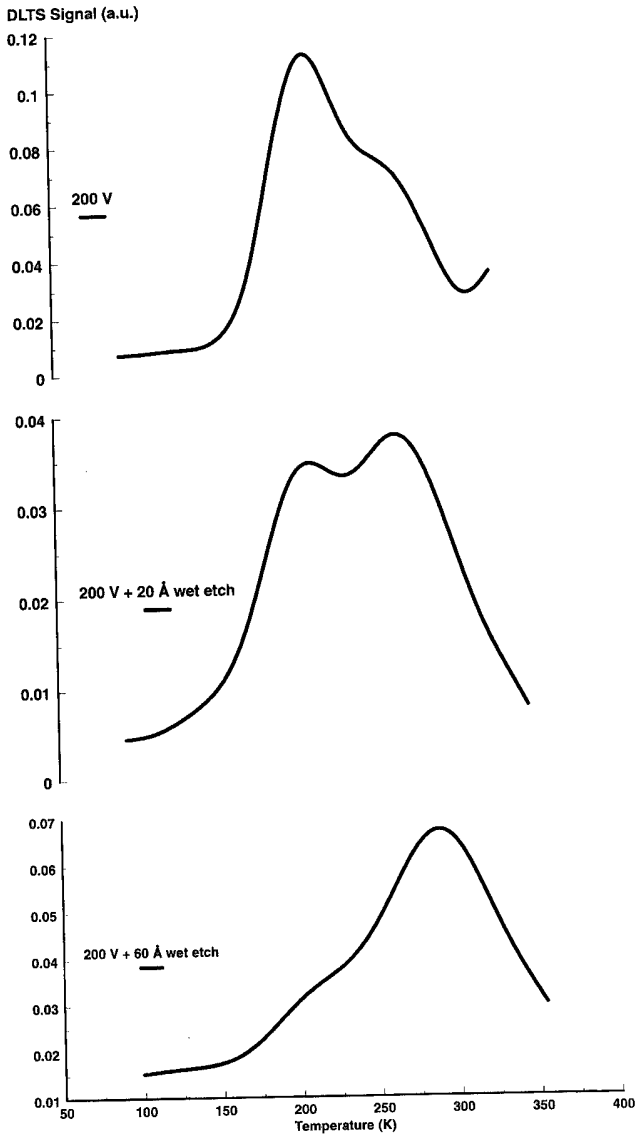


FIG. 11. DLTS spectra of samples after RIE at 200 V and wet-chemical etching.

ideality factor, when compared with a reference wet etched sample. This can be attributed to the creation of a damaged layer near the surface, that leads to a higher density of interface states (metal/semiconductor interface). As indicated by temperature measurements performed on Schottky diodes, the real barrier height of the samples etched at 200 V (low ion energy) compares well with those obtained from the room temperature $I-V$ characteristics (0.5 eV). For samples etched at a self-bias voltage of 450 V or greater, the effective barrier height is higher than the real barrier height. This effect can be attributed to the existence of an interfacial layer at the Schottky junction. It should be noted here that for the fabrication of the Schottky diodes, RIE was performed using identical etch times and gas residence times. Since hydrocarbon polymer deposition and surface damage formation increase with increasing rf power, it can be inferred that the thickness of the heavily damaged region present at the InAlAs surface increases with rf power. Structural investigations of the induced defects was not presented in this work, but we found, in the case of InP after RIE,[24] a surface damage which consists of a contamination (C, H, O) layer on top, then an In_2O_3 oxide layer and a defect layer. The interfacial layer present at the Schottky junction is probably due to this heavily damaged region. The higher ideality factor obtained on samples etched at high self-bias voltage $\geqslant$450 V (1.3, 1.4, and 1.6) actually suggests the presence of a thin insulating layer at the interface.

Taylor and Morgan[23] have examined the electrical properties of irradiated Ni/n-GaAs Schottky barriers and have shown that an excess current flows, the extent of which is dependent on the concentration of radiation induced damage centers generated at the metal–semiconductor interface. This proposal is consistent with the observed behavior of the current/voltage characteristics in our Schottky diodes. The proposed model is based on the existence of defect states near the interface whose density is enhanced by irradiation and due to the interface state field; the conduction and valence bands are bent downward at the interface. In reverse bias, the current transport mechanism is essentially a trap assisted tunneling process. When, however, trapping levels are present within the forbidden energy gap, then tunneling electrons may be captured by these levels (trap assisted tunneling process) and, if the temperature is high enough, they will be thermally emitted to the conduction band. At low reverse bias, a generation mechanism can develop through band center levels, thus resulting in a generation current; this phenomenon will be increased by radiation damage, producing a greater number of defect levels.

Let us now consider the effects of a thermal treatment. Samples etched at 200 V show an increase in the effective barrier height and reverse $I-V-T$ current after thermal annealing. This effect can be explained simply by considering the incorporation of an oxidelike insulator (containing possibly Al_2O_3) at the Schottky junction. It has been noted by Wada *et al.*[12] that oxidation of InP increases the effective barrier height. They proposed an expression of the barrier height ϕ_b which takes into account the existence of an interfacial oxide layer. This analytical model is based on the oxide charge effect which depends on the physiochemical properties of the oxide and/or the oxide–semiconductor interface and on the surface state density related to the existence of defects near the interface. Many authors[13,14] have discussed the nature of the oxide present at the InAlAs surface after reactive ion etching, and they found significant concentrations of metal oxides such as InO_x and AlO_x using Auger electron spectroscopy and x-ray photoelectron spectroscopy. They found that group V elements also form oxides like As_2O_3 and As_2O_5. In this case, a MIS Schottky junction is formed and tunneling transport through the thin insulator (oxide) occurs, leading to an increase of the current at high reverse bias.

For the samples etched at 450 and 600 V, the effective barrier height did not change after annealing. This effect can

be attributed to a highly damaged surface layer with a large density of defects induced by higher plasma self-bias. In this case, the native interfacial oxidelike insulator incorporates a high density of traps which favor trap assisted tunneling, resulting in an increased leakage current. Diodes processed at 800 V degrade after annealing.

Temperature measurements on Schottky diodes fabricated on all these samples show an effective barrier typically higher by 0.1 eV than the real barrier after thermal treatment, confirming the possible presence of an oxidelike interfacial layer after RTA. Furthermore, the etched samples show an increase of the reverse current after thermal annealing, probably by tunneling through the oxide at the interface and also by tunneling through the traps at 0.45 eV, as supported by DLTS measurements.

In order to evaluate the thickness of the damaged region of the samples for which we obtain the smallest electrical degradation (RIE at 200 V), a set of $I-V$ measurements were recorded, after step-by-step chemical etching. We found after etching a 60 Å layer, an improvement in the electrical characteristics of the Schottky diodes in terms of barrier height, ideality factor, and leakage current. These data confirm our assumption about the presence of a heavily damaged interfacial layer at the Schottky junction after RIE.

The DLTS data after RIE showed an increase of the concentration of the level E_2 (0.42 eV) which was accompanied by a broadening of the peak related to these defects. The presence of this wide peak shows that a band of defects has been created in the upper half of the band gap. This observation has been already reported in the case of GaAs after ion implantation, and it has been associated with the so-called U band.[15]

The reverse current measured on the Schottky diodes processed on the etched samples can be explained by tunneling current mechanism through these defect levels as deduced by the Richardson plot, which suggests that the trap E_2 is associated with plasma induced damage. In the same way, the concentration of trap E_2 was found to increase in InAlAs after CHF_3/SF_6 RIE of a SiN-UVCVD film deposited on InGaAs/InAlAs/InP heterostructures.[16]

The extent of the damaged region after RIE has been reported by many authors for GaAs and InP.[17] They concluded that the RIE induced damages are distributed over a region deeper than the projected ion range. This may be related to phenomena such as channeling, vacancy migration, and enhanced diffusion induced by ion bombardment.[18] On the other hand, the main sources of defects induced by plasma processing at low ion energy in InP based materials are localized in the surface region, probably as a result of preferential loss of group V elements leading to nonstoichiometric surface. Therefore, the Fermi level pinning at the surface by such defects should be taken into consideration.[19] A unified defect model has been proposed by Spicer *et al.*[19] which suggests possible correlation between the Fermi level pinning and the dominating defect centers in GaAs, especially As_{Ga} and Ga_{As}. Then, the change in Schottky barrier height is governed by the movement of the Fermi level, which is related to defect states. Our results support such assumptions, since the real barrier height obtained on Schottky diodes after RIE corresponds to the energetic position of the dominating deep levels.

The increase of the concentration in the E_2 trap after RIE and its association with plasma induced damages are confirmed using wet chemical etching after RIE: a progressive decrease of E_2 concentration is observed when etching the surface down to the value found in the reference sample.

It should be noted here that the concentration of the level E_2 after RIE does not increase with increasing self-bias, and at the same time a broadening of the corresponding DLTS spectra is observed after RIE. This suggests that the level E_2 certainly corresponds to complex defects, since it is quite different from traps created by other types of technological processes such as electron irradiation[20] which create simple defects. It has been proposed by Luo *et al.*[21] that the trap E_2 is related to an arsenic related complex defect such as di-interstitial. Then, after RIE, an increase of such complex defects (E_2) may result from the association of the impurities already present in the starting material and the created defects. Such phenomena have been reported in the case of boron implantation in GaAs,[15] where complex defects result from the association of donors and/or impurities present in GaAs together with the created defects.

The levels E_3 (0.6 eV) and E_1 (0.2 eV) behave differently. A decrease of the concentration of the level E_3 (0.6 eV) is observed after RIE when compared to the wet etched sample (dominant defect) while the level E_1 (0.2 eV) disappears after RIE.

Annealed samples show a small decrease of the concentration of the level E_2, while levels E_1 and E_3 reappear with lower concentration. However, a reduction of the broadening spectra of the level E_2 is the main effect of thermal annealing, suggesting a reduction of its energy distribution. The low decrease of the concentration of the level E_2 after RTA supports the suggestion that is related to complex defect, since simple point defects in general can be easily annealed out.

However, variations of the levels E_1 and E_3 should be compared with observations made by Pang *et al.* on GaAs after ion implantation.[22] They found a dramatic reduction of the EL2 level after Ar ion implantation, and they attributed EL2 to an antisite As_{Ga} defect or a complex involving the antisite As_{Ga} defect. Then, the lower EL2 concentration after Ar implantation, implies a reduction in the density of As_{Ga}. After RTA, the EL2 peak increases with temperature between 400 and 600 °C, but with lower concentrations when compared to their reference sample. They attributed this effect to As loss during RTA. However, in InAlAs, the mechanisms of defect formation are still unknown, but we found variations after RIE similar to those for GaAs. Then, further studies are needed to fully understand the variations of levels E_1 and E_3 after RIE and the effects of a thermal treatment on the deep electron traps.

V. CONCLUSION

CH_4/H_2 RIE induced damages in InGaAs/InAlAs/InP heterostructures have been characterized by means of $I-V$, $C-V$, and DLTS measurements. We have observed a decrease in Schottky barrier height and an increase in ideality factor for the etched samples. The corresponding DLTS spectra have shown an increase of the defect at 0.42–0.50 eV, a decrease of the defect density at 0.6 eV while the deep traps at 0.2 eV disappear.

$I-V$ measurements on samples etched at 200 V show a decrease of the real barrier height from 0.6 eV (reference sample) to 0.48 eV and a tunneling current through traps at 0.45 eV, which is found by DLTS measurements to be the dominant defect level. Annealing of these samples at 400 °C induces a small reduction of the trap density at $E_c-E_t\approx 0.4$ eV. However, the samples cannot fully recover their initial characteristics since there is only a small reduction of the defect density and possibly a modification of their depth distribution. A wet etching of these samples after RIE improves the electrical characteristics of the Schottky diodes, and we found that the thickness of the heavily damaged region extends to ~60 Å from the surface. The damages induced by plasma processing at 450, 600, and 800 V are not removed by a thermal treatment.

Finally, RIE with CH_4/H_2 at low self-bias can allow a good control of the etching depth but at the expense of a significant induced surface damage. Characterization of these RIE induced damage suggests that the use of CH_4/H_2 reactive ion etching in transistor fabrication (gate recessing) should be done at even lower self-bias (<200 V).

ACKNOWLEDGMENTS

The authors wish to thank A. Scavennec for useful discussions and guidance; J. Etrillard for fruitful discussions on RIE technique; and L. Cohausz, C. Joly, J. L. Courant, and S. Biblemont for technical assistance.

[1]H. E. G. Arnot, R. W. Glew, G. Schiavini, L. J. Rigby, and A. Piccirillo, Appl. Phys. Lett. **62**, 3189 (1993).
[2]S. Agarwala, I. Adesida, C. Caneau, and R. Bhat, J. Vac. Sci. Technol. B **11**, 2258 (1993).
[3]L. D. Nguyen, A. S. Brown, M. A. Thompson, and L. Jelloian, IEEE Trans. Electron Devices **39**, 2007 (1992).
[4]C. Lauterbach, H. Albrecht, M. Beschoner, R. Gessner, and M. Schier, Proceedings of the Third International Conference Indium Phosphide and Related Materials, Cardiff, Wales, 1993 (unpublished), p. 610.
[5]D. Lecrosnier, L. Henry, A. Le Corre, and C. Vaudry, Electron. Lett. **24**, 1254 (1987).
[6]P. Kwan, K. N. Bhat, J. M. Borrego, and S. K. Ghandi, Solid-State Electron. **26**, 125 (1983).
[7]S. M. Sze, in *Physics of Semiconductors* (Wiley, New York, 1969).
[8]S. Weiss and R. Kassing, Solid-State Electron. **31**, 1733 (1988).
[9]W. R. Hong, S. Dhar, P. K. Bhattacharya, and A. Chin, J. Electron. Mater. **16**, 271 (1987).
[10]K. Nakashima, S. Nojima, Y. Kawamura, and H. Asahi, Phys. Status Solidi **103**, 511 (1987).
[11]P. Omling, L. Samuelson, and H. G. Grimmeiss, J. Appl. Phys. **54**, 5117 (1983).
[12]O. Wada, A. Majerfeld, and P. N. Robson, Solid State Electron. **25**, 381 (1982).
[13]S. J. Pearton, W. S. Hobson, F. A. Baiocchi, A. B. Emerson, and K. S. Jones, J. Vac. Sci. Technol. B **8**, 57 (1990).
[14]N. Yoshida, M. Totsuka, J. Ino, and S. Matsumoto, Jpn. J. Appl. Phys. **33**, 1248 (1994).
[15]G. M. Martin, P. Secordel, and C. Venger, J. Appl. Phys. **53**, 8706 (1982).
[16]A. Clei, R. Palla, and J. C. Harmand (unpublished).
[17]S. J. Pearton, U. K. Chakrabarti, and F. A. Baiocchi, Appl. Phys. Lett. **55**, 1633 (1989).
[18]J. M. E. Harper, J. J. Cuomo, and H. R. Kaufman, J. Vac. Sci. Technol. **21**, 737 (1982).
[19]W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. **16**, 1422 (1979).
[20]D. Pons, A. Mircea, and J. Bourgoin, J. Appl. Phys. **51**, 4150 (1980).
[21]J. K. Luo, H. Thomas, S. A. Clark, and R. H. Williams, Proceedings of the IPRM, 1994 (unpublished), p. 363.
[22]Y. G. Wang and S. Ashok, J. Vac. Sci. Technol. B **12**, 2280 (1994).
[23]P. D. Taylor and D. V. Morgan, Solid-State Electron. **19**, 481 (1976).
[24]M. Achouche, J. F. Bresse, M. Juhel, J. Decobert, J. Etrillard, and A. Clei, Proceedings of the Tenth International Colloquium on Plasma Processes, Antibes, France 1995 (unpublished), p. 496.

Cl_2/Ar plasma etching of binary, ternary, and quaternary In-based compound semiconductors

J. W. Lee, J. Hong, C. R. Abernathy, E. S. Lambers, and S. J. Pearton[a)]
University of Florida, Gainesville, Florida 32611

W. S. Hobson and F. Ren
AT&T Bell Laboratories, Murray Hill, New Jersey 07974

(Received 30 November 1995; accepted 1 May 1996)

A simple Cl_2/Ar plasma chemistry can provide smooth, high-rate etching of InP, InAs, InGaAs, AlInAs, and InGaAsP at room temperature under conditions in which there is a balance between formation and sputter desorption of the normally involatile $InCl_3$. When the neutral/ion ratio is either too high or too low, surface roughening is apparent due either to the presence of $InCl_3$, or to preferential loss of the group V element. The etching has been investigated as a function of microwave power (600–1000 W), rf power (0–300 W), process pressure (1.5–10 mTorr), and Cl_2:Ar ratio under electron cyclotron resonance conditions. Use of N_2 or H_2, rather than Ar, as gas additives to the chlorine, did not produce smooth, stoichiometric etched surfaces. © *1996 American Vacuum Society.*

I. INTRODUCTION

In-based compound semiconductors form the basis for a variety of technologically important devices, including InGaAsP/InP diode lasers for optical communication systems operating at 1.3 and 1.5 μm,[1] InGaAs/AlInAs heterojunction bipolar transistors for high bit-rate digital applications,[2] and InAs infrared detectors.[3] One of the future directions will be continued miniaturization and higher device densities, which requires development of uniform, high-yield etching techniques for pattern transfer. In fabricating the devices mentioned above there are a wide range of etch depths needed, from relatively shallow mesas (<1000 Å) (Ref. 4) to through wafer via holes up to 100 μm deep.[5–7] Since CH_4/H_2-based plasma chemistries provide maximum etch rates of $\leqslant$1000 Å/min for InP and related materials, they are suitable only for a limited number of these applications.[8] Traditionally, Cl_2-based plasma chemistries have required concurrent sample heating in order to aid desorption of the $InCl_3$ etch product.[9–18] This heating adds additional processing complexity and needs addition of a sidewall polymer forming gas such as CH_4 to prevent undercutting.

Recently it has been reported that simple Cl_2/Ar plasmas can produce very high etch rates for In-based semiconductors at room temperature when operated under electron cyclotron resonance (ECR) conditions.[19,20] A plausible explanation is that the high ion flux produces efficient sputter desorption of the $InCl_3$ before it forms a thick selvage layer. It is necessary to more completely investigate the parameter space for this plasma chemistry to establish the usable process window. In this article we report on ECR etching of InP, InAs, InGaAs, AlInAs, and InGaAsP in Cl_2/Ar as a function of plasma composition, microwave and rf power, process pressure, and gas additive (Ar, N_2, or H_2). Acceptably smooth etched surface morphologies are obtained for specific sets of these parameters. We also confirm that the chlorine etch residue concentration is much lower on ECR Cl_2/Ar etched InP compared to that measured for samples exposed to reactive ion etch (RIE) discharges.

II. EXPERIMENT

Epitaxial layers of $In_{0.53}Ga_{0.47}As$, $Al_{0.48}In_{0.52}As$, or InGaAsP (λ=1.3 μm) lattice matched to InP were grown on semi-insulating (Fe-doped) substrates by either metal organic chemical vapor deposition[21] or metal organic molecular beam epitaxy.[22] Wafers of InP or InAs from undoped Czochralski boules, together with the epitaxial materials, were lithographically patterned with AZ5209E photoresist to produce features with dimensions between 0.5–100 μm. Etch

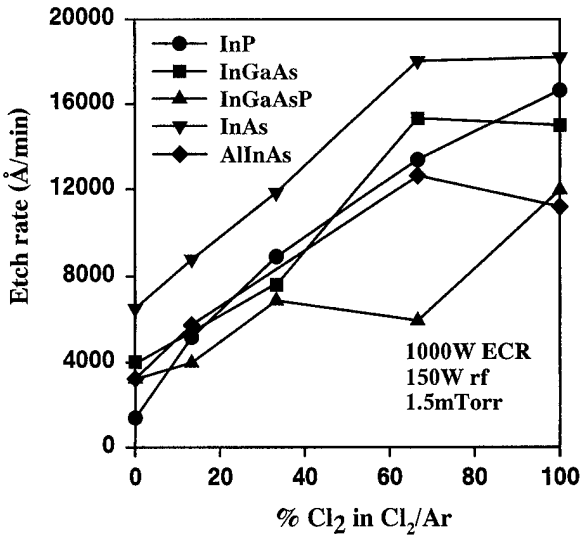


FIG. 1. Etch rates of In-based materials as a function of Cl_2 composition in Cl_2/Ar ECR discharges (1.5 mTorr, 1000 W microwave, 150 W rf, and total gas flow rate 15 sccm).

a)Electronic mail: spear@mse.ufl.edu

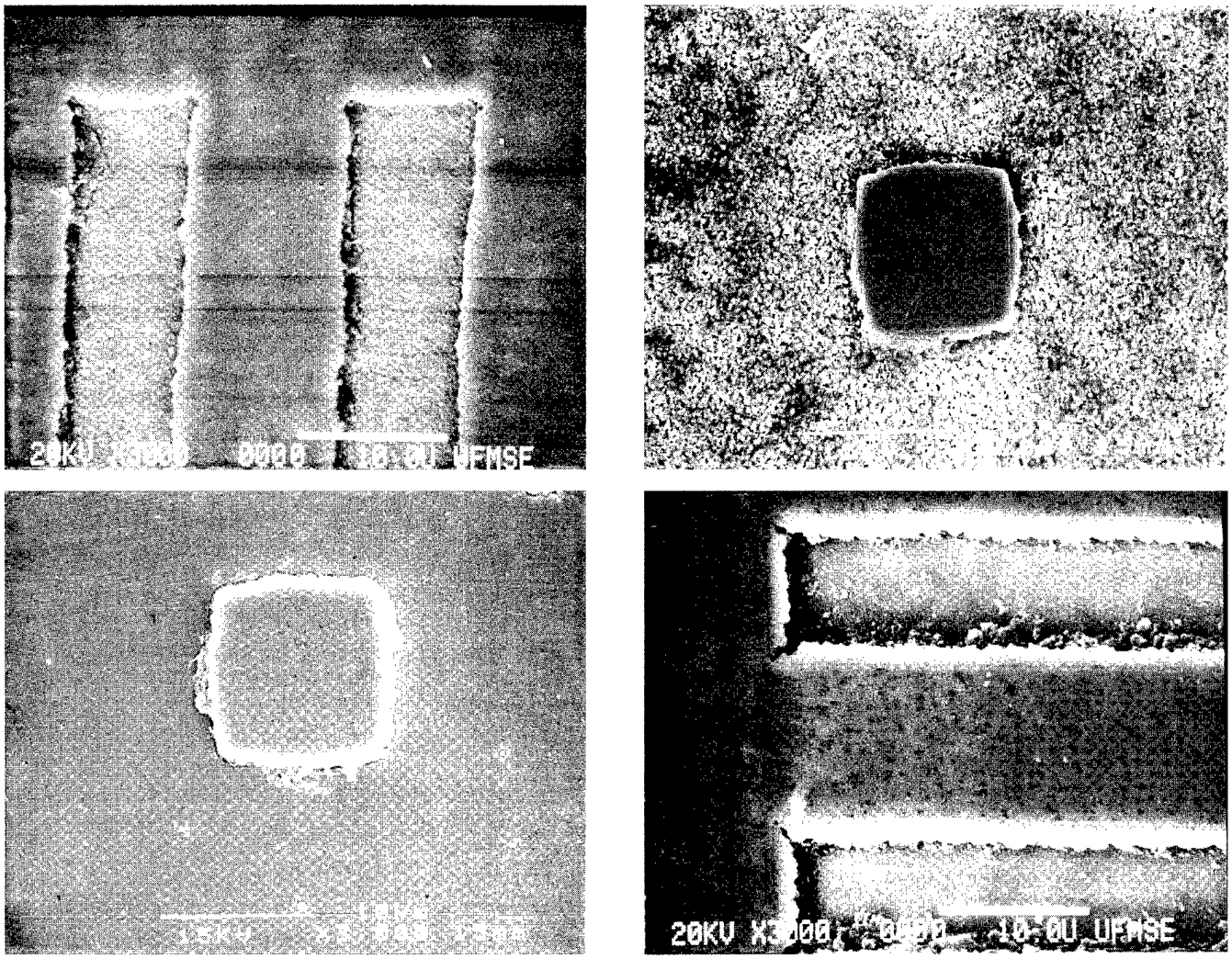


FIG. 2. SEM micrographs of features etched into InP using 1.5 mTorr, 1000 W microwave, 150 W rf discharges of pure Ar (top left), $2Cl_2/13Ar$ (top right), $10Cl_2/5Ar$ (bottom left), or pure Cl_2 (bottom right).

rates were measured by stylus profilometry of the largest features after removal of the resist in acetone. Scanning electron microscopy (SEM) and atomic force microscopy (AFM) were used to examine the surface morphology of the etched samples, while Auger electron spectroscopy (AES) was employed to monitor the near-surface stoichiometry at three spots across the 5×5 mm^2 samples.

The etching was performed in a Plasma-Therm SLR 770 system. The samples enter through a load lock and are mechanically clamped to a He backside-cooled chuck held at 25 °C that is biased with rf (13.56 MHz) power (0–300 W). We believe from observation of resist behavior that the sample temperature did not rise above 70–80 °C during the short (1 min) exposures employed. Thus, we do not believe sample heating plays a significant role in our data. The plasma is generated with application of 600–1000 W of 2.45 GHz microwave power to an Astex low profile ECR source. Electronic grade Cl_2, Ar, N_2, or H_2 were injected into this source through mass flow controllers. The process pressure was varied from 1.5–10 mTorr by throttling the turbomolecular pump that evacuates the reactor chamber. We held the rf power constant, rather than dc-induced voltage. The latter varies from approximately −60 V at 50 W rf to −220 V at 200 W rf, as measured at the sample electrode. The biases are higher under RIE conditions (−80 V at 50 W, −290 V at 200 W). The magnetic field was not used when only the bottom rf power was applied.

III. RESULTS AND DISCUSSION

Figure 1 shows the etch rates of the five different materials as a function of plasma composition in Cl_2/Ar discharges. The total gas flow rate was held constant at 15 standard cubic centimeters per minute (sccm). The general trend of higher etch rates with increased chlorine content is expected because of the greater active species concentration available for reaction with the semiconductors. Note at this low pressure (1.5 mTorr) and high microwave power (1000 W) condition, the sputter rates with pure Ar are $\geqslant$1200 Å/min for all the materials. This is a result of the high-ion flux under these conditions, but represents a severe limitation if one is trying to achieve selectivity for one material over another. At Cl_2

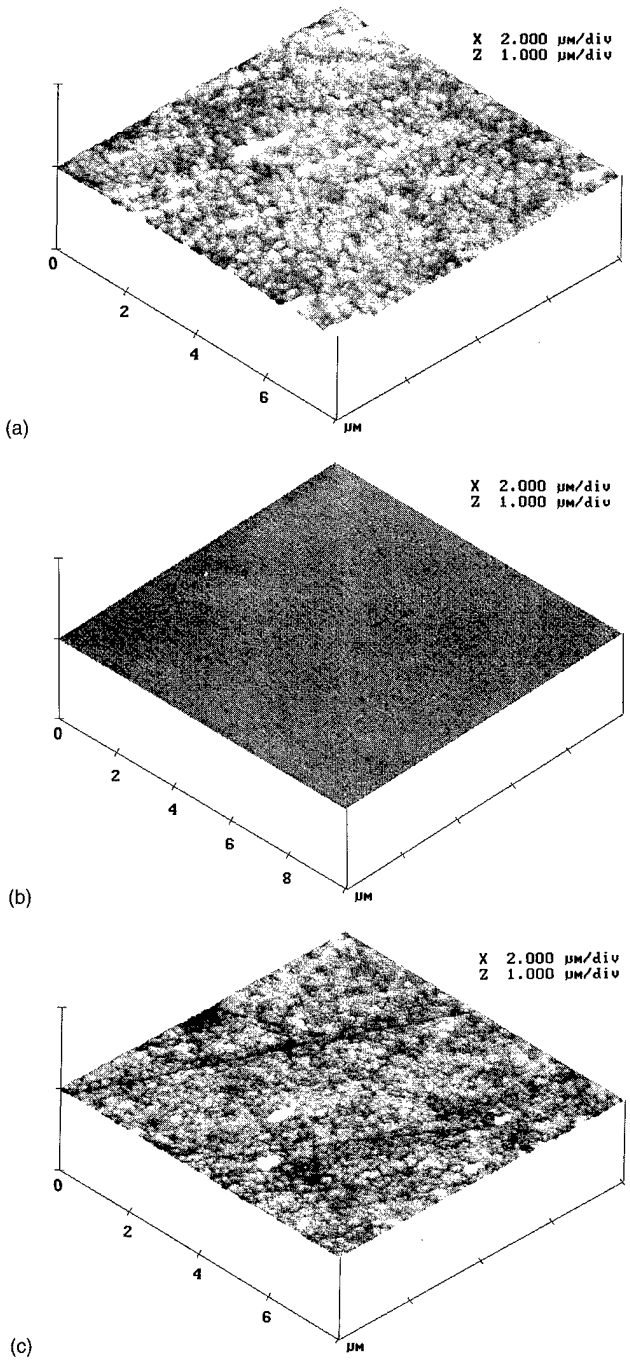


FIG. 3. AFM scans from InP samples etched in 1.5 mTorr, 1000 W microwave, 150 W rf discharges of pure Ar (a), $5Cl_2/10Ar$ (b), or pure Cl_2 (c).

percentages $\geq 67\%$ the etching was significantly isotropic, while below this value the sidewalls were basically vertical (sidewall angles $\geq 80°$).

The surface morphology of the In-based materials was a strong function of the Cl_2:Ar ratio in the discharge. As an example, Fig. 2 shows SEM micrographs of features etched into InP with 1.5 mTorr, 1000 W microwave and 150 W rf discharges of pure Ar (top left), $2Cl_2/13Ar$ (top right), $10Cl_2/5Ar$ (bottom left), or pure Cl_2 (bottom right). The morphology with pure Ar is quite rough due to preferential sputtering of P, as reported previously.[23] Similarly, with pure Cl_2 the large differential in volatility between the $InCl_3$ and PCl_3 etch products also leads to a rough morphology. However, at a 10:5 ratio of Cl_2:Ar the surface is very smooth, which confirms that it is possible to balance the removal rates of the etch products without heating the sample by optimizing the ion/neutral flux ratio. To quantify the surface roughness, AFM was performed in the tapping mode. Figure 3 shows AFM scans from InP etched in pure Ar (a), $5Cl_2/10Ar$ (b), or pure Cl_2 (c). The root-mean-square roughness measured over 10×10 μm areas was 7.3 nm for the Ar-etched materials, and 14.4 nm for the pure Cl_2 sample, while the Cl_2/Ar (5/10) sample had an rms value of only 0.64 nm. The latter is comparable to an untreated substrate. We were able to achieve this surface quality for Cl_2 flow rates between 5 and 10 sccm in the total Cl_2+Ar flow rate of 15 sccm. The near surface of these samples retained their stoichiometry, even for etch depths of several microns. Figure 4 shows an AES surface scan (top) and depth profile (bottom) from an InP sample etched for 2 min in a $10Cl_2/5Ar$, 1.5 mTorr, 1000 W microwave, 150 W rf discharge. There are several noteworthy aspects to this data. First, there are no chlorine residues remaining on the surface, suggesting that sputter-enhanced removal of the $InCl_3$ is very efficient under these conditions. In addition, only adventitious carbon and oxygen from the native oxide are present as contaminants on the surface. These result from the fact that the samples are exposed to atmosphere upon removing them from the reactor to transfer to the AES system.

Figure 5 shows the dependence of material etch rate on microwave power in $2Cl_3/13Ar$, 1.5 mTorr, 150 W rf discharges. The rates increase at higher powers due to a higher atomic density (as evidenced from monitoring the optical emission line intensity) and a higher ion flux. The transition from rough-to-smooth morphologies shifts to lower Cl_2:Ar ratios at lower microwave powers because of the need to maintain the balance between $InCl_3$ formation and removal.

The vast difference in etch rates for the five materials between the ECR mode and conventional reactive ion etching (RIE) is shown in Fig. 6, where the rf power dependence of the rates has been determined both with and without 1000 W microwave power. Again this is a result of the much higher active chlorine densities and the higher ion density ($\sim 5\times10^{11}$ cm^{-3} for ECR versus 3×10^9 cm for RIE, as obtained from microwave interferometry measurements.).[24] Indeed, under RIE conditions it is necessary to heat the InP to achieve rates >500 Å/min. Temperatures of 200–250 °C are generally employed under these conditions.

The difference between the two etch modes is not simply the rates. Figure 7 shows AES surface scans from two InP samples etched in $2Cl_2/13Ar$, 1.5 mTorr, 300 W rf discharges with either 1000 W microwave power (top) or 0 W, i.e., RIE mode (bottom). The In/P ratio is much higher in the latter sample and it has a significant amount of residual chlorine.

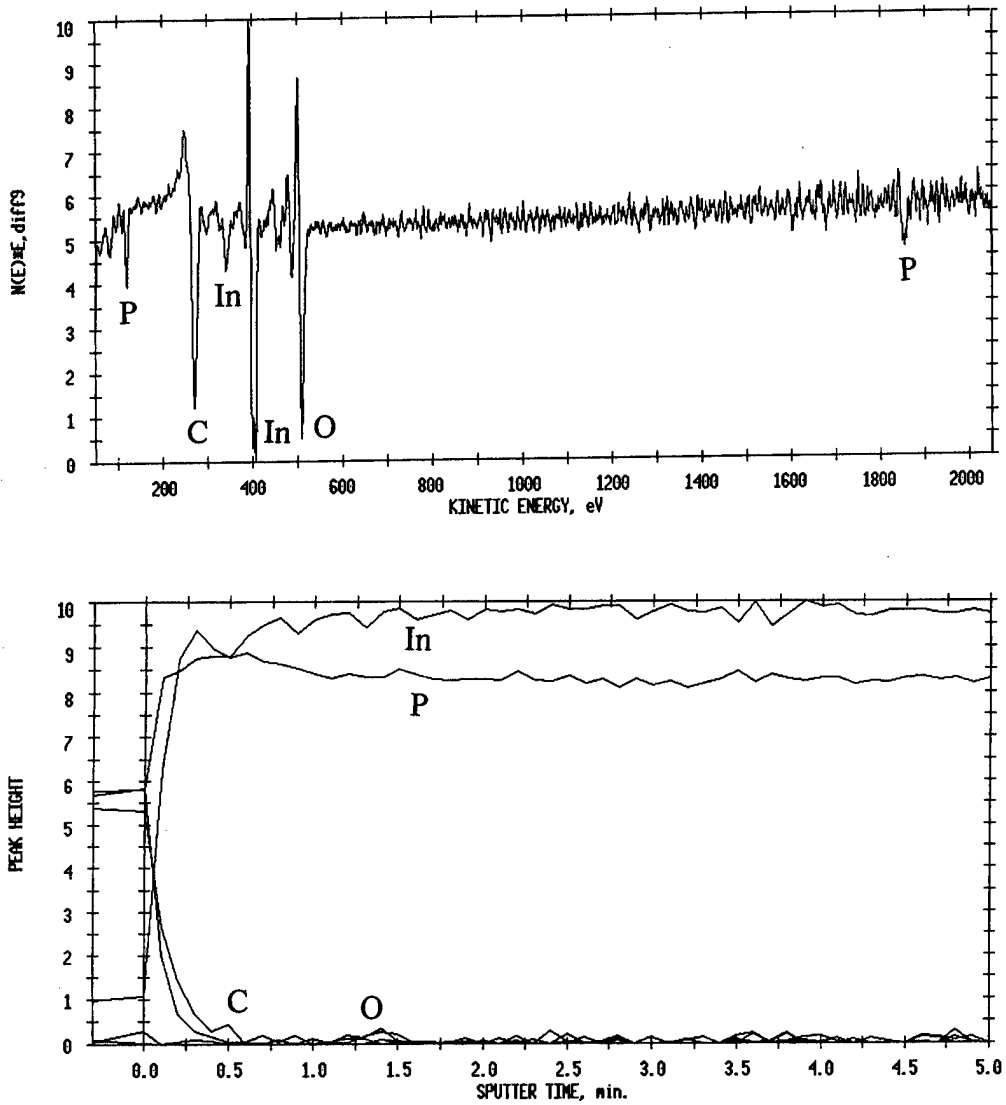


FIG. 4. AES surface scan (top) or depth profile (bottom) from InP etched in a $10Cl_2/5Ar$, 1.5 mTorr, 1000 W microwave, 150 W rf discharge.

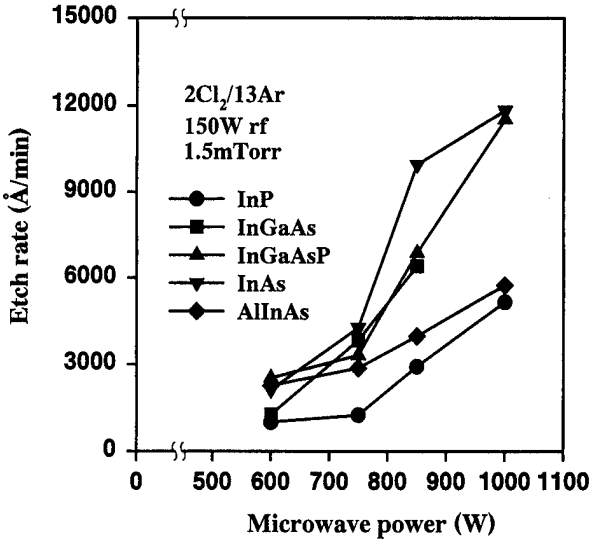


FIG. 5. Etch rates of In-based materials as a function of microwave power in 2Cl/13Ar, 1.5 mTorr, 150 W rf discharges.

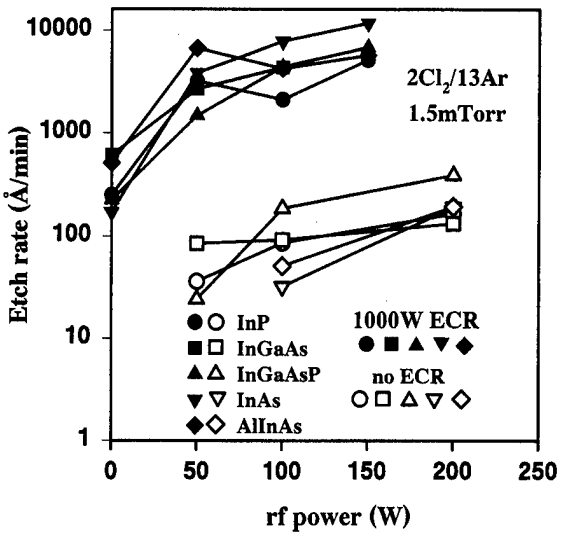


FIG. 6. Etch rates of In-based materials as a function of rf power in $2Cl_2/13Ar$, 1.5 mTorr discharges, either with or without 1000 W of microwave power.

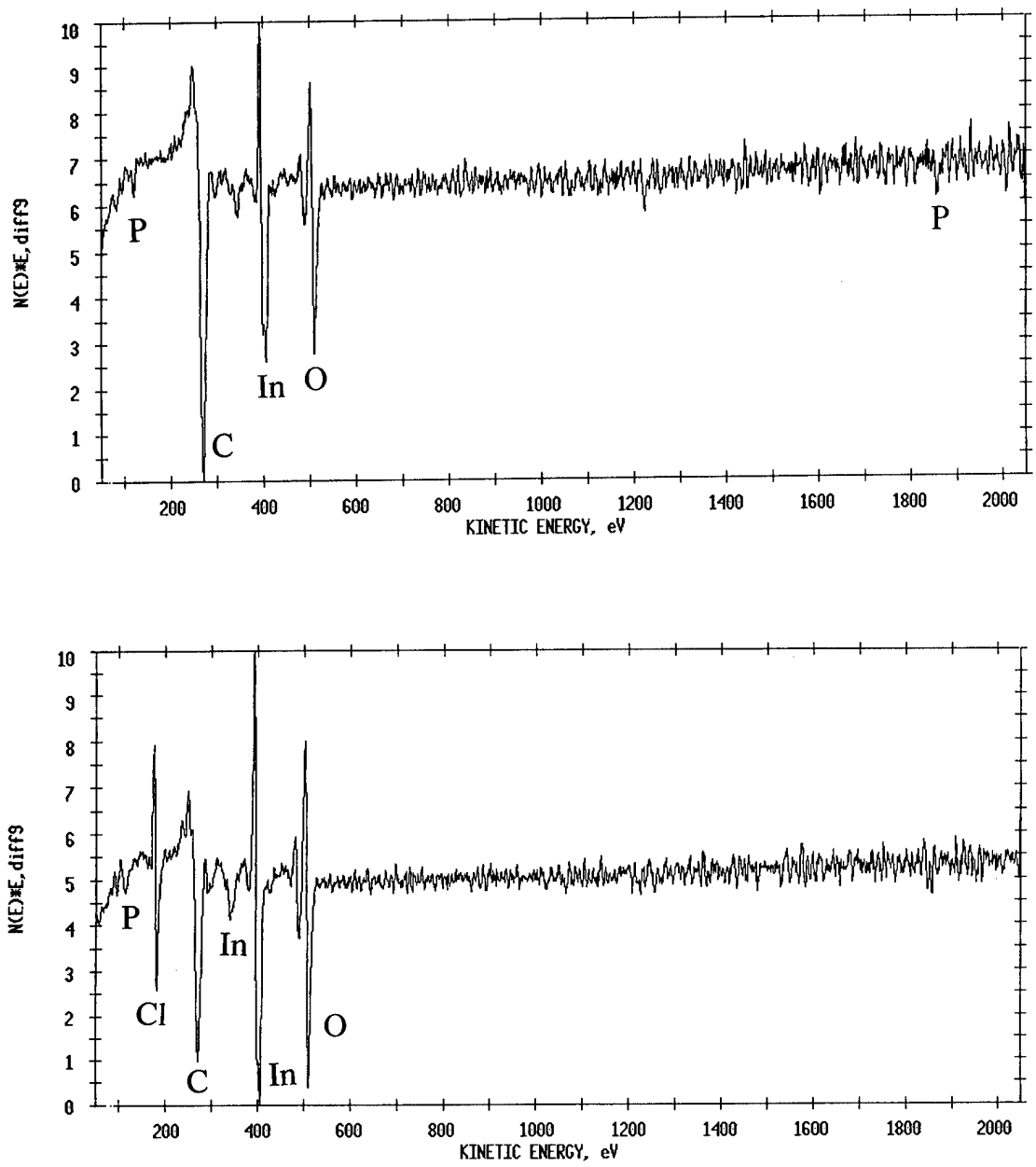


FIG. 7. AES surface scans from InP samples etched in $2Cl_2/13Ar$, 1.5 mTorr, and 300 W rf discharges with either 1000 W (top) or 0 W (bottom) microwave power.

This is consistent with the presence of an $InCl_3$ selvage layer under RIE conditions, which normally requires elevated sample temperatures to desorb. The higher ion flux under ECR conditions appears to prevent formation of this "poisoning" layer.

The pressure dependence of etch rate at fixed microwave (1000 W) and rf (150 W) powers is shown in Fig. 8. While the etch anisotropy degrades somewhat above 7.5 mTorr, the etch rates are well suited to via-hole formation processes, where rates of at least 1 μm/min are necessary.[5–7] Moreover, the rms surface roughness of InP and the other materials improves dramatically at higher pressure, as shown in Fig. 9. We again assume that for this plasma composition of $2Cl_2/13Ar$ (which gives rough morphologies at 1.5 mTorr) if one increases the neutral/ion ratio, then it is possible to balance the formation and removal of the $InCl_3$.

There has been recent interest in the use of Cl_2/N_2 mixtures for ECR etching of InP.[20,25] Smoother surfaces were obtained relative to either pure Cl_2 or Cl_2/Ar under some conditions.[25] Figure 10 shows the InP etch rate as a function of chlorine percentage in Cl_2/Ar, Cl_2/H_2 or Cl_2/N_2 at fixed pressure (1.5 mTorr, rf (150 W) and microwave (1000 W)

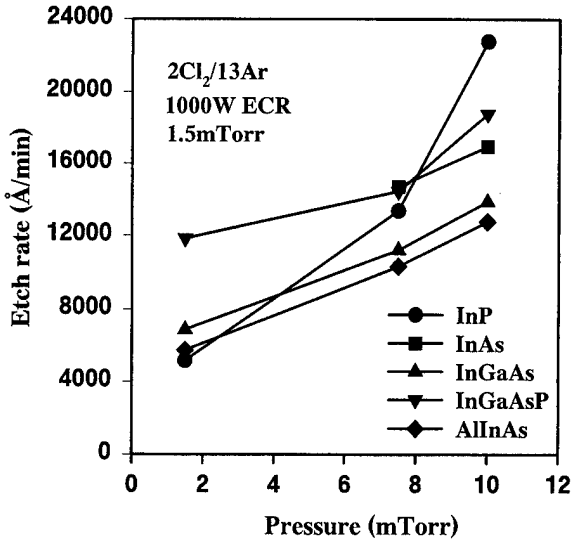


FIG. 8. Etch rates of In-based materials as a function of pressure in $2Cl_2/13Ar$, 150 W rf, and 1000 W microwave discharges.

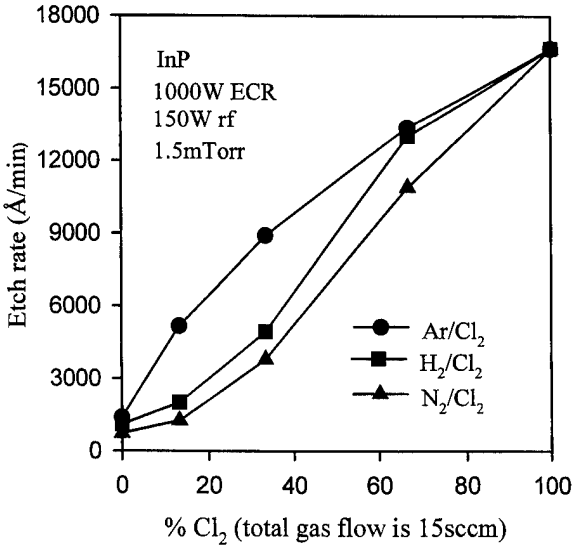


FIG. 10. Etch rates of InP as a function of Cl_2 percentage in 1000 W microwave, 150 W rf, and 1.5 mTorr discharges of Cl_2/Ar, Cl_2/N_2 or Cl_2/H_2.

power. The addition of either H_2 or N_2 reduces the etch rate relative to Ar. This appears to be due both to the lighter mass of these species reducing the sputtering yield, and to a reduction in atomic chlorine density. Figure 11 shows optical emission spectra in raw intensities from both Cl_2/N_2 (top) and Cl_2/H_2 (bottom) discharges under the same conditions. The reduction in intensity of the chlorine neutral radical lines between 700 and 850 nm was also observed by Miyakuni *et al.*[25] Since the excitation cross section of active species depends on electron temperature, which will change with source power, pressure, and other parameters, it is necessary to be cautious in inferring ground state [Cl] densities, but in each case we observed a decrease in neutral radical line intensity when measured against standard Ar lines. We did not observe InP morphologies as smooth for these two chemis-

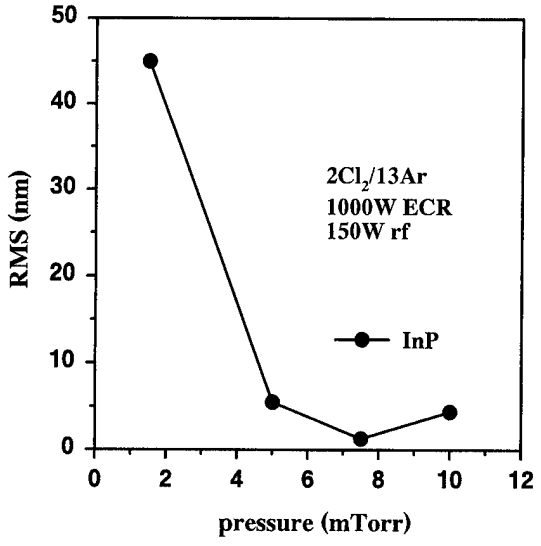


FIG. 9. rms surface roughness of InP samples etched in $2Cl_2/3Ar$, 1000 W microwave, and 150 W rf discharges, as a function of process pressure.

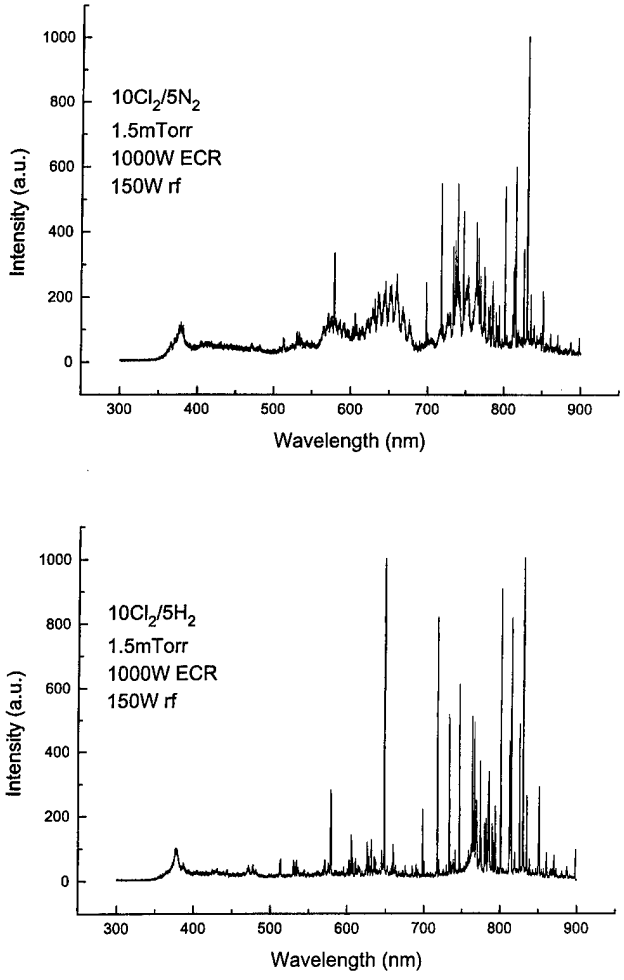


FIG. 11. Optical emission spectra from 1000 W microwave, 150 W rf, and 1.5 mTorr discharges of $10Cl_2/5N_2$ (top) or $10Cl_2/5H_2$ (bottom).

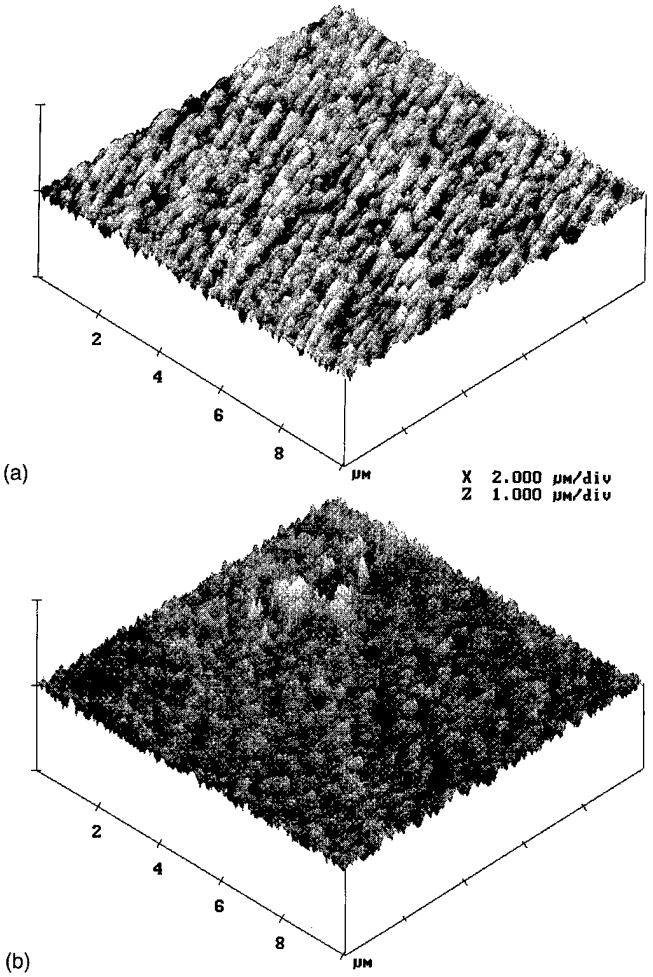


FIG. 12. AFM scans from InP samples etched in 1000 W microwave 150 W rf, and 1.5 mTorr discharges of $10Cl_2/5N_2$ (top) or $10Cl_2/5H_2$ (bottom).

tries as obtained with the simple Cl_2/Ar mixtures. AFM surface scans are shown in Fig. 12 from samples etched in 1.5 mTorr, 1000 W microwave, and 150 W rf discharges of $10Cl_2/5N_2$ (top) or $10Cl_2/5H_2$ (bottom). The rms roughness values were 31.9 and 34.7 nm, respectively, much worse than that obtained for Cl_2/Ar (0.64 nm). Clearly the addition of Ar is the best choice in terms of etch rates and surface morphology. The drawback of using H_2 or N_2 is that they add an additional chemical component to the plasma chemistry, which complicates the control of the near-surface stoichiometry.

IV. CONCLUSIONS

High etch rates near room temperature with smooth, stoichiometric surfaces are possible for In-based III–V semiconductors under ECR conditions by controlling the formation and desorption of the $InCl_3$ etch product. This is achieved at particular combinations of Cl_2-to-Ar flow rates for different pressures, microwave, or rf powers. AES surface analysis shows that under RIE conditions a thick $InCl_3$ selvage layer forms that leads to slow rates and rough morphologies.

ACKNOWLEDGMENTS

The work at UF is partially supported by a DOD University Research Initiative administered through AFOSR (Contract No. F49620-96-1-0026). The staff of the Microfabritech facility is gratefully acknowledged.

[1]See, for example, *Quantum Well Lasers*, edited by P. S. Zory (Academic, San Diego, CA, 1995).
[2]See, for example, *InP HBTs—Growth, Processing and Applications*, edited by B. Jalali and S. J. Pearton (Artech House, Norwood, MA, 1995).
[3]See, for example, F. Capasso in *Semiconductors and Semimetals, Vol. 24; Applications of Multiquantum Wells, Selective Doping and Superlattices*, edited by R. Dingle (Academic, San Diego, CA, 1987).
[4]F. Ren, Electron. Lett. **31**, 333 (1995).
[5]C. Constantine, Electron. Lett. **29**, 284 (1993).
[6]R. Khare, J. M. Brown, M. Hu, D. Pierson, M. Melendes, and C. Constantine, J. Vac. Sci. Technol. B **12**, 2947 (1994).
[7]K. Y. Hur, B. J. Guerin, and T. E. Kazior, J. Vac. Sci. Technol. B **12**, 1410 (1994).
[8]T. R. Hayes, in *InP and Related Materials-Processing, Technology and Devices*, edited by A. Katz (Artech House, Norwood, MA, 1992).
[9]S. C. McNevin, J. Vac. Sci. Technol. B **4**, 1216 (1986).
[10]C. Constantine, C. Barratt, S. J. Pearton, F. Ren, and J. R. Lothian, Appl. Phys. Lett. **61**, 2899 (1992).
[11]D. G. Lishan and E. L. Hu, Appl. Phys. Lett. **56**, 1667 (1990).
[12]G. J. Van Gurp, J. M. Jacobs, J. J. M. Binsma, and L. F. Tiemeijer, Jpn. J. Appl. Phys. **28**, 1236 (1989).
[13]C. Youtsey, R. Grundbacher, R. Panepucci, I. Adesida, and C. Caneau, J. Vac. Sci. Technol. B **12**, 3317 (1994).
[14]J. Schneider, M. Moser, and K. Affoter, *Proceedings of the 6th International Conference on InP and Related Materials* (IEEE-LEOS, Piscataway, NJ, 1994), p. 216.
[15]T. Yoshikawa, Jpn. J. Appl. Phys. **31**, 4386 (1992).
[16]M. Vernon, T. R. Hayes, and V. M. Donnelly, J. Vac. Sci. Technol. A **10**, 3409 (1992).
[17]S. W. Pang and K. K. Ko, J. Vac. Sci. Technol. B **10**, 2703 (1992).
[18]G. A. Vawter and C. I. H. Ashby, J. Vac. Sci. Technol. B **12**, 3374 (1994).
[19]K. K. Ko and S. W. Pang, J. Electrochem. Soc. **142**, 3945 (1995).
[20]F. Ren, W. Hobson, J. R. Lothian, J. R. Lopata, J. A. Caballero, S. J. Pearton, and M. W. Cole, Appl. Phys. Lett. **67**, 2497 (1995).
[21]W. S. Hobson, Mater Res. Soc. Symp. Proc. **300**, 75 (1993).
[22]C. R. Abernathy, J. Vac. Sci. Technol. A **11**, 889 (1993).
[23]S. J. Pearton, U. K. Chakrabarti, A. P. Perley, and K. S. Jones, J. Appl. Phys. **68**, 2760 (1990).
[24]S. J. Pearton, T. Nakano, and R. A. Gattscho, J. Appl. Phys. **69**, 4206 (1991).
[25]S. Miyakuni, R. Hattori, K. Sato, H. Takano, and O. Ishibara, J. Appl. Phys. **78**, 5734 (1995).

Metal stack etching using a helical resonator plasma

C. B. Labelle,[a] H. L. Maynard,[b] and J. T. C. Lee
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

(Received 7 December 1995; accepted 19 April 1996)

A low-pressure etching process for advanced aluminum metallization stacks was developed using a high-density helical resonator plasma source (Prototech model ESRF 600) mounted on a Lucas Labs cluster tool. The metallization stacks consisted of a 300 Å TiN antireflection layer on 6000 Å of Al (1% Cu) with a 1000 Å TiN diffusion barrier and a 100 Å Ti film to enhance adhesion to the underlying SiO_2. The features widths were as small as 0.45 μm. The films were etched using gas mixtures of Cl_2/BCl_3. The BCl_3 proved to be an important additive to reduce notching of the Al film at the interface between the Al and the top layer of TiN. Best feature profiles were obtained using 80–90 sccm Cl_2 and 10–20 sccm BCl_3 at the following reactor conditions: 2.0 mTorr, wafer platen temperature $T=0$ °C, 100 W rf bias power, and 1500 W source power. More anisotropic profiles are obtained by either decreasing the wafer platen temperature or increasing the rf-bias power. The photoresist is also stripped in the same process chamber using an oxygen plasma at 5 mTorr, 50 W rf-bias power, and 1500 W source power at a chuck temperature of 25 °C. Extensive application of real-time process diagnostics, including optical emission spectroscopy and full wafer interferometry, aided process development by identifying end points, etching rates, and etching rate uniformities.
© *1996 American Vacuum Society.*

I. INTRODUCTION

As the manufacturing of integrated circuits moves toward 0.25 μm design rules, the requirements for metal etching are increasingly difficult to meet using conventional parallel-plate reactive ion etchers (RIEs). An alternative to these etchers are low-pressure, high ion-density plasma sources such as helical resonators, helicons, transformer/inductively coupled plasma (TCP/ICP) sources, and electron cyclotron resonance (ECR) sources. These new plasma sources are capable of highly anisotropic etching due to their low-pressure operation (2–5 mTorr), can maintain high throughput due to the high ion densities achievable (10^{11}–10^{12} cm^{-3}), exhibit high etching rate uniformity via single-wafer processing, and permit independent control of the ion energy through wafer biasing.[1,2] For metal etching, this independent control of ion energy is essential to achieving vertical wall profiles over multilayer stacks, thus overcoming the natural proclivity of different layers to etch isotropically at different rates.

State of the art metallization layers are stacks of several metal films engineered to minimize electrical resistance and maximize thermal stability and reliability.[3] Aluminum is the conductor of choice because of its low resistivity and ability to adhere to SiO_2, but its low melting temperature, ability to dissolve Si, tendency to corrode, and low resistance to electromigration pose processing challenges.[4] The top layer is a thin (typically 100–500 Å) antireflective coating (ARC) of Ti–W or TiN to improve resist patterning on the highly reflective aluminum. The ARC layer may be oxidized by the preceding lithography, with much of the surface Ti converted to oxide. Beneath the ARC layer is the main conductor layer of aluminum. The Al often contains 0.5%–2.0% Cu to increase resistance to electromigration. Since Si is soluble in Al, 0.5%–2.0% Si is often also added to the Al film or, alternatively, a barrier layer can be added. The diffusion barrier is often TiN with an underlayer of Ti to promote adhesion of the barrier layer to the underlying oxide film. The processing challenge is to etch the metallization stack to produce vertical or slightly sloping sidewalls which facilitate conformal coating by the subsequent dielectric layer. The etching must be uniform across the wafer and must not sever lines, creating electrical opens, or leave metallic deposits across fields, creating shorts.

Gas mixtures of BCl_3 and Cl_2 are typically used to etch metallization stacks because fluorinated-aluminum species are not volatile except at extremely high temperatures. One role of BCl_3 is to scavenge free oxygen, enabling the removal of oxidized Al.[5] Copper is difficult to remove by dry etching, because none of its halides have an appreciable vapor pressure. Most metal etching processes rely on subsequent wet chemistry to remove residual copper.[3] As Cl and Cl_2 react spontaneously with Al, steps must be taken to prevent lateral etching of the Al film. The undercutting of the film can be prevented by the presence of a recombination species,[6,7] such as BCl_x ($x=1$ or 2), at or near the sidewalls to react with and neutralize Cl and Cl_2. Photoresist erosion in the plasma is also believed to play a role in anisotropic metal etching.[6]

Metallization stacks have been successfully etched in high-density plasma sources in both magnetized plasmas, such as helicon[8] and ECR reactors,[9] and in inductively coupled reactors such as the TCP.[10] We have used a helical resonator plasma source from Prototech Research[11] mounted on a two-chamber Lucas Labs cluster tool[12] to etch 0.45 μm metal patterns (0.35 μm VLSI technology) at low pressure.

[a]Current affiliation: Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.
[b]Current affiliation: Lam Research Corp., 4650 Cushing Parkway, Fremont, CA 94538; Electronic mail: helen.maynard@lamrc.com

Efforts were made to achieve a vertical profile using the simplest chemistry. Both pre- and postetch processing steps were established to prevent corrosion after metal etching and resist ashing.

II. EXPERIMENT

A. Etching platform and helical resonator plasma source

The etcher consists of a four-port central transfer chamber housing a robot arm for wafer handling, and a loadlock for loading cassettes of 100–200 mm diam wafers. The reactor consists of two sections: a main processing section where the plasma source and the wafer platen are mounted and a high conductance elbow leading to a 3000 ℓ/s magnetically levitated Osaka turbomolecular pump. The effective pumping speed is approximately 1200 ℓ/s. A throttling gate valve, installed directly above the entrance to the turbomolecular pump, controls the reactor pressure independently from the mass flow rate. A temperature-controlled capacitance manometer with sensitivity between 0.5 and 100 mTorr measures the neutral pressure during processing. Eight cartridge heaters imbedded in the anodized aluminum walls of each reactor allow the chamber walls to be baked to 100 °C. Normal operating procedure is to continuously maintain a "soft-bake" temperature of 65 °C, even during processing, to minimize adsorption of etch by-products on the reactor walls.

The wafer can be rf-biased through the wafer platen with a 600 W power supply. The wafer is mechanically clamped to the wafer chuck, and helium backside cooling removes heat from the wafer to the platen. Liquid nitrogen circulates through the wafer platen to remove the heat. The platen temperature can be controlled from −100 to 25 °C.

The helical resonator, Prototech model ESRF 600, consists of a quartz tube with an internal diameter of 150 mm that necks to 25 mm at the top where the feedgas is injected (Fig. 1). RF power at 27.12 MHz is inductively coupled into the plasma via a helical coil wound around the quartz tube. A Faraday shield is inserted between the coil and the quartz tube to minimize capacitive coupling that would sputter the quartz. Cooling water is circulated inside the coil as well as through the matching network and top metal plate. A perfluoropolyether circulates around the outside of the coil and Faraday shield in addition to cooling the quartz tube directly.

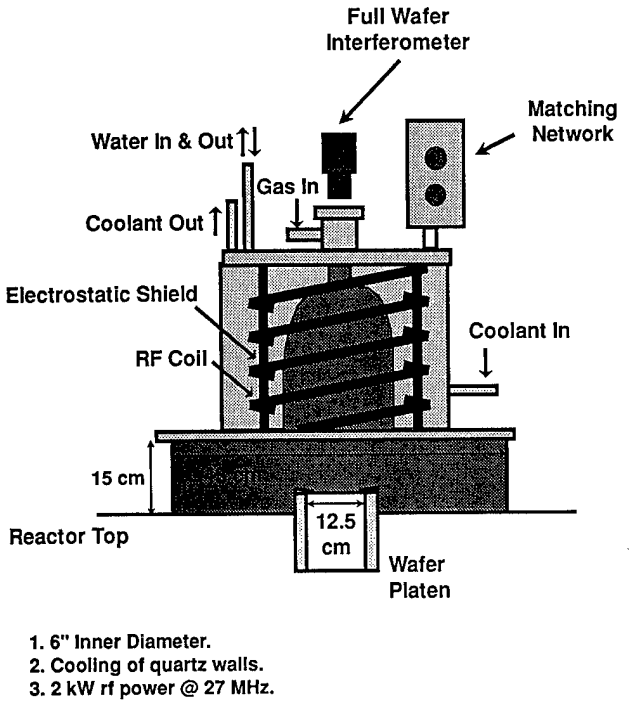


FIG. 1. Schematic of the Prototech Research 150 mm diam helical resonator plasma source, model ESHR-600 (drawn to scale).

B. Diagnostic techniques

1. Optical emission spectroscopy

A rotating-grating optical emission spectroscopy (OES) system from Monolite Instruments,[13] model MIO 6100 with a 200–850 nm photomultiplier tube detector, enables *in situ* measurement of plasma induced emission. The entire spectral range can be scanned at 12 Hz, and the resolution of the monochromator is observed to be 1–2 nm. Aluminum chloride (AlCl) emission at 261.4 nm and nitrogen (N_2) emission at 337.0 nm are simultaneously monitored during the etching at 2 Hz. The presence of the AlCl line corresponds to etching of Al while the N_2 line corresponds to etching of TiN in the multilayer metal stack described in detail below.

2. Full wafer imaging interferometer

A full wafer interferometer (FWI), model 1000-IS, from Low Entropy Systems,[14] monitors the etching in real time. Plasma emission is the light source for the interferometry while a charge-coupled device (CCD) camera acts as the detector by imaging the entire wafer, in effect turning each pixel into an independent interferometer. An interference filter (bandwidth ~10 nm) in the camera is centered at 481 nm. Ideally, one should choose the center wavelength of the interference filter to be at a wavelength of interest for the experiment, i.e., at an emission line from either a reactant or an expected product species. The filter used in this experiment was not initially intended for monitoring Al metallization etching, but it proved to be the more useful of the two available filters (481 and 750 nm). The highest resolution of the camera is on the order of 100 μm per pixel depending on the imaging optics. Fast Fourier transform analysis of the interference fringes is used to calculate the etching rates of transparent films. Etching rate uniformities can be calculated from various preselected regions on the wafer.

For opaque films, interference fringes will not be observed. However, changes in the reflected light intensity as a film clears will still provide valuable end-point information. As is discussed below, reflectometry is the mode of operation during metal etching, while interferometry is the mode of operation during resist stripping.

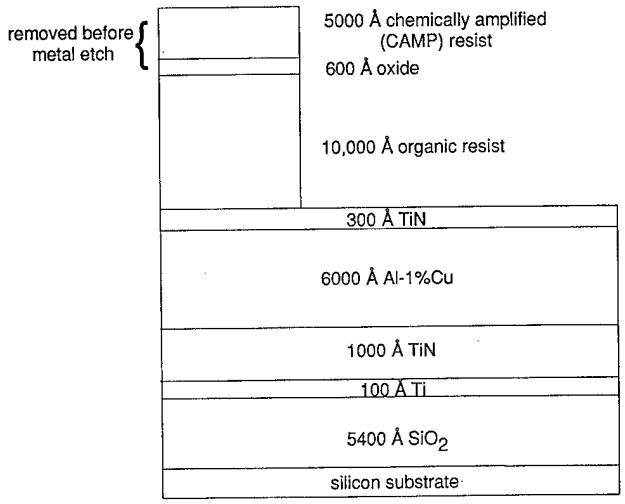


FIG. 2. Composition of metal stack prior to etching.

Finally, the full wafer interferometry system also provides optical emission capabilities, albeit only at the wavelength of the filter and integrated over the filter's bandwidth, by aligning the camera such that one of the points images a portion of the reactor that is not etching. In the Lucas Labs etcher, this is easily accomplished by choosing a point that images the wafer clamp.

C. Metal stack

Figure 2 is a cross-sectional schematic of the metal stack. The following layers are deposited on Si wafers: 5400 Å thermal SiO_2, 100 Å Ti, 1000 Å TiN, 6000 Å Al containing 1% Cu, 300 Å of TiN, and 1.0 μm of hard-baked Novolac resist. The resist is patterned using a trilevel process which deposits 600 Å of SiO_2 and 0.5 μm of chemically amplified (CAMP) resist on top of the 1.0 μm of hard-baked Novolac resist. The CAMP is exposed using deep ultraviolet (DUV) light and developed with pattern fidelity down to 0.25 μm. The CAMP patterns the oxide, which in turn patterns the hard-baked resist. The metal etching process etches through the metal stack minimizing etching into the underlying SiO_2 layer. The open area of the pattern is approximately 50% of the wafer.

D. Pre- and postetch processing

Table I summarizes the processing steps. Before etching, the samples are placed in an ethylene glycol buffered oxide etching (EG-BOE) bath, 7:1 aqueous dilution, for 3 min to remove the CAMP and thin oxide layers and to expose the layer of hard-baked resist to the plasma.

After etching through the metal stack, an O_2 plasma at 50 W rf-bias power, 50 sccm O_2, 5.0 mTorr, $T=23$ °C and 1500 W rf source power is used to remove the remaining photoresist. Removing the photoresist *in situ* prevents corrosion, which occurs when Cl_2-laden resist is exposed to air. As a final corrosion prevention step, wafers are placed in an ACT-CMI[15] bath for 10 min immediately after removal from the etcher loadlock. This bath removes any remaining trace resist and passivates the surface. In all wet chemical treatments after etching, the wafers were run through a seven-cycle rinse using de-ionized H_2O bubbled with CO_2 and then spin dried. The processing conditions and resulting etching rates are listed in Table II.

III. RESULTS AND DISCUSSION

A. Etching with pure Cl_2

1. Best profiles

The best profiles with a pure Cl_2 plasma are obtained with the following main etching conditions: 100 sccm Cl_2, 2.0 mTorr, wafer platen temperature $T=0$ °C, 100 W rf-bias power, and 1500 W source power. The total etching time is 51 s, including 5 s of overetch. The amount of overetch was intentionally kept to a minimum to study the effect of the main etching step on profile control. Figure 3 is a scanning electron microscope (SEM) micrograph of two 0.45 μm lines separated by 0.85 μm. The Al layer is undercut by approximately 150 Å on each side directly beneath the top TiN layer. All four metal layers (upper TiN, Al, lower TiN, and Ti) have vertical profiles. While the isolated regions are completely cleared, regions between nested lines need a longer overetch to completely clear the lower TiN and Ti layers. In the open areas, less than 100 Å of the underlying oxide is lost, indicating that a longer overetch will not be a problem, as a loss of up to approximately 500 Å of oxide is acceptable for 0.25 μm technology processes.

The corresponding OES traces show clear Al and TiN end points [Fig. 4(a)]. The breakthrough and end-point times are marked by a sharp rise and drop, respectively, in the emission intensity. These sharp changes indicate uniform film clearing and thus high etching rate uniformity. Traces obtained with the FWI for one point on and one point off the wafer are shown in Fig. 4(b). For the point on the wafer, only slight changes are observed at the Al layer end point and at the second TiN layer end point. In contrast, the point off the wafer shows an increase in intensity at TiN breakthrough and

TABLE I. Summary of processing steps.

1. 3 min 7:1 EG-BOE bath to remove oxide layer of trilevel resist stack.
2. Etch metal stack in helical resonator.
3. Remove resist in helical resonator with O_2 plasma using the following conditions: 50 sccm O_2, 50 W rf-bias power, 5.0 mTorr, $T=23$ C, 1500 W source power.
4. 10 min ACT-CMI bath to remove trace resist and passivate wafer surface.
5. Seven-cycle water rinse and spin dry.

TABLE II. Experimental conditions for all wafers and resulting etching rates and selectivities. The rf source power was 1500 W for all etching experiments.

	Etch conditions				Resist left after full etch (Å)	Resist etched during Al etch (Å)	Etch rates (Å/min)			Selectivity		
Run	Feed (sccm)	T (wafer platen) (°C)	Pressure (mTorr)	rf bias (W)			TiN 1	Al	TiN 2	Al:TiN 2	Al:resist	Stack:resist
21	100 Cl_2	23	2	50	3380	1500	3200	8 500	2400	3.5	2.7	1.2
24	100 Cl_2	21	2	75	2250	2250	3500	9 100	3300	2.7	2.3	1.0
26	100 Cl_2	23	2	100	4400	···	6700	10 600	3300	3.2	···	1.4
27	80 Cl_2, 20 BCl_3	23	2	100	1900	3000	5600	7 000	2800	2.5	1.7	0.98
28	45 Cl_2, 45 N_2	23	2	100	0	6000	1300	2 300	1300	1.7	0.82	0.79
29	75 Cl_2, 25 N_2	23	2	100	750	4500	4800	5 100	2500	2.0	1.2	0.85
30	100 Cl_2	−25	2	100	2250	2250	6000	10 100	4100	2.4	2.1	1.0
31	100 Cl_2	0	2	100	···	···	5400	11 100	4000	2.8	···	···
32	100 Cl_2	0	1	100	1500	3700	4800	8 100	3600	2.3	1.3	0.93
34	80 Cl_2, 20 BCl_3	0	2	100	1500	3350	7600	7 500	3000	2.6	1.7	0.93
35	80 Cl_2, 20 BCl_3	0	1	100	2800	3000	4400	6 700	2000	3.4	1.6	1.1
36	75 Cl_2, 25 HBr	0	2	100	2250	3000	4500	7 300	2800	2.6	1.7	1.0
37	100 Cl_2	0	2	100	3100	1800	5100	11 000	3800	2.9	2.8	1.1
38	80 Cl_2, 20 BCl_3	0	2	100	2250	2600	4800	7 800	3100	2.5	1.8	1.0
40	90 Cl_2, 10 BCl_3	0	2	100	3500	1500	5700	11 400	3900	2.9	3.4	1.2

a corresponding decrease at TiN end point, for both the upper 300 Å and lower 1000 Å TiN layers. This behavior at 481 nm is similar to the N_2 emission at 337 nm and may be due to a relatively weak emission of N_2 at 481.5 nm.[16] However, since the FWI integrates the optical emission intensity over a fairly wide range (10 nm), it is highly speculative to assign a source for the emission. Additionally, only by comparing the FWI signals to OES traces can we accurately identify the different breakthroughs and end points.

2. Effect of rf-bias power

Experiments run using different rf-bias powers reveal that high biases are necessary in order to obtain vertical profiles. Lower biases result in significant undercutting of the Al, and the Al layer can separate from the lower TiN layer, resulting in lifting of the metal lines. The etching rates of both Al and TiN increase with increasing rf-bias power.

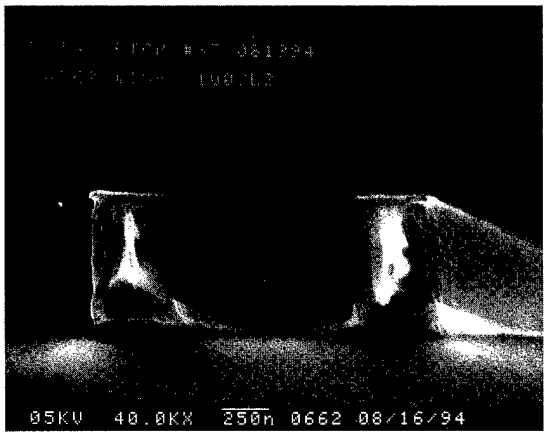


FIG. 3. SEM micrograph of metal lines etched using a pure Cl_2 plasma. Etching conditions: 100 sccm Cl_2, 2.0 mTorr, $T=0$ °C, 100 W rf-bias power, and 1500 W source power. Total etching time is 51 s. 0.45 μm wide lines separated by 0.85 μm space.

3. Effect of wafer platen temperature

Lowering the wafer platen temperature improves the feature profile, and Fig. 3 is a cross-sectional SEM of the fea-

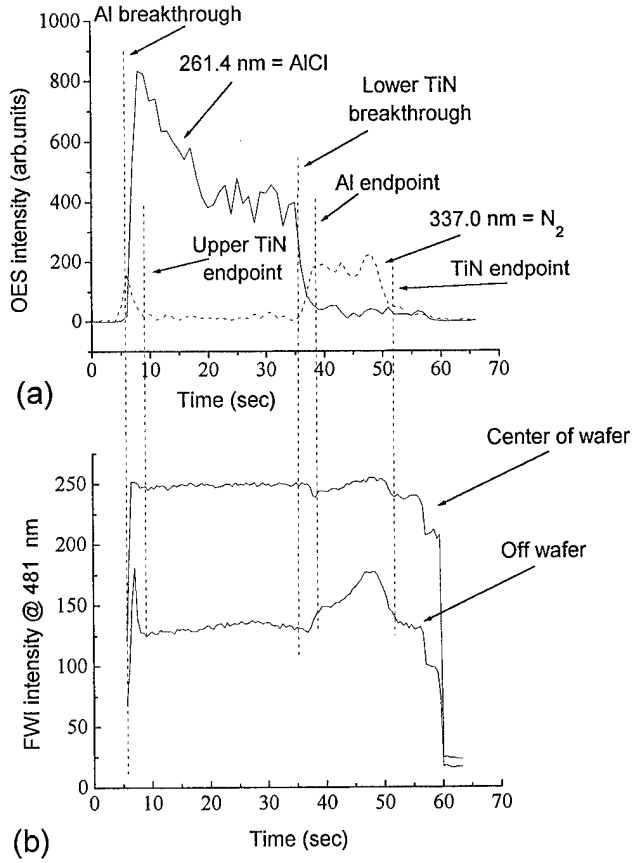


FIG. 4. Etching diagnostics for wafer etched with 100 sccm Cl_2. (a) OES signals at 261.4 nm (AlCl) and 337.0 nm (N_2); (b) FWI signals at 481 nm at points on and off the wafer.

FIG. 5. SEM micrograph of 0.45 μm wide lines etched with 100 sccm Cl_2 at the following conditions: 2.0 mTorr, 100 W rf-bias power, 1500 W source power, and $T=-25$ °C.

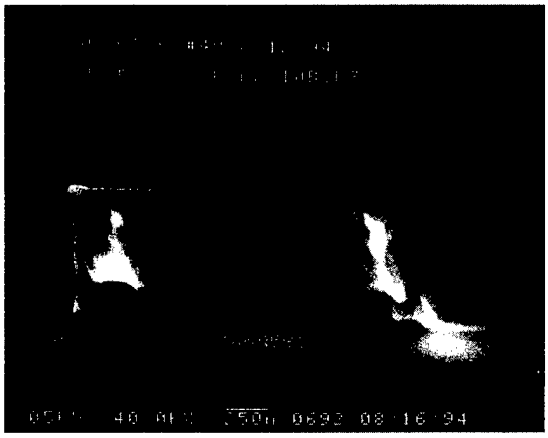


FIG. 7. SEM micrograph of 0.45 μm metal lines etched with a BCl_3-containing plasma. Etching conditions: 90 sccm Cl_2, 10 sccm BCl_3, 2.0 mTorr, $T=0$ °C, 100 W rf-bias power, and 1500 W source power. Total etching time is 48 s. Line spacing is 0.85 μm.

tures etched at 0 °C. However, it was also found that lowering the temperature to -25 °C (Fig. 5) leaves metal etch residue on the wafer surface.

Figure 6 shows the dependence of etching rate and resist selectivity on wafer platen temperature. The Al etching rate is fairly constant with platen temperature, while the TiN etching rate decreases slightly as the platen heats. At low temperatures, the increase in TiN etching rate may contribute to less Al undercutting since the overall etching time is less, allowing less time for the lateral etching of the Al layer. We also observe a decrease in Al:resist selectivity at lower temperatures, leading to a significantly increased resist consumption rate. However, the increase in the concentration of resist etch products in the plasma chemistry may contribute to the more vertical profiles obtained at lower temperatures.

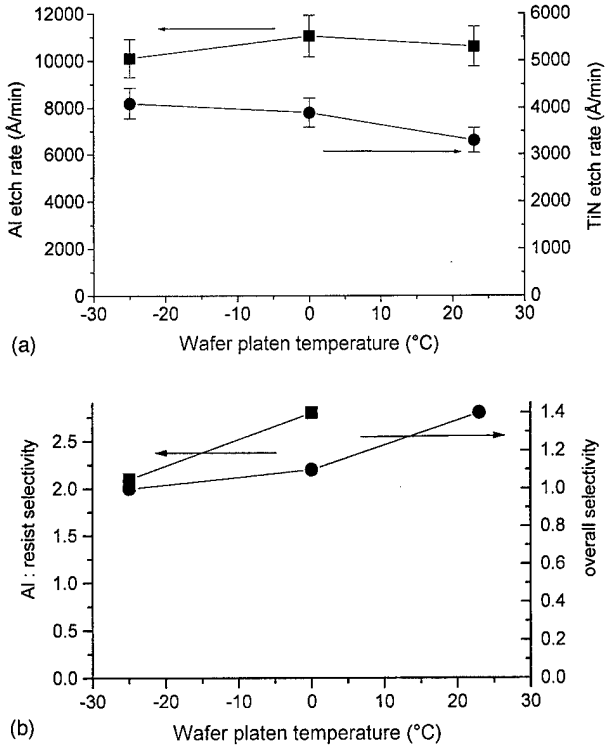


FIG. 6. Dependence of (a) Al and TiN etching rates and (b) resist selectivity on wafer platen temperature. Etching conditions: 100 sccm Cl_2, 2.0 mTorr, 100 W rf-bias power, and 1500 W source power. Error bars indicate estimates of the measurement errors.

4. Effect of plasma pressure

Decreasing the neutral pressure from 2.0 to 1.0 mTorr results in poor feature profile control. TiN and Al etching rates decrease approximately 25% with the drop in pressure, and the Al:resist selectivity drops by half.

B. Etching with BCl_3

1. Best feature profile

When using a mixture of Cl_2 and BCl_3, the best feature profiles were obtained at the following main etching conditions: 80–90 sccm Cl_2, 10–20 sccm BCl_3, 2.0 mTorr, $T=0$ °C, 100 W rf-bias power, and 1500 W source power. Aluminum undercut is almost nonexistent (Fig. 7) for the two 0.45 μm lines separated by 0.85 μm, but additional overetching is still needed to remove residual TiN and Ti between the lines. Overall, these feature profiles are as good as or better than those obtained using pure Cl_2.

Figures 8(a) and 8(b) show the corresponding OES and FWI traces, respectively, for a wafer etched with 90 sccm Cl_2/10 sccm BCl_3. Again, Al and TiN breakthroughs and end points are clear with both diagnostics. However, notice that in Fig. 4(b) the *off wafer* point has a clearer end-point signal than does the *center of wafer* point, while in Fig. 8(b) the opposite is true. Apparently, the change in gas chemistry has changed the nature of the end-point signals.

2. Effect of wafer platen temperature

When etching with a mixture of 80 sccm Cl_2 and 20 sccm BCl_3, decreasing the wafer platen temperature improves fea-

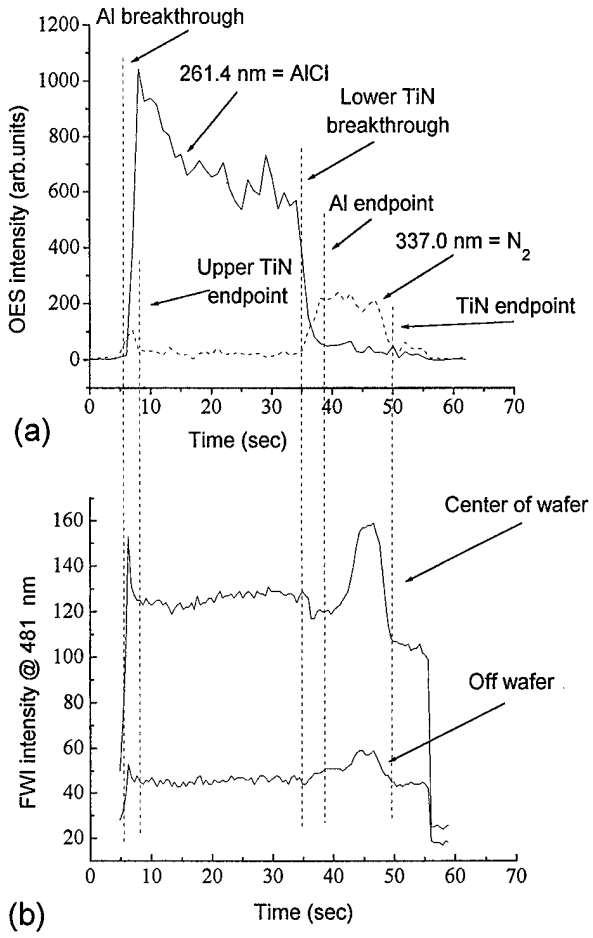


FIG. 8. Etching diagnostics for wafer etched with 90 sccm Cl_2 and 10 sccm BCl_3. (a) OES signals at=261.4 and 337.0 nm. (b) FWI signals at 481 nm at points on and off the wafer.

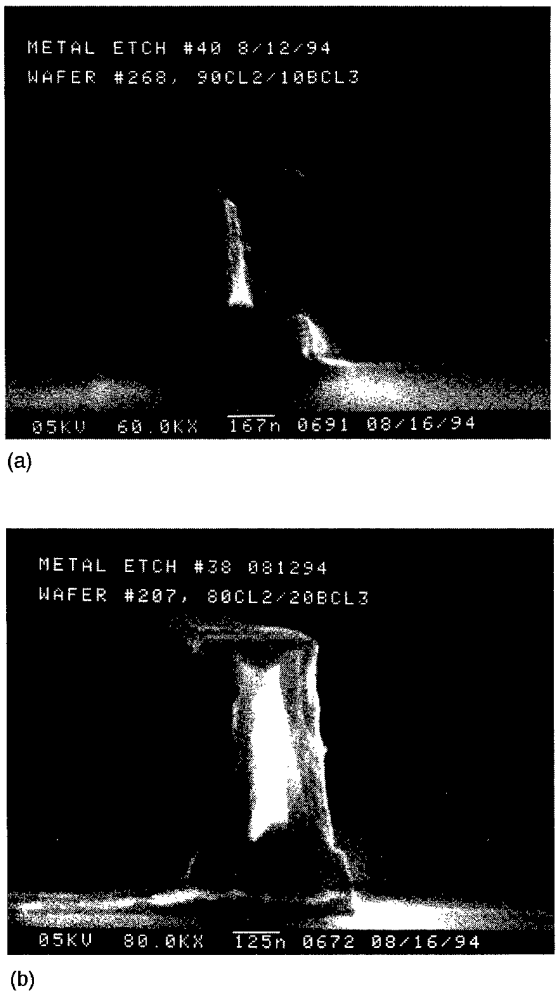


FIG. 9. SEM micrographs of metal lines etched with BCl_3-containing plasmas. Etching conditions: 2.0 mTorr, $T=0$ °C, 100 W rf-bias power, 1500 W source power, and (a) 90 sccm Cl_2, 10 sccm BCl_3, and (b) 80 sccm Cl_2, 20 sccm BCl_3.

ture profiles by reducing Al undercutting. Etching rates decrease marginally for both Al and TiN with increasing temperature. The Al:resist selectivity decreases slightly, while the overall resist selectivity increases slightly with increasing wafer platen temperature. These trends correspond to those observed when using a pure Cl_2 plasma (Fig. 6).

3. Effect of plasma pressure

For an 80 sccm Cl_2/20 sccm BCl_3 plasma, as with pure Cl_2, decreasing the neutral pressure from 2.0 to 1.0 mTorr noticeably degraded feature profiles. Both Al and TiN etching rates decrease with decreasing plasma pressure.

4. Effect of BCl_3 concentration in plasma

Feature profiles improve with the addition of BCl_3, as the cross-sectional SEMs show for wafers etched with mixtures of 10 sccm BCl_3/90 sccm Cl_2 and 20 sccm BCl_3/80 sccm Cl_2 [Figs. 9(a) and 9(b), respectively]. The Al and TiN etching rates are almost unchanged between the 0% and 10% BCl_3 plasmas, while both etching rates drop significantly as the BCl_3 concentration is increased from 10% to 20%, [Fig. 10(a)]. The same changes occur for resist selectivity [Fig. 10(b)].

Additions of BCl_3 are believed to suppress lateral etching, by the presence of subchlorinated BCl_x species ($x=1$ or 2) at or near the sidewalls tying up free Cl radicals to form BCl_3.[7,17] As more BCl_3 is added to the plasma, the Al etching rate is suppressed, and since BCl_3 etches Al much more slowly than Cl_2,[18] this leads to poor Al:resist selectivity.

Also note that RIE etching of Al alloys typically uses BCl_3-rich etchant mixtures,[19] whereas Cl_2-rich mixtures seem to be appropriate for the helical resonator. This difference in etchant concentrations may be due to the higher species dissociation found in high-density plasma sources, therefore reducing the amount of BCl_3 needed to provide a sufficient amount of BCl_x radicals. This is also true for the TCP at low pressures.[20]

C. Etching with other additives

The addition of other gases to a Cl_2 plasma has been reported to have significant effects on the feature profiles.[20]

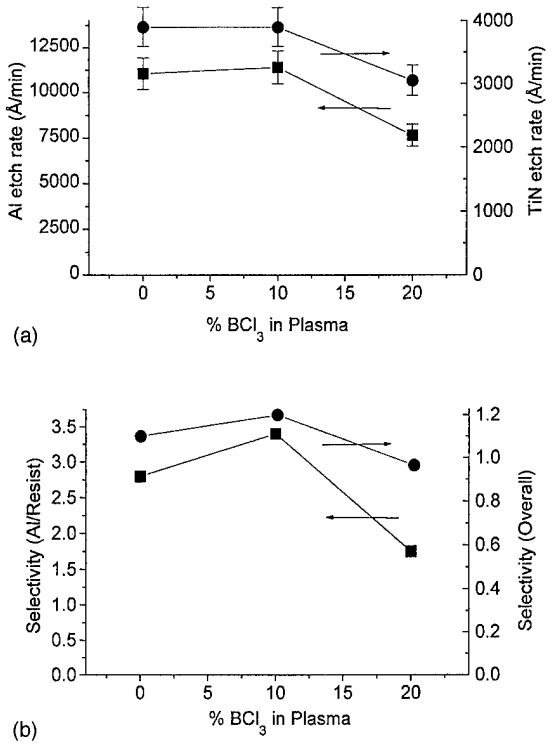


FIG. 10. Dependence of (a) Al and TiN etching rates and (b) resist selectivity on % BCl_3 in a BCl_3/Cl_2 plasma. Etching conditions: 2.0 mTorr, $T=0$ °C, 100 W rf-bias power, and 1500 W source power. Error bars indicate estimates of measurement error.

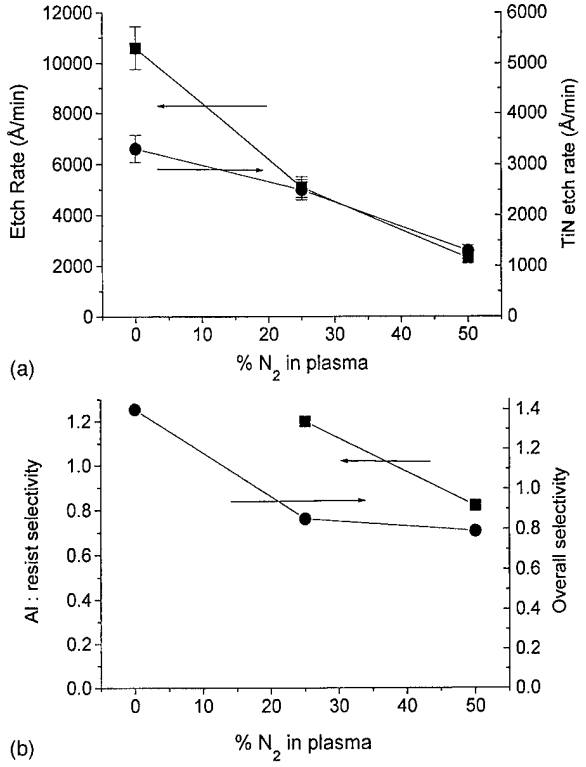


FIG. 11. Dependence of (a) Al and TiN etch rates and (b) resist selectivity on % N_2 in plasma. Etching conditions: 2.0 mTorr, $T_{\text{wafer platen}}=23$ °C, 100 W rf-bias power, and 1500 W source power.

As discussed above, BCl_3 is a helpful additive. Experiments were also run using different proportions of N_2 and HBr, with varying results.

1. Etching with N_2

Etching wafers with N_2-containing plasmas produces poor profile control and decreases both the etching rates and resist selectivities. Etching times increase dramatically from 51 s with 0% N_2 to 210 s with 50% N_2. Most of the resist etches away during the main etch. Most features uplift or are etched away, and all remaining features are extremely tapered.

Figure 11(a) shows the dependence of Al and TiN etching rates on the percentage of N_2 in the gas flow. Both etching rates decrease dramatically with increasing $\%N_2$. These decreasing etching rates account for the tapered feature profiles since, as the etching time increases, more isotropic etching can occur, and thus, the Al undercuts and tapers.

Resist selectivity decreases with increasing $\%N_2$ as well [Fig. 11(b)]. In the case of 50% N_2, the resist selectivities are extremely low, accounting for the lack of resist remaining on the surface after the main etch. With these N_2 chemistries, the plasma is actually more proficient at etching resist than it is at etching the metal stack.

2. Etching with HBr

A single trial was run using 25% HBr in the Cl_2 plasma mix. The resulting feature profiles were relatively vertical, but significant undercut of the Al layer at the interface of the Al and the upper TiN occurs (Fig. 12). The total etching time increases to 62 s, as opposed to 51 s when using pure Cl_2. Both Al and TiN etching rates as well as resist selectivities decrease with the addition of HBr to the plasma mixture.

IV. SUMMARY AND CONCLUSIONS

The best feature profiles are obtained at the following conditions: 80–90 sccm Cl_2, 10–20 sccm BCl_3, 2.0 mTorr, $T_{\text{wafer platen}}=0$ °C, 100 W rf-bias power, and 1500 W source

FIG. 12. SEM micrograph of wafer etched with HBr-containing plasma. Etching conditions: 75 sccm Cl_2, 25 sccm HBr, 2.0 mTorr, $T_{\text{wafer platen}}=0$ °C, 100 W rf-bias power, and 1500 W source power.

power. Wafers etched under these conditions have features with vertical profiles, with almost no Al undercut at the upper TiN layer. These features are slightly better than those obtained when using 100 sccm Cl_2 and the same plasma chamber conditions. The anisotropic profile achievable with a pure Cl_2 plasma in the high-density helical resonator source is better than that obtainable with pure Cl_2 in a conventional RIE plasma, but the remaining Al notch at the upper TiN interface would deem it unacceptable for a manufacturing process. A small addition of BCl_3 eliminates the Al notch, and the mixture is therefore the preferred chemistry. Independent control of the rf-bias power, in particular, also proved to be important in controlling the amount of resist erosion and to obtaining more anisotropic profiles.

Future work is intended to focus on a multistage process where different amounts of BCl_3 are added to the plasma mixture during different parts of the etch. It is hoped that through this type of processing, all Al undercut can be eliminated, and the residual TiN and Ti between pairs of lines can be cleared.

ACKNOWLEDGMENTS

The authors thank A. Kornblit, F. P. Klemens, V. M. Donnelly, and D. E. Ibbotson for many insightful discussions, K. P. Cheung and K. Collins for assistance with diagnostics, and K. V. Guinn for assistance with the helical resonator.

[1]M. A. Lieberman and R. A. Gottscho, in *Physics of Thin Films: Advances in Research and Development*, edited by J. Vossen (Academic, San Diego, CA, 1993), p. 17.
[2]I. Tepermeister, N. Blayo, F. P. Klemens, D. E. Ibbotson, R. A. Gottscho, J. T. C. Lee, and H. H. Sawin, J. Vac. Sci. Technol. B **12**, 2310 (1994).
[3]Paul E. Riley, J. Electrochem. Soc. **140**, 1518 (1993).
[4]D. B. Fraser, in *VLSI Technology*, edited by S. M. Sze (McGraw-Hill, New York, 1983), Chap. 9.
[5]S. C. McNevin, J. Vac. Sci. Technol. B **8**, 1212 (1990).
[6]W. R. Harshbarger, H. J. Levinstein, C. J. Mogab, and R. A. Porter, U. S. Patent No. 4,208,241 (1980).
[7]C. J. Mogab and H. J. Levinstein, J. Vac. Sci. Technol. **17**, 721 (1980).
[8]N. Jiwari, H. Iwasawa, A. Narai, H. Sakaue, H. Shindo, T. Shoji, and Y. Horiike, Jpn. J. Appl. Phys. **32**, 3019 (1993).
[9]S. Samukawa, T. Toyosato, and E. Wani, J. Vac. Sci. Technol. B **9**, 1471 (1991).
[10]Y. Ra, S. G. Bradley, and C.-H. Chen, J. Vac. Sci. Technol. A **12**, 1328 (1994).
[11]Prototech Research, Inc. of Tempe, Arizona.
[12]Lucas Labs, Inc. of Sunnyvale, California.
[13]Monolite Instruments of Leominster, Massachusetts.
[14]Low Entropy Systems of Somerville, Massachusetts.
[15]ACT-CMI is a proprietary organic etchant produced by Advanced Chemical Technologies of Allentown, PA.
[16]R. W. B. Pearse and A. G. Gaydon, *The Identification of Molecular Spectra* (Chapman and Hall, New York, 1963).
[17]R. W. Light, J. Electrochem. Soc. **130**, 2225 (1983).
[18]D. A. Danner, M. Dalvie, and D. W. Hess, J. Electrochem. Soc. **134**, 669 (1987).
[19]D. H. G. Choe, C. Knapp, and A. Jacob, Solid State Technol. **38**, 165 (1995).
[20]A. Kornblit (personal communication).

Microstructural stability of ohmic contacts to $In_xGa_{1-x}N$

A. Durbha, S. J. Pearton,[a] C. R. Abernathy, J. W. Lee, and P. H. Holloway
University of Florida, Gainesville, Florida 32611

F. Ren
AT&T Bell Laboratories, Murray Hill, New Jersey 07974

(Received 7 November 1995; accepted 1 May 1996)

The microstructural properties and interdiffusion reactions of some common metallization schemes (Au/Ge/Ni, Ti/Pt/Au, WSi_x, and AuBe) on GaN and $In_{0.5}Ga_{0.5}N$ have been examined using scanning electron microscopy and Auger electron spectroscopy. The objective of this study was to investigate thermally stable and uniform contacts to both materials. The WSi_x contacts were found to possess excellent thermal stability and retained good structural properties at annealing temperatures as high as 800 °C on GaN. These contacts on $In_{0.5}Ga_{0.5}N$ had a minimum specific contact resistivity of 1.48×10^{-5} Ω cm^2 and an excellent surface morphology following annealing at 700 °C. The increase in contact resistance observed at higher temperatures was attributed to intermixing of metal and semiconductor. By contrast the Ti/Pt/Au and Au/Ge/Ni contacts showed lower stabilities, but also showed little Au penetration into the underlying GaN. AuBe contacts had the poorest thermal stability of all four schemes investigated, with substantial reaction with GaN occurring even at 400 °C. The WSi_x contact appears to be an excellent choice for application in high temperature GaN electronics. © *1996 American Vacuum Society.*

I. INTRODUCTION

Owing to their large band gaps, mechanical and thermal stability, III–V nitrides are very attractive for high temperature electronics and blue and UV optoelectronic device applications. Improved material properties have recently led to a variety of devices being demonstrated.[1] Blue light-emitting devices (LEDs)[2,3] and metal–semiconductor field-effect transistors (MESFETs) have been successfully fabricated.[4] However, forming low resistance, thermally stable, and uniform ohmic contacts to a wide band gap semiconductor, such as GaN with a band gap about 3.4 eV, constitutes a major obstacle to the furtherment of nitride based devices. These devices provide high efficiency and acceptable reliability only if their contacts are stable and have low resistance ohmic characteristics.

In an earlier attempt to achieve ohmic contacts on GaN epilayers, Foresi *et al.*[5] used Al and Au contacts with 575 °C anneal cycle. However, the specific contact resistivity of these contacts was relatively poor (10^{-3} Ω cm^2). Khan *et al.*[3] used Ti/Au to contact *n*-type GaN and measured a contact resistance of 7.8×10^{-4} Ω cm^2 after annealing at 250 °C for 30 s. Nakamura *et al.*[2] have used Au (and later Au/Ni and Al) as *p*- and *n*-type contacts respectively in their LED structures. While the contact resistances were not reported in their LED structures, an operating voltage of 4 V and 20 mA forward bias is clear evidence that reasonable contact resistances were obtained. Recently Lin *et al.*[6] have obtained extremely good ohmic contacts on *n*-type GaN layers grown on sapphire substrates. Using Ti/Al metallization scheme, they were able to obtain specific contact resistivities as low as 8×10^{-6} Ω cm^2 after annealing at 900 °C for 30 s. Lin *et al.*[7] also demonstrated a novel ohmic contact scheme to GaN using an InN/GaN short-period superlattice (SPS) and an InN cap layer. Ohmic contact resistivities as low as 6×10^{-5} Ω cm^2 were achieved even without any postannealing.

An ohmic contact study of four standard metallization schemes in III–V technology: Ti/Pt/Au, Au/Ge/Ni, WSi_x, and AuBe/Au, has been undertaken in this work. Also, we investigated a scheme which had an InGaN layer on top of GaN. The InGaN has a lower bandgap than GaN and should produce lower contact resistance. Ultimately, we would like to use such a scheme for improved contact properties on devices, much as InAs is used on GaAs. WSi_x was used for metallization to the InGaN layer. Electrical characterization of the contacts was done using standard transmission line measurements (TLMs) and materials characterization included scanning electron microscope (SEM) and Auger electron spectroscopy (AES).

II. EXPERIMENT

The nominally undoped GaN and $In_{0.5}Ga_{0.5}N$ was grown on GaAs at 800 °C using $(CH_3)_3Ga$ and an electron cyclotron resonance-plasma generated N_2 flux in a metalorganic molecular beam epitaxy system. The samples predominantly consisted of the cubic phase, with typical x-ray full width at half-maximum (FWHM) of 300–500 arcsec. The *n*-type doping level in the ternary was $\sim10^{19}$ cm^{-3}, while the GaN was typically $\leq10^{17}$ cm^{-3} due to the presence of native shallow donors. At this point, it has been impossible to grow *p*-type InGaN at InN mole fractions above 0.07.

Four different metallization schemes common in III–V technology were investigated, i.e., Ti/Pt/Au, Au/Ge/Ni, AuBe, and WSi_x. The purpose of this was to establish their relative thermal stabilities on GaN. Once this was established, we wanted to use the most stable on InGaN, which as

[a]Electronic mail: spear@mse.ufl.edu

mentioned earlier is our intended final structure for device applications. The contact metals were deposited using two techniques, namely, electron beam evaporation and sputtering. The deposition sequences and the layer structures were as follows: (1) Ti–250 Å/Pt–500 Å/Au–1000 Å; (2) Ge–50 Å/Ni–50 Å/Au––Ge–268 Å––132 Å/Au–1100 Å; (3) AuBe–200 Å/Au–1000 Å, and (4) WSi_x–1000 Å. Based on initial studies of the stability of these systems, WSi_x contacts having thickness of 1000 Å were deposited on an $In_{0.5}Ga_{0.5}N$ sample. The Ti/Pt/Au contacts have proven to yield excellent contact resistance (about 10^{-7} Ω cm^2) on extremely heavily doped III–V semiconductors. The AuBe based metallization is a standard alloyed ohmic contact and produces excellent contact resistance on heavily doped *p*-type materials ($\leqslant 10^{-6}$ Ω cm^2). While we did not investigate *p*-type material in this work, it is still worth establishing the stability of this system. The Au/Ge/Ni contact is also a standard alloyed contact and produces very low contact resistances to heavily doped *n*-type III–V materials (about 10^{-7} Ω cm^2). Typically, alloyed contacts exhibit problems such as poor thermal stability, spiking, and poor reproducibility and reliability. The above three metallizations were performed using electron beam evaporation in a Temescal system. Prior to loading in an UHV system, the wafers were rinsed in a solution of H_2O:NH_4OH=20:1 for 1 min to remove the native oxide from the GaN samples in order to ensure uniform contact properties upon subsequent deposition and alloying. After a pump down to a pressure of about 2×10^{-7} Torr and *in situ* Ar^+ ion beam sputtering for oxide removal, the metallization schemes were evaporated with an electron gun.

The WSi_x ($x\sim0.45$) contact was deposited using Ar^+ assisted sputtering. Prior to loading the samples in the deposition chamber, a 1 min rinse, in a H_2O:NH_4OH (20:1) solution was done after which the contacts were deposited by sputtering with an Ar discharge at a bias voltage of 90–100 V. The refractory metal contacts have traditionally shown excellent thermal stability and are used as a gate in the processing of self-aligned MESFET devices. The electrical measurements were performed on WSi_x contacts deposited over InGaN layers on GaN. The contact metal was deposited onto square openings (100×100 μm^2) linearly spaced (with intervals of 2–16 μm) in a photoresist layer. Subsequently, InGaN mesas were etched to give the required one-dimensional current flow. In these samples, the GaN was semi-insulating and only patterning of the InGaN was necessary. In real device structures, one would employ conducting GaN (such as for channel layers in transistors) and then TLM measurements would require patterning of the GaN as well. After the metallizations were performed, each sample was split into pieces and thermally annealed at temperatures from 400 to 900 °C for 30 s in a N_2 ambient using rapid thermal annealing technique (AG Heatpulse 410 System.)

III. RESULTS

A. Scanning electron microscopy

The Au/Ge/Ni samples were annealed at temperatures of 400, 500, 600, and 700 °C. The as-deposited sample and that heat treated at 500 °C, as seen from their micrographs in Fig. 1, are rather smooth due to the fact that Ni overplate is used (the micrograph of the sample annealed at 400 °C is not included because its surface morphology is not very different from the as-deposited sample and the sample annealed at 500 °C). The Au/Ge/Ni sample heat treated at 600 °C showed, signs of degrading. The surface was no longer smooth and as the sample annealing temperature was increased further to 700 °C the SEM micrograph at a magnification of 10 000 reveals many shallow small pits. Comparing the micrographs of all the samples shows that 500 °C does not really effect the surface morphology whereas heat treatments at 600 and 700 °C do degrade the sample morphology.

The Ti/Pt/Au samples were heat treated at temperatures of 500, 700, 800, and 900 °C for 30 s. The morphology of the as-deposited sample (shown in Fig. 2) is quite smooth. The sample that was annealed at 500 °C showed signs of degradation and at 700 °C the surface became more degraded. The formation of small islands was more apparent and their size and density increased at 800 °C. The maximum temperature for annealing these contacts would be around 450–500 °C, which is approximately about the temperature used for obtaining low temperature Ti/Pt/Au contacts on GaAs samples.[8]

The WSi_x contacts on GaN samples were heat treated at temperatures of 400, 700, 800, and 900 °C for 30 s. The SEM micrographs of the as-deposited sample and the sample annealed at 900 °C are shown in Fig. 3. The micrographs of the other samples annealed in the temperature range of 400–800 °C were not included because the surface morphology was similar to that of the as-deposited sample. The WSi_x contacts were also deposited on an $In_{0.5}Ga_{0.5}N$ film grown on GaAs which was further split into pieces and annealed in the same temperature range as the WSi_x/GaN system. The SEM studies performed on these samples revealed that the as-deposited sample exhibited a very smooth surface and that there was no change in the surface morphology of samples annealed at temperatures of 400 and 700 °C. The surface morphology of the samples annealed at 900 °C showed only a small amount of surface roughness. The maximum annealing temperature to obtain good surface morphology WSi_x contacts on InGaN samples would therefore be in the range 700–800 °C.

The AuBe samples were annealed at temperatures of 400, 500, 600, and 700 °C for 30 s. The as-deposited AuBe sample was characterized by a smooth surface but the sample annealed at 400 °C already showed a degree of roughness. It was clearly evident from the sample annealed at 700 °C that the surface was badly degraded. Thus, for the AuBe contacts on GaN the maximum annealing temperature should be less than 400 °C for it to retain its surface morphology.

B. Auger electron spectroscopy

The Auger survey scans of the as-deposited contacts indicated the surface layer of the metal had the usual contami-

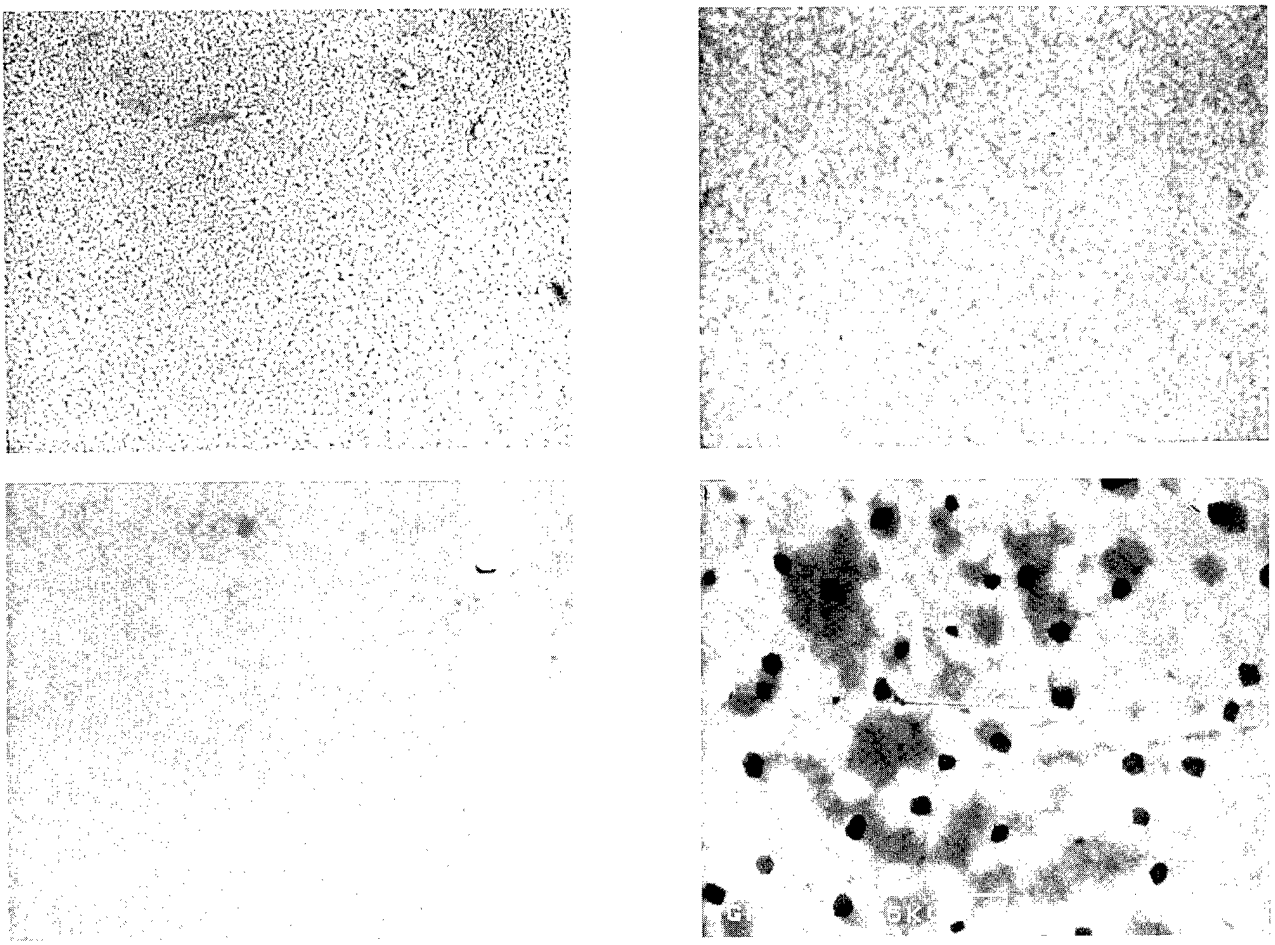


FIG. 1. Surface morphology of Au/Ge/Ni/GaN contacts (magnification ×10 000); (top left) as-deposited, (top right) annealed at 500 °C, (bottom left) annealed at 600 °C, and (bottom right) annealed at 700 °C.

nation from atmospheric O and C. These surface contaminants were quickly sputtered away during the Auger depth profiling.

The Auger depth profile of as-deposited Au/Ge/Ni contacts on GaN indicated a slight interdiffusion between Au and Ni and extensive interdiffusion between Ni and Ge. There was also moderate interdiffusion between Au and Ge. There was also some indiffusion between Ni, Ge, and the GaN, with outdiffusion of Ga into the Ni and Ge layers. The penetration of Au into the GaN was insignificant. The suggestion of interdiffusion of Ni and Ge agrees with results of Au/Ge/Ni contacts on GaAs samples.[8]

The Auger depth profile of the as-deposited Ti/Pt/Au sample seen in Fig. 4(a) did not show much interdiffusion between the metals compared to the Au/Ge/Ni as-deposited sample. There was a slight interdiffusion between Au and Pt and moderate interaction between Ti and Pt, with extensive interaction between Ti and Ga. There was essentially no interaction between Au and GaN. In comparison to the as-deposited sample, the sample heat treated at 600 °C showed extensive interdiffusion between Ti and Pt [Fig. 4(b)]. There appeared to be more Ga outdiffusion into the metal layers in the heat treated sample compared to the as deposited sample. TiN is a semimetal,[9] but we do not know whether TiGaN forms in these samples—no evidence was found for such a phase in double crystal x-ray diffraction measurements.

The depth profiles of the as-deposited and annealed samples of WSi_x shown in Figs. 5(a) and 5(b) indicated that the 900 °C annealed sample showed more intermixing of the metal and GaN at the interface compared to the as-deposited sample. A similar interdiffusion was reported in the WSi_x/InGaAs system,[10] although it can be seen that the interdiffusion in the WSi_x/GaN structure is much less compared to the metallization schemes described above.

Since the AuBe contacts exhibited poor thermal stability from the SEM photographs we did not perform AES studies on them.

C. Electrical measurements

The inference drawn from the analysis of the SEM and AES results of the different metallization schemes was that the WSi_x contact was more thermally stable compared to the other schemes and hence this metallization scheme was em-

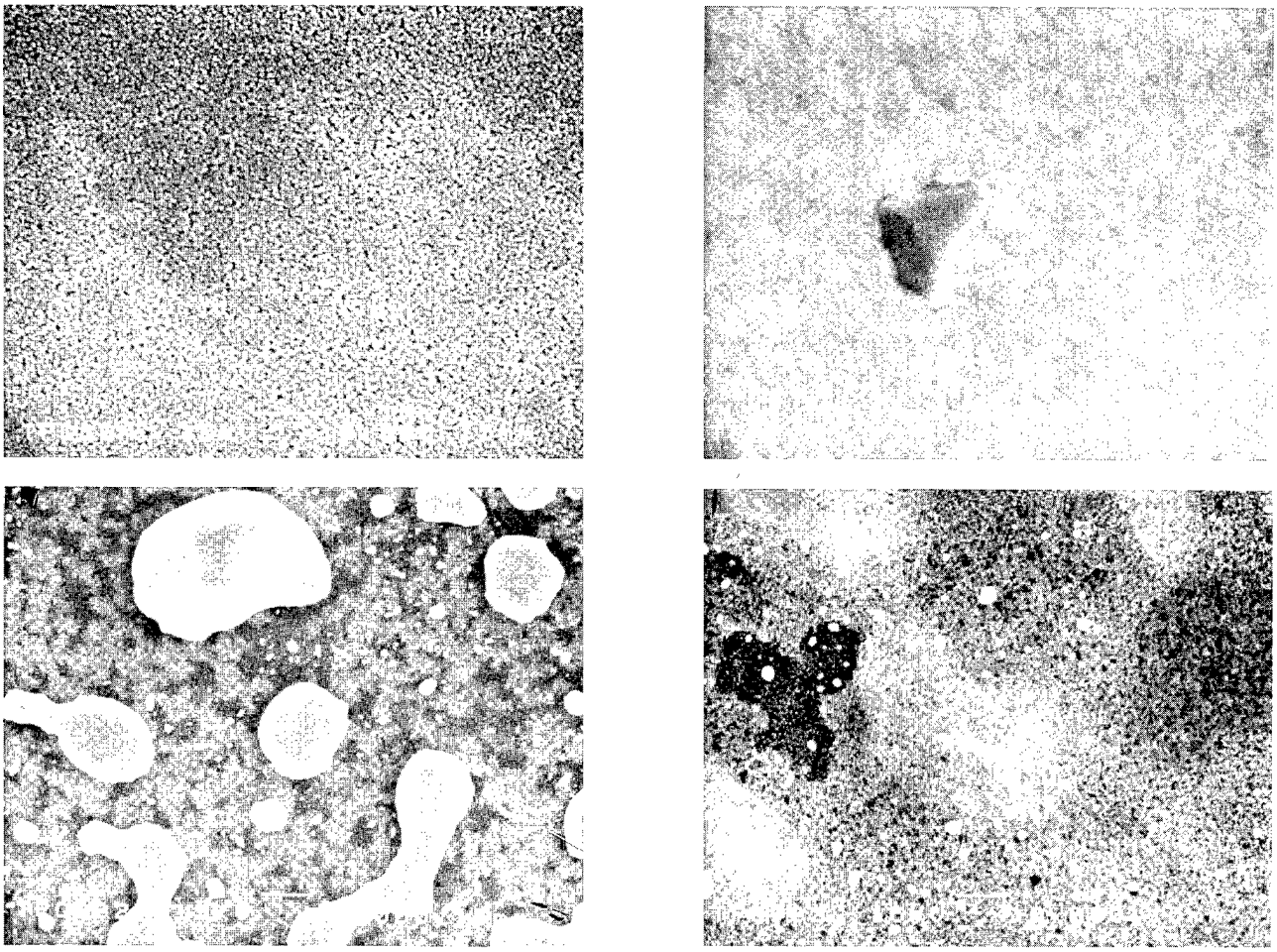


FIG. 2. Surface morphology of Ti/Pt/Au/GaN contacts (magnification ×10 000); (top left) as-deposited, (top right) annealed at 700 °C, (bottom left) annealed at 800 °C, and (bottom left) annealed at 900 °C.

ployed on $In_{0.5}Ga_{0.5}N$ to study lower band gap contacts. The samples were annealed in the temperature range of 400–900 °C.

The contact and sheet resistance of the contact can be derived from a plot of measured resistance versus gap spacing. The method of least squares fit was used to obtain a linear fit to the experimental data, which typically yielded correlation coefficient of 0.9972 or better.

The specific contact resistivity (ρ_c) is calculated from a measurement of the effective contact resistance (R_c), the contact width (W), and the transfer length (L_T):

$$\rho_c = R_c W L_T .$$

The specific contact resistance dropped from 5.95×10^{-5} $\Omega\,cm^2$ on the as-deposited sample to 3.92×10^{-5} after annealing at 400 °C, and reached a minimum value of 1.48×10^{-5} $\Omega\,cm^2$ after annealing at 700 °C (Fig. 6). The trends are similar to those of WSi_x on InGaAs.[10]

IV. DISCUSSION

First, we discuss the result for Ti/Pt/Au metallization, since this has been one of the more common schemes investigated. It has been previously noted that Ti/Au contacts on GaN[6] have large contact resistance even after thermal annealing. These contacts generally suffer from the problem of spiking into the underlying semiconductor upon annealing.[11] To mitigate this problem, Pt, which is a very good diffusion barrier has been used between Ti and Au to prevent Au spiking. In our work, a similar kind of observation was made. The AES studies of the as-deposited sample did not show intermixing of metals and the SEM studies confirmed these results since the sample surfaces were smooth. From the AES studies of the sample annealed at 600 °C, the AES profile confirms that the Pt barrier has prevented Au spiking into the underlying GaN. This solid phase reaction between Ti and GaN to form TiN can help the contact operate through a tunneling mode. From the AES depth profile, it can be seen that for the sample annealed at 600 °C, Ga has penetrated through most of the Ti film. According to Stareev *et al.*,[8] who made similar observations of Ti/Pt/Au contacts on GaAs, the Ga accumulation at the Ti/Pt boundary causes an underestimate of the redistribution of As atoms and, therefore, they performed complementary studies on the species behavior at Ti/Pt interface using RBS spectra of Ti/Pt samples deposited on SiO_2. No experimental verification of Ti–Pt intermixing reaction even for an annealing temperature up to 600 °C was

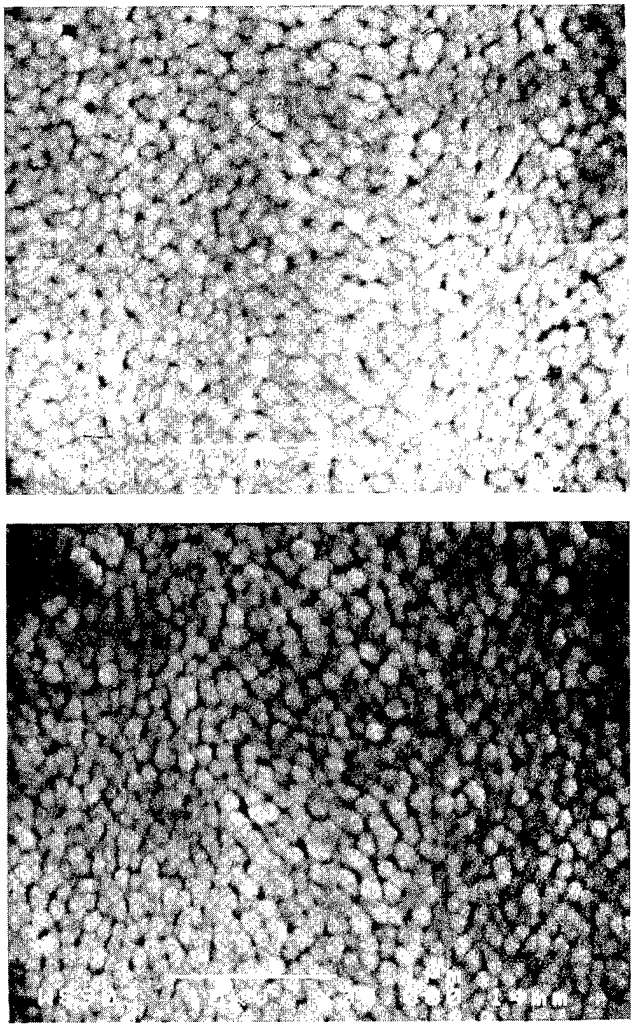


FIG. 3. Surface morphology of WSi_x/GaN contacts (magnification ×30 000); (top) as-deposited, and (bottom) annealed at 900 °C.

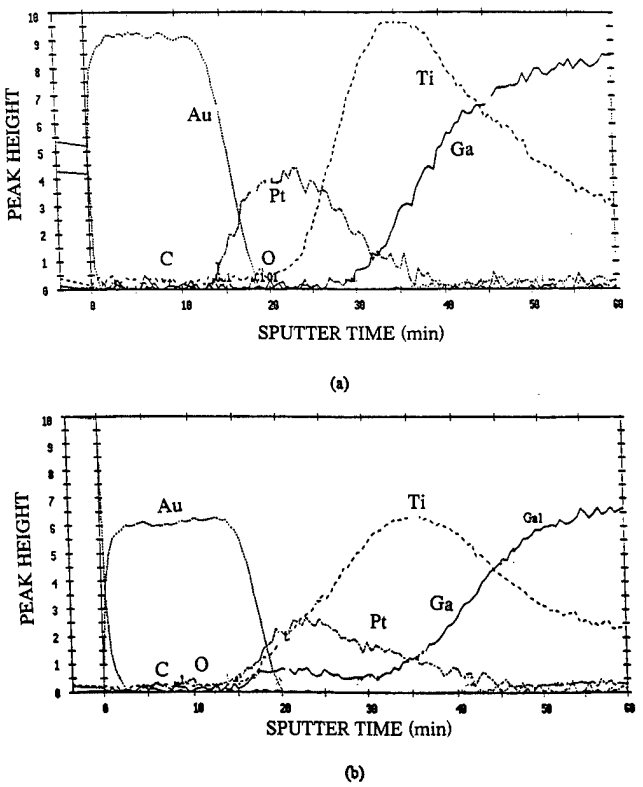


FIG. 4. AES depth profiles of TiPtAu contacts on GaN (a) as-deposited and (b) after annealing at 600 °C for 30 s.

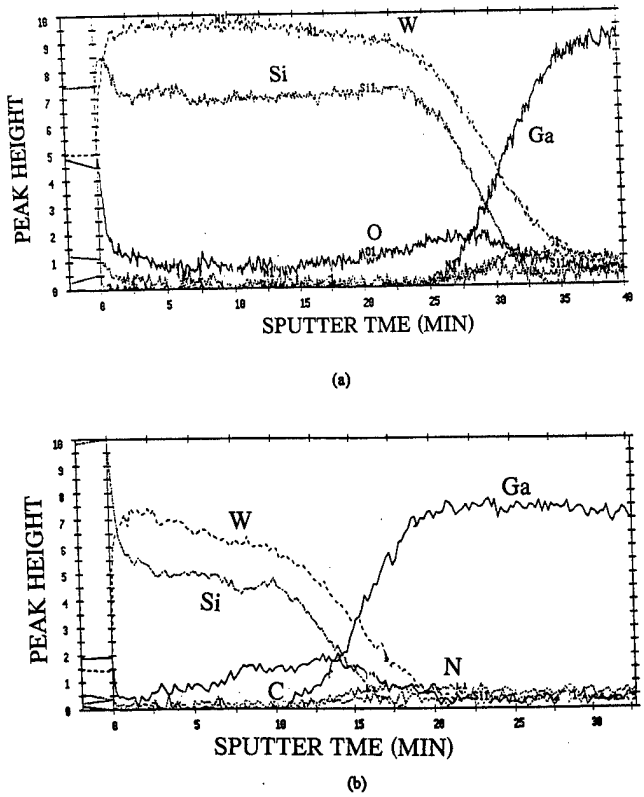


FIG. 5. AES depth profiles of WSi_x contacts on GaN (a) as-deposited and (b) after annealing at 900 °C for 30 s.

observed. Moreover, the accompanying ternary intermixing reaction of T–Pt–Ga at the Ti/Pt boundary takes place only if the Ti film is entirely saturated with As. That means the changes at the Ti/Pt boundary observed in the Pt/Ti/GaAs system after annealing at 600 °C would be attributed to participation of As in a complicated ternary reaction with Ti and Pt. Thus, in our case if we suppose N is extracted from the GaN without decomposing GaN, then an accumulation of N vacancies would be created in the GaN near the metallurgical junction. Since N vacancies most likely act as donors, this region would be heavily doped *n*-GaN which provides the configuration needed for tunneling contacts. Thus annealing temperatures of about 500 °C would be optional for obtaining thermally stable Ti/Pt/Au ohmic contacts but temperatures higher than that may lead to extensive reaction at the interface which implies degrading of the sample.

Turning to the Au/Ge/Ni and AuBe metallization schemes, we find their extensive reaction with GaN at relatively low temperatures makes them unsuitable for applications such as electronic devices operating in the 300 °C range. By contrast, WSi_x provides fairly stable contact to

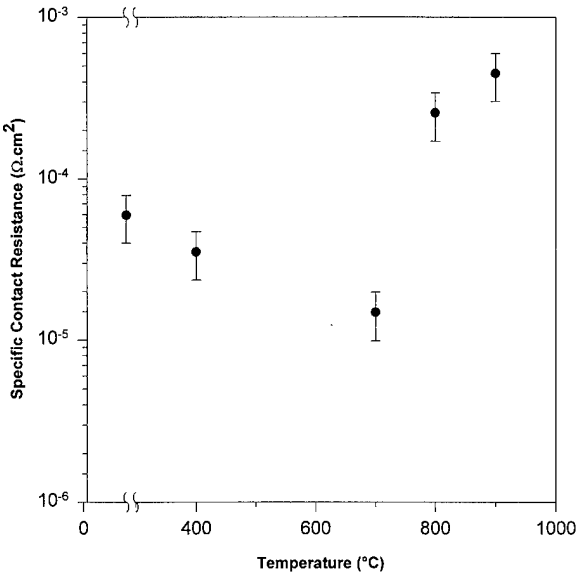


FIG. 6. WSi_x specific contacts resistivity on $In_{0.5}Ga_{0.5}N$ vs annealing temperature.

both GaN and InGaN, with decent resistances after annealing at 700 °C on the latter material, and no visible change in morphology even at 900 °C. Although we have no long-term aging results, there is generally a correlation between good thermal stability and reliability since both are dependent on minimizing metal–semiconductor reactions. As expected from past data, on other III–V materials, it appears that refractory metals are promising candidates for stable contacts on nitrides.

V. CONCLUSIONS

It was evident from the microstructural and interdiffusion studies of the as-deposited and annealed samples of all these metallization schemes that the WSi_x contacts exhibited the best thermal stability and retained good structural properties at annealing temperatures as high as 800 °C on GaN. Processing of FET devices involve a high temperature annealing step for implant activation, typically ≥900 °C. The stability of the $WSi_x/In_{0.5}Ga_{0.5}N$ contacts are high enough to allow dopant activation in heterojunction bipolar transistor or FET structure processing. The contact degradation at higher annealing temperature was related to increases in sheet resistance, which in turn resulted from the degradation of the metal–semiconductor interface.

The AuBe/GaN contact system showed poor thermal stability while the Au/Ge/Ni/GaN contact system was thermally stable up to temperatures of 500 °C. The Ti/Pt/Au contacts were found to remain substantially stable at 450–500 °C. There was not penetration of Au into GaN in this scheme, owing to the Pt diffusion barrier. These contacts exhibited better structural properties than sample Ti/Au contacts deposited on GaN.

The scope of future work in the field of WSi_x/InGaN system is probably to experiment with the In concentration to lower the band gap of InGaN. The contact resistance could be measured as a function of annealing temperature, and the conduction mechanisms and thermal stability established, since the contact resistance should improve for higher InN concentrations.

ACKNOWLEDGMENTS

The work at UF is partially supported by an NSF Grant No. DMR (94-21109), an AASERT Grant No. USARO (J. M. Zavada) and a DOE URI administered through ONR (N00014-92-J-1895). The staff of the Microfabritech facility is greatly appreciated.

[1]H. Amano, M. Kito, X. Hiramatsu, and I. Akaski, Jpn. J. Appl. Phys. **28**, L2112 (1989).
[2]S. Nakamura, T. Mukai, and M. Seno, Jpn. J. Appl. Phys. **30**, L1998 (1991).
[3]M. A. Khan, J. N. Kuznia, A. R. Bhattarai, and D. T. Olson, Appl. Phys. Lett. **63**, 1786 (1993).
[4]M. A. Khan, A. Bhattarai, J. N. Kuznia, and D. T. Olson, Appl. Phys. Lett. **63**, 1214 (1993).
[5]J. S. Foresi and T. D. Moustakas, Appl. Phys. Lett. **82**, 2859 (1993).
[6]M. E. Lin, Z. Ma, F. Y. Huang, Z. F. Fan, L. H. Allen, and H. Morkoç, Appl. Phys. Lett. **64**, 1003 (1994).
[7]M. E. Lin, F. Y. Huang, and H. Morkoç, Appl. Phys. Lett. **64**, 2557 (1994).
[8]G. Stareev, H. Kunzel, and G. Portmann, J. Appl. Phys. **74**, 7344 (1993).
[9]N. Wittmer, P. Ketorius, J. W. Mayer, and M. A. Nicolet, Solid State Electron **20**, 433 (1977).
[10]A. Lahav and M. Genut, Mater. Sci. Eng. B **7**, 231 (1990).
[11]B. M. Henry, A. E. Staton-Beran, V. Sharma, M. A. Crouch, and S. S. Gill, Mater. Res. Soc. Symp. Proc. **240**, 431 (1992).

Patterned eutectic bonding with Al/Ge thin films for microelectromechanical systems

Bao Vu and Paul M. Zavracky
Northeastern University, Boston, Massachusetts 02115

(Received 20 September 1995; accepted 31 May 1996)

In this article, we report our results using the aluminum/germanium eutectic to create high quality patterned bonds between two silicon dice. The bonds are formed using thin metal layers and with essentially no pressure applied. We have measured bond strength by fabricating and bonding patterned dice. Pull tests were conducted and the force required to separate the bonds was measured and found to be about 1.6×10^7 Pa. When bonds break, portions of the substrate are removed. Testing of the hermiticity of the bond demonstrated that leak rates below the detection limit of the leak tester (10^{-9} sccs) are possible. © *1996 American Vacuum Society.*

I. INTRODUCTION

Many microelectromechanical systems (MEMS) that are fabricated using bulk micromachining consist of several dice that are bonded together to create a single unit. Examples include pressure sensors, accelerometers,[1] interferometers, and microvalves. Over the past several years, a number of unique bonding techniques have been developed including electrostatic bonding, wafer bonding, low temperature glass bonding, diffusion bonding, and alloy bonding.[2] Alloy bonding has the advantage of providing a means for simultaneously creating an electrical interconnection and a mechanical bond.[3]

The use of alloy bonding by the semiconductor industry is not new. Thin foils of low melting point gold alloys are used to bond dice to packages. These gold alloys have high strength and produce a good bond that provides both electrical and thermal conductivity to the package.[4] For MEMS the application of foils has limited usefulness because the bond layer cannot be lithoghraphically patterned.

In an alloy bonding process, the metal constituents of the alloy are deposited as thin films with different thicknesses. The bonding temperature can be minimized by choosing the eutectic composition. Three bonding modes can be achieved. In the first, the constituents of the alloy are deposited on two separate substrates such that one substrate contains one component of the binary alloy and the other contains the second component. These are held in close contact with each other, usually under pressure. When heated, solid state diffusion occurs forming a bond between the two surfaces at a temperature below the melting point of the eutectic alloy (diffusion bonding).[5] In the second approach, a binary eutectic alloy is chosen.[6] The components of the eutectic alloy are deposited on two substrates either individually on separate wafers as before or in pairs on each substrate. When the eutectic melting point is reached, a liquid of the eutectic composition is formed. Joining occurs when the liquid surfaces of the parts combine. Surface tension may help to draw the parts together. Upon solidification, the individual elements forming the eutectic separate into a solid containing regions of each. If the ratio of the deposited materials is adjusted away from the eutectic then, upon exceeding eutectic temperature, melting at the interface begins. More material is consumed by the melt as long as the eutectic composition can be maintained. As one of the components becomes depleted, the liquidus point effectively increases and the mixture partially solidifies. In thermal equilibrium, the Lever law can be used to determine the ratio of compounds in the liquid and solid states. Since diffusion in the liquid state is many times faster than in the solid state, more rapid joining with less pressure applied is possible for eutectic bonding. In the third method, a binary system consisting of a low and a high melting point material that also contains a high melting point compound is sought. Peritectics often satisfy these requirements. A layer of the low melting point material is deposited on a substrate which is formed from the higher melting point material or has a thick layer on its surface. When heated, liquid-phase bonding occurs. The liquid reacts with the surface, dissolving the high temperature material until the composition of the liquid matches the liquidus at the bond temperature and the mixture solidifies. Continued heating promotes the formation of the high melting point phase through solid state diffusion. This combination of both types of bonding is referred to as solid-liquid diffusion bonding (SLID).[7]

Little has been written on patterned alloy bonding for MEMS applications. In this article, we describe the results of a series of experiments conducted to optimize eutectic alloy bonding for MEMS. Our work concentrates on the aluminum/germanium binary system. We deposit layers of aluminum and germanium on two substrates such that each substrate contains both Al and Ge at the appropriate thicknesses to create the eutectic alloy. The eutectic composition is targeted in order to obtain the lowest temperature bond. An underlying objective of this work was to isolate the bond layers from the substrate with an insulator (silicon dioxide or silicon nitride) rather than to consume the substrate.[8,9] Dice with layers deposited on their surfaces are brought together and heated to form a bond. We have varied the total film thickness at two levels and formed bonds at several temperatures and times. Both pull tests and leak tests were performed. They indicate that the Ge/Al eutectic is suitable for application to MEMS.

II. SAMPLE PREPARATION

In this work, careful attention was given to the preparation, processing, and fabrication of the samples. In particular, exposure to air can cause the alloy constituents to oxidize. In the Al/Ge system, aluminum readily forms a tenacious oxide that is difficult to remove, whereas germanium oxide can be stripped off in a dilute HF. For this reason, germanium is deposited after aluminum in a single deposition.

In our process, we deposit thin films of each layer individually. This permits precise control over the ratio of the elements forming the binary alloy. The layers are deposited successively in a single pumpdown, patterned photolithographically, and etched. After resist stripping, the wafers are cleaned. Just before bonding, the surface oxide on the germanium is removed in HF and the wafers or die are brought together and clamped. The samples are then heated above the eutectic temperature and held for a short period of time (several minutes). Upon cooling, the samples can be tested to determine the properties and the quality of the bond.

Prime silicon ⟨100⟩ 3 in. wafers were used without regard to their resistivity. All the wafers were subjected to a standard pre-oxidation clean in a modified process originally formulated by Kern and Puotinen at RCA.[10,11] Thermally grown SiO_2 was used as a diffusion barrier on the Si surface and as a masking layer for Si etching. A wet oxidation was performed at 1000 °C for 1.5 hours. The thickness of the oxide was measured with a Nanospec Optical Film Thickness Measurement System and was typically about 8000 Å.

To achieve an Al and Ge bond at 424 °C, the wt % ratio of each component is selected to match that of the eutectic. For a given thickness of aluminum, the germanium thickness can be calculated as follows:

$$t_{Ge}=(t_{Al}\times \text{ wt }\%_{Ge}\times\rho_{Al})/(\rho_{Ge}\times \text{wt }\%_{Al}),$$

where, t_{Al} is the thickness of the aluminum, wt $\%_{Ge}$ is the weight percentage of the germanium, wt $\%_{Al}$ is the weight percentage of the aluminum, ρ_{Al} is the density of the aluminum, and ρ_{Ge} is the density of the germanium. Therefore with 1 μm of aluminum, 0.573 μm of germanium is required to obtain the eutectic alloy (53 wt % germanium).

The adhesion of deposited thin films used in the bonding process must be excellent. In an ideal bond, the bond layer should be stronger than the materials being bonded. Poor adhesion of the bond layers to silicon dioxide could lead to reduced bond strength. To promote the adhesion of Al and Ge thin films a 10 minute clean in a fresh mixture of DI/H_2O_2/NH_4OH (5:1:1 by) at a temperature of 80 °C followed by a 5 minute DI water rinse and drying in a rinser dryer was used prior to thin film deposition.

The thin films were deposited by *e*-beam evaporation. Prior to deposition the chamber was evacuated to less than 5×10^{-6} Torr. The substrate temperature was raised to 250 °C to improve the adhesion of Al to the wafers. The germanium layer was deposited on top of the aluminum at about 100 °C. This temperature was chosen to insure minimal interaction of the layers prior to bonding. A calibrated quartz crystal monitor was used to determine deposition thickness *in situ*.

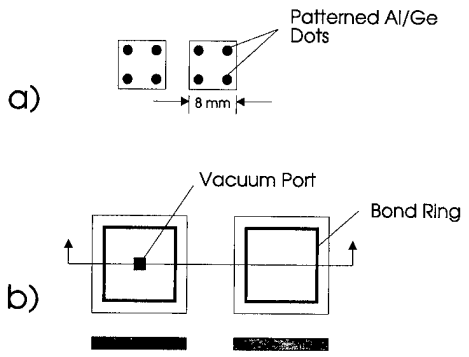


FIG. 1. Two dice shown in (a) are prepared as described in the text and diced. The dice are placed face to face and aligned using their edges. A similar procedure is used for hermetic testing with the die shown in (b).

After the metals were deposited, the wafers were patterned photolithographically. The Ge layer was etched in a solution of 50:2 H_2O_2/NH_4OH and the aluminum in phosphoric, acetic, nitric (PAN) at 50 °C. Two types of samples were prepared. For bond strength measurements, square dice with four circular dots at the edges as shown in Figure 1(a) were patterned. Experiments were conducted with 1.25 mm dots and 0.5 mm dots. A second study of the hermiticity of the bonds was conducted with square annular bond rings prepared on two separate and different dice. One die contained just the bond ring and the second die contained both a bond ring and an anisotropically etched hole at its center as shown in Figure 1(b).

A 25% tetramethyl ammoniumhydroxide [(TMAH), $(CH_3)_4NOH$] solution is used to anisotropically etch silicon.[12] The advantage of TMAH is its high selectivity to Si over SiO_2 (about 4000 at 90 °C[13]). Therefore 8000 Å SiO_2 layer easily survives the TMAH etch of the silicon substrate. The solution of TMAH is heated to 100 °C to obtain an etch rate of about 1.1 μm/min. With 416-μm-thick silicon wafers, the time required to completely etch through is estimated at 6 hours. The TMAH etching is terminated after a total etching time of 5 hours and 49 min to obtain the square hole depth of about 385 μm, leaving a thin diaphragm that is breakable with tweezers when desired.

III. EXPERIMENTAL PROCEDURE

A. Sample clamping and heating

Sample bonding is a simple matter of aligning two dice, clamping them together, and heating them to above the eutectic temperature. Since a dicing saw is used to separate the samples, the dots are aligned to the edges of the die with an accuracy of a few microns. Alignment can be performed using the die edge. Infrared alignment is not necessary. Sample alignment was checked using an infrared microscope.

Clamping is accomplished in two ways. The first employs a simple set of clips. Clips were prepared from some refractory metals including titanium and molybdenum. The clips

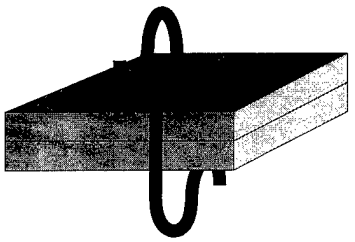


FIG. 2. Two dice clamped together with refractory metal springs.

are placed at the edges of the die as shown in Figure 2. Full wafers were bonded with just three clips. No other means of applying pressure to the samples was used. In our tests of individual die bonding, we found that the use of only a single clip can lead to poor bonds, especially when bonding is conducted at one atmosphere. Therefore the results presented in this article for atmospheric bonding were all performed with multiple clips.

The second clamping method provides both alignment and clamping in a single fixture. The fixture was fabricated in graphite. It consists of two pieces as shown in Figure 3. The lower piece contains a well in which the die can be placed and aligned using the sides of the graphite part. A second graphite lid is placed on top of the samples and a refractory spring is used to clamp the fixture. The thickness of the lid is equivalent to the thickness of the base unit to ensure uniform heating of the sample from the top and the bottom. Thorough testing was not performed to determine if uniform heating is indeed necessary for a good bond. Early experiments using a rapid thermal annealer as the heat source (heating from only one side), however, produced poor results.

Samples were heated in two ambients; in vacuum and in a reducing atmosphere. For vacuum bonding, a horizontal Pyrex bell jar was fitted with a heating apparatus constructed from stainless steel sheet. This apparatus contains two graphite meander strip heaters separated by about 1.5 in. (Figure 4). The heaters are about 6 in. square. Two quartz 1/8-in.-diameter rods are used to support the sample in the center between the two heaters. A thermocouple is suspended between the two heaters and in close proximity to the sample to be bonded. The correlation between the thermocouple and optical temperature readings of the heaters agrees to within 20 °C. The two meander heaters create a blackbody cavity in

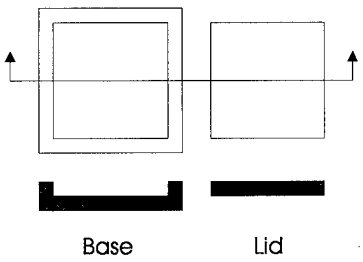


FIG. 3. A drawing of the graphite clamping fixture.

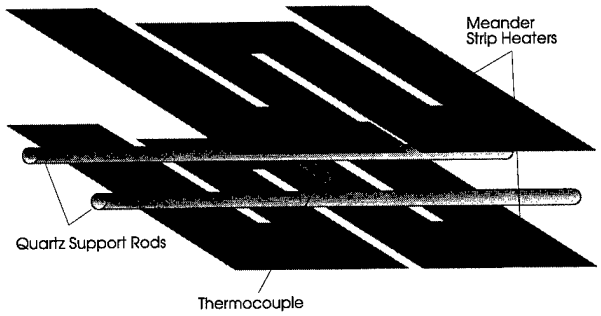


FIG. 4. Drawing of the vacuum bonding system showing two meander strip heaters, quartz support rods, and the location of the thermocouple. These parts are surrounded by a stainless steel heat shield and installed in a vacuum system.

which the sample's temperature rises and comes into equilibrium with its surroundings. This vacuum bonding method should be suitable for devices requiring hermetic sealing.

The second method involves the use of a standard tube furnace (Thermco Minibrute) in which forming gas is introduced. The samples are placed at the front of the tube that is subsequently capped. Forming gas (N_2/5%H_2) is purged for a time equivalent to twenty complete volume changes. Care is taken to minimize the exposure of the sample to high temperatures before the atmosphere inside the tube has settled. A single hole in the end cap allows the sample to be pushed into the tube via a push rod and allows the purge gas to escape.

Initial bonding tests yielded poor results even when conducted at temperatures as high as 575 °C. We attribute this to poor cleanliness of the metal-bond surface layer. The principal source of concern is the surface of the Ge thin film. After photolithography, chemical cleans are employed to remove photoresist from the wafer surface. The wafer is then rinsed in water and dried. We speculate that during this procedure a hydrous oxide film forms on the Ge surface. This layer can be subsequently removed in a dilute HF solution. Some degree of reoxidation prior to sample insertion in the furnace is likely to occur.

B. Bond nature

Figures 5 and 6 and show the results for bonding with and

FIG. 5. Optical micrograph showing a portion of the bond ring from a sample that was bonded with no special surface preparation. The result is a poor bond with only scattered areas that are partially bonded.

FIG. 6. Photomicrograph of a sample that was carefully cleaned prior to bonding. Although the sample was not well aligned, large bonded regions are apparent. In these regions, the silicon has either transferred from the mating sample or was pulled away from this sample during bond separation.

without a pre-bond clean. The bonding appears to be more uniform after cleaning as expected. The big black areas shown in Figure 6 are missing pieces of the bulk Si substrate that has been removed during sample separation.

Bonding experiments were conducted on die containing four dots as shown in Figure 1(a). The deposited layers consisted of 1 μm of aluminum and 0.57 μm of germanium. Experiments were conducted at one atmosphere in a tube furnace. All dice were dipped in a solution of 50:1 H_2O:HF for 10 seconds, rinsed briefly, and blown dry in dry nitrogen prior to being clipped together. Samples from each experiment were broken apart by separating the die with a knife edge. Three bonding regions can be qualitatively identified. An *incomplete* bond would be one for which the sample either fell apart after removing the spring, or required only a modest effort to break apart. Under the microscope, an incomplete bond showed little or no evidence of the removal of silicon from the substrate. In a *partial* bond, frequent occurrences of silicon removal after die separation were observed. As much as 70% of the bond area contained either a void, indicating transfer of material from the substrate to the mating chip, or a chip of silicon, indicating silicon transfer from the mating chip. In a *complete* bond, the original dot has been completely replaced by a void or a chip.

Bonding was performed at 475, 525, and 575 °C. At 475 °C bonding was incomplete after 30 min and improved marginally at 45 min. At 525 °C complete bonding could be achieved in 30 min. Increasing the temperature to 575 °C did not improve the results significantly. These results are reported in Table I.

A second set of samples was prepared with 2 μm of Al and 1.14 μm of Ge on Si/SiO_2 wafers (Table II). These were processed and prepared using the procedures described for the previous experiments. Complete bonding could be achieved at 475 °C after only 10 min. Samples that were placed in the furnace at 575 °C for 5 min did not bond, probably because they did not reach thermal equilibrium.

TABLE I. Bonding results with 1 μm Al and 0.57 μm Ge (average of five samples).

Temperature (°C)	Bond time (min)	Bond nature
475	15	Incomplete
475	30	Incomplete
475	45	Partial
525	15	Incomplete
525	30	Complete
575	5	Incomplete
575	15	Partial

TABLE II. Bonding results with 2 μm Al and 1.14 μm Ge (average of five samples).

Temperature (°C)	Bond time (min)	Bond nature
474	10	Complete
474	20	Complete
575	5	Incomplete
575	10	Complete

C. Pull tests

Pull tests were conducted in a simple apparatus that was constructed for this purpose. Bonded samples were mounted to sample holders using epoxy. The sample holders consisted of 1/4 in. rods approximately 1 in. long and that had an eyelet centered at one end and were machined flat at the other. The sample mounting took place in a *V* block that was used to ensure alignment between the two sample holders. Mounted samples were loaded into a fixture (Figure 7) containing a piston at one end in line with a strain sensor at the other. Steel cable was used to connect the eyelets on either side of the sample to the piston and strain gauge. Air pressure applied to the piston provided the force to pull the samples apart. The strain gauge was calibrated using dead weights.

Tests were conducted on samples patterned with dots as shown in Figure 1(a). Samples with 1 mm dots would not separate even at a force of 25 N. Five samples bonded with 0.5 mm dots separated at an average force of 12.39 N and a standard deviation of 2.35 N. Therefore, the yield pressure, assuming complete bonding over the entire area, was about 1.6×10^7 Pa. This pressure is equivalent to one-tenth of the yield strength of aluminum. After the samples had been separated, they were examined under the microscope (see Figure 8). In all cases, the substrate had been removed from one side of the bond. This suggests that the silicon substrate yielded before the bond.

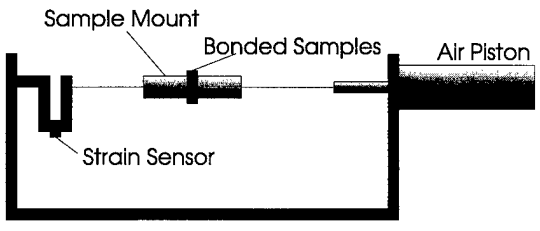


FIG. 7. Schematic of the pull tester used in this work.

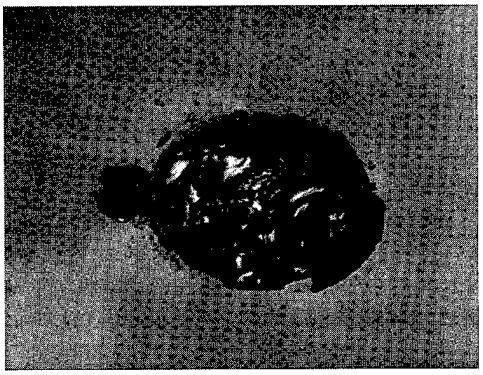


FIG. 8. Optical micrograph of a 0.5 mm dot after pull testing. The black material is silicon that has been removed from the mating piece.

D. Bond hermiticity

For many MEMS applications, it is necessary to create vacuum sealed cavities. To test Al/Ge bonds for hermiticity we created samples in which one die has an anisotropically etched port through which the cavity between two bonded dice can be evacuated. Using a helium leak tester, a series of bonded samples was tested. The connection of the sample to the leak detector is shown schematically shown in Figure 9. The sample is simply placed on an O-ring that seals the sample to the leak port. The system is capable of detecting helium down to a leak rate of 10^{-9} sccs. An internal calibrated leak was used to check the operation of the detector. It should be noted that, at this sensitivity, a typical bulk micromachined device with a volume of 10^{-3} cm^3 would reach one atmosphere in two weeks.

Twenty samples were prepared for testing. These were approximately evenly split for a 2×2 experimental matrix designed to test the influence of time and temperature. Times ranged from 10 to 30 minutes and the temperature was set at 525 and 575 °C. For samples with 1 μm Al and 0.57 μm Ge bonding at 525 °C and 30 minutes produced leaks at about 5×10^{-4} sccs. Table III details these results. Increasing the temperature to 575 °C produced samples with leak rates below the detection limit. Sixty percent of the samples prepared with 2 μm Al and 1.14 μm Ge and tested at 525 °C for 30 minutes were below the detection limit (see Table IV). All of the thicker samples bonded at 575 °C had undetectable leak rates.

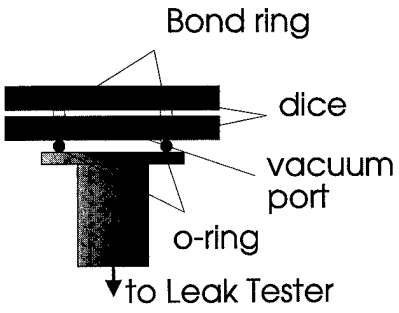


FIG. 9. Cross section of the test die used for leak testing.

TABLE III. Leak test results with 1 μm Al and 0.57 μm Ge.

Temperature (°C)	Time (min)	Standard leak rate (sccs)
525	30	Undetectable
525	30	10-3
525	30	5×10^{-4}
525	30	7×10^{-4}
525	30	5×10^{-4}
525	30	Over the range
575	10	10^{-3}
575	30	Undetectable
575	30	Undetectable
575	30	Undetectable
575	30	Over the range
575	30	Over the range

IV. OBSERVATIONS

Several observations can be made about these results. Very strong bonds can be made using the Al/Ge eutectic alloy system. This was clearly demonstrated by pull tests. One may question why the silicon yielded before the bond broke. A possible weakness in the test was the use of multiple dots patterned in each die that was bonded. Clearly the strain in the silicon has a maximum at the dot periphery. A second contributing factor is that the epoxy used to mount the samples provides enough elasticity that the sample deformation is greater than it would be for a more rigid mounting scheme. In this case, the silicon deformation added stress to the silicon surrounding the bond causing it to yield.

A second observation is that there is a significant effect of time and temperature on bond quality. In general, longer bonding times and higher temperatures produced better bonds. It is also observed that thicker films bond more quickly and at lower temperatures than thin bonds. In our work, 2 μm was considered a practical limit for a single layer and results in a total bond thickness of 6.8 μm for the aluminum/germanium system. A final observation is that there appeared to be only negligible difference between the quality of bonds formed in vacuum and those in reducing ambient.

The time and temperature dependence can be explained in part by postulating a surface oxide layer, as has been reported in other alloy systems.[14,15] At the interface between the two bond metal layers, the two surface layers of Ge are in

TABLE IV. Leak test results with 2 μm Al and 1.14 μm Ge.

Temperature (°C)	Time (min)	Standard leak test (sccs)
525	20	Undetectable
525	20	3×10^{-7}
525	20	3×10^{-7}
525	30	Undetectable
525	30	Undetectable
525	30	1.5×10^{-7}
575	20	Undetectable
575	20	Undetectable

contact. In handling and in exposure to oxygen in the atmosphere, a thin layer of germanium oxide grows. In addition, any oxide trapped during film deposition accumulates at the surface of the molten metal. The presence of the oxide inhibits the merger of the two molten bond layers. Bonding does occur at pinholes and cracks in this thin oxide layer that form as the bond is heated. For incomplete bonding, a pattern of spots where bonding occurred is surrounded by unbonded regions. By capillary action and with time, the bond area increases as the oxide is pushed aside or is dissolved into the molten material. Evidence for this is seen in the partially bonding wafers where a larger percentage of the area has been bonded. The eutectic bonding process could be improved by finding a way to avoid the growth of native germanium oxide or by causing it to dissolve into or flow off the melt during bonding. To avoid problems from oxides trapped in the films during deposition, one could pre-melt the alloy. After the initial melt, the surface is cleaned in preparation for final bonding.

A possible contributing factor to explain the higher bonding temperatures required to achieve a good bond is a variation in the ratio of the thicknesses of the germanium and the aluminum. If the ratio is offset to either side of the eutectic, the liquidus temperature will increase. We first note that the dual *e*-gun system used for this work had a rotating substrate holder. The guns were placed at equal distances from the center of the system and on a line passing through the center. Therefore, evaporation from either gun is essentially equal. Both depositions are performed sequentially and therefore the thickness variation across the wafer is essentially identical for both depositions. Therefore, the alloy composition is constant across the wafer, and only its total thickness varies. The only other source of variation is in the ratio of the thicknesses. In other words, while the alloy is the same everywhere on the wafer, it may not be the eutectic alloy. We observe that, around the eutectic composition, the liquidus can be approximated as a straight line with a slope

$$\left.\frac{dT_l(\text{wt }\%_{\text{Ge}})}{d\ \text{wt }\%_{\text{Ge}}}\right|_{\text{wt }\%_{\text{eu}}^{\text{Ge}}} = s_{\text{Ge}},$$

where T_l (wt $\%_{\text{Ge}}$) is the liquidus, wt $\%_{\text{eu}}^{\text{Ge}}$ is a weight percent of Ge in a near eutectic alloy composition that is slightly Ge rich, and S_{Ge} is the slope of the liquidus.

It should be noted that the quartz crystal monitor measures the deposited mass, or the product of the density, thickness, and area. The ratio of measurements from two sequential depositions removes the area and we define Γ_{alloy} as the ratio $\rho_{\text{Ge}} \cdot t_{\text{Ge}} / \rho_{\text{Al}} \cdot t_{\text{Al}}$. Then it can be shown that

$$\frac{d\ \text{wt }\%_{\text{Ge}}}{d\Gamma_{\text{alloy}}} = \frac{1}{(\Gamma_{\text{alloy}}+1)^2},$$

and therefore

$$\left.\frac{dT_l(\Gamma_{\text{alloy}}^{\text{Ge}})}{d\Gamma_{\text{alloy}}^{\text{Ge}}}\right|_{\Gamma_{\text{eu}}} = s_{\text{Ge}} \cdot \frac{1}{(\Gamma_{\text{eu}}+1)^2},$$

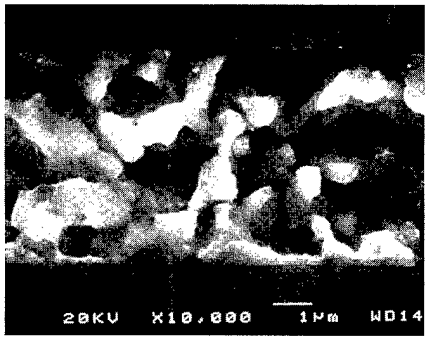


FIG. 10. A SEM cross section of a complete bond. To enhance the contrast between the aluminum and the germanium, the germanium was etched in HF.

where $\Gamma_{\text{alloy}}^{\text{Ge}}$ is a ratio that's slightly germanium rich, and Γ_{eu} is the eutectic ratio.

The slope of the liquidus with increasing Ge concentration is about 15 °C/% for the aluminum/germanium system. This slope is about double the slope of the liquidus for the Ge poor alloy. From the above we can estimate that even if our control of the thickness ratio is in error by 10% (5% for each deposition), the increase in the liquidus will be 37 °C.

When complete bonding occurs, the dice cannot be broken at the original bond line. We have cross sectioned well bonded dice and examined the interface using a scanning electron microscope (SEM). An micrograph is shown in Figure 10. The silicon dioxide layers on both die are clearly visible as the dark thin films. The bond layer shows two differently shaded regions. Analysis of these regions with x-ray dispersion indicates that the light region is an aluminum rich alloy and the dark layer is germanium rich. This result is characteristic of the solidification of eutectic alloys.[16]

Aluminum/germanium bonding can be used for electrical interconnection. In our early efforts, however, we observed a large number of devices that were shorted to the substrate. To determine the origin of the shorts, we fabricated samples that consisted of an oxidized silicon substrate, a metal layer, a LTO capping layer, and the bond layer. The bond layer was patterned into dots for probing; these were subsequently heated to emulate a bonding cycle. We found a large number of shorts between the dots and the underlying metal layer. Upon closer examination, it was determined that the Al/Ge alloy diffused through the oxide. After capping the LTO with silicon nitride, the tests were conducted without any shorting observed.

V. CONCLUSIONS

Eutectic bonding using Al and Ge was investigated. Samples were prepared with individual layers of both metals with thicknesses adjusted to create the eutectic when fully melted. Several observations were made. Proper surface preparation of the layers is essential for good bonding. Our process leaves the germanium as the upper layer that can be cleaned prior to bonding with a dilute HF dip and rinse step. Second, the bonding does not take place when the melting

point is reached. In fact, we have found that it is necessary to exceed the eutectic melting point by about 100 °C to ensure a reliable bond. Third, the bond does not form immediately upon reaching the bond temperature. A time/temperature dependence exists with the quality of the bond improving as both increase. Finally, the bond improves as the layer thicknesses increase. Bonds can be achieved with 1 μm of Al and 0.57 μm of Ge but, by doubling the thickness of both, the quality of the bond is improved and bonding can be conducted at a lower temperature. Using the techniques described in the article, we have achieved high strength bonds that yield above 1.6×10^7 Pa. We also found that hermetic sealing using eutectic bonding is feasible. All the bonds tested were patterned photolithographically and formed with essentially no external pressure.

ACKNOWLEDGMENTS

The authors would like to thank Keith Warner for his help in setting up experimental equipment needed to conduct these experiments and for many helpful suggestions. They would also like to thank Nick McGruer for his critical review of this document. Finally, they thank Jianchao Wang for his careful observations and detection of the substrate shorting problem. Funding for this work was supported in part by the Jet Propulsion Laboratory.

[1]P.M. Zavracky, F. Hartley, N. Sherman, T. Hansen, and K. Warner, The 7th International Conference on Solid-State Sensors and Actuators, Yokohama, Japan, 7-10 June, 1993 (unpublished).

[2]W.H. Ko, J.T. Suminto, and G.J. Yeh, Bonding Techniques for Microsensors, *Micromachining and Micropackaging of Transducers,* edited by C.D. Fung, P.W. Cheung, W.H. Ko, and D.G. Flemming (Elsevier Science, Amsterdam, 1985).

[3]R.F. Wolffenbuttel and K.D. Wise, in Ref. 1.

[4]G.S. Matijasevic, C.C. Lee, and C.Y. Wang, Thin Solid Films **223**, 276 (1993).

[5]C.C. Lee, C.Y. Wang, and G. Matijasevic, IEEE Trans. Components, Hybrids, Technol. **16**, No. 3 May (1993).

[6]A.-L. Tiensuu, M. Bexell, J.-Å. Schweitz, L. Smith, and S. Johansson, Sensors Actuators A **45**, 227 (1994).

[7]L. Bernstein, Electromechan. Soc. **113**, 1282 (1966).

[8]A.-L Tiensuu, J.-A. Schweitz, and S. Johansson, 8th International Conference on Solid-State Sensors and Actuators, Stockholm, Sweden, June 26–29, 1995 (unpublished).

[9]A.P. Lee, D.R. Ciarlo, P.A. Krulevitch, S. Lehew, J. Trevino, and M.A. Northrop, in Ref. 8, pp. 368–371.

[10]W. Kern and D.A. Puotinen, RCA Rev., June 187 (1970).

[11]W. Kern, Semicond. Int., April, p. 94 (1984).

[12]U. Schnakenberg, W. Benecke, and P. Lange, Transducers 91, International Conference on Solid State Sensors and Actuators, June 24–28, pp. 815–818.

[13]O. Tabata, R. Asahi, H. Funabashi, and S. Sugiyama, Transducers 91, International Conference on Solid State Sensors and Actuators, June 24–28, pp. 811–814.

[14]C. Y. Wang, Y.C. Chen, and C.C. Lee, IEEE Trans. Components, Hybrids, Manuf. Technol. **16**, Dec. (1993).

[15]T.P.L. Li, E.L. Zigler, and D.E. Hillyer, Proceedings of the 22nd IEEE International Reliability Physics Symposium, 1984 (unpublished) pp. 167–174.

[16]See, for example, C.R. Barrett, W.D. Nix, and A.S. Tetelman, *Principles of Engineering Materials* (Prentice–Hall, Englewood Cliffs, NJ, 1973).

Simulation of three-dimensional refractory metal step coverage over contact cuts and vias

M. K. Sheergar and T. Smy[a)]
Department of Electronics, Carleton University, Ottawa, Ontario K1S 5B6, Canada

S. K. Dew and M. J. Brett
Department of Electrical Engineering, University of Alberta, Edmonton, Alberta T6G 2G7, Canada

(Received 6 July 1995; accepted 1 May 1996)

In very large scale integration metalization processes, thin diffusion barrier metal films deposited over vias and contacts are used to inhibit the diffusion of the top metal layer into the underlying junctions. These thin films are normally sputter deposited and their reliability is crucial to die yield and device lifetime. As wafer sizes increase, die at the wafer edge receive a directionally asymmetrical sputter flux, leading to nonuniform and asymmetrical film coverage over vias and contacts. Two three-dimensional (3D) extensions of the two-dimensional (2D) film growth program SIMBAD are presented. The first extension (quasi-3D) assumes a radially symmetrical adatom flux distribution and would be appropriate for the simulation of film deposition on regions near the central axis of a radially symmetrical target. Due to the assumption of a symmetrical flux distribution, the resulting film growth is symmetrical. The second extension (interpolated-3D) does not assume that the flux is symmetrical and is intended for the simulation of asymmetrical film growth off-axis. Both extensions provide a 3D surface of the deposited film and the microstructure and profile of 2D slices through the via or contact. In order to determine the accuracy of the two models, simulations were compared to W films sputter deposited over contacts having an aspect ratio of 3:1. The models were tested for a substrate placed directly under the center of the target. The interpolated-3D model was also tested for asymmetrical fluxes by placing the substrate off-axis by 2.2, 5.0, and 8.1 cm. The comparisons with the experimental results indicate that both the surface profiles and the microstructure of the deposited films were predicted well by the two models. The interpolated 3D model was very successful at predicting deposition profiles for off-axis deposition flux. © *1996 American Vacuum Society.*

I. INTRODUCTION

The trend in the microelectronics industry is for a continued reduction in device feature sizes in order to increase the level of integration and the device performance.[1] This has led to the requirement of diffusion barrier layers to reduce material interactions leading to, for example, shallow junction spiking. The deposition of barrier layers is usually done by sputtering, which is a well established process in the microelectronics industry, for a wide variety of materials.[2]

A number of factors determine the effectiveness of a barrier layer. It is important that the deposited film be as uniformly deposited as possible over the topography to be coated. For sputtering processes, the uniformity of the film is primarily determined by the shadowing of the adatom flux by the initial topography and the depositing film. The final film profile is therefore strongly determined by the angular adatom flux distribution. In addition to the surface profile of the film, the internal microstructure of the film plays a large role in determining the effectiveness of the diffusion barrier. Sputtered thin films exhibit a particular microstructure,[3] which is determined to a large extent by the melting point of the material being deposited, the angular distribution of the adatom flux, and the orientation of this flux to the local surface normal. This microstructure is important in determining a variety of film properties, including local density and diffusion barrier permeability.

Traditionally, the development of sputtering processes has been dominated by experimental optimization of the film properties. However, due to increased complexity of the integrated circuit structures and the tighter tolerances, this method is becoming prohibitively expensive and time consuming. It is therefore important to develop process simulators to model and help optimize these processes. Previous modeling can be split into two distinct groups: string algorithm models that represent the surface as a connected set of line segments[4–6] and assume the internal film to be homogeneous, and microstructural models which provide a depiction of the internal film structure in addition to the film profile.[7,8]

Predictions from the reported models have been extensively compared to scanning electron micrographs (SEM) of cross sections of films deposited over trenches. A large number of physical effects have been incorporated, including nonunity sticking coefficients, ion bombardment effects, and high temperature effects.[6,9,10] In addition, Monte Carlo flux transport models[11] have been coupled with the growth simulator to study the effects of sputter geometries, deposition conditions, and the use of collimators.[7]

The modeling of sputter deposition of refractory barrier layers has primarily been two dimensional (2D) and conse-

[a)]Electronic mail: tjs@doe.carleton.ca

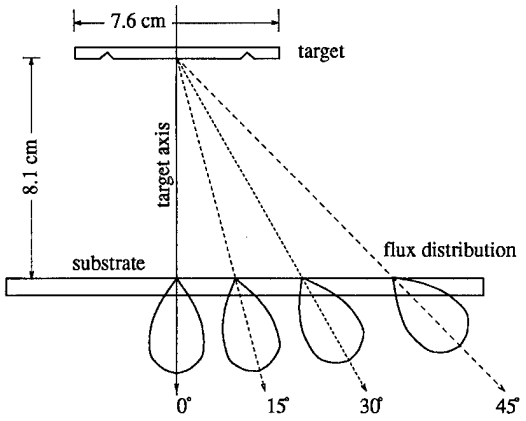


FIG. 1. Incident sputter flux distribution at various angles off the central axis of the target.

quently limited to deposition over long features such as trenches. In this situation the three-dimensional (3D) flux can be projected onto a 2D plane and the growth simulation can be done in 2D. A number of situations arise that are inherently 3D. These include deposition into vias/contacts and off-axis deposition. For these cases both the film surface and the flux incident on this surface must be represented in three dimensions. At present there has been some initial work on 3D film growth,[12–17] however, important 3D effects have often been neglected. For example, the effect of the asymmetrical sputter flux distribution present at the wafer edge needs to be addressed (see Fig. 1). This asymmetrical flux distribution will cause asymmetries in the film profile and also effect the local microstructure of the film. Both of these effects can seriously affect the reliability of the metallization at the edge of the wafer and must, therefore, be minimized. In order to simulate 3D film growth of thin films over very large scale integration (VLSI) topography it is necessary to have a 3D angular flux distribution and 3D model of the flux shadowing due to the topography.

In general, it might be felt that a full 3D representation of the film is also needed, but this is often very inefficient and unnecessary. For example, to represent a $(2\ \mu m)^3$ structure with 10 nm resolution requires $\sim 10^7$ elements resulting in hundreds of megabytes of memory usage and a correspondingly long simulation time. Furthermore, in order to quantitatively interpret the film depiction, selected 2D slices are usually taken through the structure, thereby ignoring most of the 3D representation. If possible, it is far more efficient to represent the film just in those selected 2D slices provided the 3D effects discussed above can still be accounted for. (For example, the same scenario above would only require only 4×10^4 elements per 2D slice.)

This article describes two such algorithms for providing important 3D information (including internal microstructure) using efficient 2D film representations. To accomplish this, two programs are used. SIMSPUD (Simulation of sputter flux distribution) is a three-dimensional Monte Carlo simulation of the transport of the sputtered material to the substrate.[11] This program is used to obtain the 3D spatial and angular distributions of the depositing sputter flux. SIMBAD (simulation by ballistic deposition) is a 2D thin film growth simulator used for the simulation of the deposition of the metals by sputtering, bias sputtering, and chemical vapor deposition.[18] As discussed above, two 3D extensions for SIMBAD will be presented. The first extension (quasi-3D) exploits cylindrical symmetry to reduce the system dimension to two. If the flux distribution and substrate topography have this symmetry (for example, a round contact or via hole), then a 3D position (ρ,ϕ,z) in cylindrical coordinates can be reduced to a 2D one (ϕ,z) suitable for use in SIMBAD. However, when the substrate feature or flux distribution is not cylindrically symmetrical (for example, inside a damascene structure or far from the axis of deposition), the quasi-3D approach is no longer valid. Under these conditions, the more general SIMBAD extension (interpolated-3D) can be used.

The interpolated-3D algorithm uses a small number of vertical slices at different angles to represent the key views of the 3D features. Between these slices, a smooth interpolation of the film surface is used to properly account for 3D shadowing effects.

Both of these SIMBAD extensions are described in greater detail below. In addition, predictions from both models will be compared to scanning electron micrograph (SEM) cross sections of W films deposited over circular contact cuts.

II. MODELING

A. Sputter flux transport—SIMSPUD

SIMSPUD is a three-dimensional Monte Carlo vapor transport simulation.[11] In this model, individual sputter flux particles are followed from the time they are emitted from the target, through collisions in the sputter gas, to their final deposition onto the substrate. The position, angle, and energy of each particle is tracked in order to collect the statistics on the corresponding sputter distribution. Particles are emitted from the target at a random position which reflects the erosion profile of the sputter target. The emission angular distribution can be input from experimental results or alternatively, a cosine distribution can be used. A hard sphere scattering model gives scattering angles, but an energy-dependent collision cross section is used to determine collision events. Both 2D and 3D statistics are collected for any specified region of the substrate.

B. Film growth—SIMBAD

1. Two-dimensional simulation

The basic SIMBAD program is a two-dimensional Monte Carlo simulator developed to model the growth and microstructure of thin films deposited over topography.[7,9,10,18] The basic model accounts for the effects of vapor flux shadowing and surface diffusion by the adatoms. These are the most important factors in determining the step coverage and the columnar microstructure at low to moderate temperatures. In SIMBAD, the film is represented by an aggregation of 10^4–10^5

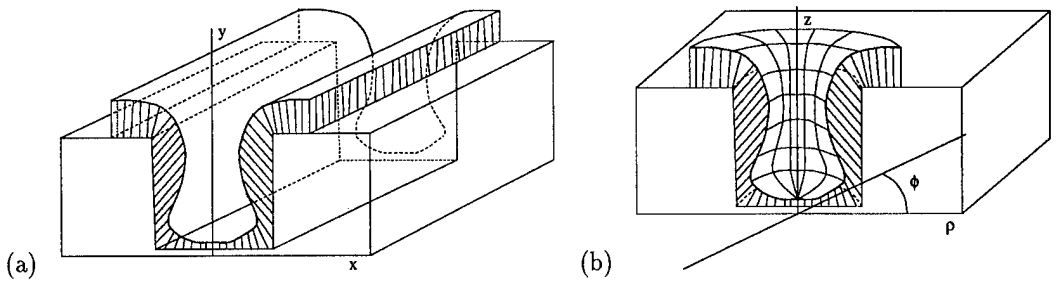


FIG. 2. (a) Schematic drawing of a standard two-dimensional SIMBAD simulation of a trench and (b) a quasi-3D SIMBAD simulation of a round contact assuming rotational symmetry.

identical hard disks. Each of these disks represents the statistically averaged behavior of a large number of film atoms (typically 1000 per disk). These disks are serially launched from just above the growing film and follow straight line trajectories until they strike the substrate or the film. The disks aggregate into the film after a short surface diffusion to minimize a surface curvature-dependent chemical potential. The combination of ballistic shadowing and short range surface diffusion successfully accounts for the formation of the columnar thin film microstructure characteristic of refractory metals at typical deposition temperatures.

The angular distribution computed by SIMSPUD for sputter flux arriving at the substrate is used by SIMBAD to generate the initial disk trajectories. The interaction between this distribution and the substrate topography is largely what determines film coverage. Without extensions, the basic 2D SIMBAD model is suitable for simulating deposition in a long feature such as trench [see Fig. 2(a)] This does first require the projection of the 3D flux distribution from SIMSPUD into a plane perpendicular to the trench.

2. Quasi-three-dimensional simulations

The simulation of a circular via or contact positioned under the center of cylindrically symmetrical target can be performed by exploiting cylindrical symmetry (quasi-3D simulation). As described above, positions and trajectories in 3D coordinates (ρ,ϕ,z) can be mapped back to the 2D representation through a simple rotational projection: $(\rho,z)\rightarrow(x,z)$. Consequently, the cross section of any slice through the contact center will be essentially identical. This is depicted in Fig. 2(b). An additional consideration is mass conservation during radial motion. For example, the effect of any particle impinging at radius ρ must be divided over the area it represents, $\rho d\rho d\phi$. This leads to a $1/\rho$ weighting factor, which, for a discrete model like SIMBAD, is implemented by adding disks with a probability proportional to $1/\rho$. The singularity at $\rho=0$ is handled by truncating the probability equal to 1 at some small $\rho=\rho_0$. This leaves a small notch of width $2\rho_0$ in the center of quasi-3D simulations which can be minimized at the expense of simulation efficiency.

Another limitation of the algorithm is that surface diffusion proceeds in 2D due to the $1/\rho$ weighting limitations. However, for the relatively small diffusion characteristic of refractory materials, this is not a serious deficiency. The main limitation of quasi-3D is, of course, that often the cylindrical symmetry is broken by, for example, deposition near the wafer edge where the peak of the flux distribution would be shifted toward the wafer center. In this case, the more general interpolated-3D algorithm would be needed.

3. Interpolated-three-dimensional simulations

As discussed above, a full three-dimensional simulation of the microstructure of a sputter deposited film would be computationally taxing because of the need to store a great deal of internal information. However, a model which only represents detailed internal information in selected critical locations and just stores the gross surface information everywhere else would require much less computation and memory. The SIMBAD interpolated-3D model takes this latter approach and computes the gross 3D surface features by smoothly interpolating between a number of high resolution, highly detailed 2D slices. Each of these slices is vertical and passes through the center of the simulation. The basic simulation geometry for two slices is shown in Fig. 3, although any number could be used to increase the interpolation accuracy. The interpolated 3D surface is then used to determine the film shadowing effects and the 3D profile of the deposited film. The 2D slices provide highly detailed microstructural depictions through representative sections of the film.

Figure 4 shows a simplified cross section of two slices through the center of the contact, displaced by 90° in angle. The surface profiles in each slice are divided into an equal number of parts. As the profile of the film along these two slices are not symmetrical, the length of each of these divisions will not be equal. The interpolated surface between these two orthogonally displaced slices are shown by curved lines. Figure 4 shows the linear (in cylindrical coordinates) interpolation between two corresponding points, one on each of the two slices. The aim is to achieve a smooth transition from point P_1 to P_2. This was done by dividing the angle between these two points (which is 90° in this case) in the x-y plane into the number of equal angles each denoted by ϕ, as shown in the figure. The distance ρ (as shown in the figure) is then calculated by using the relation

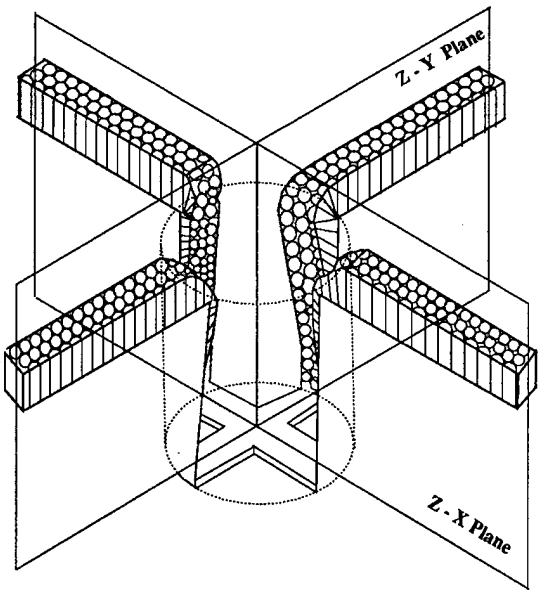


FIG. 3. A schematic depiction of two 2D slices used for an interpolated 3D simulation.

$$\rho = y_1 - y_1 \frac{\phi}{90^\circ} + x_1 \frac{\phi}{90^\circ}. \tag{1}$$

The height of the surface at the position r is then calculated using

$$z = z_1 - z_1 \frac{\phi}{90^\circ} + z_2 \frac{\phi}{90^\circ}. \tag{2}$$

Likewise, all the line segments between points P_1 and P_2 are interpolated and joined. The interpolation was repeated for all the points on the two slices to give the interpolated profile of the deposited film. Each of the quadrilaterals formed was then divided into triangles as shown. Similar interpolations were carried out between all the slices to give the cylindrical profile of the contact as shown in Fig. 8. The interpolated surface is then used to determine the film shadowing for subsequent deposition.

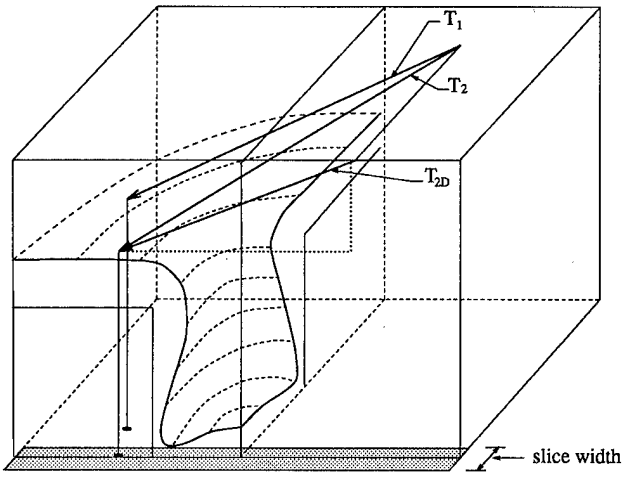


FIG. 5. Disk trajectories, shadowing effect and transformation of 3D to 2D trajectory for 2D SIMBAD deposition.

The simulation of the growth of the surface involves the generation of an initial 3D surface from the 2D slices of the initial topography. A number of disks are then deposited onto the 3D surface using a 3D flux distribution. The trajectories of these disks are followed until they collide with the surface. If the trajectory strikes the 3D surface at a 2D slice, the disk is deposited into that slice. Figure 5 is the depiction of the trajectories of disks launched from points above the substrate. T_1 and T_2 are two such trajectories. It can be seen that as T_1 falls outside the boundary of the slice (5% of the total width), it is therefore discarded. On the other hand T_2 falls within the boundary of the slice and is therefore accepted. The trajectory T_2 which is a 3D trajectory is then projected

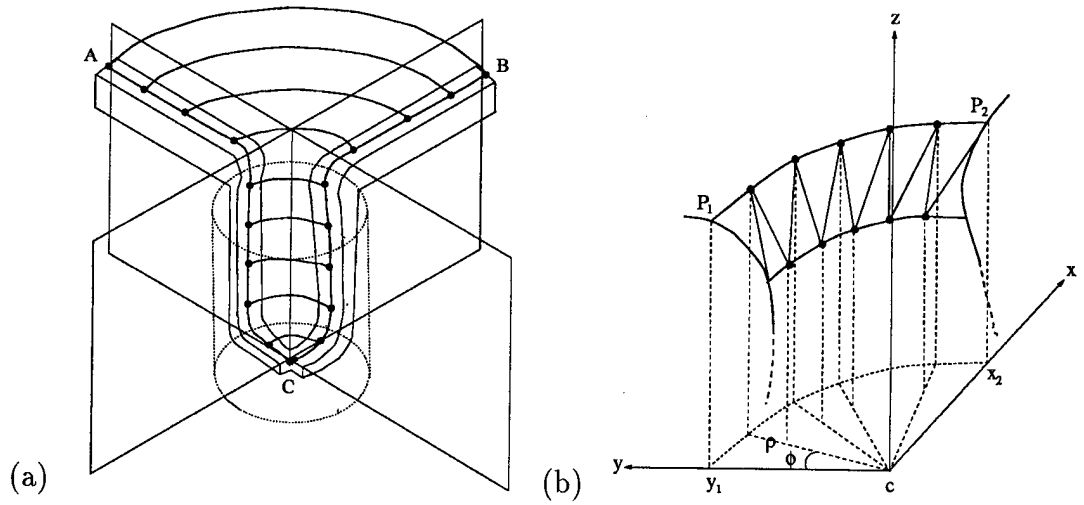


FIG. 4. (a) Cross section of the two orthogonal slices through the center of the contact and (b) 3D interpolation between slices.

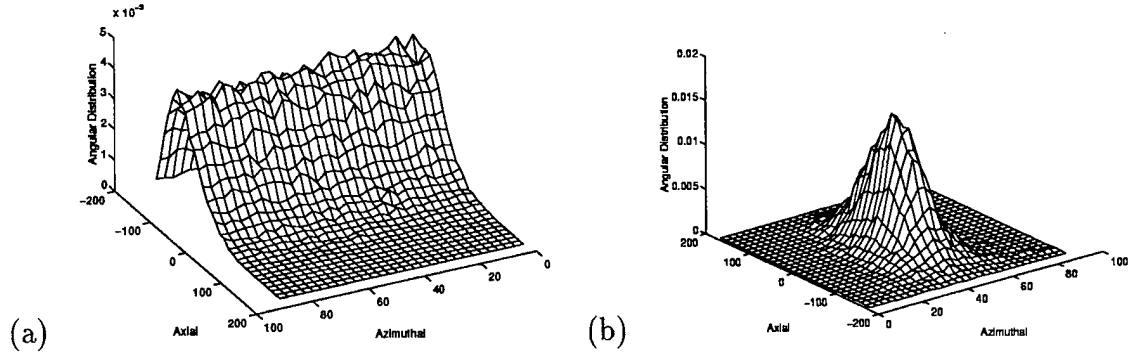


FIG. 6. Three-dimensional adatom flux at the wafer surface for a point (a) directly below the target and (b) off-axis by 8.1 cm.

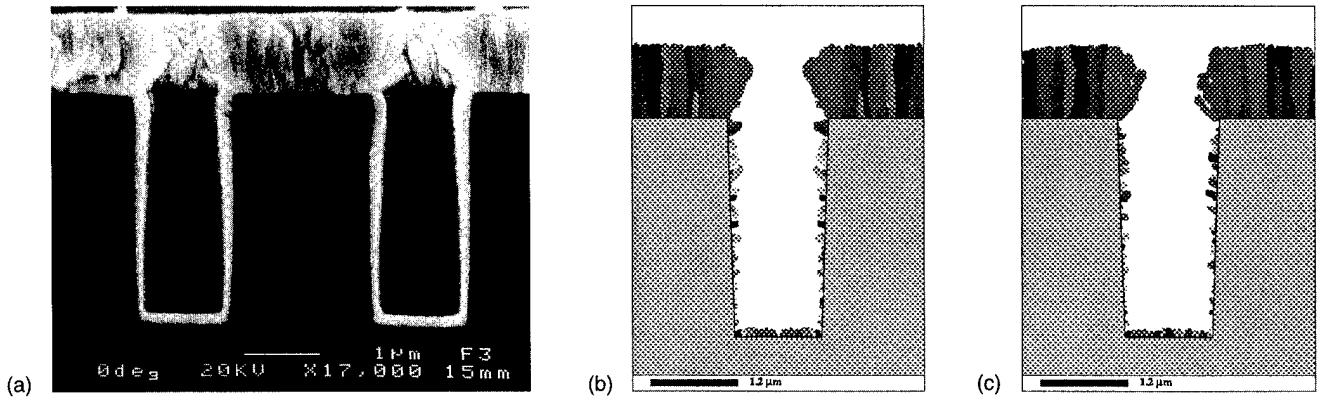


FIG. 7. SEM (a) of W film deposited directly under the target and the corresponding quasi-3D (b), and interpolated 3D (c) simulations.

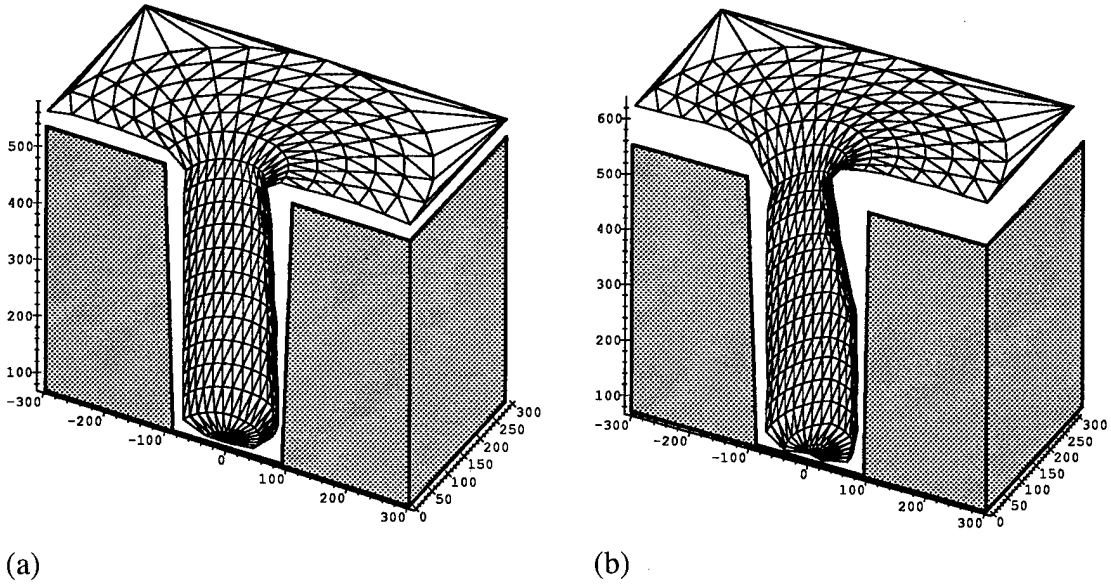
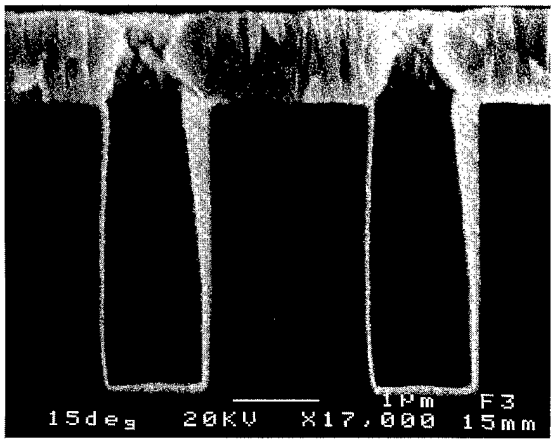
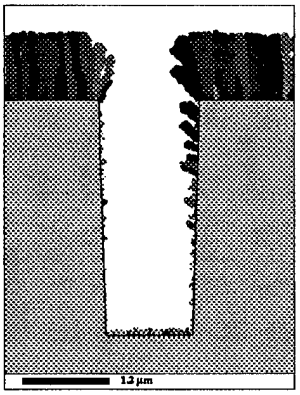


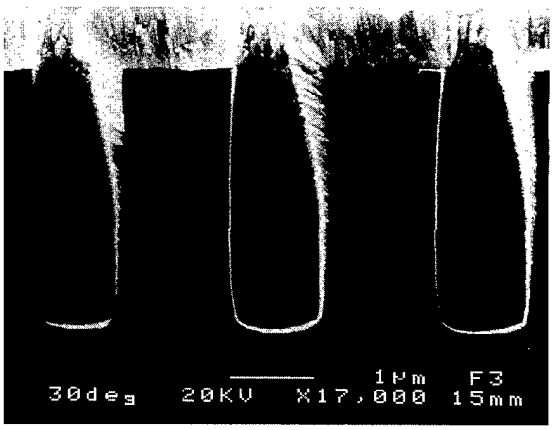
FIG. 8. Consecutive depictions of the interpolated 3D surface of a film deposited off-axis after (a) 0.35 μm and (b) 0.7 μm of growth.



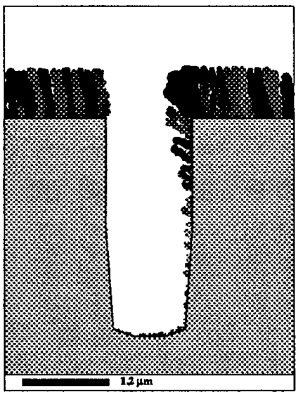
(a)



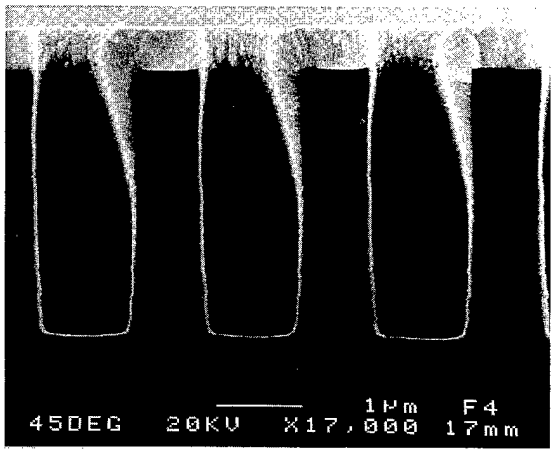
(d)



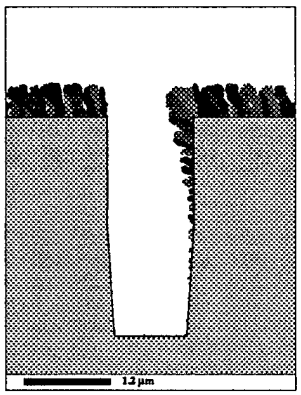
(b)



(e)



(c)

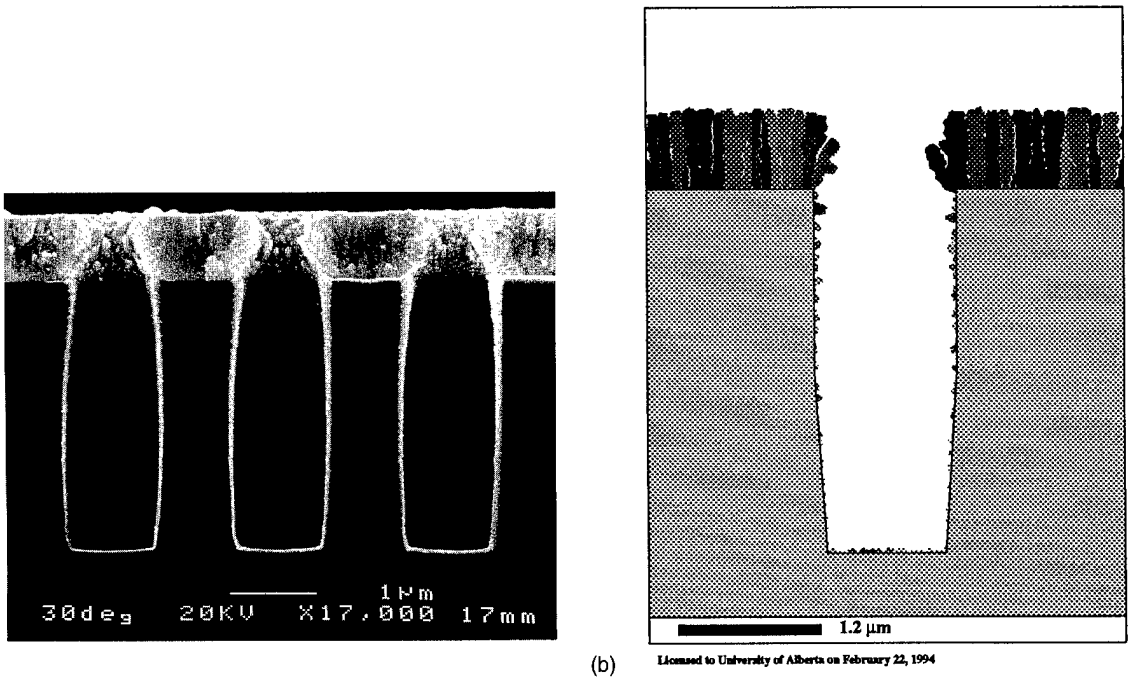


(f)

FIG. 9. SEM photographs of W films deposited over contacts off-axis by (a) $x=2.2$ cm ($\theta=15°$), (b) $x=5.0$ cm ($\theta=30°$), and (c) $x=8.1$ cm ($\theta=45°$). Corresponding SIMBAD simulations (d)–(f) are shown to the right of each SEM.

onto the 2D slice in y-z plane as T_{2D} and deposited using the simple 2D SIMBAD model. Using this procedure a number of disks are deposited into the 2D slices, and then a new, updated 3D surface is generated. In order to maintain accuracy the 3D surface must be updated frequently with respect to the film growth.

The accuracy of the simulated profile also depends on the number of slices considered. The higher the number of



(a) (b)

FIG. 10. SEM (a) and corresponding simulation (b) of deposition on the y axis (at right angles to the off-axis flux) for the case of $x=5.0$ cm ($\theta=30°$).

slices, the smaller the region to be interpolated and higher will be the accuracy of the simulated film. The simulation time is linearly proportional to the number of slices. It will, however, still be much faster than a fully 3D model. A typical film growth simulation on a Sun Sparc 20 runs in under 20 min.

The model does have a number of limitations. It does not incorporate diffusion in the tangential direction, so it can be expected to only be strictly correct for small surface diffusion lengths. The sticking coefficient is assumed to 1.0, and this will limit the type of processes and materials that can be simulated. Lastly, the simple interpolation scheme for the 3D surface limits the initial topography to cylindrical or eliptical geometries.

III. RESULTS

To determine the accuracy of the simulations, calculated results were compared to experimental films. High purity tungsten was dc sputtered from a 7.5-cm-diam target using a planar magnetron source at 100 W of constant power with nominally -325 V target bias. The deposition pressure was 11 mTorr of argon. The base pressure was about 5×10^{-7} Torr. The substrate profile consisted of square contacts in sizes from 0.8 to 2.0 μm wide and with aspect ratios of up to 3 formed in thermal oxide on silicon wafers. The substrates were placed at 0, 2.2, 5.0, and 8.1 cm away from the central axis of the target. This corresponds to deposition at 0°, 15°, 30°, and 45° respectively, for a target to substrate plane distance of 8.1 cm. After the deposition, the wafers were cleaved through the center of the contacts in two orthogonal directions for SEM cross-section analyses.

For the simulations, 3D angular sputter flux distribution statistics were collected using SIMSPUD for the situations described above. This information was then input to the appropriate 3D version of SIMBAD to simulate the deposition of films over a contact. Figure 6 depicts two of the 3D angular distributions for the W flux from a magnetron source and collected at 0° and 45° positions using SIMSPUD. As expected, the flux collected at the point directly below the target is radially symmetric, showing a ring of flux at constant axial angle determined by the position of the race track on the target. The off-axis position displays a significantly asymmetric flux due the small size of the target. This target was chosen as its small size would emphasize the effect we are trying to model.

Figure 7(a) depicts the SEM cross-section micrograph of a film deposited over contacts placed directly under the center of the target. The poor film coverage on the bottom and side walls is a result of the high aspect ratio. The film coverage is symmetrical (left and right) for both the bottom and the sidewalls, as expected. Figures 7(b) and 7(c), respectively, present the quasi-3D and interpolated 3D profiles of the film obtained from SIMBAD. As the angular distribution of material flux is radially symmetrical directly under the center of the target, the quasi-3D and the interpolated 3D simulations are more or less identical.

Figure 8 shows a three-dimensional view of the surface of the films deposited at an angle of 30° over contacts for two different thicknesses. These two figures show the interpolated 3D surface used in the simulation for the flux shadowing. The interpolated surface can be seen to be smooth and well behaved.

Figures 9(a)–9(c) are the SEM cross-section micrograph of the off-axis films deposited on wafers placed at 2.2, 5.0, and 8.1 cm from the central axis of the target, corresponding to 15°, 30°, and 45°, respectively. The poor step coverage

observed in Fig. 7 is again apparent; however, the nonuniformity of the bottom and side wall coverage of the contacts is also clearly visible. The farther the wafer is from the central axis of the target, the higher the degradation in the bottom and sidewall coverage of the contacts. The nominal film thickness also tapers off as the substrate is moved farther away from the axis of the target. Figures 9(d)–9(f) are the corresponding interpolated 3D computer simulations. The film profiles predicted using this model very closely resemble the experimental ones. This is so for the microstructure as well as the film profiles. (The SEM image appears more dense along the sidewalls, but this is partly an artifact of the finite depth of focus of the SEM. The SIMBAD images are thin slices.)

Figure 10 shows the SEM cross-section micrograph for the perpendicular slice (along the y axis) of the film deposited on wafers placed at 5.0 cm from the target axis (30°) and the corresponding 3D interpolated simulation. Here, the film coverage at the bottom and the side walls are more uniform, indicating a symmetrical flux distribution in this plane. Once again, the 3D interpolated simulation closely predicts the profile and microstructure of the experimental film.

IV. CONCLUSIONS

In this article, we present a simulation package for the efficient prediction of 3D sputtered film growth over VLSI topography. Simulation of the 3D flux transport from the target to the substrate is performed using the Monte Carlo simulator, SIMSPUD. Two models of 3D film growth are investigated. The first assumes cylindrical symmetry of the adatom flux and substrate topography and is useful for predicting film coverage and microstructure in the region of the wafer close to the central axis of the sputter system. The second model is more general and accommodates an asymmetrical sputter flux in order to predict asymmetrical film growth. This second model would be particularly useful for simulation of asymmetrical film coverage over vias and contacts near wafer edges.

Both film growth models are based on the 2D simulation package, SIMBAD. As SIMBAD produces microstructural information as well as film profile evolution, the two models produce microstructural depictions of the simulated film through specific 2D slices. In order to determine the usefulness and accuracy of the two models, W films were deposited over contact cuts in SiO_2 and cross-sectional SEMs were compared to simulation predictions. The models were found to very closely predict both the film profile and the internal microstructure for films directly under the target and also for films located well away from the central axis of the sputter system. These results establish the utility of these simulation techniques for designing and optimizing sputtering systems for future, more demanding integrated circuit metalization.

[1]P. Singer, Semicond. Int. **17**, 57 (1994).
[2]R. Burggraaf, Semicond. Int. 28 (1990).
[3]J. A. Thornton, J. Vac. Sci. Technol. A **4**, 3059 (1986).
[4]A. R. Neureuther and E. W. Scheckler, IEEE Trans. Comput.-Aided Design Integ. Circuits Sys. **13**, 219 (1994).
[5]J. McVittie, J. Rey, L.-Y. Cheng, A. Bariya, S. Ravi, and K. Saraswat, Extended Abstracts, *TECHCON '90* (Semiconductor Research Corporation, San Jose, 1990), p. 16.
[6]T. S. Cale, G. B. Raupp, J. Vac. Sci. Technol. B **8**, 1242 (1990).
[7]D. Liu, S. K. Dew, M. J. Brett, T. Janacek, T. Smy, and W. Tsai, J. Appl. Phys. **74**, 1339 (1993).
[8]L. J. Friedrich, S. K. Dew, M. Brett, and T. Smy, Thin Solid Films **266**, 83 (1995).
[9]S. K. Dew, T. Smy, R. N. Tait, and M. J. Brett, J. Vac. Sci. Technol. A 9, 519 (1991).
[10]S. K. Dew, T. Smy, and M. J. Brett, IEEE Trans. Electron Devices **39**, 1599 (1992).
[11]S. K. Dew, T. Smy, and M. J. Brett, Jpn. J. Appl. Phys. 33, 1140 (1994).
[12]F. H. Baumann, R. Liu, C. B. Case, and W. Y.-C. Lai, in *Proceedings of the International Electron Devices Meeting '93* (IEEE, New York, 1993), p. 861.
[13]F. H. Baumann and G. H. Gilmer, in *Proceedings of the International Electron Devices Meeting '95* (IEEE, New York, 1995), p. 89.
[14]D. S. Bang, J. P. McVittie, K. Ç. Saraswat, Z. Krivokapic, J. A. Iacoponi, and J. Gray, in *Proceedings of the International Electron Devices Meeting '95* (IEEE, New York, 1995), p. 97.
[15]Z. Hsiau, E. C. Kan, J. P. McVittie, R. W. Dutton, in *Proceedings of the International Electron Devices Meeting '95* (IEEE, New York, 1995), p. 101.
[16]S. Tazawa, F. Leon, G. Anderson, T. Abe, K. Saito, A. Yoshii, and D. Scharfetter, in *Proceedings of the International Electron Devices Meeting '92* (IEEE, New York, 1992), p. 173.
[17]H. Yamada, T. Shinmura, Y. Yamada, and T. Ohta, in *Proceedings of the International Electron Devices Meeting '94* (IEEE, New York, 1994), p. 553.
[18]R. N. Tait, S. K. Dew, T. Smy, and M. J. Brett, J. Appl. Phys. **70**, 4295 (1991).

Liner conformality in ionized magnetron sputter metal deposition processes

S. Hamaguchi[a] and S. M. Rossnagel[b]
IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598

(Received 9 November 1995; accepted 19 April 1996)

The conformality of thin metal films (liners) formed on high-aspect-ratio trench structures in ionized magnetron sputter deposition processes is studied numerically and experimentally. The numerical simulator (SHADE) used to predict the surface topography is based on the shock-tracking method for surface evolution. The simulation results are in good agreement with experimentally observed thin-film topography. It is shown that combination of direct deposition and trench-bottom resputtering results in good conformality of step coverages and the amount of the resputtering needed for the good conformality is almost independent of trench aspect ratios. © *1996 American Vacuum Society.*

I. INTRODUCTION

The reliable formation of densely packed interconnect metal lines with low resistance is a challenging task in computer-chip manufacturing. The most advanced technologies currently under development require filling of interconnect lines with the aspect ratio (i.e., ratio of depth to lateral dimension) as high as three. Conventional deposition methods such as chemical vapor deposition (CVD)[1] and collimated magnetron sputter deposition[2–4] may not provide an effective means to fill such high-aspect-ratio trenches and vias in manufacturing environments.

Prior to the metal filling process into dielectric trenches and vias, lining (i.e., deposition of a thin metal layer) is often employed to provide adhesion layers or diffusion barriers for materials to be deposited in the subsequent processes. Such a thin metal film (liner) must be sufficiently conformal, i.e., uniform in thickness on sidewalls and bottoms of trenches (or vias), so that the adhesion or diffusion-barrier characteristics are also uniform on the surface.

Depositing conformal liners into high-aspect-ratio trenches and vias is also a difficult task in low-temperature processes. For metal species to reach the bottom, the metal flux incident upon the wafer surface must be anisotropic. An anisotropic flux, however, does not deposit the material onto vertical sidewalls directly. Weaker anisotropy of the incoming flux, on the other hand, may deposit the material on the sidewalls but tends to deposit more near the top edges of the trench, resulting in nonconformal deposition. This effect is routinely observed with collimated sputtering of Ti and TiN diffusion barriers.[4]

The ionized magnetron sputter deposition process[5,6] can circumvent this dilemma by using its ability to resputter the deposited metal from the bottom of the trench. The ionized magnetron sputter deposition system is schematically depicted in Fig. 1. This is a combination of the conventional magnetron sputter system and an inductively coupled plasma generator. The metal atoms are sputtered from the magnetron cathode and then ionized near the center of the chamber by the high-density Ar plasma generated by the induction coil. The bias voltage on the wafer then controls the directionality and bombardment energies of the metal and Ar ion fluxes incident upon the wafer surface.

If the ion bombardment energy is sufficiently high, the deposited metal on the bottom of the trench may be resputtered and then redeposited on the sidewalls. By controlling the bias voltage on the wafer, one can adjust the amount of resputtering to form a conformal metal liner over the trench structure.

II. EXPERIMENTAL SETUP AND SIMULATION MODEL

As shown in Fig. 1, the experimental system essentially consists of the magnetron and multiple-turn radio frequency induction (RFI) coil. The magnetron power may be typically varied up to 10 kW and 30 kW on 200 mm and 300 mm target cathodes. The RFI coil can be powered up to 3–5 kW at 13.56 MHz. The dc-bias voltage applied to the wafer holder may be varied up to −200 V with substrate currents in the several ampere range. In typical experimental conditions, the Ar pressure is varied from 1 to 35 mTorr. When a large metal ion flux is desirable, the high-pressure (e.g., 35 mTorr) operation is employed. When the pressure is high, the sputtered-atom mean free paths decrease and thus the relative ionization level increases. Since the collision frequency in the sheath region (about 0.1 mm thick) above the wafer holder is still small, the increased pressure does not affect the directionality of ion fluxes.

The trench structures used for the experiments are typically less than 1 μm deep and have various aspect ratios. The trenches are typically made of etched SiO_2 on Si. Metals that can be used for lining in ionized magnetron sputter metal deposition processes include Ti, TiN, Ta, TaN, Cu, Cr, AlCu, AlN, W, and WN.

The numerical simulator used to predict film topography is called SHADE (shock-tracking algorithm for deposition and etching). It uses the shock-tracking method[7] for its surface evolution algorithm and has been applied to various etch and deposition processes.[8–10] For the two-dimensional simula-

[a]Electronic mail: satoshi@watson.ibm.com
[b]Electronic mail: rossn@watson.ibm.com

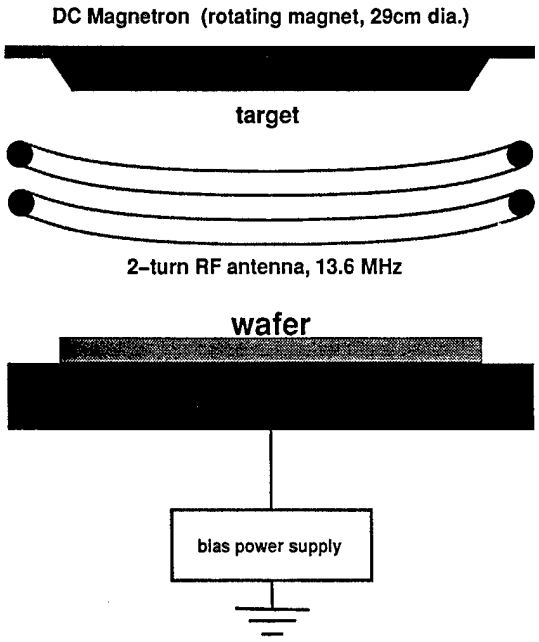


FIG. 1. Schematic diagram of the experimental apparatus.

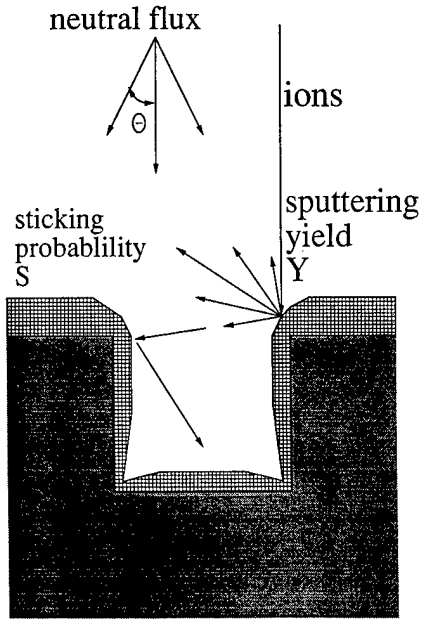


FIG. 2. Schematic diagram of particle transport.

tion, the surface is represented by a sequence of nodes (string) and the fluxes of Ar ions, metal ions, and neutral atoms are calculated at each node. Since the nodes are connected by straight line segments, there are at least two surface normals at each node: surface normals associated with left-side and right-side line segments. The shock-tracking method allows multiple surface normals at each nodal point in addition to these two surface normals. The multiple surface normals may be regarded as an approximation of the corner point (node) by an infinitesimal polygonal surface. Since the etch rate (or sputtering yield) of the material surface depends on the angle formed by the the incident ion flux and surface normal,[11] the multiple surface-normal representation provides an accurate description of surface evolution. Details of the shock-tracking method may be found in Ref. 7.

Figure 2 schematically shows the transport of neutral and ion species near the surface. Since the length scales of the trench are sufficiently smaller than the mean free paths, collisionless transport is a valid assumption for all the gas-phase species. As indicated in Fig. 2, the metal and Ar ion fluxes are assumed to be unidirectional and vertical to the substrate top surface. Since the value of the ion temperature measured in electron volts is sufficiently smaller than the bias voltage, the angular distribution of the ions is insignificant. On the other hand, the neutral metal flux has a large angular distribution. For the sake of simplicity, we assume that the incident neutral metal has an uniform angular distribution with cutoff angle Θ. For a low-pressure operation, where ions sputtered from the target cathode in the magnetron region undergo virtually no collisions, this cutoff angle is determined by the tool geometry, i.e., the ratio of the distance between the target and wafer to the radius of the target.[10] Here we chose $\Theta=52°$ based on our tool geometry[10] for the simulation presented in this article.

When an energetic ion hits the surface, it sputters the surface material with sputtering yield Y. The sputtering yield, which is defined as the number of ions sputtered from the surface per incident ion, depends on the angle θ formed by the incident flux and surface normal. Figure 3 shows the angle dependence of the etch rate [which is proportional to $Y(\theta)\cos\theta$] that is used for the numerical simulation. The rate is normalized to one at $\theta=0$. The film topographies ob-

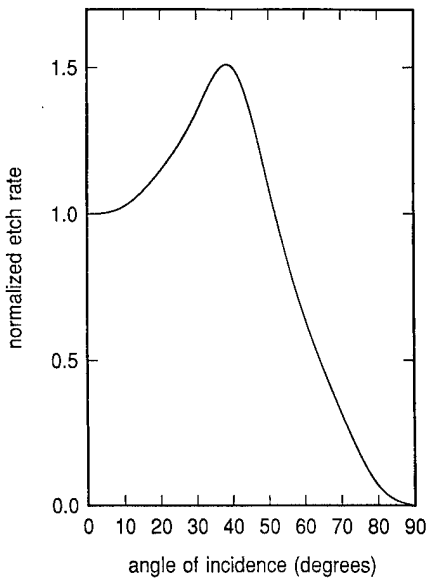


FIG. 3. The dependence of the normalized etch rate as a function of the slope angle (i.e., angle of ion beam incidence) used for the simulation.

tained from the numerical simulation with this etch rate function are in good agreement with experimental observations, as will be shown later.

Metal atoms sputtered by the incident ions are assumed to have sufficiently low energies so that they do not resputter the surface material when then reach other surface sites. Instead, such atoms and the neutral metal atom flux directly coming from the plasma source region are adsorbed on the surface. The probability that an adsorbed atom stays on the surface indefinitely is called the sticking coefficient $\mathscr{S}$. For Cu, $\mathscr{S}=1$ is known to be a good approximation. Those atoms that are reemitted (i.e., desorbed or resputtered) from the surface are assumed to have the cosine angular distribution about the surface normal:

$$f^R = f_\nu(|\mathbf{v}|,\mathbf{x})\cos^{\nu+1}\theta, \tag{1}$$

with $\nu=0$, where f^R is the velocity distribution function of the reemitted atoms at position $\mathbf{x}$ on the boundary curve, and θ is the angle between the velocity vector $\mathbf{v}$ and the surface normal $\hat{\mathbf{n}}$ [Fig. 4(a)]. If $\nu>0$ or $-1<\nu<0$, the function above represents over-cosine or under-cosine distributions, respectively.

The normal component of the incoming flux at position $\mathbf{X}$ on the surface is related to the velocity distribution function $f(\mathbf{v},\mathbf{X})$ as

$$\mathscr{F}(\mathbf{X}) = \int_{\hat{\mathbf{N}}\cdot\mathbf{v}<0} \hat{\mathbf{N}}\cdot\mathbf{v} f(\mathbf{v},\mathbf{X}) d\mathbf{v},$$

where $\hat{\mathbf{N}}$ is the unit vector normal to the surface at $\mathbf{X}$, and the integration is taken over the half velocity space ($\hat{\mathbf{N}}\cdot\mathbf{v}<0$). The outgoing flux is defined in a similar manner with the integration over the other half space $\hat{\mathbf{N}}\cdot\mathbf{v}>0$.

The normal component of the outgoing flux from the surface, i.e., the flux of reemitted metal atoms, is the sum of the desorption flux and resputtering flux:

$$\mathscr{F}^{\text{out}}(\mathbf{X}) = (1-\mathscr{S})\mathscr{F}^{\text{in}}(\mathbf{X}) + \mathscr{F}^{\text{etch}}(\mathbf{X}), \tag{2}$$

where $\mathscr{F}^{\text{in}}(\mathbf{X})$ is the incoming flux of (low-energy) metal atoms and therefore the first term $(1-\mathscr{S})\mathscr{F}^{\text{in}}$ represents the total desorption flux. The second term represents the resputtering flux due to the ion bombardment (etching). The incoming neutral metal flux $\mathscr{F}^{\text{in}}(\mathbf{X})$ may be split into the two components

$$\mathscr{F}^{\text{in}}(\mathbf{X}) = \mathscr{F}^R(\mathbf{X}) + \mathscr{F}^D(\mathbf{X}), \tag{3}$$

where the first term $\mathscr{F}^R$ represents the incoming neutral metal flux that is originated (reemitted) from other sites on the surface and the second term $\mathscr{F}^D$ represents the neutral metal flux arriving directly from the plasma source. Using the reemission distribution function Eq. (1), we may relate the desorption flux $\mathscr{F}^R(\mathbf{X})$ to the outgoing metal atom flux $\mathscr{F}^{\text{out}}(\mathbf{x})$ at every position $\mathbf{x}$ on the surface as

$$\mathscr{F}^R(\mathbf{X}) = \frac{A_{\nu+4}(\nu+2)}{\pi} \int_{\mathscr{B}} ds\, K_\nu(\mathbf{X},\mathbf{x})\mathscr{F}^{\text{out}}(\mathbf{x}), \tag{4}$$

where

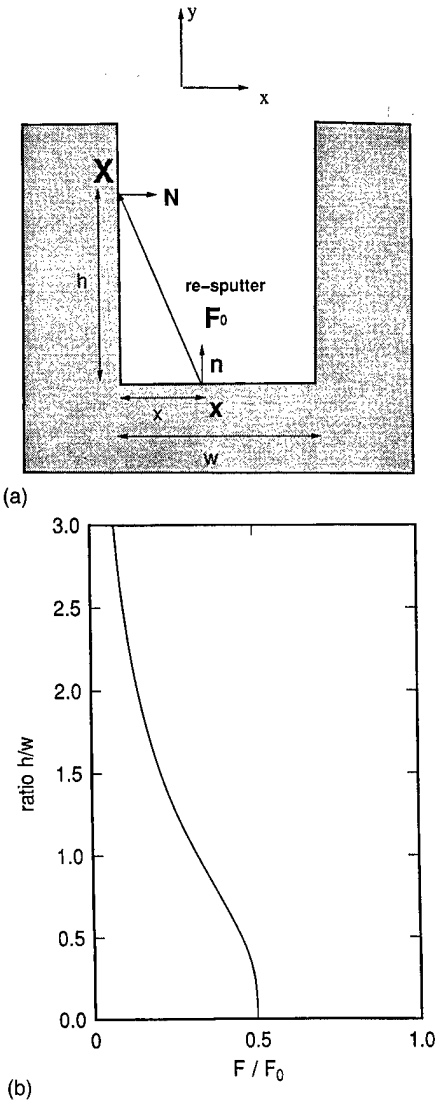


FIG. 4. (a) The coordinate system. (b) The flux ratio $\mathscr{F}/\mathscr{F}_0$ (which is proportional to the film thickness deposited by resputtering) as a function of the normalized height h/w of the trench wall.

$$K_\nu(\mathbf{X},\mathbf{x}) = g(\mathbf{X},\mathbf{x}) \frac{(\hat{\mathbf{R}}\cdot\hat{\mathbf{N}})|\hat{\mathbf{R}}\cdot\hat{\mathbf{n}}|^{\nu+1}}{R} \tag{5}$$

is the integration kernel, the coefficient $A_{\nu+4}$ is given by

$$A_{\nu+4} = \Gamma\left(\frac{\nu+3}{2}\right)\Gamma\left(\frac{1}{2}\right)/2\Gamma\left(\frac{\nu}{2}+2\right),$$

and $\mathscr{B}$ denotes the surface boundary. Note that for the cosine distribution ($\nu=0$), $A_4=\pi/4$. In Eq. (5), the function g is the visibility factor (i.e., $g=1$ if the points $\mathbf{X}$ and $\mathbf{x}$ are on the line of sight, and $g=0$ otherwise), $\mathbf{R}=\mathbf{x}-\mathbf{X}$, $R=|\mathbf{R}|$, $\hat{\mathbf{R}}=\mathbf{R}/R$, and $\hat{\mathbf{n}}$ is the unit vector normal to the surface at $\mathbf{x}$.

Substituting Eqs. (2) and (3) into Eq. (4) results in an integral equation for $\mathscr{F}^R(\mathbf{X})$ for given source terms $\mathscr{F}^{\text{etch}}(\mathbf{X})$ and $\mathscr{F}^D(\mathbf{X})$. More details of the gas-phase transport, including the derivation of Eq. (4) and actual expressions for the source terms $\mathscr{F}^{\text{etch}}(\mathbf{X})$ and $\mathscr{F}^D(\mathbf{X})$, may be found in Ref. 10.

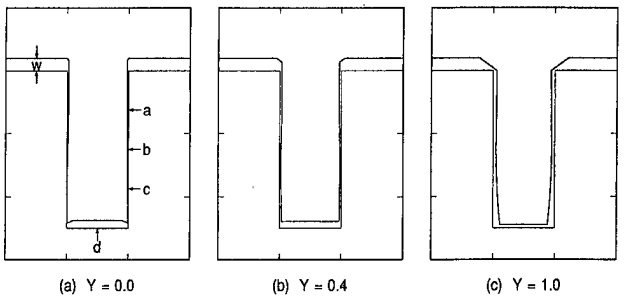


FIG. 5. Profiles of the deposited films obtained from numerical simulations. The aspect ratio of each trench is 2.5. The ratio of the ion to neutral flux on the top surface is 1:1. The sputtering yields at zero incident angle $Y(0)$ are (a) 0, (b) 0.4, and (c) 1.0.

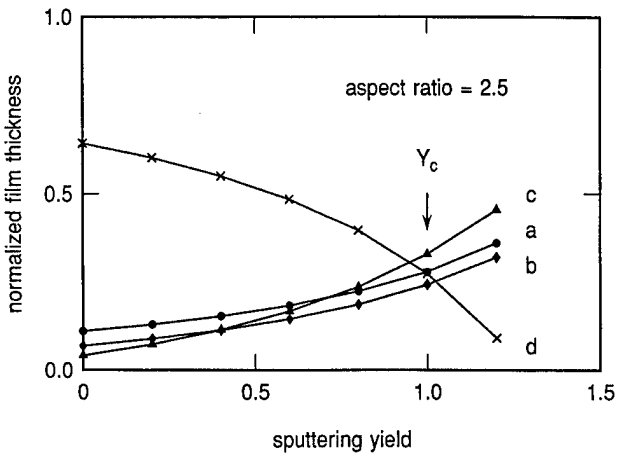


FIG. 6. The ratio of film thickness at each position [a, b, c, and d indicated in Fig. 5(a)] to the top film thickness w as a function of the sputtering yield $Y(0)$ under the conditions used in Fig. 5.

III. NUMERICAL AND EXPERIMENTAL RESULTS

Before proceeding to numerical simulations, we briefly discuss resputtering on trench bottoms. To estimate the amount of resputtering from a trench bottom, we first ignore sidewall metal deposition due to direct neutral fluxes from the plasma source. In this case, as shown in Fig. 4(a), the flux received at height h on the sidewall is only related to the resputtered outgoing flux $\mathcal{F}_0$ from the bottom through Eq. (4). The resputtering flux $\mathcal{F}_0$ is assumed to be uniform over the entire trench bottom. Assuming that the reemission has the cosine distribution [i.e., $\nu=0$ in Eq. (1)], we obtain from Eq. (4),

$$\mathcal{F}=\frac{\mathcal{F}_0}{2}\int_0^w dx\,\frac{(\hat{\mathbf{R}}\cdot\hat{\mathbf{N}})|\hat{\mathbf{R}}\cdot\hat{\mathbf{n}}|}{R}$$

$$=\frac{\mathcal{F}_0}{2}\int_0^w dx\,\frac{hx}{(\sqrt{x^2+h^2})^3} \tag{6}$$

$$=\frac{\mathcal{F}_0}{2}\left(1-\frac{h}{\sqrt{w^2+h^2}}\right),$$

where we take the x axis along the bottom surface, and w represents the width of the trench bottom. Figure 4(b) shows the resputtered flux received on a vertical sidewall as a function of the ratio h/w. Since the deposition rate is proportional to the flux $\mathcal{F}$, Fig. 4(b) also represents the relative thickness of the deposited thin film on the vertical sidewall.

For an infinitely wide trench (i.e., $w=\infty$, aspect ratio is 0), the magnitude of the resputtered flux becomes independent of the height h and is given by $\mathcal{F}=\frac{1}{2}\mathcal{F}_0$. This is not surprising since the limit of $w\to\infty$ with fixed h is equivalent of the limit of $h\to 0$ with fixed w. (Note that $\mathcal{F}$ depends only on the parameter h/w.) As is evident in Fig. 4(b), the flux variation is small for small h/w. This result indicates that the smaller the aspect ratio is, the more uniformly the films are deposited on the sidewalls.

In numerical simulations that we shall present momentarily, the neutral metal flux from the source is also taken into account. Since metal atoms sputtered from the target are not generally 100% ionized by the RFI coil, the effect of the neutral flux cannot be ignored. In typical operations, the ratio of the metal neutral flux to the metal ion flux is about 1:1.[10] As mentioned earlier, the metal neutral flux is assumed to have an isotropic distribution with a cutoff angle 52°. The ion flux is a delta-function beam, perpendicularly incident upon the horizontal substrate surface.

Figure 5 shows profiles of deposited thin metal films (liners) under three different sputtering yields $Y=Y(0)$ at angle 0 (horizontal surface); (a) $Y=0$, (b) 0.4, and (c) 1.0. The aspect ratio of the initial trench is 2.5 and the sticking coefficient $\mathcal{S}$ is assumed to be one. The value Y is actually the *effective* sputtering yield, i.e., the number of sputtered atoms per incident *metal* ion, regardless of whether the sputtering is due to metal ions or Ar ions. If sputtering by the Ar ion flux is negligible, then Y is the Cu-on-Cu sputtering yield.

Figure 6 shows the step coverage (i.e., ratio of the film thickness at a particular position to the flat-surface film thickness w) at four different sites indicated in Fig. 5(a) as functions of the effective sputtering yield Y. The positions a, b, c are at $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$ of the trench depth, and d is at the center of the trench bottom. Since the ion to neutral metal flux ratio is 1:1, the film is deposited on the top surface only when $Y<2.0$.

As may be seen in Fig. 5(c), the sidewall film thickness is not necessarily a monotonic function of the height, unlike the case of Fig. 4(b). At position a, a significant portion of the incoming metal flux arrives directly from the plasma source, rather than the resputtering from the bottom, and thus the film thickness is slightly larger than that at position b. When the sputtering yield Y is high [e.g., Fig. 5(c)], a significant portion of the incoming metal flux at position c is due to the resputtering from the trench bottom, and thus the film thickness is again slightly larger than that at position b. Therefore, the film topography on the sidewall tends to be slightly bowed (thicker near the top and bottom and thinner in the middle) when Y is high. However, for sufficiently thin films, such as those used in liner technology, this variation of the sidewall-film thickness is not significant.

The critical sputtering yield Y_c, which we define as the Y

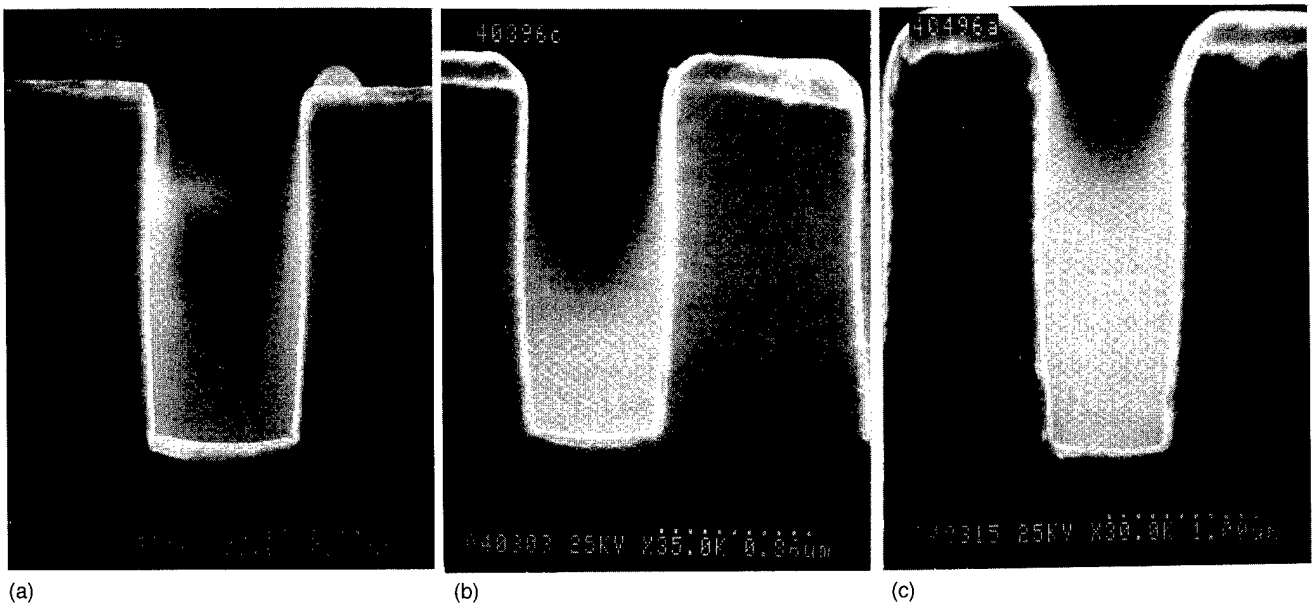


FIG. 7. SEM pictures of Cu deposition under the conditions similar to Fig. 5. The applied bias voltages are (a) 12, (b) 50, and (c) 100 V.

value at which the step coverages at b (middle of the sidewall) and d (middle of the bottom) are equal, gives a good guidance for the film conformality. Under the conditions used in Fig. 5, $Y_c = 1.0$.

Experiments were done with Ti, TiN, AlCu, and Cu over the range of 0.5–4 aspect ratios and good conformal step coverages were obtained in all these cases. Figure 7 shows typical results of such systems under conditions similar to those used in Fig. 5. The bias voltages used in these experiments were 12, 50, and 100 V for Figs. 7(a), 7(b), and 7(c), respectively. Cu films are shown in Fig. 7 since the higher brightness of Cu in scanning electron microscopy (SEM) photographs most clearly delineates the film topography.

Figure 8 shows simulation results of liner deposition on trenches of aspect ratio 1. The conditions used here are the same as those used in Fig. 5. Figure 9 shows the step coverage (i.e., normalized film thickness) at four different positions indicated in Fig. 8(a) as functions of the sputtering yield Y. The positions a, b, c are at 1/4, 1/2, and 3/4, respectively, of the trench depth, and d is at the center of the trench bottom, as before. It is seen in Fig. 9 that the critical sputtering yield for conformality is again $Y_c \simeq 1$. Evidently, Y_c depends weakly on the aspect ratio A for the range $1 \leqslant A \leqslant 2.5$ studied here. Thus the ionized magnetron sputter metal deposition system with an appropriate sputtering yield Y can form conformal metal step coverages over trenches with a wide range of aspect ratios in a single process.

IV. CONCLUDING REMARKS

Numerical simulations together with experimental results show that conformal deposition of metal liners can be achieved by simultaneous deposition and resputtering in ionized magnetron sputter metal deposition processes. The ionized magnetron sputter deposition process has two significant advantages over the conventional magnetron sputter deposition process. One is the directionality (or anisotropy) of the

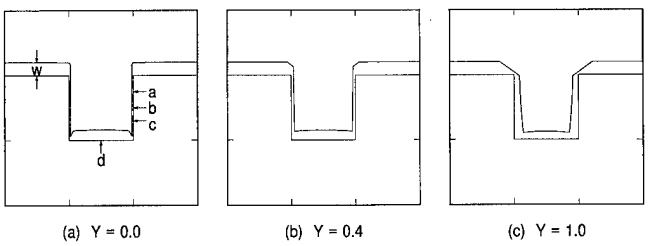


FIG. 8. Profiles of the deposited films obtained from numerical simulations. The aspect ratio of each trench is 1. Other conditions are the same as those used in Fig. 5. The sputtering yields at zero incident angle $Y(0)$ are (a) 0, (b) 0.4, and (c) 1.0.

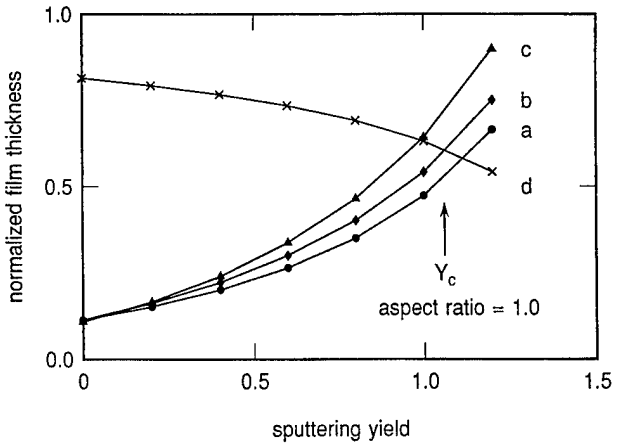


FIG. 9. The ratio of film thickness at each position [a, b, c, and d indicated in Fig. 8(a)] to the top film thickness w as a function of the sputtering yield $Y(0)$ under the conditions used in Fig. 8.

metal deposition flux due to the presence of highly ionized gas-phase metal and the applied bias voltage on the substrate surface. With high ionization of metal in the plasma region, a large metal ion flux with high directionality may be extracted from the source to the substrate without damaging the substrate. The other advantage is its ability to control the ion flux energy incident upon the substrate surface. The amount of resputtering from the substrate surface can be easily controlled by the bias voltage.

It is also shown that, by selecting an appropriate bias voltage to control surface resputtering, we are able to form conformal step coverages over various trenches with different aspect ratios in a single process. In this method, metal atoms resputtered from the bottom of the trench are effectively redeposited on trench sidewalls. As the bias voltage (i.e., sputtering yield on the sample surface) is increased, the film thickness on the sidewalls increases, whereas the film thickness on the trench bottom decreases. The appropriate bias voltage (i.e., critical sputtering yield) is the one that gives similar film thicknesses on sidewalls and trench bottom. Since the critical sputtering yield Y_c is almost independent of trench aspect ratios, one can conformally line trenches with a wide variety of aspect ratios in a single process.

[1]A. V. Gelatos, A. Jain, R. Marsh, and C. J. Mogab, MRS Bull. **19**, 49 (1994).
[2]S. M. Rossnagel, D. Mikalsea, H. Kinoshita, and J. J. Cuomo, J. Vac. Sci. Technol. A **9**, 261 (1991).
[3]S. K. Dew, D. Liu, M. J. Brett, and T. Smy, J. Vac. Sci. Technol. B **11**, 1281 (1993).
[4]R. Joshi and S. Brodsky, Proceedings of the 1992 VLSI Multilevel Interconnect Conference, June 1992 (unpublished), p. 235.
[5]S. M. Rossnagel and J. Hopwood, Appl. Phys. Lett. **63**, 3285 (1993).
[6]S. M. Rossnagel and J. Hopwood, J. Vac. Sci. Technol. B **12**, 449 (1994).
[7]S. Hamaguchi, M. Dalvie, R. T. Farouki, and S. Sethuraman, J. Appl. Phys. **74**, 5172 (1993).
[8]S. Hamaguchi and M. Dalvie, J. Vac. Sci. Technol. A **12**, 2745 (1994).
[9]S. Hamaguchi and M. Dalvie, J. Electrochem. Soc. **141**, 1964 (1994).
[10]S. Hamaguchi and S. M. Rossnagel, J. Vac. Sci. Technol. B **13**, 183 (1995).
[11]P. Zalm, Surf. Interface Anal. **11**, 1 (1988).

High-vacuum versus "environmental" electron beam deposition

Albert Folch,[a] Jordi Servat, Joan Esteve, and Javier Tejada
Facultat de Física, Universitat de Barcelona, 08028 Barcelona, Spain

Miquel Seco
Facultat de Química, Universitat de Barcelona, 08028 Barcelona, Spain

(Received 17 October 1995; accepted 11 April 1996)

Electron beam deposition (EBD) provides an inexpensive way to fabricate nanostructures of various materials in a scanning electron microscope (SEM). However, the purity of metals deposited from an organometallic precursor gas is impaired by simultaneously deposited carbon coming both from the organometallic molecule and the residual contamination gas in the SEM chamber. We discuss carbon-contamination EBD in a standard high-vacuum SEM and compare it to EBD of Au in an environmental SEM (ESEM). The ESEM allowed us to perform "environmental" EBD (E-EBD), i.e., EBD in the presence of an environmental gas (1–10 Torr) in addition to the organometallic precursor gas. We built a simple device that contains a reservoir for the organometallic precursor and goes on the sample stage of the ESEM. With this device we were able to highlight the advantages of E-EBD over conventional, high-vacuum EBD. We discuss the basic chemical reactions underlying the E-EBD process. © *1996 American Vacuum Society.*

I. INTRODUCTION

Conventional deposition techniques currently used in the fabrication of micron or submicron structures (e.g., microelectronics devices) do not allow the deposition of materials on a very small area of the substrate. Instead, materials are deposited over the whole substrate and the pattern is defined afterwards by removing material from the adjacent areas. The creation of a given pattern involves several tedious steps (such as masks, resists, photo- or electron-beam lithography, lift-off, and/or selective etching) which result in costly processes and generally require planar substrates. The cost is usually justified by the large number of devices produced in parallel. However, alternative techniques are sought for many applications where a few devices suffice. Also, it is particularly important for microsensor engineering to develop patterning techniques that are compatible with nonplanar substrates.

Electron beam deposition (EBD) is an inexpensive technique suitable for the fabrication of structures of different sizes, shapes, and materials in the submicron or nanometer scale. In the presence of a surrounding gas, a finely focused electron (e^-) beam is observed to cause the deposition of material from the gas only in the area irradiated by the beam. In EBD, materials are patterned and deposited *simultaneously*. This one-step process is comparatively inexpensive. Since it is maskless and resistless, it has the advantage that it allows patterning on nonplanar substrates. On the other hand, since it is induced locally, it is time-consuming. Another disadvantage concerns the purity of the deposits, as will be discussed later.

The observation of EBD is not new. The formation of carbonaceous films on surfaces upon electron bombardment in vacuum haunted researchers for more than 30 years.[1] Ennos[2] identified specimen contamination in the electron mixcroscope produced by the interaction between the bombarding e^- and the organic molecules adsorbed on the specimen surface. Broers *et al.*[3] were the first to deposit nanometerscale carbon-contamination structures in a scanning transmission electron microscope (STEM). We describe our studies on carbon-contamination EBD in Sec. II. By allowing a small pressure of an organometallic vapor in the chamber of a scanning electron microscope (SEM), Matsui and Mori[4] deposited 0.15-μm-wide metallic lines. The principle of EBD is sketched in Fig. 1. The dissociation of some or all of the gas molecules that happen to adsorb on the irradiated spot results in the formation of a deposit consisting of the nonvolatile remains from that dissociation.[5–8] The deposition rate depends on the total e^--beam current density at the beam spot, on the effective adsorption rate of the precursor gas molecules onto the substrate, and on the probability for an electron to cause EBD of a molecule. In addition, the beam current density depends not only on the beam spot size but also on the rate of emission of secondary and backscattered electrons, which also contribute to the process. Recently, EBD has been demonstrated with a scanning tunneling microscope (STM) operated in field-emission mode.[9] For our work we used SEMs because of their flexibility and availability. We developed an instrumentally simple device to perform EBD without needing to modify the SEM chamber for introduction of the precursor gas. We call this device the "EBD cell" (Sec. III). We used it to demonstrate "environmental" EBD (E-EBD). This approach has several advantages over conventional EBD (Sec. IV). The growth process is analyzed in Sec. V.

[a]Author to whom all correspondence should be addressed. Present address: Microsystems Technology Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts; Electronic mail: afolch@mit.edu

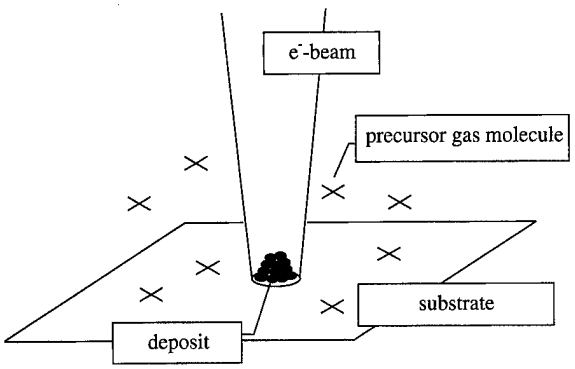


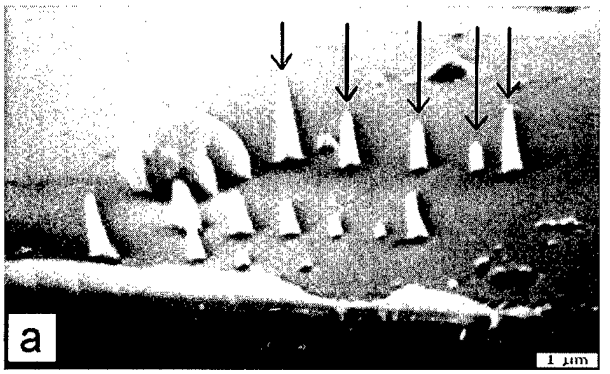
FIG. 1. Principle of EBD. A focused e^- beam induces a chemical change on the precursor gas molecules that are adsorbed on the irradiated spot.

II. CARBON-CONTAMINATION EBD

One may straightforwardly carry out an EBD experiment by means of a standard SEM operated at low beam currents (~10–50 pA) and high beam energies (~15–30 keV). These parameters are generally not critical. A flat sample with a thin film of sputtered gold to avoid charging will generally suffice. A rough sample will make the observation of the submicron-sized deposits difficult. A requirement is that the e^--beam be well focused. Stopping the e^- beam on a given spot for a few minutes will typically cause the growth of a submicron-diameter column or tip. The growth rate varies greatly from one SEM to another, probably due to different residual gas pressures and/or composition. If the growth rate is very slow, beam instabilities and/or drifts will prevent the formation of a fine column.

We deposited columns and lines on a thin film of Au (~1000 Å) on Si with an SEM operating at a base pressure of 5×10^{-5} Torr. A microfabricated trench in the Si substrate served as a landmark for easy location of the deposits. A set of tips is shown in Fig. 2(a). A nearby dust speck helped to finely focus the e^- beam. The tips were grown for periods of time varying from 30 s to 10 min, and with beam currents and acceleration voltages ranging 10–30 pA and 15–40 kV, respectively. The apparent conical shape can be attributed to instabilities of the e^- beam that effectively spread the e^- beam spot, such as noise in the deflection electrodes, mechanical vibrations, and thermal drifts. The capability to write lines is demonstrated in Fig. 2(b).

An interesting observation is that these contaminants can be present even at lower pressures, as in a scanning Auger microanalysis (SAM) system operating at a base pressure of 10^{-9}–10^{-10} Torr. On an indium-tin-oxide (ITO) sample previously cleaned by Ar^+ sputtering (0.5 min, 20 μA/cm^2, 2 keV) we exposed a 75 μm × 50 μm area by raster scanning the e^- beam of the SAM for 16 hours. After this period of time, the exposed area appeared blackened in an electron micrograph and its Auger analysis revealed the presence of carbon only. The black appearance can be explained by the particularly low secondary electron emission rate of carbon. The thickness of the deposit has to be at least several nanometers in order to screen the Auger emission of the indium, tin,



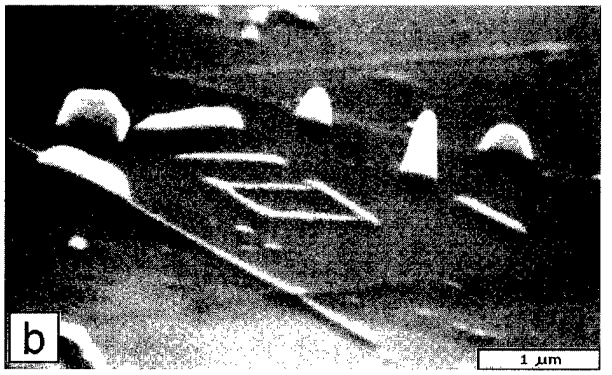


FIG. 2. Illustrative examples of carbon-contamination EBD. The white scale bar represents 1 μm. To deposit a tip, the beam is focused on the desired spot for a certain amount of time. (a) A set of tips is grown near a dust speck. The series of 5 tips in a row indicated by arrows was grown for periods of time of (from left to right) 10, 5, 2.5, 1, and 5 min. Beam current and energy were 30 pA and 40 keV, respectively. (b) The four 1-μm-long lines forming the square were deposited in 4 min each with a single line scan (4 min/μm) at 8 pA and 15 keV. The two broader, taller lines were deposited under the same beam conditions but 6 times slower (24 min/μm).

and oxygen in the underlying ITO substrate. Therefore, the amount of carbon present on any given sample is overestimated because the analyzing e^- beam simultaneously causes deposition of carbon. This problem poses a challenge to the characterization of very small deposits. Furthermore, the composition of a spot or column deposited with a stationary e^- beam cannot be analyzed with an e^- beam of similar resolution, since the analysis is comparatively slow and is severely affected by drifts. Usually, since the efficiency of Auger electron excitation is very low, high beam currents (~1 nA) are used in Auger analysis as compared to EBD (~10 pA), which results in bigger beam spot sizes. A partial solution to this problem is to analyze large deposits made by rastering the e^- beam across a large area (~100 μm^2).

III. ORGANOMETALLIC EBD: THE "EBD CELL"

The nature of the precursor gas determines the composition of the deposits. Metals, for instance, can be deposited from organometallic precursor gases.[4–8] Unfortunately, the purity of the deposits is affected by undesired deposition of carbon coming not only from the residual contamination

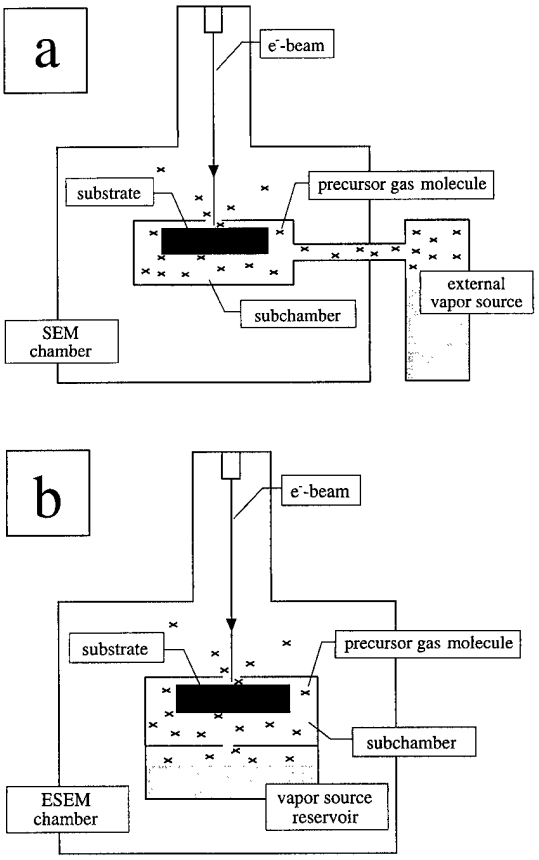


FIG. 3. Schematic of (a) the usual subchamber setup for EBD, and (b) our "EBD cell" setup. The environmental gas, not depicted for clarity, fills up the whole ESEM chamber.

gases, as discussed in Sec. II, but also from the organometallic precursor compound. Unfortunately also, since high vacuum is required for operation of the SEM, the precursor gas must be allowed inside the SEM chamber in small quantities. This results in low growth rates. Generally, the gas is introduced from an external source into a subchamber (containing the sample) within the SEM chamber [see Fig. 3(a)].[4–6] Some authors used a gas nozzle directed onto the sample.[7] The SEM chamber must be modified to accommodate a gas feedthrough, which may represent a serious inconvenience.

It is important to realize that high precursor gas pressures are desirable not only to obtain high growth rates but also to minimize carbon-contamination incorporation. At very low precursor gas pressures carbon deposition dominates and results in grossly impure deposits. Therefore, there is an essential incompatibility between the high vacuum requirements of conventional SEMs and the desired purity of the deposits.

To avoid hardware modifications to the SEM, we made a small device that acts as an *internal* vapor source and goes on the SEM sample stage. We call this device an "EBD cell" [see Fig. 3(b)]. The EBD cell is a 3.5-cm-diameter two-chamber assembly: one subchamber for the sample and a reservoir for the precursor material. Three screws hold the two chambers together and a sealing Viton O-ring in between. The vapors of the precursor material leave the reservoir through a fixed small aperture (which we call the "reservoir aperture"), enter the sample subchamber and escape through another fixed aperture (the "sample aperture") that exposes the sample to the e^- beam. Thus some gas molecules adsorb on the sample spot that is being irradiated by the beam. The apertures were TEM diaphragms (with hole diameters ranging from 5 to 500 μm) glued with silver paint on 1-mm-diameter holes machined on top of each chamber. In practice, the choice of aperture sizes is limited by (a) the desired growth rate, (b) the sample area accessible by the beam, (c) the amount of precursor material given its vaporization rate, and (d) the vacuum requirements of the SEM. Since the apertures are removable, this is straightforwardly determined by trial and error. Variable apertures, pressure gauges and an electrode gauge could be implemented for a precise control and *in situ* variation of the growth rates.

To eliminate the vacuum constraints we used an environmental SEM (ESEM).[6] With the ESEM it is possible to image in the presence of up to 20 Torr of certain gases. The detection of secondary electrons for imaging is possible due to the highly ionized state of the environmental gas above the beam spot. We used Ar, H_2O 2O, and a mixture of 80% Ar and 20% O_2 (Ar/O_2) as environmental gases.

The performance of the cell was first optimized by experimenting with EBD of Fe from iron pentacarbonyl, $Fe(CO)_5$, a liquid with a vapor pressure of 3 Torr at 20 °C.[11] Then we deposited Au at room temperature from $Au(CH_3)_2$(hexafluoroacetylacetonate), a liquid with a vapor pressure of 700 mTorr at 20 °C (hereafter referred to as DMGHFAC). We emphasize that both the organometallic precursor, emanating from the reservoir, and the environmental gas (Ar,H_2O, or Ar/O_2), filling up the whole ESEM chamber, were present during deposition.

In all the experiments reported here, the EBD cell was used in combination with the ESEM, but it has been used successfully in high vacuum in a standard SEM as well.[11] A few droplets of DMGHFAC last ~30 min in the reservoir at environmental pressures of ~1–3 Torr if the reservoir aperture and the sample aperture are 50 μm and 1 mm in diameter, respectively. This yields growth rates of around 40 nm/s.[8] A key feature of the EBD cell performance is that at our environmental gas pressures (~1 Torr) the precursor gas molecules leave the subchamber by diffusion, rather than ballistically as in previous subchamber setups in vacuum. This yields low evaporation rates of the precursor compound while allowing a large sample area exposed to the e^- beam.

IV. ENVIRONMENTAL EBD OF HIGH-ASPECT-RATIO NANOSTRUCTURES

The concept of environmental EBD (E-EBD), i.e., EBD in the presence of an environmental gas in addition to the precursor gas, can be advantageous for several applications. This process was recently applied to reduce carbon incorporation in EBD of Au.[8] The environmental gas was either Ar, H_2O, or Ar/O_2. The presence of Ar/O_2 or H_2O reduced the C content of the deposits (as measured with Auger electron

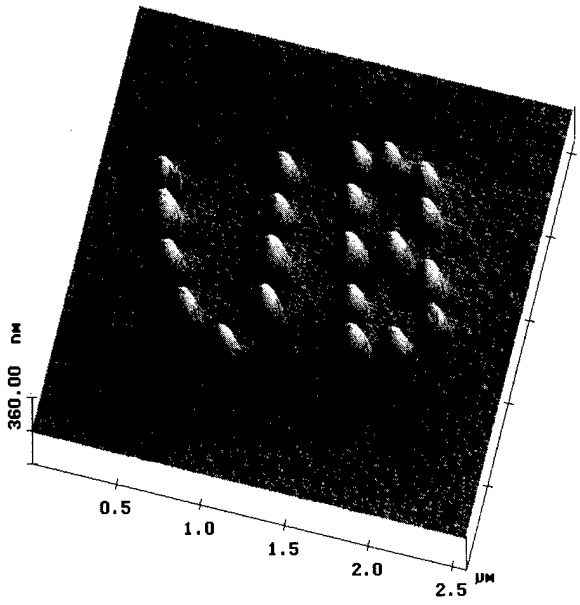


FIG. 4. AFM image of EBD columns deposited on Si from DMGHFAC vapor at 30 keV. The columns appear convoluted with the AFM tip, as manifested by identical sidewall features on each column.

spectroscopy) whereas Ar did not. Moreover, the Au content increased with increasing H_2O or Ar/O_2 pressure, whereas it was significantly affected by changes in Ar pressure. The highest Au contents (50%) were achieved under 10 Torr of Ar/O_2. With H_2O vapor, the overall contents were smaller (up to 20% under 3 Torr). Very low Au contents ($\sim$2–3 %) were systematically observed in the absence of environmental gas. This indicated a large amount of hydrocarbon residual gases in the ESEM chamber and highlighted the accomplishment of 50% Au composition by means of E-EBD.

The concept of E-EBD can also be applied to create high-aspect-ratio nanostructures. Since the ESEM is able to operate at high environmental gas pressures (0–20 Torr), it is inherently compatible with higher precursor gas pressures than standard high-vacuum SEMs. This can yield very high growth rates and aspect ratios. We achieved linewidths of $\sim$100 nm despite the high gas pressures. An example is shown in Fig. 4. A pattern of columns reading "UB" was deposited in 5 s/column from DMGHFAC on Si under 2 Torr of Ar/O_2 and with a 30 keV beam. The columns have an average height of 180 nm, which yields a growth rate of 36 nm/s. The widths ($\sim$100 nm) were measured with an SEM since in the atomic force microscope (AFM) image the columns appear wider than they really are due to convolution with the AFM tip. Since every column has the same rough features on its right sidewall, one can conclude that it is a feature present on the AFM tip itself. This suggests that arrays of EBD columns could be used as calibration tests to know the exact shape of an AFM tip.

The same procedures needed to create column structures such as those in Fig. 4 are applicable when the substrate is not flat. We were able to deposit >4-μm-high, 0.1-μm-thick columns from DMGHFAC on top of commercial microfabricated AFM tips. Indeed, the unique capability of EBD for patterning nonplanar substrates caught the attention of STM and AFM researchers very early because sharp, high-aspect-ratio tips are critical in several applications.[12] Two examples are shown in Fig. 5. The deposition parameters were the same as in Fig. 4. The tips are not perfectly straight because the deposition of each tip was done with several short exposures ($\sim$5 s) to allow re-focusing of the beam between exposures. This was done to avoid broadening of the column. Tip heights are (a) 4 μm and (b) 1.5 μm.

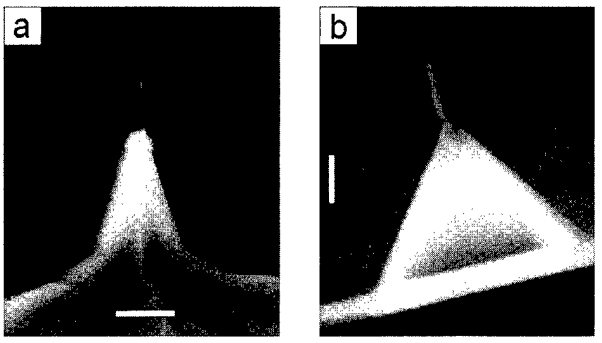


FIG. 5. EBD columns deposited on top of pyramidal AFM tips. The white scale bar represents (a) 5 μm and (b) 1 μm.

Last but not least, with E-EBD it is possible to fabricate nanostructures on bulk insulators, such as the silicon nitride pyramid in Fig. 5(b), because in the ESEM the sample is continuously discharged through the ionized environmental gas.

V. DISCUSSION

It is important to note that the ESEM itself is not essential to the concept of EEBD. Similarly, we believe that the addition of a reactive gas to reduce carbon incorporation during EBD may as well be implemented in high vacuum. One could, for example, direct two nozzles to a sample in a high-vacuum SEM, one transporting the organometallic compound and another one transporting pure O_2. The composition of the deposits is probably determined by the ratio of the flows or partial pressures of the two gases. Furthermore, we expected electron scattering at our high environmental gas pressures to broaden the EBD features. However, we observed no dependence between environmental gas pressures $<$5 Torr and the width of the EBD features, which was routinely 100 nm or less (see Figs. 4 and 5). This is in accordance with the ESEM's surprisingly high resolution. Danilatos[13] has suggested that the beam is self-focused by a concentric plasma sheath, and its associated electrostatic field, around the beam. As noted above, in any case, lower pressures of similar mixtures of gases could be used in a high-vacuum EBD setup to deposit features at a slower rate and higher resolution. We stress that the concept of E-EBD is applicable to STM-EBD as well.

The composition of the deposits is virtually independent of the e^- beam energy for the values typically used in SEM imaging ($\sim$10–40 keV). This observation rules out the pos-

sibility that substrate heating by the impinging electrons plays a significant role in the EBD mechanism, as the heating effect increases with increasing energy. Essentially, EBD of a gaseous molecule is the rupture of one or more of the chemical bonds within the molecule by one or more high-energy electrons ($>$ 1 keV), resulting in a solid deposit. Typical bonds in an organic molecule such as C-C, C-H, or C-O have energies of 3.25, 3.43, and 4.31 eV, respectively. Hence a high-energy electron cannot be used to selectively rupture the chemical bonds present in an organic molecule. The dynamic role of the continuously changing surface has not been investigated yet to our knowledge.

We explain our results on E-EBD by noting that highly reactive species are generated by the dissociation of the precursor gas molecules adsorbed on the substrate as well as by the ionization of environmental gas molecules above the surface. Thus, simultaneously with deposition, C in the deposits can be desorbed as CO or CO_2 molecules if an environmental gas supplies the oxygen ions or molecules necessary for that reaction to occur. Ar/O_2 is more effective than H_2O at doing so due to its richer presence of O at the compared pressures. That is, a total pressure of 10 Torr of Ar/O_2 (80%/20%) represents a partial pressure of 2 Torr of pure O_2. Thus the density of O atoms in 10 Torr of Ar/O_2 is twice as high as that in 2 Torr of H_2O, and accordingly it yields superior Au contents. Pure O_2, which could not be used for safety reasons,[14] should yield higher Au contents at similar environmental gas pressures. Finally, in our opinion the inert nature of Ar satisfactorily explains that Ar does not create a reactive environment.

A key observation is that in our experiments on organometallic EBD with DMGHFAC there is no trace of F or O in the Auger analysis of the deposits for standard e^- beam current and acceleration voltage deposition parameters (see Ref. 8). Since high-energy electrons break all bonds unpreferentially, the above observation suggests that the organometallic molecule is fully dissociated into its atomic constituents by the e^- beam. Rather than two or more inert molecules, the outcome of the dissociation is a set of highly reactive ions, probably monatomic, the recombination of which produces the observable deposit along with volatile compounds. The creation of F_2 and O_2 molecules, for example, satisfactorily explains the absence of F and O in the Auger analysis of the deposits. In general, any reducing species present in the environmental plasma (such as H^+ from H_2O) will also contribute to their desorption. In the case of Ar, it is probably a small partial pressure of other residual gases (such as H_2O) that provides the necessary reducing agents. The Au:C:O:F:H atomic ratios in the DMGHFAC molecule are 1:7:2:6:8. Although H cannot be detected with Auger spectroscopy, the above arguments should also hold true for H. For this reason, we expect the deposit to be composed of only Au and C in a Au:C ratio of 1:7 (12.5% Au) *at least*. Therefore, in EBD of Au with DMGHFAC, Au contents smaller than 12.5% must be attributed to simultaneous carbon-contamination deposition from residual hydrocarbon gases in the SEM chamber.

Probably, in high vacuum the organometallic compound does not fully dissociate under certain low-electron-density conditions. Indeed, assuming full dissociation the most likely recombination favors the creation of two desorbing CO molecules, other recombinations being far less likely. Accordingly, it should contain up to 16.7% Au (Au:C=1:5) *at most*. However, higher Au contents have been reported by others.[5,6] This suggests that in high vacuum the molecule is not fully dissociated and that bigger species are desorbed. We could not investigate ranges of pressure $<$0.1 Torr with the ESEM because of its inaccuracy in pressure control. We explain the higher degree of dissociation observed in E-EBD as compared to conventional high-vacuum EBD by noting that in E-EBD the precursor gas molecules adsorbed on the e^- beam spot are bombarded by electrons as well as by reactive ions.

VI. CONCLUSIONS

We have described an innovative setup for EBD that requires no modifications to the SEM chamber. It consists of a small subchamber with a built-in reservoir for the precursor compound, thus eliminating the need for a feedthrough for introduction of the precursor gas from outside the SEM chamber. With this device we had previously introduced the concept of environmental EBD (E-EBD). The composition of Au deposits can be substantially improved by a judicious choice of the environmental gas and the growth rates can be much higher than in high-vacuum EBD, which yields high-aspect-ratio nanostructures. Here we discussed its advantages and analyzed the beam-induced reaction mechanisms of E-EBD as compared to high-vacuum EBD.

ACKNOWLEDGMENTS

The authors thank the Ministerio de Educación y Ciencia (Spain) and Direcció General de la Recerca (Catalonia, Spain) for partial support of this research. The EBD cell was machined by D. Binagui and J. Santana. A commercial SAM (operated by C. H. Peters) and ESEM at the Massachusetts Institute of Technology was used. The authors acknowledge fruitful comments by M. S. Wrighton, A. J. Garratt-Reed, M. Frongillo (M.I.T.), R. Fontarnau (Serveis Científico-Tècnics, U.B.), E. Lora-Tamayo (Centre Nacional de Microelectrònica, Bellaterra), and M. Salmeron (Lawrence Berkeley Laboratory). The authors are indebted to S. Cantin (Instrumat, S.A.) for loaning a Nanoscope III (Digital Instruments) to the U.B., and to J. Melngailis (M.I.T.) for providing the gold compound.

[1]R. L. Stewart, Phys. Rev. **45**, 488 (1934); J. H. L. Watson, J. Appl. Phys. **18**, 153 (1947); J. Hillier, J. Appl. Phys. **19**, 227 (1948); K. M. Poole, Proc. Phys. Soc. London B **66**, 542 (1953).
[2]A. E. Ennos, Brit. J. Appl. Phys. **4**, 101 (1953); **5**, 27 (1954).
[3]A. N. Broers, W. W. Molzen, J. J. Cuomo, and N. D. Wittels, Appl. Phys. Lett. **29**, 596 (1976).
[4]S. Matsui and K. Mori, Jpn. J. Appl. Phys. **23**, L706 (1984).
[5]H. W. P. Koops, R. Weiel, D. P. Kern, and T. H. Baum, J. Vac. Sci. Technol. B **6**, 477 (1988).
[6]K. L. Lee and M. Hatzakis, J. Vac. Sci. Technol. B **7**, 1941 (1989).
[7]K. T. Kohlman-von Platten, J. Chlebek, M. Weiss, K. Reimer, H. Oertel, and W. H. Brünger, J. Vac. Sci. Technol. B **11**, 2219 (1993).

[8]A. Folch, J. Tejada, C. H. Peters, and M. S. Wrighton, Appl. Phys. Lett. **66**, 2080 (1995).
[9]R. M. Silver, E. E. Ehrichs, and A. L. de Lozanne, Appl. Phys. Lett. **51**, 247 (1987); M. A. McCord, D. P. Kern, and T. H. P. Chang, J. Vac. Sci. Technol. B **6**, 1877 (1988).
[10]G. D. Danilatos, Micron Microsc. Acta **14**, 307 (1983).
[11]A. Folch, Ph.D. thesis, Universitat de Barcelona, Spain, 1994.
[12]Y. Akama, E. Nishimura, A. Sakai, and H. Murakami, J. Vac. Sci. Technol. A **8**, 429 (1990); D. Keller, D. Deputy, A. Alduino, and K. Luo, Ultramicroscopy **42–44**, 1481 (1992).
[13]G. D. Danilatos, Adv. Electron. Electron. Phys. **71**, 109 (1988).
[14]The hydrocarbon oil of our ESEMs mechanical pump could ignite in contact with pure O_2. A perfluorinated polyether oil should be used instead to pump pure O_2.

Simulations of metal thin film thermal flow processes*

Hung Liao and Timothy S. Cale[a)]
Department of Chemical, Bio and Materials Engineering, Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6006

(Received 22 December 1995; accepted 19 April 1996)

We use EVFLOW, a thin film thermal flow process simulator, to simulate the thermal flow of metal thin films in axisymmetric contacts/vias on patterned wafers, where two dimensional surface evolution is appropriate. The flow processes considered are (1) Al and Au thin film laser melting processes, for which process temperatures are higher than the metal melting points, and (2) conventional Al thermal anneal processes (400–550 °C). Surface tension induced mass transport in the flowing films is calculated by solving the Navier–Stokes and continuity equations for incompressible fluids. Our simulations of Au film profiles during laser melting process predict the experimentally observed trends in void formation and collapse with feature geometry and deposited film thickness. In laser melting processes, the void shrinks and vanishes, resulting in fully filled contacts and planarized metal surfaces. Al film evolution in the thermal anneal process is modeled using a surface layer melting model. The thickness of the liquidlike flowing layer is assumed to be 0.02 μm in this work. The trends in the formation of voids with feature geometry and deposited film thickness predicted by EVFLOW reflect experimental observations. Voids may form during the thermal anneal processes. The surface layer melting model as used in this work will not predict the collapse of voids. © *1996 American Vacuum Society.*

I. INTRODUCTION

Fabrication of multilevel interconnect structures used in modern integrated circuits is becoming increasingly challenging as the device density and the number of interconnect layers increase with decreasing device dimensions. Tungsten chemical vapor deposition (CVD) plug processes remove deposited film conformality problems in contacts and vias which are present in conventional sputter deposited (PVD) aluminum, and are widely used in production.[1,2] However, tungsten CVD processes are more complex, more costly, and tend to have more particulate issues than PVD aluminum processes.[3,4] In addition, it is preferable to form interconnects and plugs simultaneously.

Metal reflow technologies, such as Al thermal flow after low temperature sputter deposition, has been investigated as promising methods for forming plugs in submicron contacts and achieving planarized surfaces.[5–9] Heating the entire wafer above the melting point of Al (its alloys, in practice) (~660 °C), either in a conventional furnace or by rapid thermal annealing, is undesirable because of potential interactions of Al with other materials. Among the methods being studied in order to minimize the thermal budget for the reflow are (1) heating with pulsed lasers ($\sim 10^{-6}$ s) and (2) annealing at temperatures between 400 and 550 °C. Pulsed lasers allow flowing metal films locally and have been extensively investigated.[10–13] So far, excimer lasers such as XeCl excimer lasers have been used in commercial systems.[14] Laser reflow of a variety of Al alloys (Al–1% Si, Al–1% Si–0.5% Cu, and Al–1% Si–0.1% Ti) using an excimer laser showed excellent planarization for all topologies, including submicron contacts.[15] Mukai, Sasaki, and Nakamo[16] detected the presence of undesired voids in contacts with aspect ratios greater than 1, beneath fully planarized Al surfaces following excimer laser irradiation. Bernhardt *et al.*[17] and Marella, Tuckerman, and Pease[18] also found voids in Au coated trench structures following excimer laser irradiation. Those voids could be eliminated using higher fluence laser irradiation. Park *et al.*[7] reported that submicron contacts with high aspect ratio (>1) were completely filled using conventional thermal annealing under high vacuum conditions. Electrical tests with fully processed complementary metal oxide semiconductor (CMOS) devices showed no degradation of contact resistance or junction leakage current and improved electromigration resistance. For higher aspect ratio contacts, the final flowed films showed planarized metal surfaces with voids inside.[7]

This article reports on our computer simulations of laser melting and conventional annealing of metal films in axisymmetric contacts. The initial deposited metal film and substrate profiles either represent experimental SEMs or from deposition simulations using EVOLVE,[19] which has shown quantitative agreement with experimental data in previously reported work [for example see Refs. 20 and 21]. Metal films are first cold sputter deposited (either experimentally or virtually) in contacts and then the simulator EVFLOW[22] is used to simulate the flow process. We first discuss our simulation results for Al film evolution during laser melting, focusing on void deformation; the simulated as-deposited Al film has a large void inside. The effect of contact aspect ratio on metal film evolution during laser melting processes is demonstrated by simulating Au film evolution in two contacts, with aspect ratios of about 1 and about 0.5. Digitized as-deposited Au film and substrate profiles are input to EVFLOW, in addition to the required fluid properties. We then discuss our simulation results for Al thermal annealing, which we

*Published without author corrections.

a)Electronic mail: cale@enws165.eas.asu.ed

model using a "surface layer melting model" which is similar to the one presented by Hirose, Kikuta, and Yoshida.[23] The influences of initial feature geometry and as-deposited film thickness on void formation during thermal annealing of uncollimated and collimated deposited Al films are then discussed. A limitation of the surface layer melting model is highlighted.

II. SURFACE DIFFUSION AND SURFACE LAYER MELTING MODELS

The key to understanding thermal annealing under the melting point of a metal is that not all atoms are equivalent in nature; those present at the surface are energy rich due to having fewer neighbors than those in the bulk. Annealing of metals has been described using two kinds of models: curvature driven surface diffusion models and surface melting models. In the surface diffusion models, surface atoms diffuse along the surface due to gradients in surface curvature. Cale *et al.*[20] developed a method to estimate surface diffusivity as a function of temperature using film profiles as deposited in commercial sputter deposition equipment. The comparisons between EVOLVE predictions are in good agreement with experimental profiles.

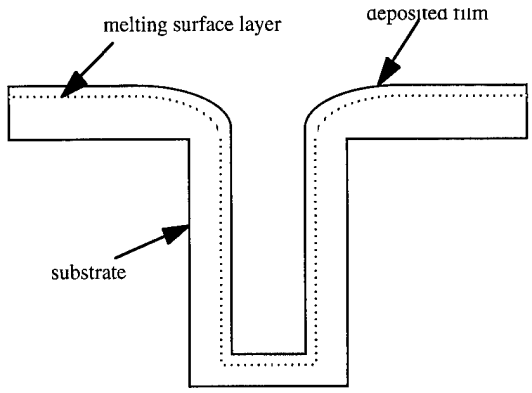


FIG. 1. The principle of the surface melting model for the Al thermal anneal process simulations.

Surface melting models for Al anneal processes are based on experimental and simulation studies of surface disordering of crystal Al surfaces. Stoltze, Norskov, and Landman[24] used molecular dynamics to simulate the behavior of Al(110) and Al(111) surfaces as they approach the melting point. They found a definite disordering of the outermost layers of Al(110) up to 200 °C below the bulk melting point, and the

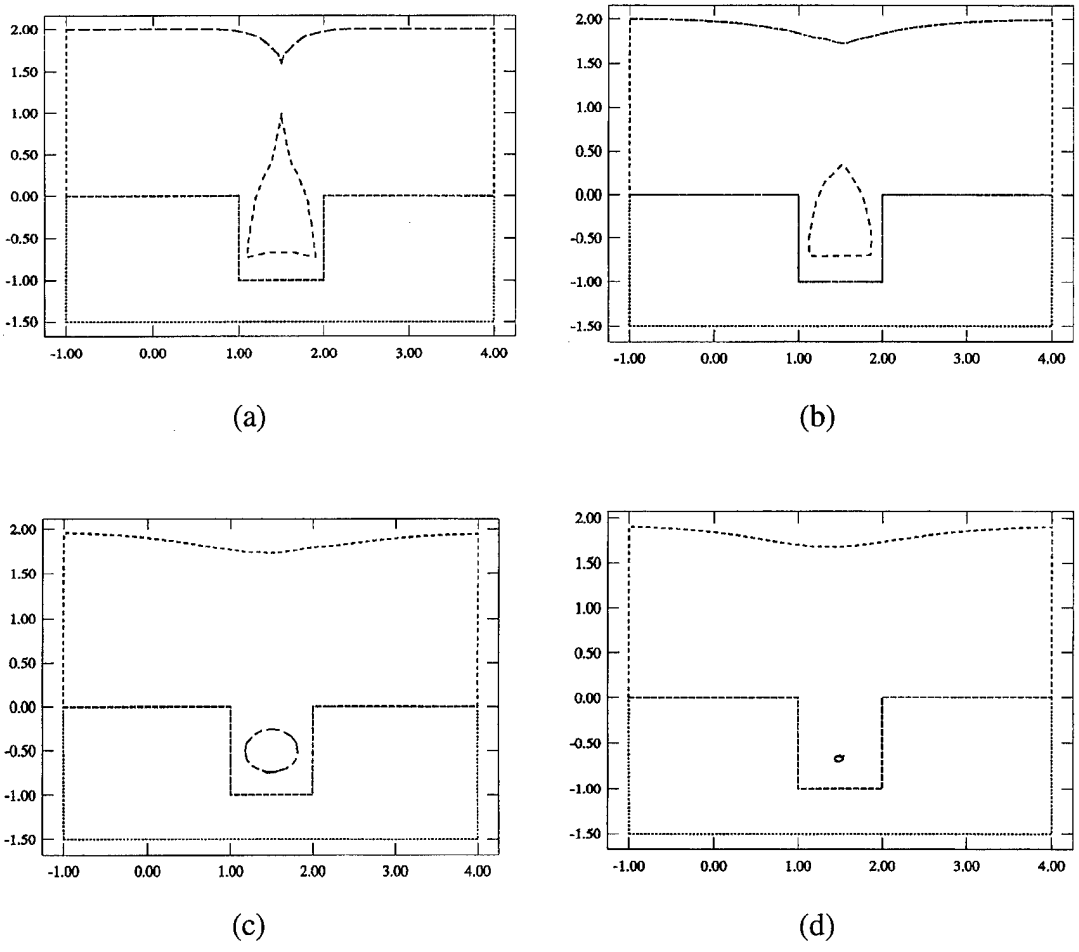


FIG. 2. (a) 2 μm as-deposited Al film in a 1×1 μm contact. (b)–(d) Al film profiles at selected times during a laser melting process.

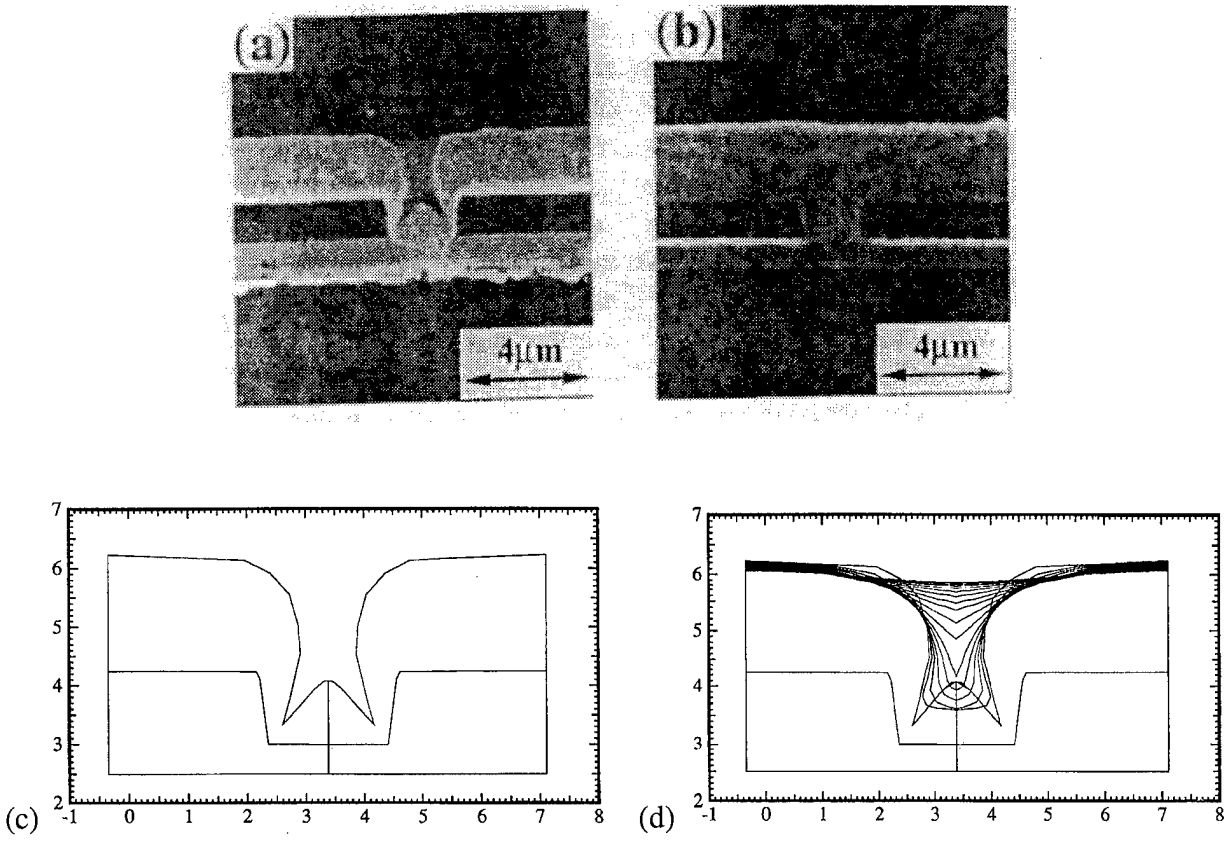


FIG. 3. (a) Sputter deposited Au film in a contact with aspect ratio about 0.5. (b) Au film profile after irradiation with a 3.5 J/cm^2, 600 ns laser pulse. (c) Simulated Au topography during the laser melting process.

close-packed (111) surfaces showed much weaker disordering than the (110) surfaces. Denier *et al.*[25] experimentally studied the thermal disordering of Al(110) and Al(111) surfaces. Medium energy ion scattering (MEIS) measurements indicated a melting induced disordering of the Al(110) surfaces, but not on the (111) surfaces at temperature about 423 °C. The number of disordered monolayers grows as the melting point is approached. Molecular dynamics simulations, scattering experiments, and transport measurements have shown that the diffusion coefficients of these mobile layers are very large and are comparable to the mobility of a bulk liquid at its melting point.[26,27] Hirose, Kikuta, and Yoshida[23] proposed that the mass transport of Al occurs only in a thin surface layer during conventional thermal anneal processes. Mass transport inside the melting layer was calculated using a fluid dynamics model. They estimated the thicknesses of the flow layers by matching the anneal time at each temperature; the thickness diverges as the melting point is approached. Figure 1 illustrates the principle of the surface melting models. The surface melting model is used in this work to simulate Al film evolution in conventional thermal anneal processes.

III. EVFLOW

EVFLOW is described in detail by Liao and Cale;[28] it consists of EVOLVE, a physically based low pressure deposition/etch process simulator and FLOW, a thin film thermal flow process simulation program. The conservation law based surface movement routines in EVOLVE are used to update the location of the free surface and void profiles according to the local velocity vectors. The flow process is modeled as an incompressible, surface tension driven, free-boundary problem, and is an extension of the model presented by Bornside *et al.*[29] EVFLOW is capable of simulating thin film thermal flow processes in infinitely long trenches with arbitrary vertical cross sections and cylindrical contacts/vias; i.e., situations for which feature surfaces can be represented in two dimensions. The full dimensionality of the transport is taken into account.

In EVFLOW, the flow process is divided into a number of time steps, and the location of the free surface and void is determined as a function of time; i.e., the pseudo-steady-state approach is used. Within each simulation time step, EVFLOW generates a mesh in the flowing film, and solves the steady state version of the dimensionless Navier–Stokes equations for incompressible fluids using a penalty function finite element method.[30] The penalty function method is usually considerably more economical in terms of computer memory and execution speed than the general mixed order interpolation formulation.[30] A difficulty with the use of this approach when applied to the surface layer melting model is discussed in Sec. V.

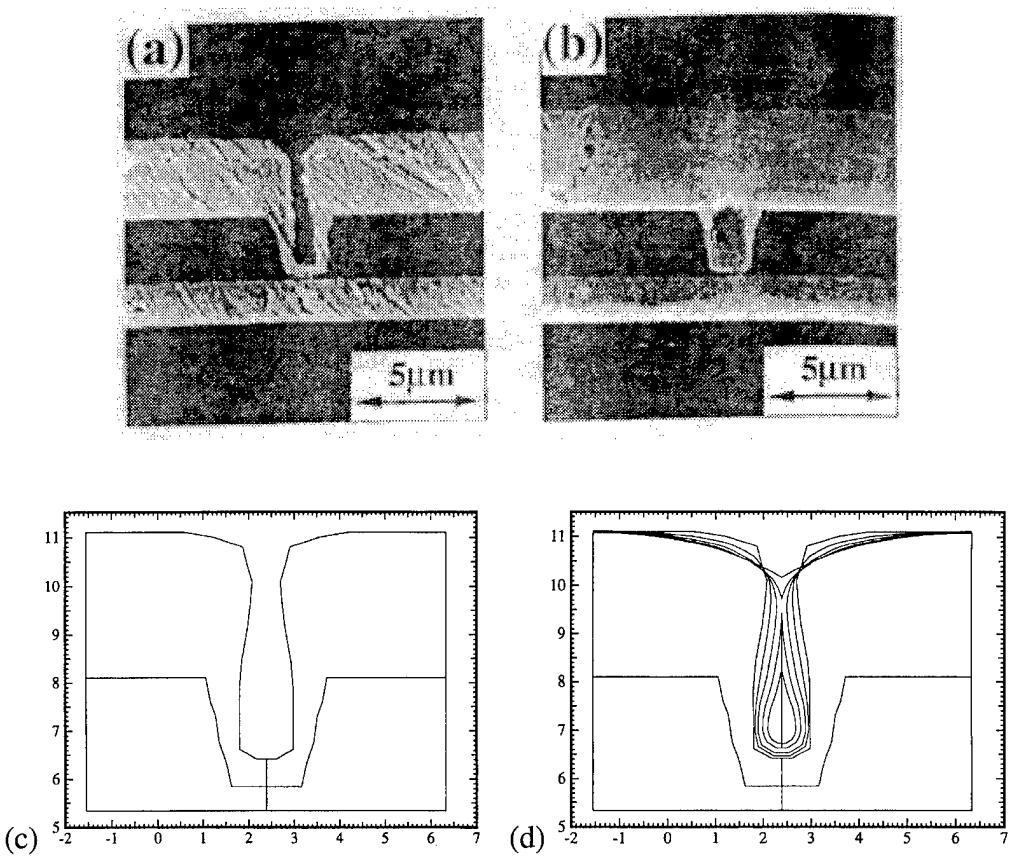


FIG. 4. (a) Sputter deposited Au film in a contact with aspect ratio about 1. (b) Au film profile after irradiation with a 3.8 J/cm^2, 600 ns laser pulse. (c) Digitized substrate and as-deposited Au film profiles. (d) Simulated Au topography during the laser melting process.

IV. EXPERIMENTS AND SIMULATIONS OF LASER MELTING

In order to illustrate film evolution during laser melting processes, simulations begin with a simulated sputter deposited 2 μm Al film in a 1×1 μm idealized contact. Figure 2(a) shows the Al film predicted by EVOLVE assuming that the Al sticking factor is 1 during the sputter deposition. There is a large void inside the film as would be expected. The values of surface tension and viscosity of Al used in the flow simulation are 914 dyne/cm and 0.013 P, respectively.[31] The film evolution calculated by EVFLOW is illustrated in Fig. 2(b), 2(c), and 2(d). The void shrinks and finally vanishes during the laser melting process, which results in totally filled contact and planarized topography. In other words, the laser melting process can lead to 100% filled contact and planarized topography if enough laser fluence is provided. The experimental results of Bernhardt *et al.*[17] and Marella, Tuckerman, and Pease[18] support this conclusion; i.e., no voids were found in high aspect ratio samples if high enough laser fluences (and hence temperature) were used to reflow the metal. The metal viscosity decreases exponentially with increasing temperature and the surface tension decreases linearly,[31] which results in shorter characteristic film leveling times at high temperatures.[28] It should be noted that in this simulation and those discussed in the following, the simulation domain contains a large volume of metal above the vias. As the via fills, the decrease in metal height above the flat surface is small, which is supported by experimental results.

Marella, Tuckerman, and Pease[18] studied via/contact hole filling with Au, Cu, and Al–1% Cu films using a flashlamp-pumped dye laser (600 ns pulse duration and λ=490 nm) and an excimer laser (35 ns and λ=308 nm). Test structures containing either metal-to-metal vias or metal-to-silicon contacts were prepared, starting with silicon wafers as substrates. All metal films were formed by sputter deposition, while all SiO_2 layers were deposited by a low pressure CVD process. Vias and contacts were exposed and dry etched into the SiO_2. Thin (100–300 Å) $Ti_{0.3}W_{0.7}$ films were sputter deposited to provide adhesion between the SiO_2 and the metal films. The thickness of the metal and insulator layers and the diameter of the vias/contacts were varied in order to produce a number of different as-deposited metal profiles. Metal plugs before and after laser irradiation were inspected by SEM.

The laser melting cases simulated were for Au films, for which Marella, Tuckerman, and Pease[18] provided the most experimental information; e.g., film profiles before and after melting. Experimentally, the smaller the aspect ratio of the contact, the easier it is to fill. Figures 3(a) and 3(b) show experimental Au film profiles in a contact with an aspect ratio of about 0.5, before and after irradiation with a 3.5 J/cm^2, 600 ns laser pulse. The profile evolution predicted by

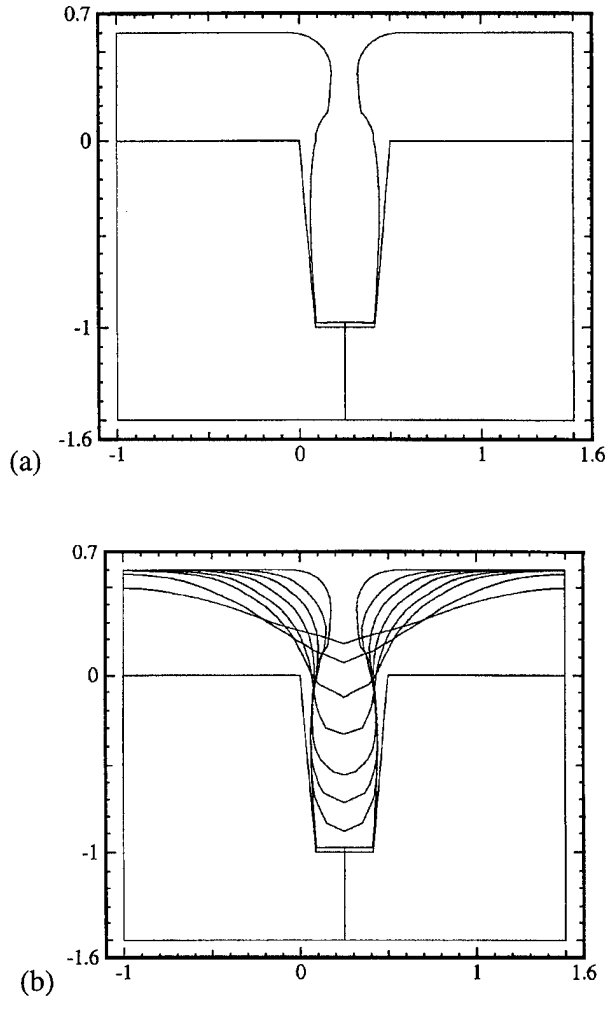


FIG. 5. (a) EVOLVE simulated 0.6 μm sputter deposited Al film in a 0.5×1 μm contact with 15° tapered sidewall. (b) EVFLOW prediction of the Al profiles during the anneal process.

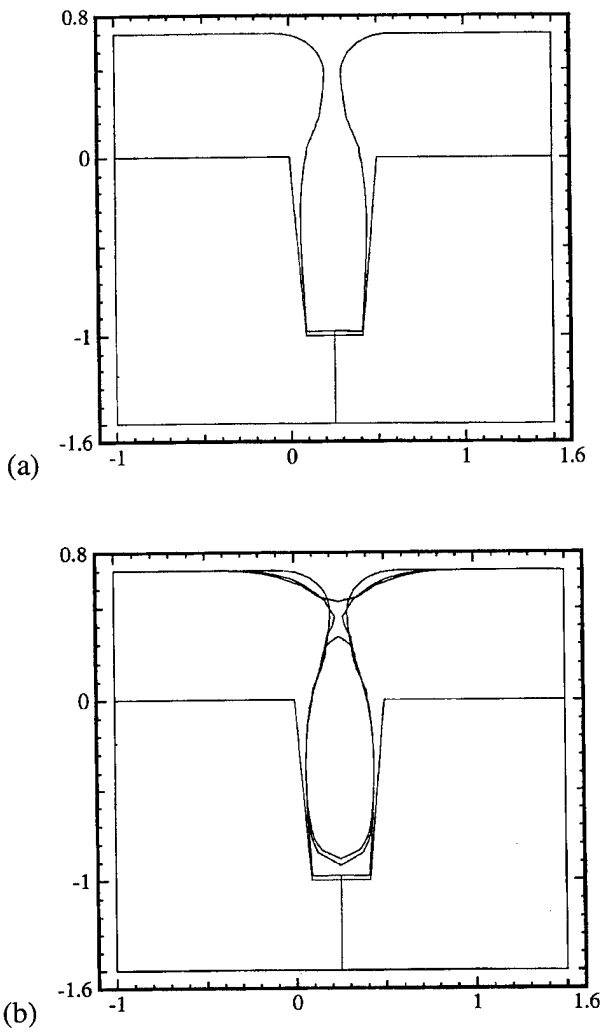


FIG. 6. (a) EVOLVE simulated 0.7 μm sputter deposited Al film in a 0.5×1 μm contact with 15° tapered sidewall. (b) EVFLOW prediction of the Al profiles during the anneal process.

EVFLOW, using substrate and as-deposited Au profiles as shown in Fig. 3(c) as input, are shown Fig. 3(d). Because EVVLOW is limited to axisymmetric features, the average of the left and right hand sides of the Au film in the SEM is used to represent the deposited film profile. The viscosity and surface tension of Au at the melting point are used in the simulations and are 0.05 P and 1140 dyne/cm.[31] The processing time in EVFLOW is set to 600 ns; i.e., the time-temperature trajectory is assumed to be an ideal step up to the melting point and a step down 600 ns later. The contact is filled and the surface is planarized.

Figure 4(a) shows an as-deposited Au film in a tapered contact with an aspect ratio about 1. The final film profiles after irradiation with a 600 ns, 3.8 J/cm^2 laser pulse is shown in Fig. 4(b). The film and substrate profiles shown in Fig. 4(c) are input to EVFLOW. The values of Au viscosity and surface tension used are the same as in the above example. Figure 4(d) shows the EVFLOW calculated time evolution of the Au film topography, which illustrates that the void is formed at the beginning of the process, and collapses afterward. The surface is becoming more planarized with time. The exposure time or intensity of the laser pulse is not large enough to eliminate the void. By increasing the flow time somewhat, the void will disappear. The predicted void is smaller than that of the experiment and does not have the same shape. Nevertheless, we feel that the simulation results support the model used. In order to predict film profile evolution more accurately, better estimates of the time–temperature trajectories and fluid properties are required. At this time, it is satisfying to predict trends correctly.

The mechanism of void formation during the laser melting process can be explained using the concept of surface tension driven mass transport, which is proportional to the local curvature. For vias, consider the curvatures in the plane of the paper (as seen in cross sectional views) and the curvatures in the "horizontal" plane, which are given by the local hole radii of the opening (which is also seen in cross sectional views). For wider contacts (smaller aspect ratios), the three-dimensional (3D) curvatures at the top edges are smaller than those for narrower contacts, for a given film thickness. At the via mouth, the curvature in the plane of the

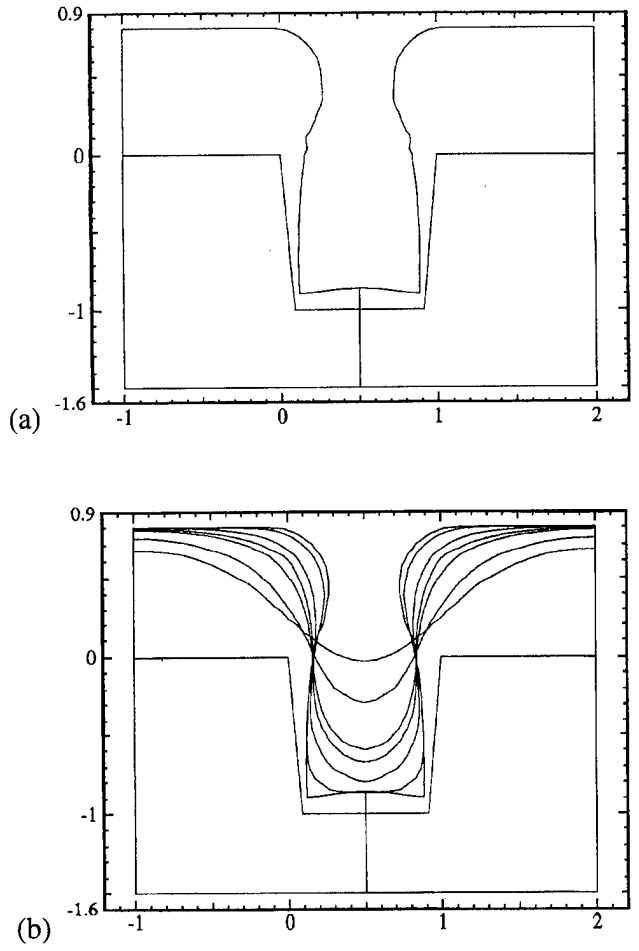


FIG. 7. (a) EVOLVE simulated 0.8 μm sputter deposited Al film in a 1×1 μm contact with 15° tapered sidewall. (b) EVFLOW prediction of the Al profiles during the anneal process.

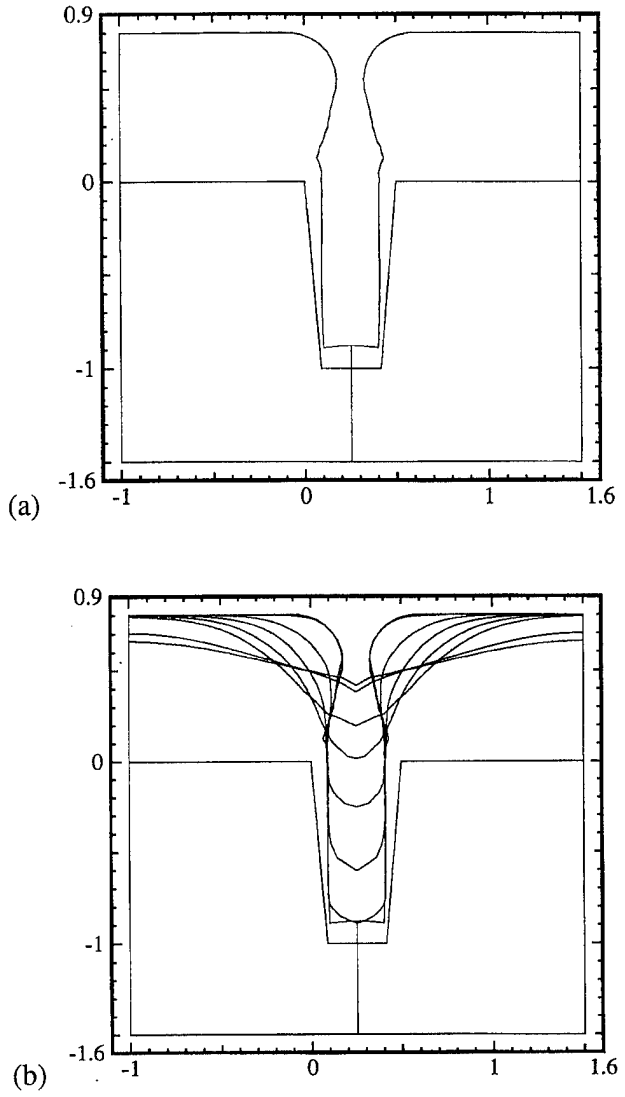


FIG. 8. (a) EVOLVE simulated 0.8 μm collimated sputter deposited Al film in a 0.5×1 μm contact with 15° tapered sidewall. (b) EVFLOW prediction of the Al profiles during the anneal process.

paper drives the film into the via. The curvature in the horizontal plane drives the film to pinch off and form a void. Thus, the films are more likely to pinch off before filling the feature in higher aspect ratio features, because the curvature in the horizontal plane is larger (the films are closer to closure). Also, the amount of movement needed to close the mouth is lower. Once formed, the void collapses for the duration of the melt process.

V. SIMULATIONS OF THERMAL ANNEALING

EVFLOW is used to predict the evolution of Al film profiles during thermal anneal processes, using cold sputter deposited Al film profiles predicted by EVOLVE as input. The thickness of the melted surface layers is assumed to be 0.02 μm. The actual thickness of the melting surface layer may be much thinner than 0.02 μm (Refs. 23–25); however, we chose 0.02 μm because of limitations of computer memory and execution speed. The values for surface tension and viscosity used in the simulations are 1077 dyn/cm and 420 P, as used by Hirose, Kikuta, and Yoshida.[23]

Figure 5(a) shows a simulated sputter deposited 0.6 μm Al film in a 0.5×1 μm contact with a 15° tapered sidewall. The contact is filled without a void after a simulated thermal anneal process [Fig. 5(b)]. If the deposited film thickness is increased to 0.7 μm, however, the contact is found to have a void left in it, as shown in Figs. 6(a) and 6(b). This can be explained, as for the case of laser melting, by (1) the higher surface curvature and (2) the film is closer to closure, for the 0.7 μm case relative to the 0.6 μm film. Once the void is formed, it will tend to become spherical. Because the surface melting model does not include transport in the bulk of the films, the void volumes will stay the same volume while becoming spherical. In order to predict void shrinkage, the model would have to include transport through the bulk of the film (see Ref. 32 for one approach).

Based on the above analysis, the key to preventing void formation is to prevent closure of the feature mouth by keeping a large enough opening. This implies that contacts with smaller aspect ratios and thinner as-deposited films are easier to fill than those with larger aspect ratios and thicker films. As shown in Figs. 7(a) and 7(b) a simulated 0.8 μm Al film

can completely fill a 1×1 μm contact after reflow. One way to decrease the likelihood of void formation for a given deposited film thickness and feature size is to use collimated sputter deposition. Figures 8(a) and 8(b) show that a simulated 0.8 μm collimated sputter deposited Al film fills a 0.5×1 μm contact after thermal annealing. The opening at the top of the feature mouth is larger than that of conventional sputter depositions at a given film thickness. Thus, a larger film thickness can be deposited without void formation during reflow.

These trends were obtained by Hirose, Kikuta, and Yoshida[23] in analyzing film topography during Al thermal anneal processes; however, the processing times to closure are significantly different. By adjusting the melted layer thickness to match the experimental simulation times of 30–180 s,[7] Hirose, Kikuta, and Yoshida[23] estimated it to be about 0.005 μm. However, the processing time in the simulation results presented above are of the order of 10^{-2} s. This is due to the larger melted layer thickness used in this work.

A limitation of the penalty function formulation became obvious during the simulations of thermal anneal processes; the penalty factor cannot be 0. The continuity equation modified for the penalty function formulation in axisymmetric contacts is

$$\epsilon p+\left(\frac{1}{r}\frac{\partial}{\partial r}(ru)+\frac{\partial v}{\partial z}\right)=0, \tag{1}$$

in which p is pressure, ϵ is the penalty factor (typically $10^{-9}\leqslant\epsilon\leqslant10^{-4}$), and u and v are velocities in the r and z directions, respectively. Consider a void inside a contact with a thin melted surface layer. The steady state of the void is a sphere, and the volume inside the sphere should be the same as the initial void volume (mass balance). At the steady state, velocities (u,v) should be zero with the surface tension induced forces balanced by pressure. However, the modified continuity equation in the penalty formulation leads to zero pressure if u and v are identically zero, since the penalty factor cannot be 0. Using the penalty function formulation, the void will continue to shrink in size. Because of this limitation, EVFLOW is only used to simulate the evolution of the free surfaces during the Al thermal anneal processes after void formation. This situation does not exist in the laser melting processes, in which the films totally melt and the steady state is a fully planarized surface and a completely filled via.

VI. CONCLUSIONS

Trends in reflow behavior and void formation during laser melting and thermal anneal processes calculated by EVFLOW agree with experimental trends. Decreasing the curvatures of the deposited films at the feature mouths reduces the possibility of void formation during both the laser melting and thermal anneal processes. The methods of reducing curvature include reducing deposited film thickness and using collimated sputter deposition techniques. The voids formed in laser melting processes can finally disappear, if enough laser intensity or pulse duration is provided. Due to the limitations of the surface melting model used to simulate thermal annealing, voids will not disappear. The model needs to be extended to include a route to void closure in order to match experimentally observed film and void profile evolution.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support for this work from the Semiconductor Research Corporation and the National Science Foundation.

[1]D. M. Brown, B. Gorowitz, P. A. Piacente, R. J. Saita, R. H. Wilson, and D. W. Woodrcoff, Tech. Dig. IEDM 66 (1986).
[2]C. Kaana, W. Cote, J. Cronin, K. Holland, P-I. Lee and T. Wright, Tech. Dig. IEDM 209 (1987).
[3]K. Sugain, T. Shinzawz, S. Kishida, H. Okabayashi, and Y. Murao, in *Proceedings of the International VLSI Multilevel Interconnection Conference*, 1992, p. 463.
[4]R. Liu, K. P. Cheung, W. Y-C. Lai, and R. Heim, in *Proceedings of the International VLSI Multilevel Interconnection Conference*, 1992, p. 329.
[5]M. Taguchi, K. Kogama, and Y. Sugano, in *Proceedings of the International VLSI Multilevel Interconnection Conference*, 1992, p. 219.
[6]H. H. Hoang, F. S. Chen, M. Zamanian, G. A. Dixie, C. C. Wei, and F. T. Liou, in *Proceedings of the International IEEE VLSI Multilevel Interconnection Conference*, 1991, p. 414.
[7]C. S. Park, S. I. Lee, J. H. Park, J. H. Sohu, D. Chin, and J. G. Lee, in *Proceedings of the International IEEE VLSI Multilevel Interconnection Conference* (IEEE, New York, 1991), p. 326.
[8]H. Nishimura, T. Yamada, and S. Ogawa, in *Proceedings of the International IEEE VLSI Multilevel Interconnection Conference*, 1991, p. 170.
[9]H. Ono, Y. Ushika, and T. Yota, in *Proceedings of the International IEEE VLSI Multilevel Interconnection Conference*, 1990, p. 76.
[10]D. B. Tuckerman and R. L. Schmitt, in *Proceedings of the International IEEE VLSI Multilevel Interconnection Conference* (IEEE, New York, 1985), p. 24.
[11]D. B. Tuckerman and A. H. Weisberg, IEEE Electron Device Lett. **EDL-7**, 1 (1986).
[12]R. Mukai, N. Sasaki, and M. Nakamo, IEEE Electron Device Lett. **EDL-8**, 76 (1987).
[13]W. Spiess and H. Stack, J. Vac. Sci. Tech. B **7**, 127 (1989).
[14]E. Ong, H. Chu, and S. Chen, Solid State Technol. **34**, 63 (1991).
[15]B. Woratchek, P. Carey, M. Stolz, and F. Bachman, Appl. Surf. Sci. **43**, 264 (1989).
[16]R. Mukai, N. Sasaki, and M. Nakamo, IEEE Electron Device Lett. **EDL-8**, 76 (1987).
[17]A. F. Bernhardt, R. J. Contolini, D. B. Tuckerman, and A. H. Weisberg, in *Laser and Particle Beam Chemical Processes on Surfaces*, edited by A. W. John, G. L. Loper, and T. W. Sigmon (Materials Research Society, Pittsburgh, 1989), p. 559.
[18]P. F. Marella, D. B. Tuckerman, and R. F. Pease, Appl. Phys. Lett. **56**, 2625 (1990).
[19]EVOLVE is a low pressure deposition process simulator developed by T. S. Cale at Auburn State University and Motorola, Inc. with funding from the Semiconductor Research Corporation and the National Science Foundation.
[20]T. S. Cale, M. K. Jain, C. J. Tracy, and R. L. Duffin, J. Vac. Sci. Technol. B **11**, 311 (1993).
[21]M. K. Jain, T. S. Cale, C. J. Tracy, and R. L. Duffin, in *Proceedings of the 9th International VLSI Multilevel Interconnection Conference* (IEEE, New York, 1992), p. 447.
[22]EVFLOW is a thin film thermal flow process simulator developed by H. Liao under the direction of T. S. Cale at Auburn State University, with funding from the Semiconductor Research Corporation.
[23]K. Hirose, K. Kikuta, and T. Yoshida, Tech. Dig. IEDM 557 (1994).
[24]P. Stoltze, J. K. Norskov, and N. Landman, Phys. Rev. Lett. **61**, 440 (1988).
[25]A. W. Denier van der Gon, R. J. Smith, J. W. Gay, D. J. O'connor, and J. F. van der Veen, Surf. Sci. **227**, 143 (1990).
[26]T. Nguyen, P. S. Ho, T. Kwok, C. Nitta, and S. Yip, Phys. Rev. B **46**, 6050 (1992).
[27]M. Bienfait, Surf. Sci. **272**, 1 (1992).

[28]H. Liao and T. S. Cale, Thin Solid Films **253**, 419 (1994).
[29]D. E. Bornside, R. A. Brown, S. Mittal, and F. T. Geyling, Appl. Phys. Lett. **58**, 1181 (1991).
[30]C. A. J. Fletcher, in *Computational Techniques for Fluid Dynamics* (Springer, Berlin, 1991), Vol. II.
[31]C. J. Smithells and E. A. Brands, in *Metals Reference Book* (Plenum, New York, 1978).
[32]T. Smy, S. K. Dew, M. J. Brett, W. Tsai, M. Biberger, K. C. Chen, and S. T. Hsia, in *Proceedings of the 11th International VLSI Multilevel Interconnection Conference*, 1994, p. 371.

Nonlinear analysis of the $I-V$ characteristics in Ti/Si and $TiSi_2$/Si Schottky diodes

J. Pérez-Rigueiro, C. Jiménez, R. Pérez-Casero, and J. M. Martínez-Duart
Departamento de Física Aplicada, C-XII, Universidad Autónoma de Madrid, 28049 Madrid, Spain

(Received 25 September 1995; accepted 4 May 1996)

Ti/Si and $TiSi_2$/Si diodes have been prepared by direct current (dc) sputtering and rapid thermal processing. Their $I-V$ characteristics from 77 K up to 300 K have been analyzed using a nonlinear least squares method to fit the experimental data. All the experimental data have been fitted using two exponentially dependent currents and a series resistance. Two parameters are fitted for each current: a saturation current and an exponential parameter. As is discussed in the article, one current can be assigned to the thermoionic current, since the exponential parameter fits well to the theoretical value q/kT and therefore a temperature independent Schottky barrier height can be calculated. The second current, which yields an exponential parameter independent of temperature, cannot be assigned to a mechanism of direct tunneling through the barrier. As a result, we have proposed a trap-enhanced tunneling mechanism to explain this current. © *1996 American Vacuum Society.*

I. INTRODUCTION

The rectifying behavior of metal–semiconductor contacts has been known for over a century, and the first theories to explain this effect date back to 1930s.[1] This classical analysis led to the well known expression for solid state diodes, $I=I_{th}[\exp(qV/nkT)-1]$ where I_{th} is the thermionic saturation current, V is the applied voltage, n is the ideality coefficient, and the other constants have their habitual meaning. This expression has been widely used in the analysis of Schottky diodes mainly because a plot $\ln(I)$ vs V can be fitted using the linear least squares method obtaining I_{th} and n.

Several attempts have been made to give n a clear physical meaning. For example, it was soon realized that the presence of a series resistance (R_s) with the Schottky diode increases the ideality coefficient. This led to a series of related methods[2] to calculate R_s. Considering the effect of R_s, a modified voltage given by $V_{mod}=V-I\cdot R_s$ can be used in the diode equation, leading to values of n closer to 1.

Although the linear analysis including the ideality coefficient is still used,[3,4] the possibility of using a nonlinear least squares method provides a new way of performing the fitting of the $I-V$ characteristics.[5] In this article we apply such a method to Ti/Si and $TiSi_2$/Si Schottky diodes prepared by direct current (dc) sputtering and rapid thermal processing.

II. NONLINEAR FITTING OF $I-V$ CHARACTERISTICS

The aim of the nonlinear analysis is to resolve each current present in the Schottky diode. Since each current has to be related to a transport mechanism of electrons or holes through the barrier, this approach allows a deeper physical insight into the problem. The nonlinear fitting of $I-V$ characteristics of Schottky diodes by Donoval *et al.*[5] supposes four mechanisms which contribute to the $I-V$ characteristics at low temperatures: thermionic emission, generation-recombination, tunneling, and leakage current.

In this work we adopt a nonlinear least squares method to fit the experimental $I-V$ curves of Ti/Si and $TiSi_2$/Si diodes measured at temperatures from 77 to 300 K. In our approach we make no assumption about the possible mechanisms for current transport, except for the presence of a thermionic emission current. The experimental data are fitted to the function:

$$I=\sum_{i=1}^{N}(I_i\cdot(\exp[\alpha_i\cdot(V-I\cdot R_s)]-1), \tag{1}$$

where I and V are the measured intensity and voltage and R_s (series resistance), I_i (saturation current i), and α_i (exponential factor i) are fitting parameters. The nonlinear least squares fitting was made by means of a Levenberg–Marquardt algorithm[6] using the minimum number of exponential functions.

III. EXPERIMENT

Titanium thin films (~1000 Å thickness) were deposited on p-type boron-doped silicon (100) using a dc magnetron sputtering system. The concentration of dopants in the silicon was estimated from $C-V$ measurements[7] to be $N_A\sim5\times10^{15}$ cm^{-3}. Prior to the deposition the samples were dipped in HF (5%) to eliminate the native oxide. The geometrical definition of the diodes was made using a lift-off technique. The area of the diodes was about 1 mm^2.

$TiSi_2$/Si diodes were prepared by annealing the Ti/Si contacts in an N_2 atmosphere at 700 °C for 20 s in a rapid thermal processing system. Aluminium backcontacts were deposited on both the Ti/Si as deposited samples and the $TiSi_2$/Si samples. To improve the electrical characteristics of the contact the samples were annealed at 400 °C for 5 min.

The samples were subsequently introduced in a cryostat and forward $I-V$ characteristics were measured from 77 K up to 300 K in steps of 25 K, using a HP4140B pA meter.

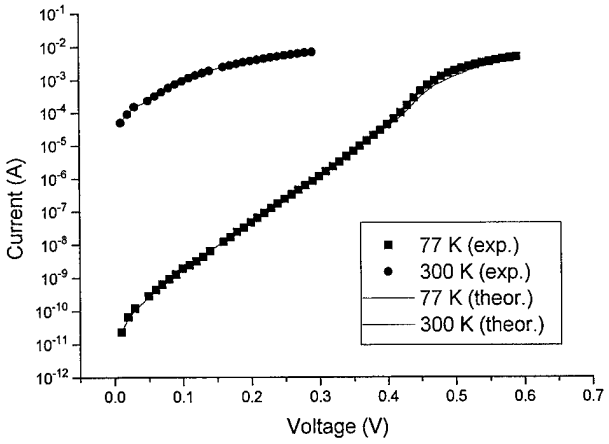


FIG. 1. Experimental and calculated $I-V$ characteristics of $TiSi_2$/Si and Ti/Si diodes at 77 and 300 K.

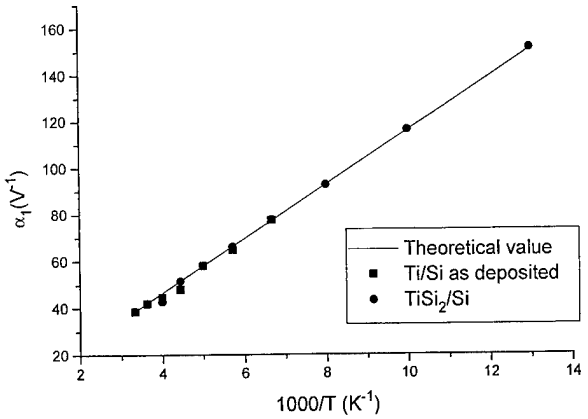


FIG. 2. Comparison of the fitted parameter α_1 with its theoretical value q/kT.

IV. RESULTS AND DISCUSSION

The experimental data of Ti/Si and $TiSi_2$/Si diodes could be fitted at all temperatures (77–300 K) using only two exponential functions, one of them corresponding to the thermionic emission current. The physical meaning of the other exponential term will be discussed below. This means that five parameters had to be fitted: two saturation currents (I_1 and I_2), two exponential factors (α_1 and α_2) and the series resistance. Figure 1 compares the experimental and theoretical characteristics of the $TiSi_2$/Si diode at 77 and 300 K.

The quality of the fitting can be measured using the parameter r^2, which is equivalent to the parameter r^2 of the linear least squares fitting. In this case however, r^2 is zero for an ideal fitting. The fittings of the $TiSi_2$/Si diode had an $r^2<1$ for all temperatures. The fittings of the Ti/Si diode had $r^2<2$ except for the measurements at 77 and 100 K, which yielded $r^2=16.9$ and $r^2=13$, respectively. No evident improvement was obtained by increasing the number of exponential functions of the fitting or even by including a linear function (simulating a leakage current); so the data concerning these temperatures should be taken cautiously.

The only previous physical assumption in this study is the presence of a thermionic emission current. This assumption allows an autoconsistency test for the method since the exponential factor α_1, corresponding to the thermoionic emission current, can be compared with its theoretical value q/kT.[8]

Figure 2 compares the variation of the parameter α_1 as a function of temperature with the theoretical values of q/kT. The excellent agreement found between both values indicates that this fitting is equivalent to taking $n=1$. Thus, the saturation current I_1 can be assigned to the thermoionic emission current. From the value of the saturation current, the Schottky barrier height ϕ can be calculated as $\phi=-kT/q\cdot\ln(I_{th}/A^*ST^2)$, where S is the area of the diode and A^* is the Richardson constant, which is taken as $A^*=30$ A/cm^2 K.^{2} The calculated Schottky barrier heights are shown in Fig. 3.

In a previous work, the Schottky barrier height was calculated at room temperature.[9] A clear difference in the barrier height of the as deposited Ti/Si ($\phi\sim0.55$ V) and the treated $TiSi_2$/Si ($\phi\sim0.52$ V) diodes was reported. The values now reported at 300 K in Fig. 3 coincide with those obtained in the previous work. At low temperatures the Schottky barrier height remains approximately constant, showing a slight increase in its value with decreasing temperature.[10]

The evolution of the series resistance with temperature is shown in Fig. 4. The series resistance decreases from 300 K down to 150 K where it shows a minimum, and then increases its value from this temperature down to 77 K. The behavior down to 150 K can be explained by the decrease of the scattering mechanisms dependent on temperature (i.e., phonons). Below 150 K the intrinsic conduction becomes steadily more important. The concentration of dopants has been calculated from the $1/C^2$ versus V plots. The values obtained were $N_A=6.9\times10^{15}$ cm^{-3} for the Ti/Si diode and $N_A=4.7\times10^{15}$ cm^{-3} for the $TiSi_2$/Si diode. These values are in good agreement with the difference of the series resistance shown between both diodes.

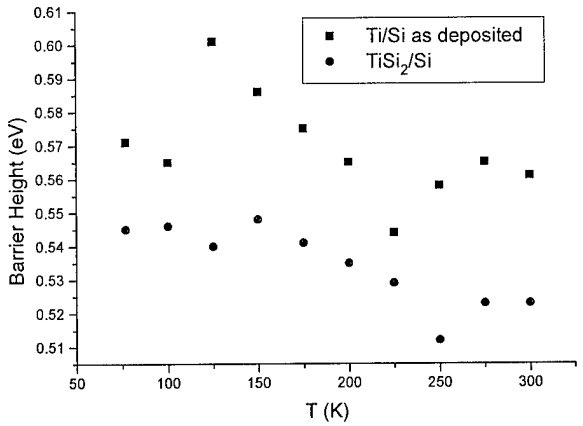


FIG. 3. Evolution of the calculated Schottky barrier height with temperature.

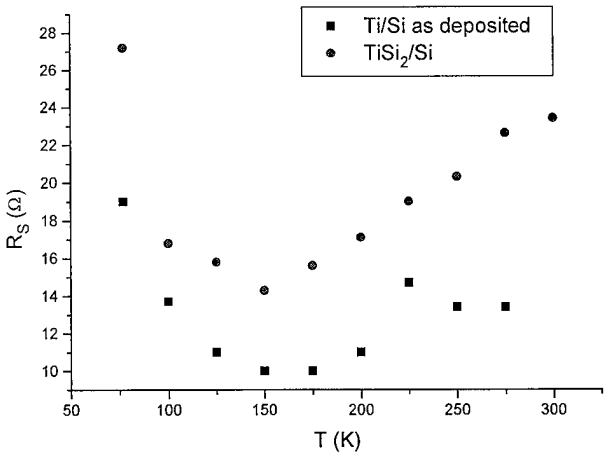


FIG. 4. Evolution of the calculated series resistance with temperature.

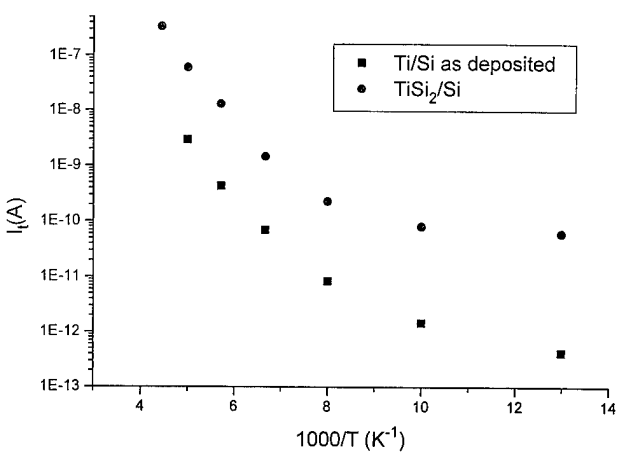


FIG. 6. Evolution of the calculated saturation current I_2 with temperature.

The exponential parameter and the saturation current of the second exponential term in Eq. (1) are shown in Figs. 5 and 6, respectively. The exponential parameter remains approximately constant with temperature; therefore this current cannot be assigned to the generation-recombination mechanism, since its exponential parameter depends on temperature as $q/2kT$. The obvious candidate seems to be a tunneling current. Two tunneling currents through a Schottky barrier have been proposed: the thermionic field and the field current, which have been discussed in the classical papers by Padovani and Stratton[11] and Crowell and Rideout.[12] According to these works the thermoionic-field current is typical of Schottky contacts on moderately doped ($N<10^{17}$ cm^{-3}) semiconductors. This current presents an exponential behavior with voltage, with an exponential parameter given by

$$E_0=E_{00}\coth\left(\frac{E_{00}}{kT}\right), \tag{2}$$

where $E_{00}=q\hbar(N_A/2\epsilon)^{1/2}/(2m^*)^{1/2}$.

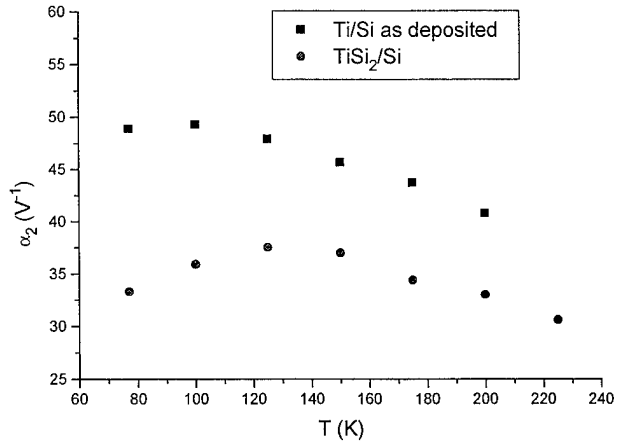


FIG. 5. Evolution of the calculated exponential parameter α_2 with temperature.

In this frame Eq. (2) was used to calculate the exponential parameter for the Ti/Si, diode taking m^*, the mean effective mass for holes in silicon, as $m_h^*=(m_{hh}^{*3/2}+m_{lh}^{*3/2})^{2/3}$. Values of 56.9 and 150 V^{-1} were obtained for α_2 at 200 and 77 K, respectively. Although these values are in the same order of magnitude as the values obtained from the fitting, no significant variation with temperature is observed in the values obtained from the fitting. Therefore the second experimental current cannot be assigned to the thermionic-field current.

On the other hand, the field current has been described for degenerate semiconductors with its Fermi level within the valence band (for a p-doped semiconductor).[11,12] Although this is not the case, the field current shows in the case of degenerate semiconductors an exponential behavior with voltage with an exponential parameter which is independent of temperature.

It is known that there is a different behavior of the exponential parameter of both tunneling currents with temperature. This is due to the fact that, whereas in the field current the maximum number of electrons tunneling through the barrier has an energy independent of temperature, the energy at which the maximum number of electrons tunnel in the thermionic-field current depends on a compromise between the statistical distribution of electrons with energy (increasing with decreasing energy) and the transmission coefficient through the barrier (decreasing with decreasing energy). Therefore, the temperature independent exponential parameter can be assigned to a tunneling process at a fixed energy above the Fermi level.

A possible mechanism to explain the behavior observed in our samples is the trap enhanced tunneling through the Schottky barrier.[13] According to this mechanism the electrons do not tunnel directly into the metal. The tunneling consists of a two-step process: first, the electrons tunnel from the conduction band into traps in the depletion layer; then from these traps the electrons finally tunnel into the metal. Since the traps are situated at fixed energies above the Fermi level, the electrons preferentially tunneling into the metal will have this energy independent of the temperature. The

enhanced tunneling mechanism can be formally described using an exponential function similar to the thermionic-field emission, but the numerical value of the exponential parameter is smaller in the case of the enhanced tunneling mechanism. This is in agreement with the experimental data, since the values of α_2 are smaller than the theoretical value 56.9 eV^{-1} at 200 K.

Figure 6 shows the variation of the saturation current I_2 with temperature. At relative high temperatures log (I_2) versus $1/T$ can be fitted to a straight line. This is in agreement with the enhanced tunneling mechanism, since the amount of current will be dependent on the number of electrons which have the appropriate energy. The statistical distribution in energy of these electrons corresponds to a Boltzmann distribution for moderately doped semiconductors. It can also be observed in Fig. 6 that the saturation current for the $TiSi_2$/Si diodes is greater than for the Ti/Si ones, in agreement with their higher density of traps.[14]

At low temperatures (100–77 K) I_2 shows a different behavior, which can be due to the presence of a leakage current through the diode. Except for the Ti/Si diode no attempt has been made to resolve this additional current. This is caused by the low value of the r^2 parameter obtained from the original analysis, which does not allow to discriminate clearly whether an additional current is physically significant or not.

V. SUMMARY AND CONCLUSIONS

The forward $I-V$ characteristics of two different Schottky diodes: Ti/Si and $TiSi_2$/Si have been analyzed by a nonlinear method. We have shown in this work that the $I-V$ characteristics can be fitted with two exponentials, if a series resistance is considered. The first exponential is attributed to the thermionic emission current. This hypothesis is supported by the good agreement between the exponential parameter and its theoretical value (q/kT). The fact of fitting the exponential parameter is equivalent to using an ideality coefficient equal to one. In this case, however, the ideality coefficient can be considered an option of the fitting, instead of an integral part of it, since the fitting can be improved by including more functions and so, the use of the ideality coefficient is considered more an autoconsistency test of the method than part of the fitting itself.

The value of the barrier height remains approximately constant at different temperatures. This implies that the difference between the barrier height of the Ti/Si and $TiSi_2$/Si diodes is a characteristic of the materials and is not due to the appearance of new conduction mechanisms.

The assignment of the second current is more controversial. It cannot be assigned to the generation-recombination current nor to the direct tunneling current through the diode. The characteristics of this current are similar to the enhanced tunneling current, and therefore we have assigned the current to the enhanced tunneling mechanism.

The possibility of performing nonlinear fitting to $I-V$ characteristics allows a more direct insight into the physics of the problem.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Autonomous Community of Madrid for financial support under Contract No. C169/91. One of the authors (J.P.-R.) would also like to acknowledge the Spanish Research Office (CICyT) for granting a Research Assistantship.

[1]W. Schottky, Naturwissenshaften **26**, 843 (1938).
[2]H. Norde, J. Appl. Phys. **58**, 5052 (1979).
[3]J. Nakamura, H. Niu, and S. Kishino, Jpn. J. Appl. Phys. **32**, 699 (1993).
[4]T. Ishida and H. Ikoma, J. Appl. Phys. **74**, 3977 (1993).
[5]D. Donoval, M. Barus, and M. Zdimal, Solid-State Electron. **34**, 1365 (1991).
[6]K. M. Brown and J. E. Dennis, Numerische Mathematik **18**, 289 (1972).
[7]E. H. Rhoderick and R. H. Williams, *Metal–Semiconductor Contacts* (Clarendon, Oxford, 1988).
[8]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).
[9]J. Pérez-Rigueiro, C. Jiménez, R. Pérez-Casero, and J. M. Martínez-Duart, Thin Solid Films **246**, 172 (1994).
[10]M. O. Aboelfotoh and K. N. Ku, Phys. Rev. B **34**, 34 (1986).
[11]F. A. Padovani and R. Stratton, Solid-State Electron. **9**, 695 (1966).
[12]C. R. Crowell and V. L. Rideout, Solid-State Electron. **12**, 89 (1969).
[13]G. H. Parker and C. A. Mead, Appl. Phys. Lett. **14**, 21 (1969).
[14]T. Yoshida, H. Kawahara, and S. Ogawa, J. Appl. Phys. **76**, 5272 (1994).

The effect of the *Ti* glue layer in an integrated Ti/TiN/*Ti*/AlSiCu/TiN contact metallization process

L. Ouellet and Y. Tremblay
Mitel S.C.C., 18, Boulevard de l'aéroport, Bromont, Québec JOE 1L0, Canada

G. Gagnon, M. Caron, J. F. Currie, and S. C. Gujrathi
École Polytechnique de Montréal, Groupe de Recherche en Physique et Technologie des Couches Minces, C. P. 6079, Succ. Centre-ville, Montréal, Québec H3C 3A7, Canada

M. Biberger
Varian Associates, Palo Alto, California 94303-1025

(Received 27 October 1995; accepted 3 June 1996)

A Ti/TiN bilayer deposited under an *Ti*/AlSiCu/TiN interconnect is generally used at the contact level to stabilize the contact resistance of small diameter contacts as well as to prevent junction spiking during postdeposition anneals. An air break followed by a furnace or a RTP anneal is generally required after the deposition of the Ti/TiN bilayer to improve the barrier stability. This three step interconnect scheme results in a low net effective throughput process if the same metallization system is used twice or, alternatively, in a high cost-of-ownership process if two metallization systems are used, a first for the Ti/TiN barrier and a second for the *Ti*/AlSiCu/TiN interconnect. In this article, it is shown that collimation permits the deposition of very thick TiN barriers at the bottom of high aspect ratio contacts which allows contact metallization integration in a single metallization system. In addition, it is found that the use of a *Ti* glue layer between the barrier and the AlSiCu contact interconnect to promote the filling of these high aspect ratio contacts with Al alloys results in a yield loss of n^+ and p^+ contact chains and therefore should be avoided. Finally, it is demonstrated that an integrated Ti/TiN/AlSiCu/TiN contact metallization with excellent electrical properties is possible with collimated TiN barriers. © *1996 American Vacuum Society.*

I. INTRODUCTION

Reactively sputtered TiN is the most popular barrier for the contact metallization with aluminum alloys since it provides a chemically stable barrier with low defect density and controlled properties. This process has nevertheless a major drawback: the bottom coverage of reactively sputtered TiN drops to less than 10% when the aspect ratio of straight wall contact holes exceeds 1.2. This means that a 100 nm thick TiN only provides 10 nm of TiN at the bottom of 0.8 μm diam straight wall contacts etched through a 1.0 μm thick oxide.

Collimation of TiN was introduced few years ago to push the limits of the contact metallization with aluminum alloys as well as to permit the contact metallization with chemical vapor deposited tungsten (CVD W) in the subhalf micron processes.[1] Varian's patented 1.5:1 cds® Ti/TiN collimation process[1–6] permits a 100 nm thick TiN to provide 60 nm of TiN at the bottom of the same 0.8 μm diam straight wall contacts etched in 1.0 μm thick oxide. This means a 500% improvement over reactively sputtered TiN without collimation.

The 1.5:1 cds® TiN films can be obtained with collimation in a nitrided target mode of operation, thus providing so called TiN(B) films, or alternatively in a non-nitrided target mode of operation, thus providing so-called TiN(C) films of equally high quality: a bulk resistivity of 40 $\mu\Omega$ cm, a bulk density of more than 95%, a high compressive stress, a deep gold appearance and a [N]/[Ti] stoichiometric ratio of 1.

The improved bottom coverage and film properties of these two types of processes certainly increase the performance of the TiN barrier and could permit an integrated contact metallization with aluminum alloys at small geometry without the use of an air break and of an *ex situ* furnace or RTP anneal. This article looks at such an integrated metallization process with AlSiCu alloy.

II. EXPERIMENT

A. Contacts shape

A 1.5 μm mixed signal CMOS process involving two levels of polysilicon and two levels of aluminum alloy interconnects is used to verify the performance of a Ti/TiN/Al-1.0 wt %Si-0.5 wt %Cu/TiN metallization process in the M2000 cluster tool. An inorganic 2.0 wt %P SOG is used to planarize the upper capacitor polysilicon. This results in a local accumulation of 1.4 μm of oxide in the contact area of the RTP activated 0.2 μm deep n^+/p^- and p^+/n^- junctions where straight wall 1.0, 1.2, 1.5, and 1.8 μm diam contacts are etched. The junctions were prepared by implantation of As and B (using BF_2 molecular ions) as N and P dopants, respectively. The implanted energies and doses were 80 keV and 5.8×10^{15} at/cm^2 for As and 50 keV and 2.1×10^{15} at/cm^2 for BF_2 (the corresponding energy of the B atoms being ~11 keV). After the implantation, the dopants were activated at a temperature of 1065 °C for 10 s.

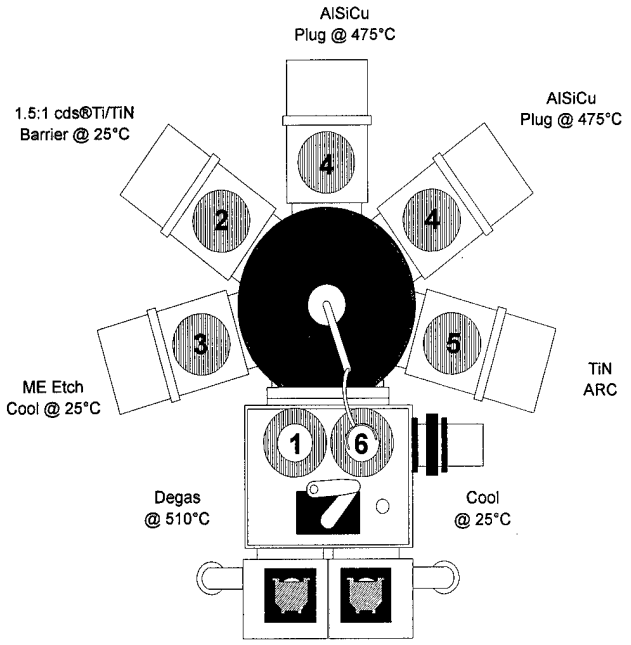


FIG. 1. Layout of the different modules composing Varian M2000 cluster tool. The numbers drawn on each module corresponds to the sequence followed during the metallization: (1) degas module, (2) collimated TiN module, (3) etch module used as a cooling module, (4) AlSiCu module, (5) collimated TiN module, and (6) cooling module.

B. Wafer process sequence

After a 60 s etch in a mixture of NH_4F:HF (50:1) buffer oxide etch solution, the wafers are loaded in a Varian M2000 cluster tool configured as illustrated in Fig. 1. The wafer movement in the system is described according to the encircled number in this figure. The wafer is first degassed for 75 s at about 510 °C in the degas station (1). While hot, the wafer is then transferred into the 1.5:1 cds® Ti/TiN module (2) onto a vacuum isolated heater operated at a setpoint of 25 °C. The wafer remained at a temperature of about 400 °C during the 150 s necessary to make the Ti/TiN or Ti/TiN/*Ti* deposition. The wafer is moved after the deposition to an etch module (3) where it is cooled for a duration of about 150 s by a chuck operating at a setpoint of 25 °C. The cooling is possible using a flow of backside argon between the wafer and the heater. The wafer is subsequently transferred in one of the two AlSiCu modules (4) operating at a heater setpoint of 475 °C where a 400 nm cold/400 nm hot two steps AlSiCu planarization (TSP) film is deposited within about 300 s. The wafer is finally transferred after the Al deposition to a cool station (6) where it is cooled within about 75 s. Each individual module that are part of this M2000 cluster tool is isolated and cryopumped. After the patterning of the interconnects, the remaining standard backend process steps up to wafer level testing are carried out.

C. Optimization of the various films thicknesses

The timing in each module is such that an integrated metallization is possible with a productivity of about 18 wafers per hour (WPH). This timing forces the collimated Ti/TiN or Ti/TiN/*Ti* barrier to be deposited within 150 s, even at the end of life of the collimator,[1–4] and limits the thickness of the bottom Ti contact layer, of the inner TiN barrier layer, and of the upper *Ti* glue layer if used.

Since a minimum thickness of the Ti contact layer is required to ensure good contact resistance to n^+, to p^+, to gate polysilicon and to capacitor polysilicon and since a *Ti* glue layer is used in some of the investigated cases, an upper limit exists for the thickness of low deposition rate TiN(B) and for the high deposition rate TiN(C).

Therefore, the thickness of the Ti contact layer under an arbitrary selected 60 nm thick TiN(B) barrier was varied in order to determine this minimum Ti contact layer thickness. No *Ti* glue layer was deposited and the barrier was covered by 800 nm of AlSiCu followed by 40 nm of TiN antireflective coating. The thickness of the Ti contact layer was fixed using the results of contact resistance measurements. A second experiment was then performed using the optimum thickness for the Ti contact layer to investigate the effect of the *Ti* glue layer on a certain number of *splits* which all maintain a productivity of 18 WPH at the end of life of the collimator. Two reference wafers without barriers were also included in this study. Table I gives the set of splits prepared for this experiment.

D. Interfacial reaction measurements

Three types of blank substrates were also used to investigate the robustness of the tested barriers. These substrates

TABLE I. Metallic multilayers used to study the effect of the *Ti* glue layer on the electrical properties of devices. The left to right ordering of the columns corresponds to the sequential order of the deposition of the various layers. The last two rows of this table corresponds to the two reference wafers included in the study.

Ti contact layer thickness (nm)	TiN barrier type	TiN barrier thickness (nm)	Ti glue layer thickness (nm)	AlSiCu thickness (nm)	TiN ARC thickness (nm)
10	[C]	95	0	800	40
10	[C]	80	10	800	40
10	[B]	30	10	800	40
10	[B]	35	0	800	40
0	···	0	0	800	40
0	···	0	0	800	40

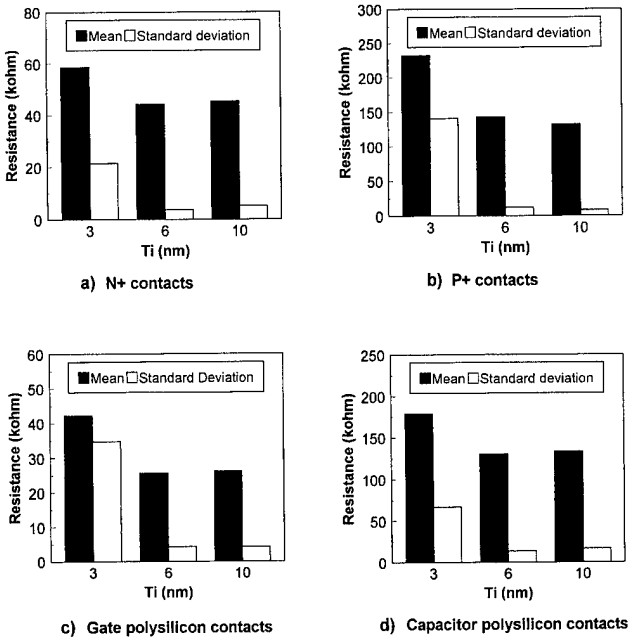


FIG. 2. Contact resistance results of n^+ (a), p^+ (b), gate polysilicon (c) and capacitor polysilicon (d) contact chains as a funtion of the Ti contact layer thickness. The black bar gives the average and the white bar the standard deviation.

were a thermal oxide, an n-type single-crystal silicon free from native oxide and a N-doped and RTP activated polycrystalline silicon. The latter substrate is known to easily dissolve in aluminum.

The tested barriers were deposited over each type of substrate and under a 100 nm AlSiCu layer and a 40 nm TiN film combination. Sheet resistance maps of 49 points per wafer of the resulting wafers were measured before and after a 60 min 450 °C anneal in order to detect any change in the sheet resistance which would result from an interfacial reaction with any of the three types of substrate or with the AlSiCu alloy. In this part of the study, two wafers were measured for each tested barrier/substrate combination.

E. Junction leakage measurements

The tested junctions have a surface of 160 000 μm^2 and a perimeter of 1600 μm. There are 27 tested junctions per size of contact per wafer. Two types of contacts were used to connect the junctions: (1) 6400 contacts of 2.0 μm$\times$2.0 μm and (2) 1600 contacts of 4.0 μm$\times$4.0 μm.

The junctions were tested at a reverse bias of 5 V in total darkness and at room temperature. In order to be used in low power CMOS devices, a leakage current density of about 5 nA/cm^2 of junction which corresponds to a total leakage current of about 10 pA for this size of a junction is desired. According to our specifications, this leakage is independent of the number of connecting contacts, but if one contact leaks the complete junction leaks.

F. Contact chain measurements

$n^+(p^+)$ diffusion contact chains having contact sizes of 1.0, 1.2, 1.5, and 1.8 μm, respectively, were measured. Each chain, located in a large p-well (n-well), has 200 contacts, has 100 n^+ (p^+) diffusion links and has 99 AlSiCu links. Similar gate polysilicon and capacitor polysilicon contact chains (no 1.2 μm capacitor polysilicon contact chains) were also measured. Located in a large p well, each chain has 200 contacts, 100 polysilicon links, and 99 AlSiCu links. As for the junction leakage measurements, there is 27 chains per type of contact per wafer.

III. RESULTS

A. Optimum Ti contact layer thickness

Figure 2 shows the average contact chain resistance and the standard deviation of the tested 1.5 μm n^+, p^+, gate polysilicon and capacitor polysilicon contact chains. High contact resistance with large standard deviation is obtained for the 3 nm thick Ti contact layer for all types of contact structures. However, lower contact resistance with small standard deviation is measured for the 6 and 10 nm thick Ti contact layer. Furthermore, the yields vary between 60% and 70% for all types of contact chains containing the 3 nm Ti contact layer while it is 100% for the other two thicknesses.

B. Effect of a *Ti* glue layer

Figure 3 shows the percentage yield of the tested 1.0 and 1.8 n^+, p^+, gate polysilicon, and capacitor polysilicon contact chains before and after annealing at 450 °C as a function of the deposited barrier structure. This is a measure of the number of good contacts on a wafer. The graphs in Fig. 3 clearly show that the yield is much less than 100% when a 10 nm thick *Ti* glue layer is present on the n^+ and p^+ contact chains. In fact, it is zero for the 1.0 μm contact chains in both cases before and after annealing. For the gate polysili-

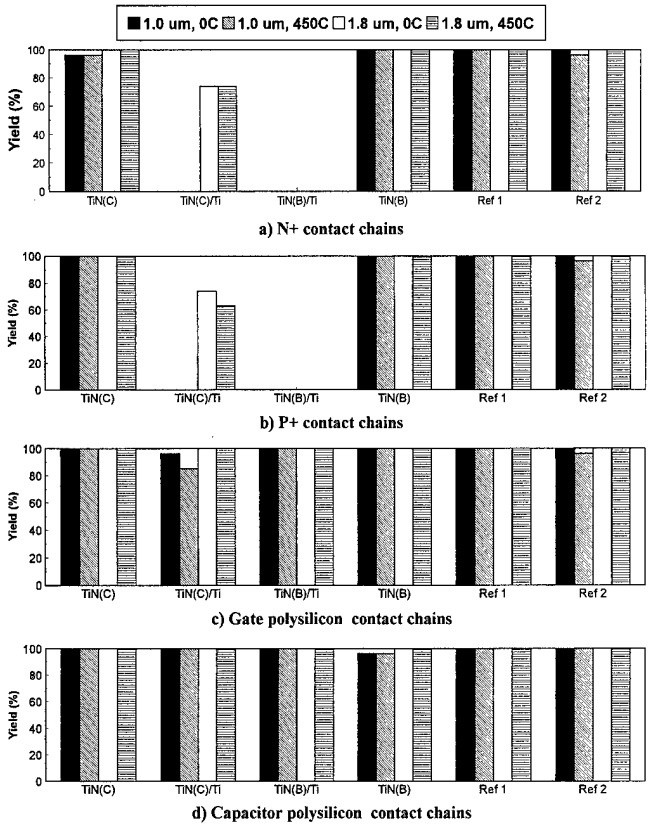


FIG. 3. Graphs of the percentage yield for n^+ (a), p^+ (b), gate polysilicon (c) and capacitor polysilicon (d) contact chains as a function of the barrier type deposited on the wafer for two different contact sizes before and after annealing at 450 °C.

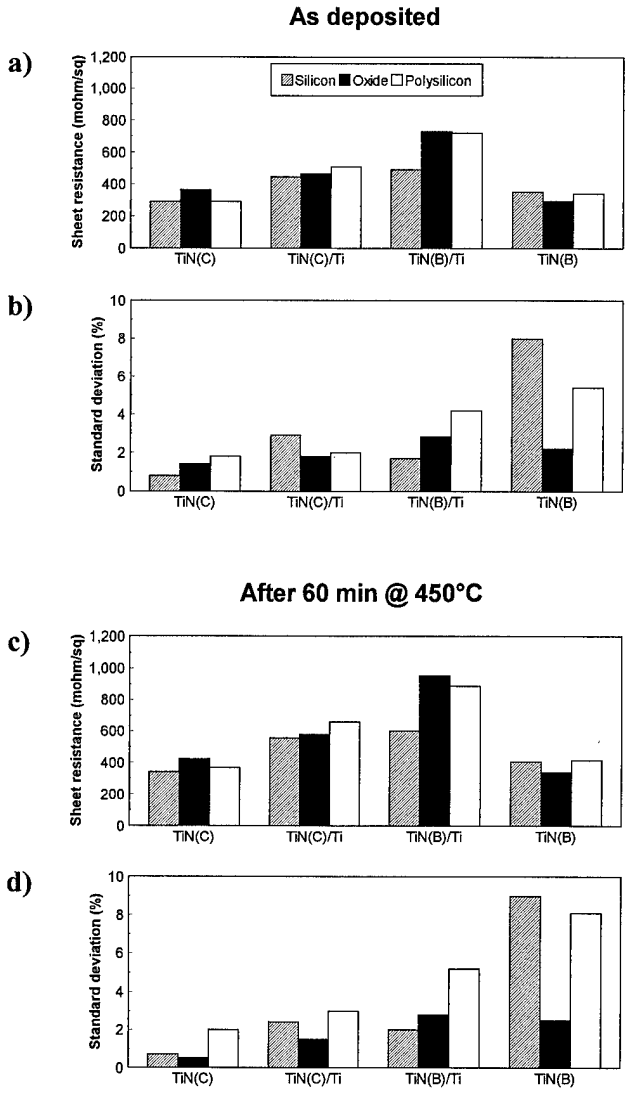


FIG. 4. Average sheet resistances [(a) and (c)] and standard deviations [(b) and (d)] before and after annealing at 450 °C for barrier with and without a *Ti* glue layer deposited on three different substrates: silicon (hatched bar), oxide (black bar), and polysilicon (white bar).

con, a slight decrease in the yield is measured for the 1.0 μm contact chains having a TiN[C]/*Ti* glue layer before and after annealing. No important yield loss is observed for the capacitor polysilicon chains. It is very clear that the use of a Ti glue layer in an integrated Ti/TiN/*Ti*/AlSiCu/TiN process sequence on the n^+ and of the p^+ contact chains should be prevented. A reaction of the *Ti* glue layer with the Al during the deposition of the alloy may be responsible for such a total yield loss.

C. Interfacial reaction measurements

Figure 4 shows the average sheet resistance and the standard deviation results of the four tested barriers deposited over the three types of substrates. Table II gives the sheet resistance for each of the compounds used in this part of the study. If the metallic multilayers keep their integrity, it is possible to estimate the expected sheet resistance value from the known thickness and resistivity of each individual film and substrate given in Table II. Since all the individual films are connected in parallel, the net sheet resistance of each wafer should be very close to the AlSiCu sheet resistance of 300 mΩ/□.

Before annealing, Fig. 4(a) clearly shows that both TiN[C] and TiN[B] barrier layer without a *Ti* glue layer maintains an average sheet resistance around 300 mΩ/□.

TABLE II. Thickness, resistivity, and sheet resistance of all the layers and substrates used in the electrical measurements.

Layer/substrate	Thickness (nm)	Resistivity ($\mu\Omega$ cm)	Sheet resistance (Ω/□)
TiN ARC	40	40	10.0
AlSiCu	100	3.0	0.30
TiN[C] barrier	95	40	4.2
TiN[C] barrier	80	40	5.0
TiN[B] barrier	35	65	18
TiN[B] barrier	30	65	21
Ti contact layer	10	80	80
Ti glue layer	10	80	80
Si substrate			1.5
PolySi			30
SiO_2			very high

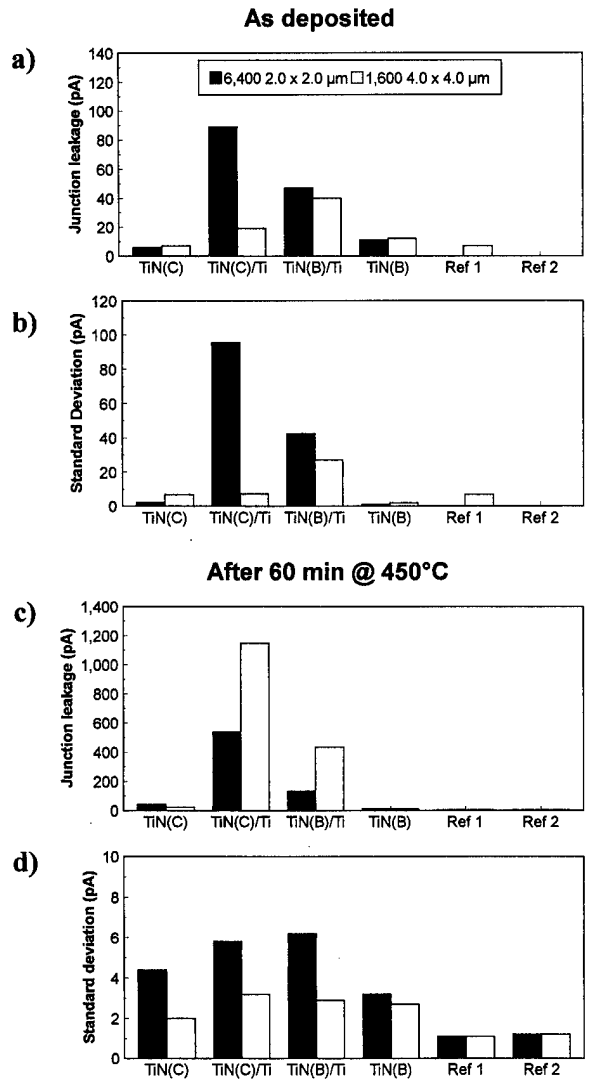


FIG. 5. $p+/n-$ junction leakage currents [(a) and (c)] and standard deviations [(b) and (d)] before and after annealing at 450 °C for barrier with and without a *Ti* glue layer and reference wafers measured on two different types of contact: 2.0×2.0 μm^2 (black bar) and 4.0×4.0 μm^2 (white bar).

However, only the TiN[C] barrier shows small standard deviation of the sheet resistance while for the TiN[B], one measures the largest standard deviation on most substrates. When a *Ti* glue layer is present, the sheet resistance vary between 400 and 750 mΩ/□ on all substrate types. After the hour long 450 °C annealing, the sheet resistance worsens for all barrier type as can be seen in Fig. 4(c). However, the increase in the sheet resistance is small for TiN[C] and TiN[B] barrier without a glue layer and large for the barriers with the *Ti* glue layer. The standard deviation remains quite similar before and after the annealing.

D. Junction leakage measurements

Figure 5 shows the average p^+/n^- junction leakage and the standard deviation results of the four tested barriers and of the two reference wafers without barrier. Before annealing, Fig. 4(a) shows that the barriers without a *Ti* glue layer have a leakage current similar to the leakage of the first reference wafer. Furthermore, the standard deviations are small. For the barriers covered with a glue layer, the leakage is many times higher than the desired value of 10 pA with a standard deviation of the same order of magnitude than the average value. The junction leakage of the wafers having a *Ti* glue layer increases by more than one decade after a 60 min thermal stress at 450 °C while a low average value is obtained for the barriers without a Ti glue layer. However, the standard deviations are approximately the same for all barrier types, typically a few pA higher than the measured standard deviation of the reference wafers.

E. Contact resistance measurements

Figure 6 shows the average and the standard deviation of the contact resistance of (a) 1.0 μm n^+ chain, (b) 1.2 μm n^+ chain, (c) 1.0 μm p^+ chain, and (d) 1.2 μm p^+ chain before and after a 450 °C-60 min thermal stress. Only the barriers without *Ti* glue layer are analyzed along with the two reference wafers since all previous measurements had shown that a *Ti* glue layer worsens the electrical properties of a device.

For the n^+ chains, it is seen that before annealing the average resistance of about 10 kΩ is the same for the barrier and reference wafers with a standard deviation around 1 kΩ. After annealing, the barrier wafers have a slightly lower resistance than the reference wafers for both contact sizes. The standard deviations follow the same trend. In the case of p^+ chains, the average resistance for the barrier wafers is already lower than the average for the reference wafers before annealing. This difference is even more striking for the standard deviations. After annealing, the lower resistance of the barrier wafers compared with the reference wafers is still observed but not so strongly for the 1.0 μm contact size. However, the standard deviations are very different between wafers with and without barrier. The above results are also observed for the 1.5 and 1.8 μm contact chains but the differences are less apparent.

In Fig. 7, the average and the standard deviation of the contact resistance of (a) 1.0 μm gate polysilicon chain, (b) 1.5 μm gate polysilicon chain, (c) 1.0 μm capacitor polysilicon chain, and (d) 1.5 μm capacitor polysilicon chain before and after a 60 min thermal stress at 450 °C are presented.

As is the case for the n^+ and p^+ chains, the gate and capacitor polysilicon chains have a better average resistance when a barrier is present on the wafers, especially after the annealing. However, the difference between a barrier and a reference wafer is less apparent for the gate polysilicon chain for both sizes. It can be seen that in the case of the capacitor polysilicon chains, the difference in the average resistance between the barrier and the reference wafers is smaller for the 1.5 μm contact size. In all cases, there is no significant difference between the two analyzed barriers after the 450 °C-60 min thermal stress; the thin 35 nm TiN[B] barrier gives as good results as the thicker 95 nm TiN[C] barrier.

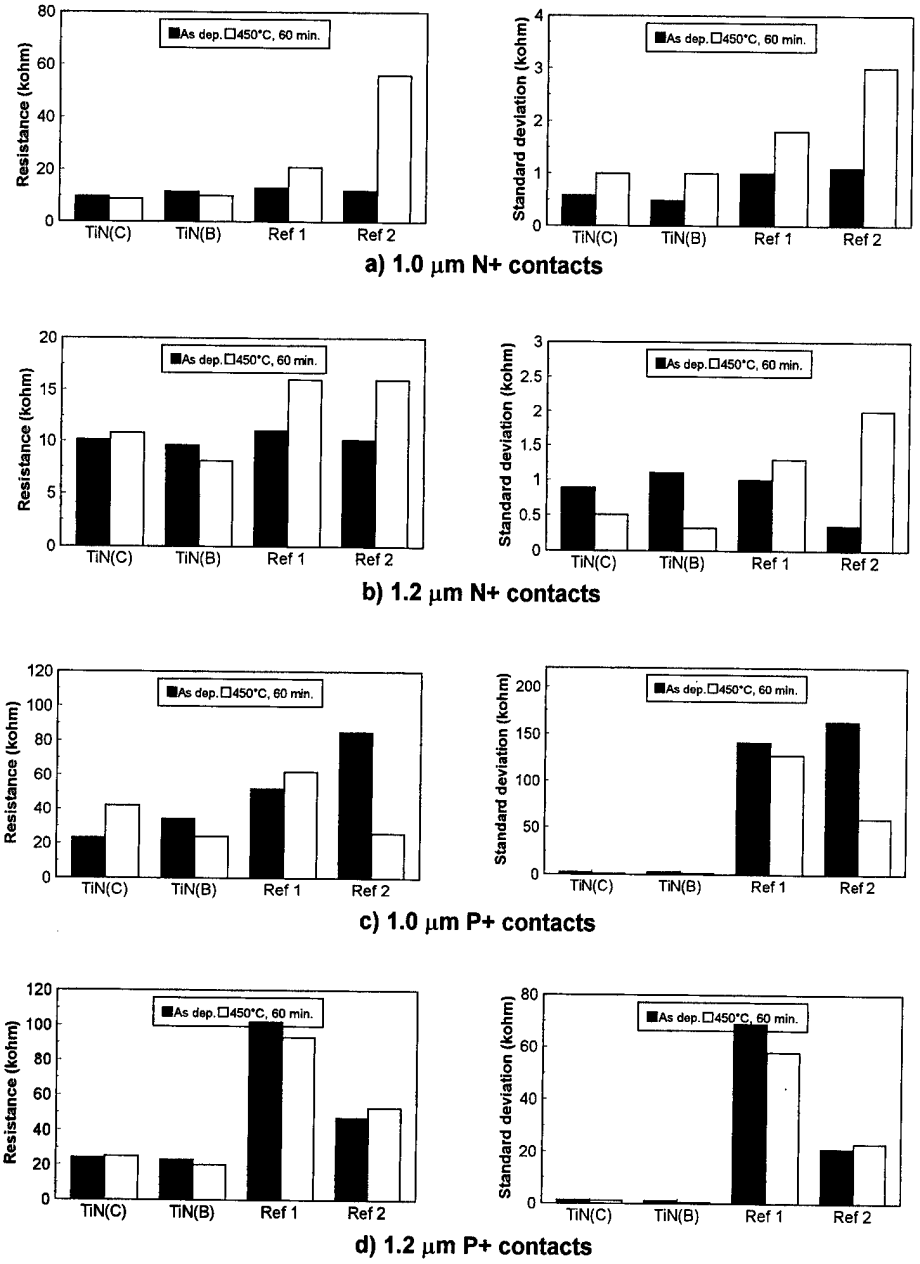


FIG. 6. Average contact chain resistances and standard deviations of 1.0 μm n^+ (a), 1.2 μm n^+ (b), 1.0 μm p^+ (c), and 1.2 μm p^+ (d) contact chains before (black bar) and after (white bar) a 450 °C annealing.

IV. DISCUSSION

A. Standard deviation vs average value

The reader must have noticed that practically all figures presented in this article show results in the form of average and standard deviation values. There are target average values which must be obtained for all the electrical characteristics measured in this work. In the industry, this target value represents the normal operating characteristic of the device. On the other hand, the standard deviation of the measurements represents a measure of the reproducibility and uniformity of the deposition process of the metallization multilayers. One has to remember that the average and the standard deviation are calculated from 27 different measurements on one wafer in the case of leakage and contact resistance while for the sheet resistance, they are determined from 49 measured sites on each of the two wafers.

For industry purposes, it is as important to attain the desired target value as well as to have tightly distributed and small standard deviations. Tightly distributed deviations indicate the excellent uniformity of a given layer throughout the wafer as well the good reproducibility of the deposition process of the metallization. These brief considerations will surely help to understand the conclusions that are reached in the following discussion.

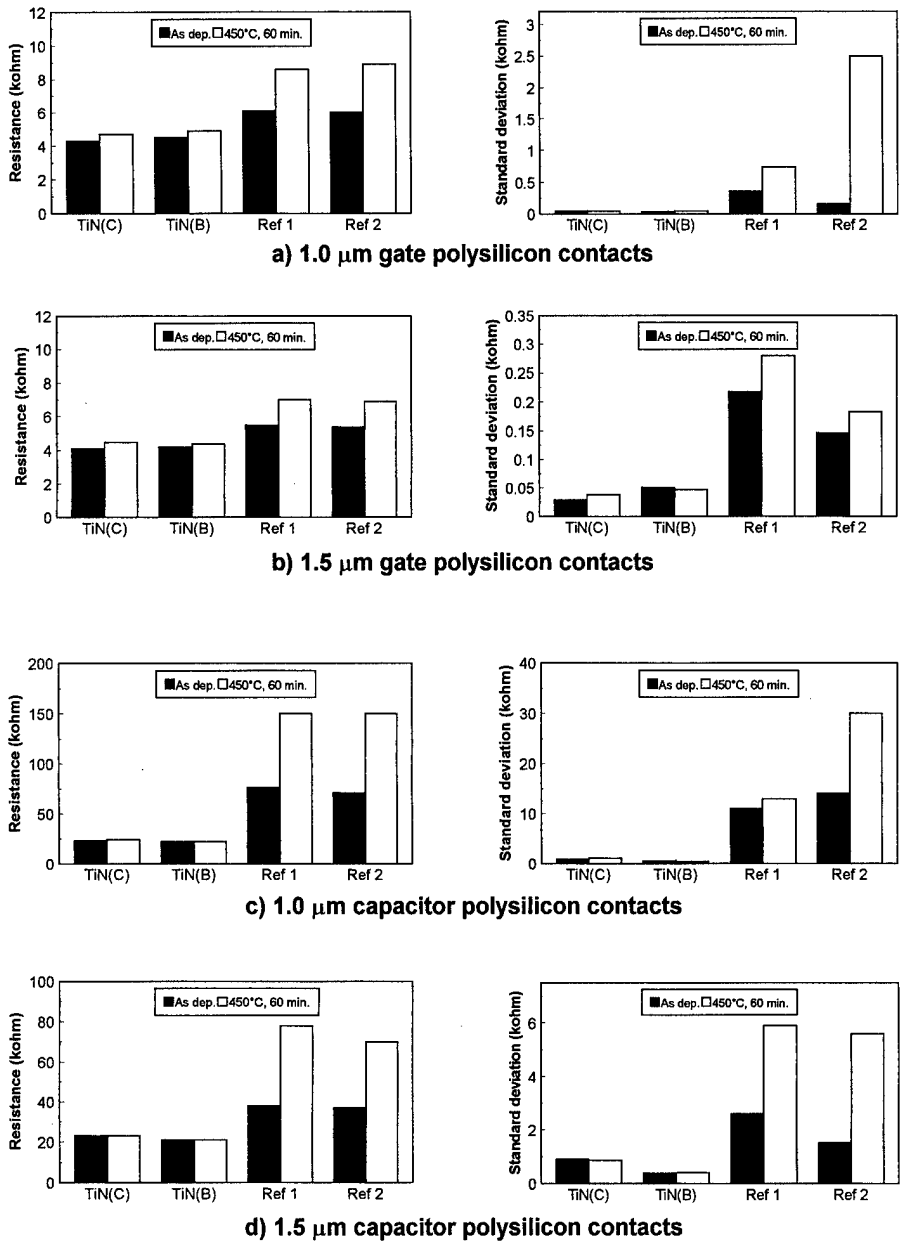


FIG. 7. Average contact chain resistances and standard deviations of 1.0 μm gate polysilicon (a), 1.5 μm gate polysilicon (b), 1.0 μm capacitor polysilicon (c), and 1.5 μm capacitor polysilicon (d) contact chains before (black bar) and after (white bar) a 450 °C annealing.

B. Ti contact layer

The results presented in Fig. 2 may be understood if a reaction between the Ti contact layer and the underlying Si substrate is occurring. The reaction product is either a TiSi or a metastable C49–$TiSi_2$ phases.[7] This hypothesis is supported by the facts that the deposition of the Ti contact layer is done on a substrate which is at a temperature of 400 °C and that the deposition of the Al alloy layer is carried out at 475 °C. Assuming that 1 nm of Ti reacts with 1 nm of Si to form 1 nm of TiSi, this latter phase will have a thickness of 3, 6, and 10 nm for the 3, 6, and 10 nm thick Ti contact layers. In the case of the C49–$TiSi_2$ phase, the Murarka and Fraser[8] rule predicts a thickness of 7.5, 15, and 25 nm for the 3, 6, and 10 nm Ti contact layers. The decrease with the thickness of the contact resistance (Fig. 2) may then be related to the silicide thickness that is being formed.

A thickness of 3 nm for a silicide corresponds to only a few atomic layers of a compound. In this case, it is very likely that the reacted layer between the barrier and the substrate is not a very uniform silicide which obviously results in a nonuniform contact, thereby increasing the contact resistance. On the other hand, at the maximal silicide thickness, i.e., 25 nm, the layer is certainly uniform and, at the same time, is sufficiently thin to avoid detrimental effect on the contact resistance due to dopant loss from the contacted substrates. This is why it was decided to use a 10 nm Ti

contact layer throughout the rest of the study.

C. *Ti* glue layer

Figures 3–5 clearly indicate that the use of a *Ti* glue layer in an integrated Ti/TiN/*Ti*/AlSiCu/TiN process sequence has a negative effect on p^+/n^- junction leakage, n^+ and p^+ contact chains resistance, as well as on sheet resistance. The results shown in Fig. 4 may provide a possible explanation of the detrimental effect of the *Ti* glue layer. It may be recalled that an average sheet resistance of 300 mΩ/□ is expected for all contacts if no reaction is taking place between any of the deposited layers. Only the barriers having the *Ti* glue layer show resistance higher than this expected value before the 450 °C annealing. A chemical reaction between the *Ti* glue layer and the AlSiCu layer for as-fabricated wafers is surely occurring during the deposition of the AlSiCu metal layer at 475 °C. In fact, Al and Ti readily react to form the $TiAl_3$ compound at temperatures as low as 400–450 °C.[9,10]

Table II indicates that only 10 nm of Ti was deposited as a glue layer. When the Ti film thickness is smaller than what is necessary to form a stoichiometric $TiAl_3$ phase with the diffusing Al, the excess Al could interact with the underlying layers.[11] The average sheet resistance of 400–500 mΩ/□ before annealing indicates that about half of the Al alloy thickness has been consumed through reaction. This indicates that the TiN barrier layer as well as the Ti contact layer beneath the *Ti* glue layer may have also reacted with the Al during the deposition of the AlSiCu layer. In the case of the TiN[B](30 nm)/Ti(10 nm) barrier, the measured sheet resistance indicates an almost complete consumption of the AlSiCu layer. Since the $TiAl_3$ byproduct has a bulk resistivity of about 50 μΩ cm, a 100 nm film would have a sheet resistance of 500 mΩ/□, value actually obtained before annealing.

As Fig. 4 shows, the sheet resistance is worse after a 450 °C annealing in the case of barrier covered with the Ti glue layer. Furthermore, Fig. 5 clearly shows that such a multilayer also has very high leakage before and after this annealing. Since the Al layer is almost fully converted into $TiAl_3$, this means that Al must have consumed the complete Ti/TiN/*Ti* multilayer and has probably penetrated the doped substrate. This, in turn, results in Si (and dopants) diffusion inside the $TiAl_3$ compound which leads to spiking problems, especially for the 1.0 μm n^+ and p^+ contacts chains which do not show satisfactory contact resistance [see Figs. 3(a) and 3(b)].

The above paragraphs indicate that the most plausible cause of the detrimental effect of the *Ti* glue layer on the electrical performance of diodes and contact chains is the reaction of the full *Ti* glue layer/TiN barrier/Ti contact layer with Al during the deposition of the AlSiCu alloy at 475 °C.

D. TiN barrier

When the metallization multilayer has no *Ti* glue layer, the results of Figs. 2–7 show the excellent electrical properties of the TiN barrier/Ti contact layer when compared with reference wafers. This means that the TiN/Ti bilayer does not react much with Al during the 475 °C deposition of the AlSiCu alloy. This is understandable since it is known that TiN is thermally stable up to 800 °C in contact with Si[12,13] and does not react much with Al up to 500 °C.[14].

As Figs. 5–7 show, the presence of the TiN barrier/Ti contact layer results in a leakage current slightly higher than the reference wafers leakage but in a much lower contact chain resistance than the reference wafers (especially after the 450 °C annealing). These results point out the necessity to use a system of barrier and contact layer in order to stabilize the contact resistance to n^+ diffusion, to p^+ diffusion, to gate polysilicon (N-doped) and to capacitor polysilicon (N-doped) during the thermal stress at 450 °C. It is generally known and predicted by the Al–Si phase diagram[15] that the aluminum doped silicon (p-type) precipitates at the bottom of these contacts. Their associated rectifying effects explain the increase of the contact resistance after thermal stresses.

The above paragraphs justify why the Ti glue layer must be avoided in the metallization multilayers and why the TiN barrier/Ti contact layers are considered to be necessary in the same metallization and have such good electrical properties.

V. CONCLUSION

The integration of a Ti/TiN/*Ti*/AlSiCu/TiN metallization sequence within a single metallization equipment giving a net productivity of 18 wafers per hour (WPH) is desirable for cost-of-ownership reasons. This sequence integration is possible with the high quality TiN barriers obtained with collimation. The two analyzed barriers, 95 nm TiN(C) and 35 nm TiN(B) give good results and the same productivity of 18 WPH. The better stability of the 95 nm TiN(C) barrier (smaller standard deviation) in the sheet resistance tests indicates a much better reaction uniformity.

The optimization of the thickness of the Ti contact layer shows that a 10 nm thick Ti contact layer is enough to ensure good contact resistance to the analyzed n^+ diffusion, p^+ diffusion, gate polysilicon, and capacitor polysilicon. The analysis of the junction leakage results, of the contact chain resistance results and of the sheet resistance results indicates that the *Ti* glue layer of an integrated Ti/TiN/*Ti*/AlSiCu/TiN metallization sequence causes an important degradation of the electrical properties of the devices before and after a 450 °C-60 min thermal stress. It was discussed that this degradation is due to a reaction between the *Ti* glue layer and the AlSiCu layer during the deposition of the alloy since the wafer is heated during part of the deposition. In fact, it is argued that the Al is in fact reacting with the underlying *Ti*/TiN/Ti/substrate during this processing step, resulting in the formation of the $TiAl_3$ compound as well as spiking in the n^+ and p^+ diffusion areas.

ACKNOWLEDGMENTS

The authors would like to thank Sylvie Picard, Stéphane Blain, Ted Darwall, Kathy Feehan, Christian Veilleux, Jules Poisson, Robert Groulx, Marc Faucher, and Gaétan Beaulieu for their continuous support during this experiment.

[1]L. Ouellet, R. Reynolds, and Y. K. Kim, Recent Advances in Metallization Technologies for ULSI Applications; Collimated Ti/TiN for 0.5 μm and 0.35 μm Technologies, 1992 Semicon Korea, Seoul, Korea.
[2]M. Biberger, L. Ouellet, and G. Tkach, Barrier Technology for 0.35 μm Device Geometries, Challenges of ULSI Interconnect Systems Symposium, San Jose, September 15, 1993.
[3]L. Ouellet, M. Biberger, and G. Tkach, The PVD Ti/TiN Solution for 0.35 μm Contact Holes with Aspect Ratio Larger than 3.0, 2nd International Symposium on Sputtering Processes, Tokyo, Japan, May 27–28, 1993.
[4]L. Ouellet, M. Biberger, and G. Tkach, Collimated Ti/TiN for Sub-0.5 μm Contacts Filled with CVD W or Al/Cu Plugs, Semicon West, San Francisco, CA, July 1993.
[5]M. Ismeurt and R. Reynolds, An In-Situ Reactively Sputtered Composite TiN Barrier for Use with Multilevel, Hot Al Fill; 0.7 μm to 0.5 μm Logic Devices, ULSI U.C. Berkeley Conference, Berkeley, CA, October 4–6, 1994.
[6]J. Schlueter, B. Jones, C. K. Huang, L. Ouellet, M. Biberger, S. Jackson, and G. Tkach, Physical Vapour Deposited Adhesion and Barrier Layers for the ULSI Era, ULSI U.C. Berkeley Conference, Berkeley, CA, October 4–6, 1994.
[7]R. Beyers and R. Sinclair, J. Appl. Phys. **57**, 5240 (1985).
[8]S. P. Murarka and D. B. Fraser, J. Appl. Phys. **51**, 393 (1980).
[9]I. Suni, M. Blomberg, and J. Saarilahti, J. Vac. Sci. Technol. A 3, 2233 (1985).
[10]Y. Inoue, S.-I. Tanimoto, K. Tsujimura, T. Yamashita, Y. Ibara, Y. Yamashita, and K. Yoneda, J. Electrochem. Soc. **141**, 1056 (1994).
[11]M. Wittmer, Appl. Phys. Lett. **37**, 540 (1980).
[12]J. M. Poitevin, D. Maheo, and G. Lemperiere, Thin Solid Films **176**, 7 (1989).
[13]C. Y. Ting and M. Wittmer, Thin Solid Films **96**, 327 (1982).
[14]L. Ouellet, Y. Tremblay, G. Gagnon, M. Caron, J. F. Currie, S. C. Gujrathi, M. Biberger, and R. Reynolds, J. Appl. Phys. **79**, 4438 (1996).
[15]M. Hansen, *Constitution of Binary Alloys* (McGraw-Hill, New York, 1958), p. 141.

Effects of oxide overlayer on thermal stress and yield behavior of Al alloy films

I.-S. Yeo,[a] S. G. H. Anderson,[b] D. Jawarani,[b] and P. S. Ho
Center for Materials Science and Engineering, University of Texas at Austin, PRC/MER Mail Code R8650, Austin, Texas 78758-1100

A. P. Clarke and S. Saimoto
Department of Materials and Metallurgical Engineering, Queen's University, Ontario, Canada K7L 3N6

S. Ramaswami and R. Cheung
Advanced Micro Devices, Sunnyvale, California 94088

(Received 11 January 1996; accepted 13 May 1996)

The stress and yield behavior of Al alloys, Al(Cu) and Al(SiCu), have been investigated using bending beam and x-ray diffraction techniques, and the effect of passivation on the stress–temperature behavior of alloys is examined. A stress analysis of trilayer (SiO_2/Al alloy film/Si) structure was formulated in order to deduce the stress of the confined metal films from bending beam measurements. The stress–temperature behavior of Al alloy films obtained from bending beam techniques was in agreement with that determined directly by x-ray diffraction techniques. The grain structure and precipitate morphology of these alloy films were studied using transmission electron microscopes. Presence of passivation leads to a change in microstructure of Al alloys, resulting in hardening for the Al(SiCu) films and softening for the Al(Cu) films when compared with the unpassivated films at room temperature. Our results reveal that oxide thickness, Al thickness, and microstructure play an important role in determining the plastic deformation behavior of the metal films, and can satisfactorily explain the contrasting behavior observed for Al(SiCu) and Al(Cu) films. © *1996 American Vacuum Society.*

I. INTRODUCTION

Metal films deposited on rigid substrates are considerably stronger than their bulk forms and show higher yield stresses during thermal cycling.[1,2] Proposed strengthening mechanisms have focused on dislocation pinning by obstacles associated with precipitates and film/substrate interface. Oxide films are commonly used to cover metal films and lines in multilevel wiring for on-chip interconnects. The overlayer passivation is expected to further strengthen the metal films and lines due to additional constraint imposing on dislocation motion. Various models have been developed to account for the yield behavior of passivated films. One specific model proposed for Pb films[3] considered the energetics of dislocation motion, taking into account the elastic strain energy of the film and the dislocation line energy at interfaces and grain boundaries, where film thickness, grain size, and overlayer thickness were parameters. This model was subsequently extended to consider the dislocation behavior and density observed by transmission electron microscope (TEM).[4] Recently, similar approaches have been applied to model the yielding behavior of Al alloy films used in integrated circuits.[5–9]

Passivation films are usually deposited at temperatures about 100–200 °C higher than those of the metal films. The elevated temperature often changes the microstructure of the metal film, for example, by grain growth or precipitate redistribution. Different microstructure can lead to change of the stress behavior of the film, particularly its yield characteristics. The passivation film and its effect on microstructure have not yet been considered in the previous studies.

In this article, we first report the results of stress measurements using bending beam and x-ray diffraction techniques of Al(Cu) films with and without passivation. These two techniques are expected to give the same results, and the experiments have been designed to include the stress of both metal film and passivation. In applying the bending beam technique to a bilayer film structure, the stress of the individual films has to be extracted from the combined deformation of the metal film and the passivation using an elastic, perfect plastic stress model.[10] Such a stress analysis is not required for x-ray techniques since it directly measures the strain of the metal film. Therefore, the x-ray results are useful for checking the bending beam results and for validating the stress analysis. After validation, one can deduce the stress of the passivated metal film from bending beam measurements.

We investigate the effects of passivation on the stress and yield behavior of two different alloy films, Al(Cu) and Al(SiCu). These two systems are chosen for our study because they are commonly used for on-chip interconnects. The results show that the passivation increases Al(SiCu) film stresses but decreases the Al(Cu) film stresses at room temperature. The contrasting behavior of the two Al alloy films can not be attributed solely to the constraint effect due to the passivation overlayer; the role of the microstructure must be

[a]Present address: Hyundai Electronics, San 136-1, Ami-ri, Budal-eub, Ichon-Si, Kyoungki-do, 467-701, Korea; Electronic mail: isyeo@ccwf.cc.utexas.edu
[b]Present address: Advanced Products Research and Development Laboratory, Motorola, 3501 Ed Bluestein Blvd, Austin, TX 78721.

taken into account. Yield behavior of the two films is analyzed by using a film strength model formulated by Chaudhari,[3] which quantifies the effect of grain size and passivation thickness on the yield strain of the film. In order to provide input for this analysis, TEM is used to determine changes in grain size and precipitate formation of the alloy films before and after deposition of the passivation. This model enables us to understand the behavior of the two Al alloy films by assessing the relative importance of microstructure versus passivation thickness in determining their yield strength.

II. EXPERIMENTAL DETAILS

Silicon wafers (625 μm thick) were oxidized (200 nm thick) to prevent reaction between metals and the Si substrate. For passivated Al(Cu) samples, a 0.6 μm thick Al(1 wt % Cu) film was sputter deposited at about 175 °C onto a Si substrate. The metal film was then encapsulated with a 0.4 μm thick phosphorus-doped silicate glass (PSG) followed by 0.7 μm plasma enhanced CVD (PECVD) SiN_x deposited at 400 °C. An Al(Cu) and a SiN_x/PSG film were also deposited on different wafers to supplement the study of the bilayer films. For passivated Al(SiCu) samples, Al(1 wt%Si–0.5%Cu) film was deposited by sputtering at 250 °C to a thickness of 0.4 μm. The tetraethyl-orthosilicate (TEOS)-based PECVD SiO_2 film (2.0 μm thick) was subsequently deposited at 400 °C. Single-layer films of Al(SiCu) and TEOS SiO_2 were deposited onto different wafers for comparisons. All samples were annealed above 400 °C to homogenize the composition of Al alloy and to provide a common point from which comparisons could be made. The back side of wafers was thinned to a thickness of ~525 μm in order to improve the sensitivity of the measurements and polished to provide a mirror surface for curvature measurements. This process step left the protective oxide only on the front side of the samples. Afterward, these samples were cut into 3×45 mm strips using a wafer dicing machine.

Thin film stresses were measured as a function of temperature by a bending beam technique, using a modified version of a system described in detail elsewhere.[11] After a stress–temperature measurement was complete, the film(s) was etched off the Si substrate to determine the curvature corresponding to zero stress. The effect of the protective oxide on the stress measurement was eliminated by measuring the radii of curvature of the etched sample with the oxide through the whole thermal cycle and subtracting the measured radii from those measured with passivation/Al films. The stress variation versus temperature observed during the first thermal cycle between 25 and 400 °C usually reflected microstructure changes occurring in the film; after the first cycle, results became reproducible for subsequent thermal cycles. The heating and cooling rate employed for thermal cycling was 5 °C/min. The stress vs temperature data shown in this article have been measured up to three times, and show comparable stress values.

X-ray measurements were carried out using a diffractometer equipped with a high-intensity rotating anode source, which was designed for fast measurements of triaxial stresses in thin, patterned films as a function of temperature. High intensity low index diffraction planes were used to reduce acquisition time. The unpassivated and passivated Al(Cu) films used for x-ray measurements were prepared under conditions identical to those used for the bending beam samples. Details of the x-ray measurements have been published earlier.[12]

III. RESULTS AND DISCUSSION

A. Al(Cu) films

Bending beam measurements were first carried out on the individual SiN_x/PSG and Al(Cu) films deposited on Si substrates. For a single-layer film, its stress is related to the measured radius of curvature of a beam by Stoney's equation[13]

$$\sigma_2 = \frac{E_1}{6(1-\nu_1)} \frac{t_1^2}{t_2 R}, \tag{1}$$

where $E_1/(1-\nu_1)$ is the biaxial modulus of substrate, t_1 and t_2 are substrate and film thicknesses, and R is the radius of curvature of the substrate measured from film side. Stresses in the SiN_x/PSG (treated as a single-layer film) and Al(Cu) film are evaluated using Eq. (1), and results are shown in Figs. 1(a) and 1(b), respectively. The stress–temperature data of SiN_x/PSG in Fig. 1(a) shows a linear dependence with small stress hysteresis, and thus reveals an elastic behavior of the film in the temperature range studied. In Fig. 1(b), the stress in the unpassivated Al(Cu) film shows an elastic behavior upon heating from room temperature to about 200 °C. Above 200 °C, stress reduction via plastic flow dominates until the end of the heating cycle at 400 °C. Upon cooling from 400 °C, the film shows a brief elastic region before undergoing plastic yield until about 150 °C. Below 150 °C, the stress–temperature curve exhibits a distinct change in slope due to film hardening. Similar results have been reported by other investigators.[5,14]

Bending beam measurements were carried out on the (SiN_x/PSG)/Al(Cu) bilayer on Si and the observed stress versus temperature behavior is shown in Fig. 1(c). The stress values shown here are the average stress $\langle\sigma\rangle$ for the composite layer determined according to the following extended Stoney's equation:

$$\sigma_2 t_2 + \sigma_3 t_3 \equiv \langle\sigma\rangle(t_2+t_3) = \frac{E_1 t_1^2}{6(1-\nu_1)R}, \tag{2}$$

where σ_2 and σ_3 are stresses in the overlayer films of thickness t_2 and t_3, $E_1/(1-\nu_1)$ is the biaxial modulus of substrate, R is the substrate radius measured from film side. Thickness of the substrate is much larger than those of the films. Bending beam techniques measure the stress of composite layer, hence a detailed stress analysis for the compos-

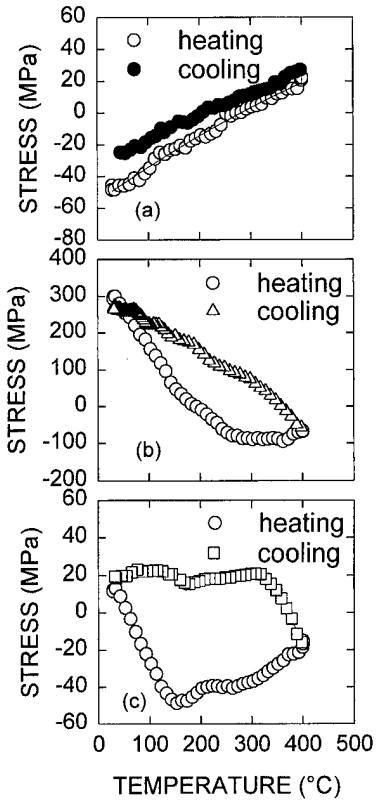


FIG. 1. Stress vs temperature curves for (a) a 0.7 μm/0.4 μm SiN_x/PSG film, (b) a 0.6 μm Al(Cu) film, and (c) a 0.7 μm/0.4 μm (SiN_x/PSG)/0.6 μm Al(Cu) bilayer structure.

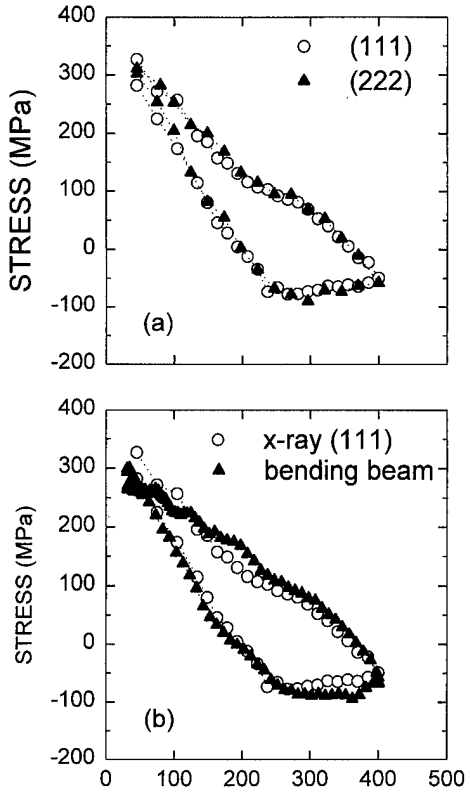


FIG. 2. (a) Stress vs temperature curve for an unpassivated Al(Cu) film obtained by x-ray diffraction techniques, stress data obtained using (111) and (222) reflections are compared and good agreement is observed (b) comparison of the stress versus temperature curve for the unpassivated Al(Cu) film obtained by x-ray diffraction techniques and bending beam techniques. Qualitative agreement as shown validates the bending beam method used for stress measurements in this study.

ite layer is required to deduce individual layer stresses, and the analysis will be presented in a later section.

The stress vs temperature behavior of the unpassivated Al(Cu) film was also measured by x-ray diffraction and the results are shown in Fig. 2(a) where stress data obtained using (111) and (222) reflections are compared. Good agreement between the corrected (111) data and the (222) data validates the alignment correction procedures developed for the x-ray measurements because (222) reflection gives higher reflection angle and more accurate stress value at the expense of measurement time.[12] Figures 1(b) and 2(a), which are compared in Fig. 2(b), show a quantitative agreement between the stresses determined by bending beam and x-ray techniques, and thus validates the equivalence of these two methods used for stress measurements. Texture of the Al films shows strong (111) orientation.[12]

Stress analyses for bilayers of elastic materials have been developed by several authors, and extended to multilayer films.[15–18] Some of these studies have concluded that stresses in multilayer films can be adequately described by a weighted average of stresses in individual films.[16,17] Since these analyses were limited to elastic materials, and may not be applicable to passivated metal films where metal film deforms plastically, we have developed a stress model for the trilayer (SiO_2/Al/Si substrate) structure, in which the interlayer film is assumed to deform in an elastic, perfect plastic manner.[10] Calculations based on this model have been performed for a 2 μm SiO_2/0.4 μm Al/500 μm Si trilayer structure, and the resulting stresses for the individual layers and the composite bilayer stresses are shown in Fig. 3. The parameters used in the calculation are given in Table I. The critical temperature for the onset of a plastic deformation is assumed to be 200 °C. The intrinsic stress values are given at 400 °C for computational simplicity, and correspond to the number necessary to shift the Al and oxide curves up or down to resemble the stresses in the elastic regime. As designed, the stress vs temperature behavior of the Al exhibits linearity between room temperature and 200 °C, whereupon further increases in temperature induce almost no change in the Al stress. (Note that the stress–temperature slope of Al film above 200 °C is not zero, since the presence of the oxide does influence stress in the Al film slightly.)

Significantly, the stress of the oxide is nearly unaffected by the presence of the plastically deforming Al layer. (However, it should be noted that plastic deformation of the Al layer is affected by the presence of the overlayer oxide.) The stress calculated for a 2 μm oxide deposited directly on Si is nearly coincident with the stress in the oxide deposited on Al, as shown in Fig. 3. Indeed, separate calculations for vari-

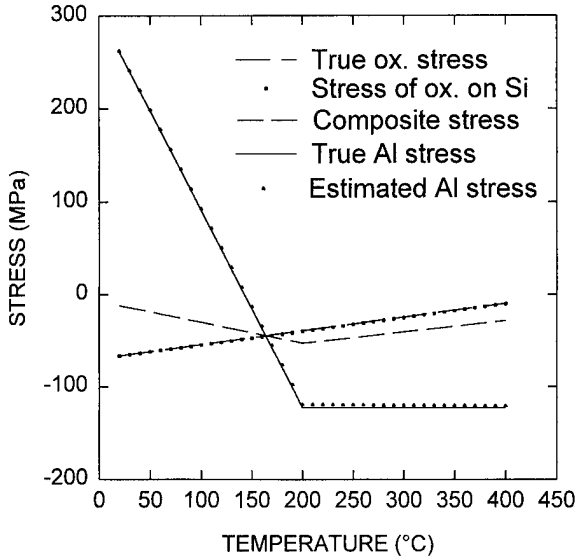


FIG. 3. Calculated stress–temperature curves for a SiO_2/Al bilayer structure on Si substrate. Al film is assumed to deform elastic, perfect plastic manner (no strain hardening) upon 200 °C.

ous combinations of oxide and metal bilayer reveal that more than an order of magnitude difference in the thickness ratio is required before the stress in the oxide on the deforming Al is significantly different from that calculated for an oxide deposited directly on Si. Thus, stresses in multilayer films with comparable respective thicknesses can be still described by a weighted average of stresses in individual films even when one of the films experiences plastic deformation. This stress superposition principle can be applicable to many stress-related problems in microelectronics, and the stress of passivated metal films can be deduced from bending beam measurements by subtracting the stresses in passivation layer.[10]

The deduced stress behavior of the passivated Al(Cu) film in the bilayer is shown in Fig. 4(a). X-ray diffraction measurements were performed to directly measure the stress of the passivated Al(Cu) film and the results are compared with the bending beam result shown in Fig. 4(a). These data show a good agreement between the results from both measurements. Since x-ray diffraction directly measures the strain (which can be converted to the stress) of the passivated metal layer, the observed agreement validates the stress analysis for the passivated metal film.

In Fig. 4(b), stresses of the unpassivated film are shown along with those of the passivated Al(Cu) film. During heating, both films exhibit an elastic regime, although the passivated film starts exhibiting plastic flow at a lower temperature (140 °C) than the unpassivated film (200 °C). Upon cooling from 400 °C, the passivated film is elastic until about 350 °C, and it deforms plastically until room temperature. Unlike the unpassivated sample, however, it does not show any significant change in slope due to hardening upon cooling below 150 °C. These results are discussed below in the light of microstructural observations.

TABLE I. Parameters used in the trilayer calculations (Ref. 25).

Material	Silicon	Aluminum	Silicon dioxide
E (GPa)	130	69	71.7
ν	0.28	0.33	0.16
Thickness (μm)	530	0.4	2.0
α ($\times 10^{-6}$)	2.14	23.4	0.94
σ^{int} (MPa)	0.0	−550	−10

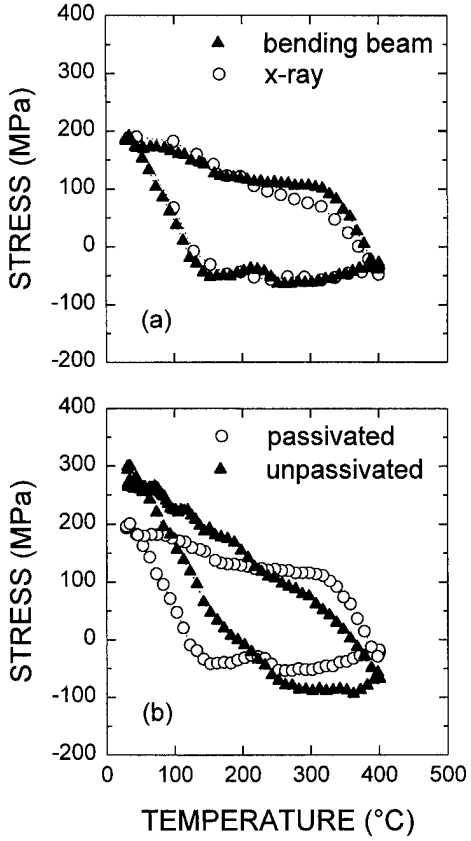


FIG. 4. (a) Comparison of passivated Al(Cu) stresses obtained by bending beam techniques and x-ray techniques, (b) comparison of the stress–temperature behavior for the passivated Al(Cu) film and for the unpassivated Al(Cu) film. The differences can be explained by a model which calculates the relative contributions of oxide thickness and grain size to the film yield stress.

B. Al(SiCu) film

The stress–temperature behavior of the individual 2 μm SiO_2 and 0.4 μm Al(SiCu) films deposited on Si are shown in Figs. 5(a) and 5(b). The SiO_2 film exhibits a smooth, nearly linear stress response as a function of temperature. With almost no hysteresis during thermal cycling up to 400 °C, our results differ from an earlier report for TEOS-based SiO_2 with a strong stress hysteresis.[19] The Al(SiCu) film behaves elastically upon heating between 25 and 220 °C as shown in Fig. 5(b). Above 220 °C, the film starts yielding in compression. Upon cooling from 400 °C, the film behaves

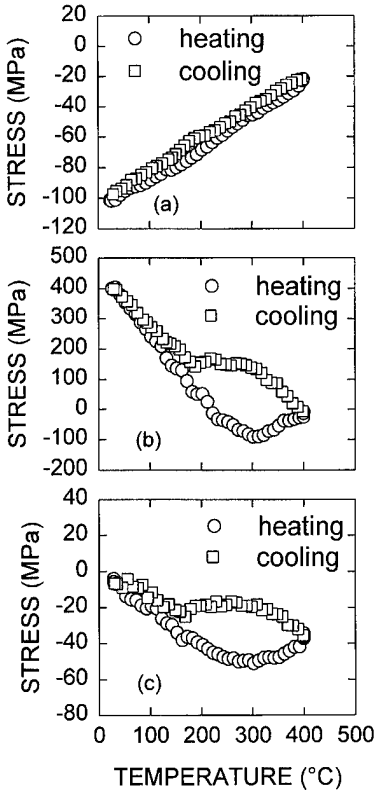


FIG. 5. Stress vs temperature curves for (a) a 2.0 μm TEOS SiO_2 film, (b) a 0.4 μm Al(SiCu) film, and (c) a 2.0 μm SiO_2/0.4 μm Al(SiCu) bilayer structure.

elastically until it reaches a stress of ~100 MPa at 330 °C, and then it begins to yield in tension. As the film is cooled below 180 °C, film strengthening occurs, and the Al(SiCu) film behaves almost elastically. Similar results for Al(SiCu) have been reported elsewhere by other groups.[5,20] In Fig. 5(c), we present a stress–temperature curve for a SiO_2/Al(SiCu) bilayer film. Note the stress–temperature curve shift toward compression because of the compressive thick oxide film.

We compare the stresses of two Al(SiCu) films, one passivated by a 2 μm oxide and the other unpassivated, in Fig. 6. The results for the unpassivated film are reproduced from Fig. 5(b). Those for the passivated Al(SiCu) have been determined by the linear subtraction procedure given in the Appendix. These two sets of data exhibit similar slopes between 25 and ~230 °C, where both oxide and metal films behave elastically. However, above ~230 °C, the two sets of data are quite different. Upon heating, the stress in the passivated film flattens out above ~300 °C while the compressive stress in the unpassivated film begins to decrease at ~300 °C continuing to 400 °C. This suggests that the presence of the passivation layer reduces the rate of plastic flow under compression in Al(SiCu) film, and leads to a higher yielding temperature for the confined film (~280 versus ~230 °C). When samples are cooled from 400 °C, both sets of data exhibit a linear range to ~300 °C, a characteristic of elastic behavior. However, between ~300 and 180 °C, the

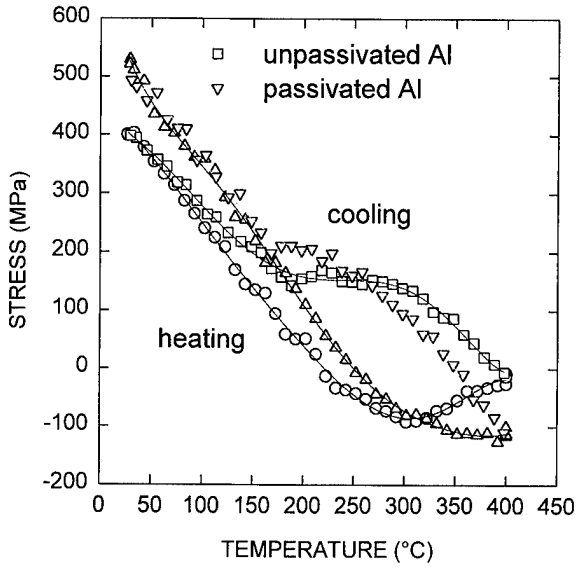


FIG. 6. Comparison of the stress–temperature behavior for a passivated 0.4 μm Al(SiCu) film and for a unpassivated 0.4 μm Al(SiCu) film. Stresses determined for passivated and unpassivated Al(Cu) films exhibit a trend opposite from those shown in Fig. 4.

two sets of data exhibit slopes distinct from one another where Al(SiCu) is yielding in tension. Below ~180 °C, both the unpassivated and the passivated Al(SiCu) films begin to harden. Our data indicate that the passivated Al(SiCu) films exhibit higher compressive stresses upon heating and higher tensile stresses upon cooling than the unpassivated Al(SiCu) films.

C. TEM examination of microstructure

Comparing with the stress measurements of their respective unpassivated films, the passivated Al(SiCu) films show a higher stress, but the passivated Al(Cu) films show a lower stress at room temperature. The contrasting behavior may not

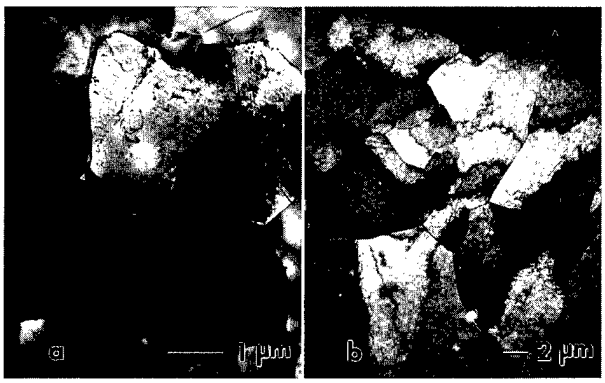


FIG. 7. Microstructure of (a) the unpassivated Al(Cu) film and (b) the passivated Al(Cu) film in plan view TEM. In (a), a dark arrow points out slip traces, and a white arrow indicates an Al_2Cu (θ phase) precipitate. In (b), dark arrow indicates a θ phase precipitate. Interestingly, no slip traces are observed despite larger average grains.

FIG. 8. Microstructure of (a) the unpassivated Al(SiCu) film and (b) the passivated Al(SiCu) film in plan view TEM. In (a), a dark arrow indicates Al_2Cu (θ phase) precipitate, and an unfilled arrow indicates a Si precipitate. In (b), slip traces are marked by dark arrows, and precipitates show similar morphology to those in the unpassivated film.

be attributed only to the effect of overlayer constraint. Data suggest that microstructural parameters are also important.

The microstructure of Al(SiCu) films and Al(Cu) films was investigated by plan view TEM after annealing at 400 °C for 30 min. For passivated samples, passivation films were removed by reactive ion etching using CF_4 gas. Samples for TEM were prepared by thinning the wafer down to 150 μm, cutting 3 mm diam disks, and dimpling to about 50 μm. The rest of the Si substrate in a dimpled region of the sample was chemically etched off using a solution of 85% HNO_3 and 15% HF. The front side Al film was protected with an etch-resistant lacquer for the disk edge and with the protective oxide which also works as an etch stopper for the center region. The sample disk was finally ion milled until an electron-transparent region was reached. A JEOL 1200 EX STEM was used for the microscopy work.

Typical microstructure of Al(Cu) and Al(SiCu) films is shown in Figs. 7 and 8, respectively. The unpassivated Al(Cu) film [Fig. 7(a)] exhibits slip traces in several grains which suggest plastic deformation. Irregular lines inside some grains are dislocations generated during ion milling. Figure 7(b) shows the corresponding microstructure for the passivated Al(Cu) film. It is remarkable that, unlike the unpassivated film, no slip traces were visible for this film although it shows much larger grains. The reason for this is not clear. The small holes seen in this film are artifacts of sample preparation procedure involving etching and ion milling. Figures 8(a) and 8(b) show the microstructure of the unpassivated and passivated Al(SiCu) film, respectively. The microstructures of the films are quite similar to each other. The observations on the microstructure of these films are summarized in Table II.

The grain size of the unpassivated and passivated films was determined from the TEM micrographs by taking the mean of the smallest and the largest dimension of each grain. Figures 9(a) and 9(b) show the grain size distribution of the Al(Cu) film and Al(SiCu) film, respectively. Several observations can be made from Figs. 9(a) and 9(b). First, passivated films have larger grains as compared to those of unpassivated films. Second, unpassivated films exhibit a log-normal grain size distribution. This suggests that the grains of the unpassivated film grow normally. It also reveals that stagnation of normal grain growth because the slopes of the distributions are steeper for larger value of grain size. Third, passivated films show a lower slope which indicates wider spread in grain size. Furthermore, the passivated Al(Cu) films apparently show a bimodal distribution. Two distinct slopes seen in the distribution correspond to a subdistribution of small and recrystallized grains which have grown normally, and another subdistribution of large grains showing

TABLE II. Summary of microscopic analysis for Al(Cu) and Al(SiCu).

Material	Average grain size	Microstructure
Passivated Al(Cu)	6.3 μm	θ precipitates, slip traces
Unpassivated Al(Cu)	1.8 μm	θ precipitates, no slip traces
Passivated Al(SiCu)	2.3 μm	Few θ precipitates, Si precipitate, few slip traces
Unpassivated Al(SiCu)	1.9 μm	Few θ precipitates, Si precipitate, few slip traces

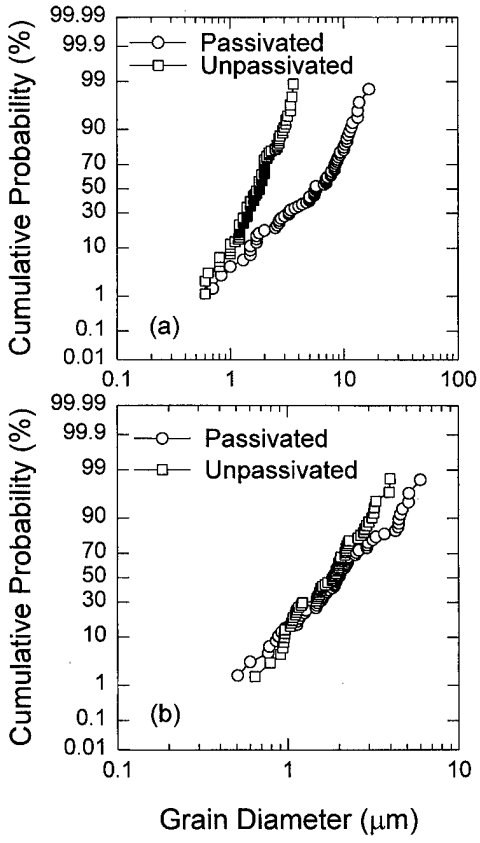


FIG. 9. Grain size distributions of the unpassivated/passivated (a) Al(Cu) film and (b) Al(SiCu) film. Passivation causes the increase in average grain size of the film.

abnormal growth of the pre-existing larger than average grains. The observed average grain size is also summarized in Table II. It should be noted that the passivated Al(SiCu) film has an average grain size similar to that of the unpassivated Al(SiCu) film, whereas the passivated Al(Cu) film has much larger grain size compared with the unpassivated Al(Cu) film.

The larger grain size in the passivated film may be attributed to two factors. First, the high temperature oxide deposition step promotes grain growth. Another contributing factor for larger grain size may be a compressive stress-driven grain growth when the Al alloy film is passivated.[21] Venkatraman *et al.*[6] showed that grain growth of an Al(Cu) film is influenced by the presence of oxide as well, i.e., grain size is significantly larger in the region where the film is supported by the layer of SiO_2 (hence is subjected to a compressive stress during heating) than in the region of the free standing film. Grain growth relaxes compressive stresses in the film, so grain growth is likely to occur under compressive stresses.[21] Since an overlayer passivation can restrict hillock growth, one of the primary mechanisms to relax compressive stress during heating, the grain growth of the passivated Al(Cu) is enhanced to relax the compressive stress.

D. Yield strengthening model for metal films

In a previous study of Pb films, a thin native oxide was found to enhance the film yield strength.[22] Since a native oxide is present over both the unpassivated and the passivated Al alloy films, the difference observed for the passivated and unpassivated films suggests that the yielding behavior depends on microstructure as well as passivation thickness. Results from TEM show distinct grain structure for the two Al alloy films. Effects of grain structure and overlayer constraint are taken into account by the film strength model formulated by Chaudhari.[3] Here, we apply this model to analyze the yield behavior observed for the Al alloy films. Since the plastic deformation of Al films can be explained by a dislocation glide mechanism through the thermal cycle,[23,24] Chaudhari's approach for the stress analysis is considered to be relevant.

In Chaudhari's model,[3] the energies per unit area for strain ϵ in the film (E_u), dislocation lines at film–substrate ($E_\perp^{fs}$) and film overlayer ($E_\perp^{fo}$) interfaces, and dislocations lying along grain boundaries ($E_\perp^{gb}$) were added together to obtain total energy (E_{tot}). By setting $dE_{tot}/d\epsilon=0$ (minimum energy criterion), the critical strain for dislocation motion (yielding) could be determined. For an isotropic film of thickness h with grains of diameter d, and covered by an overlayer oxide of thickness h', the total energy is

$$E_{tot}=E_u+E_\perp^{fs}+E_\perp^{fo}+E_\perp^{gb} \tag{3}$$

or

TABLE III. Material constants used in calculations (Refs. 25 and 26).

Material	E (GPa)	ν	μ (GPa)
Si	130	0.28	50.8
Al (at RT)	69	0.33	25.9
Al (at 316 °C)	49	0.33	18.42
SiO_2	71.7	0.16	30.9
PSG	79	0.2	32.9
SiN_x	150	0.25	60

$$\begin{aligned}E_{tot}=&Bh\epsilon^2+(\epsilon_f-\epsilon)2D\sqrt{3}\left(\frac{2}{3(1-\nu)}+\frac{1}{3}\right)\left[\ln\left(\frac{2h}{b}\right)+1\right]\\&+(\epsilon_f-\epsilon)2D'\sqrt{3}\left(\frac{2}{3(1-\nu)}+\frac{1}{3}\right)\left[\ln\left(\frac{2h'}{b}\right)+1\right]\\&+(\epsilon_f-\epsilon)\frac{hb\mu_f}{d}\frac{\sqrt{2}}{\pi}\left(2+\frac{1}{\sqrt{12}}\right.\\&\left.+\left(1+\frac{1}{\sqrt{3}}\right)\frac{1}{(1-\nu)}\right)\left[\ln\left(\frac{d}{b}\right)+1\right],\end{aligned} \tag{4}$$

where ϵ_f is the initial strain in the Al film before plastic deformation occurs, ν is the Poisson ratio of Al, $b=2.86$ Å is the length of the Burgers vector for dislocation motion along $\langle 110\rangle$ directions on a $\{111\}$ plane in Al. The constants B, D, and D' are given by

$$B=\frac{E_f}{2(1-\nu)},\quad D=\frac{b\mu_f\mu_s}{\pi(\mu_f+\mu_s)},\quad D'=\frac{b\mu_f\mu_0}{\pi(\mu_f+\mu_0)}, \tag{5}$$

where the μ_i are the shear moduli. For the material constants at room temperature and 316 °C given in Table III,[25,26] the critical strain for plastic deformation becomes

$$\begin{aligned}\epsilon_{cr}(\mathrm{RT})=&\frac{0.70}{h}\left[\ln\left(\frac{2h}{b}\right)+1\right]+\frac{0.57}{h}\left[\ln\left(\frac{2h'}{b}\right)+1\right]\\&+\frac{1.50}{d}\left[\ln\left(\frac{d}{b}\right)+1\right],\end{aligned} \tag{6}$$

$$\begin{aligned}\epsilon_{cr}(316\ ^\circ\mathrm{C})=&\frac{0.55}{h}\left[\ln\left(\frac{2h}{b}\right)+1\right]+\frac{0.57}{h}\left[\ln\left(\frac{2h'}{b}\right)+1\right]\\&+\frac{1.07}{d}\left[\ln\left(\frac{d}{b}\right)+1\right],\end{aligned}$$

where the h, h', and d are in units of angstrom. This model assumes isotropic materials and does not consider variations in grain orientation or alloy composition which are known to influence film strength.[2–4,8] It is important to note that grain size, film thickness, and passivation layer thickness all contribute to ϵ_{cr}. For constant h and d, the presence of a thick overlayer (h') will increase ϵ_{cr} although the amount may be small because of the logarithmic dependence. In contrast, if the grain size d increases due to the presence of the over-

TABLE IV. Film and passivation thicknesses and grain size determined by TEM for the different samples. These parameters predict a critical strain for yielding ϵ_{cr} at room temperature and at 316 °C. $\Delta\epsilon_{cr}$ is obtained from $\epsilon_{cr}(RT)-\epsilon_{cr}(316\ °C)$, and represents the extent of hardening during cooling.

Sample	h (Å)	h' (Å)	d (Å)	ϵ_{cr}(RT)	ϵ_{cr}(316 °C)	$\Delta\epsilon_{cr}$
Unpassivated Al(SiCu)	4000	20	18 730	2.87×10^{-3}	2.31×10^{-3}	0.56×10^{-3}
Passivated Al(SiCu)	4000	20 000	22 950	3.72×10^{-3}	3.20×10^{-3}	0.62×10^{-3}
Unpassivated Al(Cu)	6000	20	18 220	2.31×10^{-3}	1.85×10^{-3}	0.46×10^{-3}
Passivated Al(Cu)	6000	11 000	62 520	2.50×10^{-3}	2.20×10^{-3}	0.30×10^{-3}

layer, ϵ_{cr} may decrease. Since overlayer causes increase of grain size, the two terms-overlayer thickness and grain size-balance each other.

Several groups have recently discussed the interplay between pairs of the three terms in Eq. (6). For example, Doerner *et al.*[5] and Venkatraman *et al.*[6,7] have correlated changes in film strength with film thickness and grain size. Thompson[9] has modeled this behavior and derived an equation that is similar in form to the sum of the first and third terms. For a given pair of passivated and unpassivated films, the present study keeps the first term in Eq. (6) constant, which allows us to examine the interplay between the effects of overlayer oxide thickness and grain size. The critical strains for plastic deformation predicted by Eq. (6) are given in Table IV. Unpassivated films are assumed to have a 20 Å native oxide and the grain sizes listed were measured via TEM. The critical strains obtained for each of the Al films studied are reasonable, varying between 0.19% and 0.37%. We note higher yield stress for Al(SiCu) films than that of the Al(Cu) films and higher yield stress at lower temperature. This is consistent with the measurements although the obtained critical strains of Al(Cu) and Al(SiCu) cannot be quantitatively compared with each other because of the difference in alloy composition. This model also correctly predicts an increase of yield strain for passivated films at high temperature (at 316 °C) although it fails to predict a lower yield strain for the passivated Al(Cu) film than that for the unpassivated Al(Cu) film at room temperature. When we compare $\Delta\epsilon_{cr}$ of different film in the Table IV, the model predicts the hardening behavior correctly [larger hardening for passivated Al(SiCu) and less hardening for passivated Al(Cu) compared with their respective unpassivated films]. The disagreement between the measurement and the modeling for Al(Cu) films at room temperature could be attributed to the difference of reference points for the model. We conclude that Chaudhari's model successfully predicts the trend in the yield behavior observed for Al alloy films. In the Al(SiCu), grain size increase can be negligible because the grain sizes of the unpassivated and the passivated film are similar. Thus, strengthening due to overlayer confinement is more effective. While, in the Al(Cu), softening due to grain size increase is greater than the strengthening due to overlayer confinement, and a decrease in room temperature stress in the passivated Al(Cu) is observed.

IV. CONCLUSION

In this article, we have investigated the stress–temperature behavior of unpassivated and passivated Al alloy films. We developed a method to deduce the stress of the confined metal films from bending beam measurements. The bending beam results were compared with those from x-ray diffraction and a good agreement was observed. The Al alloy films show distinct stress–temperature behaviors, where passivation causes an increase in yield strength for the Al(SiCu) film but a decrease for the Al(Cu) film at room temperature. The contrasting behavior was analyzed by a film strengthening model which takes into account the effect of passivation thickness and grain structure on film yielding. This model is able to account for the yield behavior of the two alloy films during thermal cycling, and reveals two competing mechanisms: hardening due to overlayer confinement, and softening due to grain growth.

ACKNOWLEDGMENTS

This work was partially supported by the Semiconductor Research Corporation under Contract MJ-94-353 and the Texas Advanced Technology Program under Grant No. 156. Additional support from Motorola and Advanced Micro Devices is also acknowledged.

[1]M. Murakami, T.-S. Kuan, and I. A. Blech, Treat. Mater. Sci. Technol. **24**, 163 (1982), and references contained therein.
[2]W. D. Nix, Metall. Trans. A **20**, 2217 (1989), and references contained therein.
[3]P. Chaudhari, Philos. Mag. A **39**, 507 (1979).
[4]T. S. Kuan and M. Murakami, Metall. Trans. A **13**A, 383 (1982).
[5]M. F. Doerner, D. S. Gardner, and W. D. Nix, J. Mater. Res. **1**, 845 (1986).
[6]R. Venkatraman, J. C. Bravman, W. D. Nix, P. W. Davies, P. A. Flinn, and D. B. Fraser, J. Electron. Mater. **19**, 1231 (1990), and references contained therein.
[7]R. Venkatraman and J. C. Bravman, J. Mater. Res. **7**, 2040 (1992).
[8]J. E. Sanchez, Jr. and E. Arzt, Scri. Metall. Mater. **27**, 285 (1992).
[9]C. V. Thompson, J. Mater. Res. **8**, 237 (1993).
[10]S. G. H. Anderson, I.-S. Yeo, P. S. Ho, S. Ramaswami, and R. Cheung, in *Materials Reliability in Microelectronics III*, edited by K. P. Rodbell, W. F. Filter, P. S. Ho, and H. Frost [Mater. Res. Soc. Symp. Proc. **309**, 261 (1993)].
[11]S. T. Chen, C. H. Yang, F. Faupel, and P. S. Ho, J. Appl. Phys. **64**, 6690 (1988).
[12]A. P. Clarke, S. Saimoto, and P. S. Ho, in *Stress-Induced Phenomena in Metallization*, edited by P. S. Ho, C. Y. Li, and P. Totta [Am. Inst. Phys. Conf. Proc. **305**, 126 (1994)]; A. P. Clarke, Ph.D. thesis, Queen's University, Canada (1993).
[13]G. G. Stoney, Proc. R. Soc. London Ser. A **82**, 172 (1990).

[14]M. A. Moske, P. S. Ho, D. J. Mikalsen, J. J. Cuomo, and R. Rosenberg, Mater. Res. Soc. Symp. Proc. **203**, 77 (1991).
[15]Z-C. Feng and H-D. Liu, J. Appl. Phys. **54**, 83 (1983), and references contained therein.
[16]P. H. Townsend, D. M. Barnett, and T. A. Brunner, J. Appl. Phys. **62**, 4438 (1987).
[17]P. A. Flinn, D. J. Gardner, and W. D. Nix, IEEE Trans. Electron Devices **ED-34**, 689 (1987).
[18]H. C. Liu and S. P. Murarka, J. Appl. Phys. **72**, 3458 (1992).
[19]K. Ramkumar and A. N. Saxena, J. Electrochem. Soc. **139**, 1437 (1992).
[20]B. L. Draper and T. A. Hill, J. Vac. Sci. Technol. B **9**, 1956 (1991).
[21]P. Chaudhari, J. Vac. Sci. Technol. **9**, 520 (1972).
[22]M. Murakami and P. Chaudhari, Thin Solid Films **46**, 109 (1977); M. Murakami, Acta. Metall. **26**, 175 (1978).
[23]Venkatraman, Ph.D. thesis, Stanford University (1992).
[24]C. A. Volkert, C. F. Alofs, and J. R. Liefting, J. Mater. Res. **9**, 1147 (1994).
[25]R. E. Jones, Jr., IEEE Int. Relat. Phys. Symp. Proc. **87CH**, 9 (1987).
[26]A. I. Sauter, Ph.D. thesis, Stanford University (1991).

Effects of insulator surface roughness on Al-alloy film properties and electromigration performance in Al-alloy/Ti insulator layered interconnects

Hiroshi Onoda, Tadashi Narita, Kenshin Touchi, and Keiichi Hashimoto
VLSI Research and Development Center, Oki Electric Industry Co., Ltd., 550-1 Higashi-asakawa, Hachioji, Tokyo 193, Japan

(Received 22 December 1995; accepted 10 May 1996)

Effects of insulator surface roughness on the overlying aluminum alloy film properties and electromigration performance in Al alloy/Ti insulator layered interconnects have been investigated. Insulator surface roughness changes the roughness and crystallographic orientation of Al/Ti layered films formed on the insulator. In particular, the effect is prominent in the case of high-temperature sputtering of Al alloy films. A rough insulator surface deteriorates the roughness and crystallographic orientation of the Al alloy film formed on the insulator, while a smooth insulator surface improves those properties. The insulator surface roughness does not change the roughness and crystallographic orientation of the Al alloy film without a Ti underlayer. Ti roughness and crystallographic orientation are improved on a very smooth surface insulator resulting in improvement in the roughness and crystallographic orientation of the Al alloy film formed on the Ti layer. As a result, the electromigration performance in Al/Ti layered interconnects changes according to the underlayer insulator surface roughness. © *1996 American Vacuum Society.*

I. INTRODUCTION

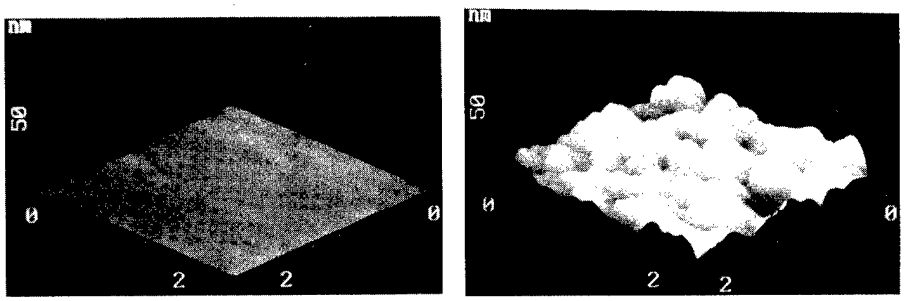
In the trend of increasing packing density of large-scale integrated circuits (LSIs), two major concerns in interconnects used in LSIs are how to obtain electrical conductance at contacts or via holes with high aspect ratios and how to obtain fine interconnects with high reliability. The former concern is partly solved by via filling with metals other than Al, for example, W, however, Al via-filling technologies such as high-temperature sputtering,[1–6] reflow process,[7,8] and high-pressure technology[9] have been investigated with respect to the low cost and simplicity of the fabrication processes. In particular, Al high-temperature sputtering is one of the most promising candidates because of its low cost without any special equipment. This technology, however, has a problem with Al surface roughness that will make the following photolithography process difficult.

As for the latter concern, there are many factors that affect the fine interconnect reliability. Those factors are metal line dimensions that determine the allowable passage current density, Al grain size, Al crystallographic orientation, and impurities contained in the Al alloy. In addition, in layered interconnects, there are other factors that make the electromigration performance complicated, such as the kinds of refractory metals and their fabrication processes, their resistivity, contact resistance to the Al alloy film, and their impact on Al grain size, Al crystallographic orientation, and impurities that will be incorporated into the Al alloy from the refractory metals. Those factors need to be optimized to obtain a highly reliable interconnect. With respect to the crystallographic orientation among those factors, Vaidya *et al.* suggested that the median time to failure (MTTF) of an Al metal line is a function of the Al grain size, the log-normal grain-size standard deviation, and the Al's preferred orientation, considering the electromigration (EM) performance and properties of the aluminum–copper (Al–Cu) alloys.[10] That is, the EM lifetime is longer in Al metal lines having a large grain size, a small grain-size standard deviation, and a strong ⟨111⟩ orientation. Kaneko *et al.* suggested that Al crystallographic orientation also affects stress-induced migration.[11] This is the result for single layer Al alloy metal lines. A lot of effort has been made to enhance the Al texture changing the underlying materials in layered interconnects.[12–16] Authors have reported that a highly ⟨111⟩-oriented Al metal line by controlling the underlying TiN preferred orientation has an excellent EM performance.[12,13] Other researchers also show the crystallographic effects of under-metal planes on Al (111) orientation in Al/Ti and Al/TiN/Ti structures changing the pretreatment and Ti thickness.[14,15] Those studies use the crystallographic relations with underlayer metals, while Toyoda *et al.* show hyper-textured ⟨111⟩-oriented Al alloy with a high EM performance on the amorphous underlayer that has a high surface energy.[16]

In this article, effects of insulator surface roughness on the overlying aluminum alloy layered interconnects have been investigated, in order to solve both problems of Al alloy surface roughness and reliability issues in Al alloy layered interconnects formed by high-temperature sputtering. It was found that insulator surface roughness changes the Al surface roughness and crystallographic orientation, in particular, in Al alloy/Ti layered interconnects formed by high-temperature sputtering. As a result, Al alloy/Ti layered interconnects formed on extremely planarized insulator have a highly reliable EM performance.

II. EXPERIMENT

In order to change the insulator surface roughness, thermal Si oxide film (500 nm), borophosphosilicate glass film (BPSG, 600 nm), and tetra-ethyl-ortho-silicate-O_3 (TEOS-O_3) based SiO_2/plasma SiO_2 double layered films were formed or deposited on 6 in. Si substrates. For the



(a) (b)

FIG. 1. The typical bird's eye view of the surface roughness on chemical mechanical polished TEOS-O_3 SiO_2/plasma SiO_2 (CMP TEOS-O_3SiO_2) (a) and nontreated TEOS-O_3SiO_2/plasma SiO_2 (TEOS-O_3SiO_2); (b) observed by AFM. The observed area is 2 μm×2 μm.

double layered films, TEOS-O_3 SiO_2 films of 800 and 1000 nm were deposited on 200 nm plasma SiO_2 on 6 in. Si substrates. Before the 800 nm TEOS-O_3 SiO_2 film deposition, a part of the samples are subjected to ethanol treatment.[17] This treatment enables one to obtain a very planarized TEOS-O_3 SiO_2 film. The TEOS-O_3 SiO_2 films of 1000 nm were subjected to chemical mechanical polishing (CMP) to remove 200 nm of TEOS-O_3 SiO_2 and planarize. The CMP was carried out with fumed silica base slurry and a pressure of 7.5 psi. The polishing rate is 196 nm/m. The remaining TEOS-O_3 SiO_2 film thickness is 800 nm after CMP. Another sample with TEOS-O_3 film (800 nm)/plasma SiO_2 film was annealed at 850 °C for 30 min for degassing before following metal film deposition. On these substrates, Ti (50 nm) and Al-0.8 wt% Si-0.3 wt% Cu alloy (500 or 700 nm) films were sequentially sputter deposited in a multichamber dc-magnetron sputtering system that has a base pressure of 2×10^{-8} Torr. Prior to the deposition, samples were subjected to preannealing at 450 °C for 2 min in the same vacuum for degassing. The substrate temperatures during deposition are 100 °C for Ti film and 220–485 °C for Al alloy films. As references, samples, on which Al alloy films were deposited without substrate heating and reflowed in the same vacuum in the temperature range of 420–485 °C, were prepared.

The surface roughness of the insulators was evaluated by atomic force microscopy (AFM). The surface morphology and reflectance of the Al alloy surface were evaluated by scanning electron microscopy (SEM) and reflectance measurement using *i* line (wavelength: 365 nm), respectively. The Ti and Al alloy crystallographic orientations and Al alloy/Ti reactions were evaluated with x-ray diffraction (XRD) analysis and transmission electron microscopy (TEM). The impurity concentration in the Al alloy film was analyzed by secondary ion mass spectroscopy (SIMS).

For EM evaluation, the Al (500 nm)/Ti (50 nm) layered metal lines 2 μm wide and 22 mm long with TiN(100 nm) as antireflection materials and plasma Si_3N_4 (800 nm) film as a passivation layer were fabricated on TEOS-O_3 film (800 nm)/plasma SiO_2 film with and without CMP treatment. The Al alloy film deposition temperature is 450 °C. They were subjected to a continuous dc current density of 3×10^6 A/cm^2 in the temperature range of 240–260 °C. Metal line resistance change is monitored during the EM tests. Finally, after the EM tests, samples were torn off from the package, and were exposed to CF_4+O_2 plasma in order to remove the passivation films and the capped TiN films, followed by optical microscope and secondary electron microscope observations.

III. RESULTS AND DISCUSSIONS

A. Effects of insulator surface roughness on Al film properties

First, in this section, the surface roughness of various insulators that were formed or deposited on Si substrates will be shown. Their effects on the surface roughness and crystallographic orientations of the overlying Al alloy layered metal film will be presented next.

Figure 1 shows the typical bird's eye view of the surface roughness on (a) chemical mechanical polished TEOS-O_3 SiO_2/plasma SiO_2 (CMP TEOS-O_3 SiO_2) and (b) nontreated TEOS-O_3 SiO_2/plasma SiO_2 (TEOS-O_3 SiO_2) observed by AFM. The observed area is 2 μm×2 μm. It is well known that CMP technology is usually used in the macroscopic planarization of interlayer dielectrics in LSI fabrication processes. In addition to the macroscopic planarization, it is clearly shown in Fig. 1 that CMP TEOS-O_3 SiO_2 has a very smooth surface down to the nm level compared to that of TEOS-O_3 SiO_2. In Table I, the roughness factors root mean square (rms) in various insulator surfaces that are calculated from observed results of AFM are summarized. Table I also includes the properties of Al films formed on those insulators. TEOS-O_3 SiO_2 means TEOS-O_3 SiO_2/plasma-TEOS SiO_2 double layer film. CMP TEOS-O_3 SiO_2 has the rms value of 0.48 nm while TEOS-O_3 SiO_2 has that of 12.20 nm. TEOS-O_3 SiO_2 films with and without CMP treatment show a remarkable difference in their roughness. Thermal oxide has the smoothest surface and TEOS-O_3 SiO_2 has the roughest surface.

The usual reflectance from samples prepared by Al high-temperature sputtering is very low because of the rough surface of deposited Al film, and this makes the alignment in the following lithography process difficult. Figure 2 shows (a) the *i*-line (365 nm) reflectance and (b) Al (111) peak inten-

TABLE I. Al alloy film properties on underlying insulators that are used for the experiment.

Underlayer insulator	Roughness (rms:nm)	450 °C deposition: Al reflectance (%)	450 °C deposition: Al grain size (μm)	450 °C deposition: Al(111) intensity (counts)	450 °C deposition: Al(111) FWHM (degree)	220 °C deposition: Al reflectance (%)	220 °C deposition: Al grain size (μm)	220 °C deposition: Al(111) intensity (counts)	220 °C deposition: Al(111) FWHM (degree)
TEOS-O_3 SiO_2	12.20	32.5	1.17	830	10.63[b]	71.1	0.56[b]	961	8.75[b]
TEOS-O_3 SiO_2 (degassed)	13.56	43.9	1.39	390					
TEOS-O_3 SiO_2 (CMP)	0.48	84.7	1.34	131 406	1.69[b]	80.8	0.46[b]	30 450	2.31[b]
TEOS-O_3 SiO_2 (ethanol treatment)	1.49	84.2	1.32	88 953					
Thermal oxide	0.15	86.7	1.30	128 701					
Borophosphosilicate glass (BPSG)	0.27	89.7	1.24	209 363	1.69[b]				1.40[b]
TEOS-O_3 SiO_2	12.20	49.3[a]	2.57[a]	4996[a]		75.1[a]		780[a]	
TEOS-O_3 SiO_2 (CMP)	0.48	51.3[a]	9.59[a]	6123[a]		82.9[a]		1758[a]	

[a]Values obtained without Ti (50 nm) underlayer.
[b]Values obtained with Al thickness of 500 nm.
All other data were taken with Al thickness of 700 nm.

sity obtained by XRD analysis from the Al-alloy (700 nm)/Ti (50 nm) structures formed on the insulators shown in Table I. Al deposition temperature is 450 °C. The reflectance depends strongly on the underlying insulator surface roughness and increases drastically as the underlying surface roughness decreases [Fig. 2(a)]. The trend in Al(111) peak intensity [Fig. 2(b)] is almost the same as that of reflectance. The Al(111) preferred orientation depends strongly on the underlying insulator surface roughness and decreases drastically as the underlying surface roughness increases.

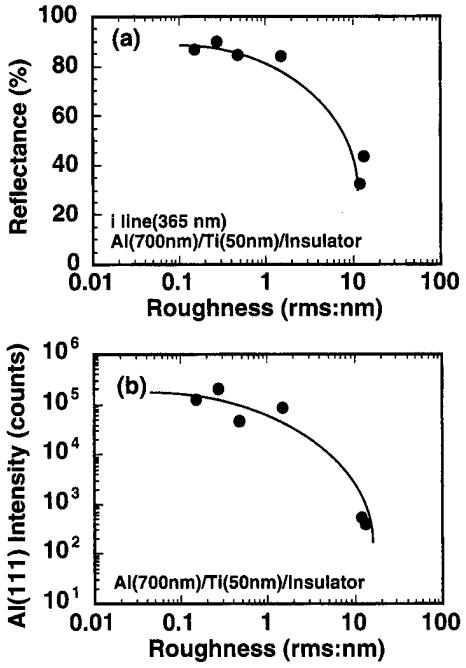


FIG. 2. The *i*-line (365 nm) reflectance (a) and Al(111) peak intensity measured by XRD from the Al-alloy (700 nm)/Ti (50 nm) structures formed on insulators which have various surface roughness shown in Table I.

Two kinds of insulators, TEOS-O_3 SiO_2 film/plasma SiO_2 double layered films with and without CMP treatment, were chosen for further experiments because only roughness of the insulator surface can be changed without the insulator structure and properties changing. Figures 3(a) and 3(b) show the XRD spectra from Al(700 nm)/Ti (50 nm)/insulator structures (a) with and (b) without CMP treatment of the underlying insulator. The values in the parentheses after Al(111) notation in Figs. 3(a) and 3(b) show Al(111) peak intensities. More than two orders of higher Al(111) peak intensity were obtained on TEOS-O_3 SiO_2 with CMP treatment compared to that on TEOS-O_3 SiO_2 without CMP treatment. The Al(200) peak can be scarcely observed only in the spectrum from the sample without CMP treatment. Although differences in unreacted Ti and reaction product $Al_3Ti(112)$ peak intensities can be seen in the spectra, this is due to the difference in the level of preferred crystallographic orientation and not due to the difference in the reaction. There is not a large difference in the reaction product thickness between Al and Ti, as shown later in Fig. 9: cross-sectional TEM observation of the samples.

Figure 4 shows the reflectance from the Al alloy (500 nm)/Ti (50 nm) formed on TEOS-O_3 SiO_2 film/plasma SiO_2 as a function of Al alloy deposition temperature. In Fig. 4, the reflectance from the samples in which Al alloy film was deposited without substrate heating and reflowed in the same vacuum at 420–485 °C is also shown. CMP treatment on the underlying insulator causes a large difference in the reflectance at elevated deposition temperatures. Without CMP treatment, the reflectance shows about 75% for the sample deposited at a low temperature of 220 °C, and it decreases drastically as the deposition temperature increases and shows a minimum at around 420 °C (about 20%). With CMP treat-

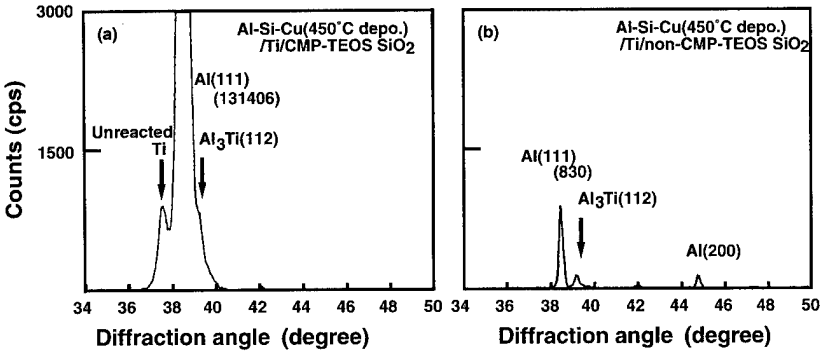


FIG. 3. XRD spectra from Al alloy/Ti/TEOS-O_3 SiO_2 structures with (a) and (b) without CMP treatment. Al alloy deposition temperature is 450 °C. The values in the parentheses after Al(111) notation in (a) and (b) show Al(111) peak intensities.

ment, the reflectance is high at a low temperature of 220 °C (about 86%), and it gradually increases as the deposition temperature increases and shows maximum also at around 420 °C (90%). On the other hand, in the case of reflow samples, the reflectance is high independent of CMP treatment, and the insulator surface roughness does not change the roughness of the deposited Al alloy surface.

Figure 5 shows the Al(111) peak intensities from the Al-alloy (500 nm)/Ti (50 nm) formed on TEOS-O_3 SiO_2/plasma SiO_2 as a function of Al alloy deposition temperature. In Fig. 5 the Al(111) peak intensities from the samples in which Al alloy film was deposited without substrate heating and reflowed at 420–485 °C are also shown. In contrast to the reflectance results shown in Fig. 4, Al(111) peak intensity from the samples without CMP treatment of the insulator shows two orders lower compared with the CMP treated samples even at the low deposition temperature of 220 °C. As the deposition temperature increases, the peak intensity decreases further and shows a minimum at around 420–450 °C. In the samples with CMP treatment, as the deposition temperature increases, the peak intensity increases and shows a maximum at around 420 °C. The maximum peak intensity in this case (Al thickness: 500 nm) is 20% higher than that of the 220 °C deposited sample. On the other hand, in the case of reflow samples with CMP treatment, the Al(111) peak shows almost the same or even better intensity as high-temperature sputtered samples, however, the peak intensity shows one order lower in the case of reflow sample without CMP treatment. This result differs from the results of reflectance shown in Fig. 4, where the reflectance does not depend on the insulator surface roughness. The underlying insulator surface roughness affects the Al alloy crystallographic orientation without changing reflectance.

The Al(111) peak intensity has been used as an index value for the degree of the texture so far, because only the (111) peak is observed in many of the samples used in the experiment. This value, however, is susceptible to receive the influence of measurement conditions. Finally in this section, the full width at half-maximum (FWHM) values of the (111) rocking curves for Al films were obtained in order to confirm the validity of Al(111) peak intensity as an index value for texture distribution.[16] The rocking curves were measured by x-ray diffraction using the Cu $K\alpha$ line by scanning an incident angle θ, while the diffraction angle 2θ was fixed for the Al(111) peak. FWHM values were obtained as functions of Al alloy deposition temperature and reflow temperature, and were shown in Fig. 6. The samples used in the measurement

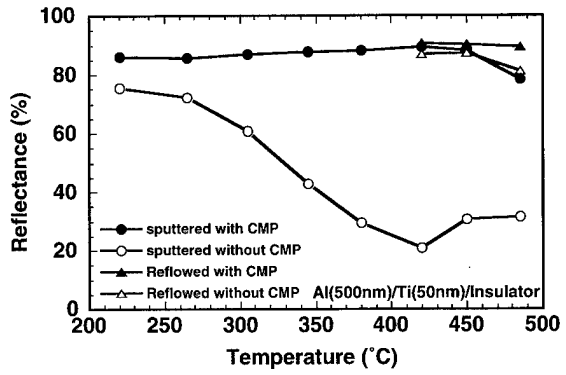


FIG. 4. The reflectance from the Al-alloy (500 nm)/Ti (50 nm) formed on TEOS-O_3 SiO_2 film/plasma SiO_2 as a function of Al alloy deposition temperature. The reflectance from the samples in which Al alloy film was deposited without substrate heating and reflowed at 420–485 °C is also shown.

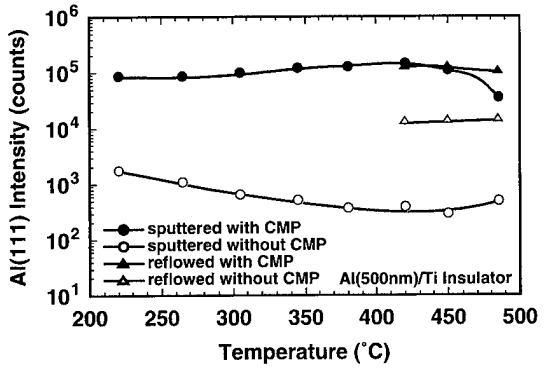


FIG. 5. The Al(111) peak intensities from the Al-alloy (500 nm)/Ti (50 nm) formed on TEOS-O_3 SiO_2/plasma SiO_2 as a function of Al alloy deposition temperature. CMP treatment on the insulators is chosen for the parameter. The Al(111) peak intensities from the samples in which Al alloy film was deposited without substrate heating and reflowed at 420–485 °C are also shown.

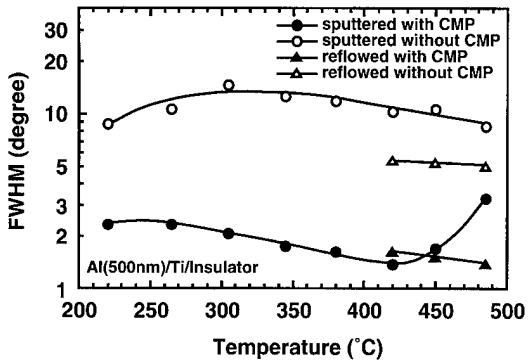


FIG. 6. FWHM values obtained from the Al-alloy (500 nm)/Ti (50 nm) formed on TEOS-O_3 SiO_2/plasma SiO_2 as a function of Al alloy deposition temperature. CMP treatment on the insulators is chosen for the parameter. The Al(111) peak intensities from the samples in which Al alloy film was deposited without substrate heating and reflowed at 420–485 °C are also shown.

are the same as those used in Fig. 5. Almost the same trends as shown in Fig. 5 are obtained, that is, FWHM value decreases as the deposition temperature increases in the samples with CMP treatment, although the change in large FWHM values obtained from the samples without CMP treatment does not correlate well with the change in Al(111) peak intensity. It seems that the measurement in the large FWHM value region is not sensitive compared with the measurement in Al(111) peak intensity, however, Al(111) peak intensity trends shown in Figs. 2(b) and 5 were confirmed to be the trends in preferred orientation.

B. Improvement mechanism in Al planarization and preferred orientation

In the previous sections, it was found that the insulator surface roughness largely affects the surface roughness and crystallographic orientation in Al alloy film deposited at elevated temperatures. In this section, the improvement mechanism of planarization and preferred orientation in Al alloy film due to the underlying insulator surface roughness will be discussed.

First, in this section, the role of the Ti underlayer for the planarization and preferred orientation should be studied. Figure 7 shows the XRD spectra from the Al alloy/TEOS-O_3 SiO_2 structures (a) with and (b) without CMP. Al alloy deposition temperature is 450 °C. The values in the parentheses after Al(111) notation in Fig. 7 show Al(111) peak intensities. When the Al alloy was deposited directly on the insulators without a Ti underlayer, the difference in Al(111) peak intensities due to the insulator roughness is not observed, although the Al(200) peak can be observed only on a non-CMP treated sample. On the other hand, as shown in Table I, the reflectance of the Al alloy film directly deposited on the TEOS-O_3SiO_2 without a Ti underlayer at a deposition temperature of 450 °C is 49.3% on a non-CMP treated insulator, while that on a CMP treated insulator is 51.3%. The Ti layer is essential for the planarization and preferential orientation of the overdeposited Al alloy film.

Figure 8 shows the XRD spectra from Ti(50 nm) TEOS-O_3 SiO_2 structures (a) with and (b) without CMP. Ti has preferred orientations of (002) and (011), and these peak intensities are 20–30-fold stronger on samples with CMP treatment than those without CMP treatment. Since Ti(002) and Ti(011) XRD peaks from the Ti powder have 20:100 intensities, Ti film in this experiment has a preferential orientation of Ti(002). This difference in Ti peak intensity and orientation change seems to induce the difference in Al alloy orientations due to the underlayer insulator roughness.

We consider the reaction between the Al alloy film and the Ti film, since Al alloy films and Ti films are easy to react[18] and to form Al_3Ti, as shown in Figs. 3(a) and 3(b). In order to obtain the visual image of the reaction, cross-sectional TEM observations of the Al/Ti layered structures deposited at 380, 420, and 450 °C were carried out and are shown in Fig. 9. At around these deposition temperatures, Al(111) peak intensity shows a maximum as shown in Fig. 5. CMP treated TEOS-O_3 SiO_2 film is used as the underlying insulator. The TEM image of 450 °C deposited on insulator without CMP treatment is also shown [Fig. 9(d)]. The reaction between Al alloy and Ti film does not proceed in a 380 °C deposited sample [Fig. 9(a)]. In the case of 420 °C deposition in which Al(111) peak intensity shows the maximum, about 60 nm reaction product layer that is considered as Al_3Ti is formed [Fig. 9(b)]. The reaction proceeds further in 450 °C deposition, and a 150–200 nm reaction product

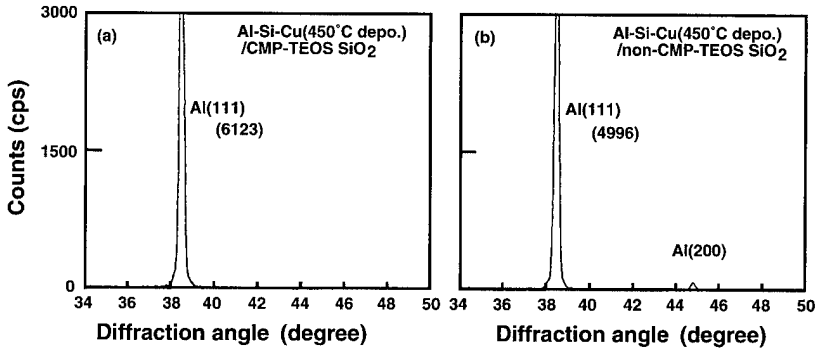


FIG. 7. XRD spectra from Al alloy/TEOS-O_3 SiO_2 structures (a) with and (b) without CMP. Al alloy deposition temperature is 450 °C. The values in the parentheses after Al(111) notation in (a) and (b) show Al(111) peak intensities.

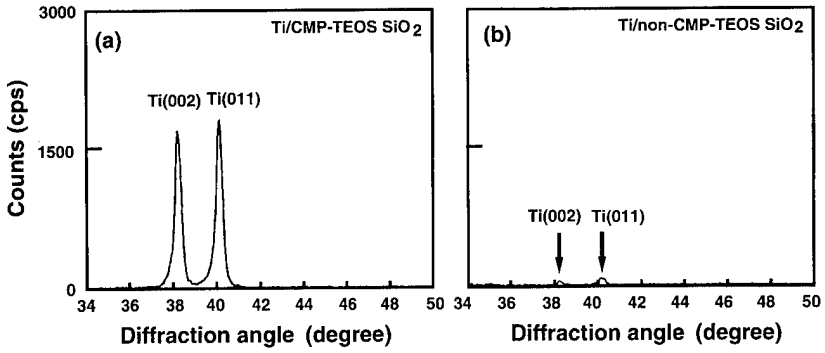


FIG. 8. XRD spectra from Ti (50 nm)/TEOS-O_3SiO_2 structures (a) with and (b) without CMP.

layer is formed. There is a large difference between CMP treated samples and non-CMP treated samples in the roughness of the interface between the metal and underlayer insulator [Figs. 9(c) and 9(d)], however, there is no difference in the thickness of the reaction layer and unreacted Ti layer. Regardless of the formation of Al and Ti reaction product, the sample on CMP treated insulator has a high Al(111) preferred orientation.

Here we consider the crystal orientation and lattice mismatch between Al, Ti, and a reaction product Al_3Ti. Figure 10 shows the schematic figures of unit cells of Al, Ti, and Al_3Ti. The Al crystal structure is a face centered cubic (fcc) structure with a lattice constant of 4.0494 Å [Fig. 10(a)]. The Ti crystal structure is hexagonal close packed (hcp) structure with lattice constants of $a=2.9503$ Å and $c=4.686$ Å [Fig. 10(b)]. The Al(111) plane is the close-packed plane in Al fcc structure and Ti(002) plane is the close-packed basal plane in the Ti hcp structure. The misfit between the close-packed planes is about 3% considering their closest interatomic distances d (Al:2.8634 Å; Ti:2.9503 Å). The misfit between the Al(111) plane and the Ti(011) plane is not so simple, however Shibata *et al.* suggested from their experiment and analysis that Al(111) preferred orientation improves on the Ti(011) plane.[14] The crystal structure of Al_3Ti is a body centered tetragonal (bct) structure with lattice constants of $a=3.84$ Å and $c=8.61$ Å.[19,20] The unit cell of this structure is considered to be two stacked unit cells of Al fcc structures and deformed, neglecting the atom difference between Al and Ti [Fig. 10(c)]. Therefore, the $Al_3Ti(112)$ plane could be regarded as the close-packed (111) plane of fcc structure for a first-order approximation. The misfits between *the close-packed plane* and Al(111) and Ti(002) are 1% and 4%, respectively [the closest interatomic distance and the second

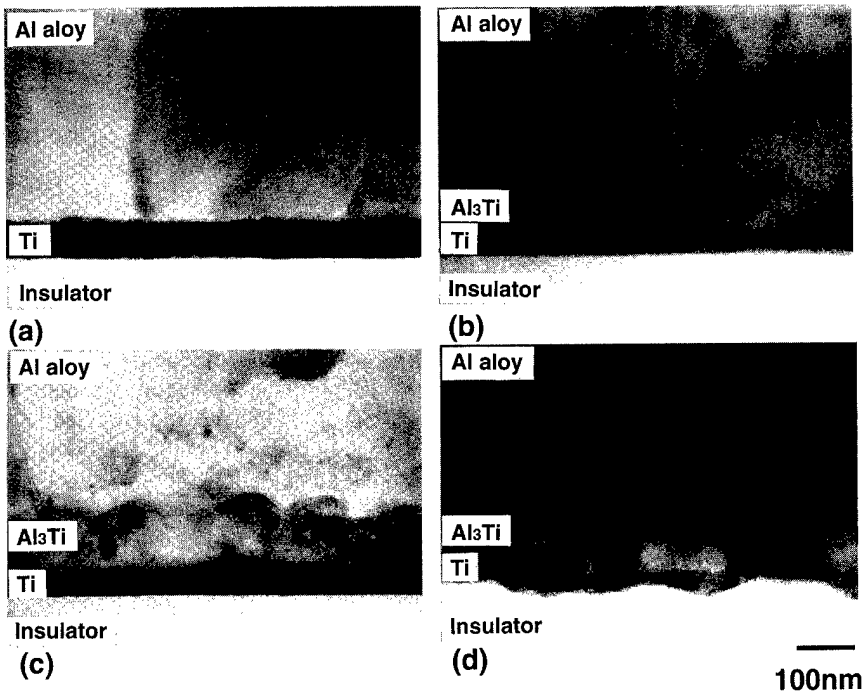


FIG. 9. The cross-sectional TEM observations of the Al/Ti layered structures deposited at (a) 380, (b) 420, and (c) 450 °C. TEOS-O_3 SiO_2 film is used as the underlying insulator. The TEM image (d) of 450 °C deposited on CMP treated insulator is also shown.

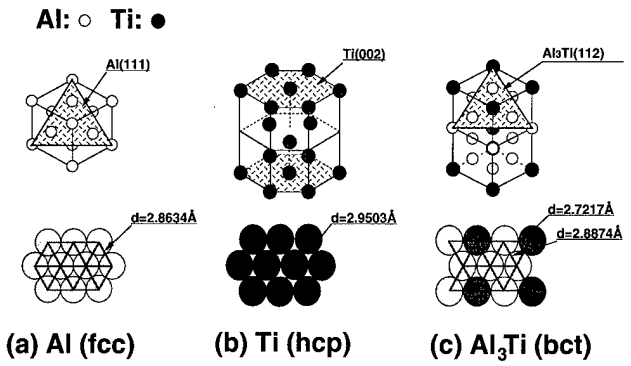


FIG. 10. The schematic figures of unit cells of Al, Ti, and Al_3Ti.

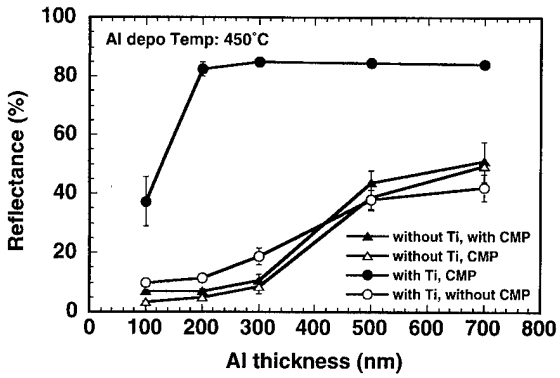


FIG. 11. The reflectance change as a function of Al alloy thickness (100–700 nm) deposited at 450 °C. The Al alloy was deposited on Ti (50 nm) on TEOS-O_3 SiO_2 or directly on the insulator with and without CMP treatment.

closest interatom distance in the $Al_3Ti(112)$ plane are 2.7217 Å and 2.8874 Å, respectively].

According to the above consideration, three crystal structures have the potential of having an epitaxial relation at least as

Al(111)//Ti(002),

Al(111)//Al_3Ti(112)//Ti(002).

The above discussion concentrates on the crystallographic orientation. Here, an experimental result will be introduced in order to discuss the planarization in Al film. Figure 11 shows the reflectance change as a function of the Al alloy thickness (100–700 nm) deposited at 450 °C. The Al alloy was deposited on Ti(50 nm) on TEOS-O_3 SiO_2, or directly on the insulators with and without CMP treatment. The reflectance from the sample with a Ti underlayer and CMP treatment shows over 80% and stays constant in the case of Al alloy thickness over 200 nm. On the contrary, the reflectances from the other sample are less than about 50% even at the Al thickness of 700 nm, although the reflectance shows a gradual increase as the Al thickness increases. The Al surface morphologies observed by SEM are shown in Fig. 12. Figure 12 shows the cases with Ti underlayer, and Al thickness ranges from 20 to 200 nm. Without CMP treatment, Al film shows the semispherical surface, and uniform film cannot be obtained even at a deposited Al alloy thickness of 200 nm.

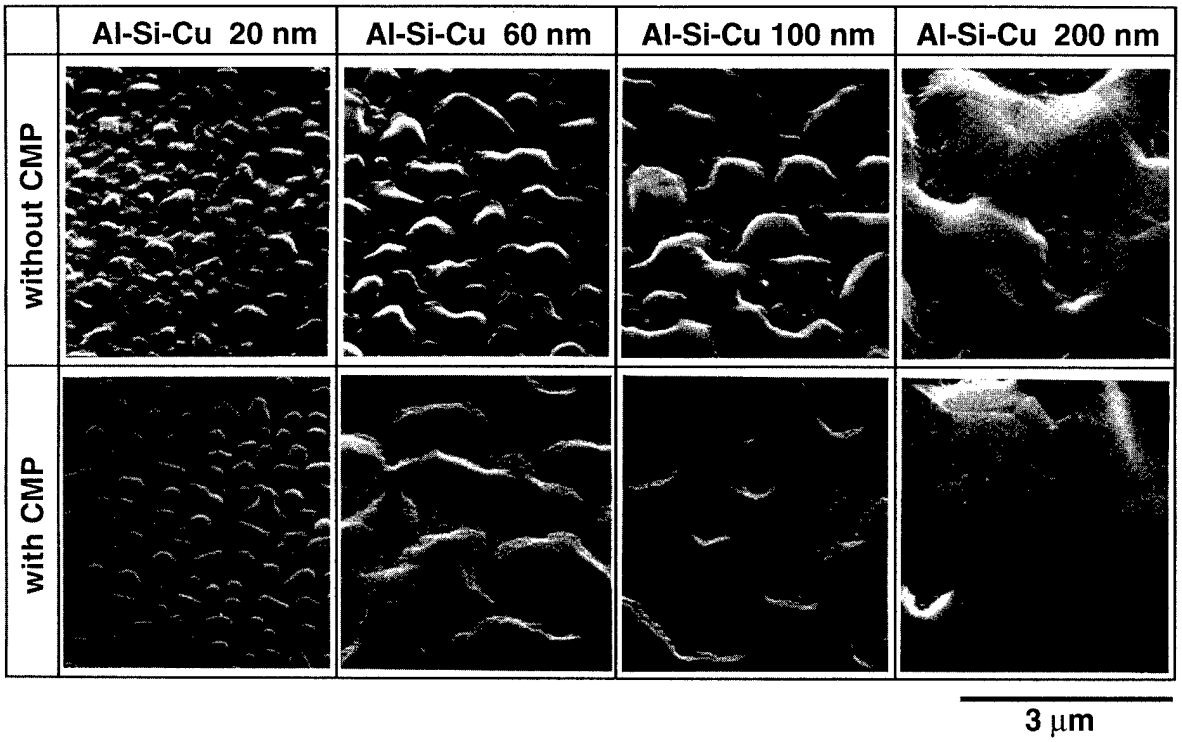


FIG. 12. SEM observation of the Al surface morphologies with a Ti underlayer. Insulator CMP treatment is chosen as a parameter. Al thickness ranges from 20 to 200 nm.

On the other hand, with CMP treatment, the deposited Al alloy film shows a smooth plateau surface, and a uniform film can be obtained over Al thickness of around 200 nm. It is considered that the fact that the sample with a Ti underlayer and insulator CMP treatment has a planarized deposited structure and uniform film over Al thickness of around 200 nm is the reason why the reflectance of the samples shows over 80% and stays constant in the case of Al alloy thickness over 200 nm.

Here, on the basis of the above discussion, how the underlying insulator surface roughness contributes to the improvement of the surface roughness and crystal preferred orientation in the Al/Ti layered structure will be summarized. Ti has a preferred crystal orientation of (002) on the smooth insulator. This is due to the preferred deposition of Ti(002) orientation perpendicular to the underlying insulator surface minimizing the surface energy. When the Al alloy film is deposited on the Ti film, Al atoms fix at Ti crystal sites in the initial stage of deposition. Hereafter, basically, the layer growth of Al alloy film occurs epitaxially, and a smooth and uniform film forms at the initial stage of Al deposition. Moreover, in the case of high-temperature sputtering, through the high Al atom migration on the substrate and taking over the underlying Ti crystal information, the highly preferred orientation in Al alloy film occurs, forming the reaction product Al_3Ti which also has an epitaxial relation with Al and Ti. On the other hand, in the case of the rough insulator, the deposited Ti also orients perpendicular to the insulator surface, however the insulator surface roughness prevents the coherence in the Ti crystal orientation. The initial Al film on the underlying Ti also has an epitaxial growth, however the crystal orientation of the film has random directions. In the case of high-temperature sputtering, through high Al atom migration, the Al growth in each direction occurs and enhances the roughness of the deposited Al film.

When the Al alloy film is deposited directly on the insulator without Ti, the layer growth occurs at lower deposition temperatures and Al surface planarizes gradually even on the rough insulator surface. This speculation is supported by the relatively high reflectance of the Al alloy film deposited on a rough insulator surface at low deposition temperature (see Fig. 4). In the case of high-temperature deposition, however, the island growth occurs because the SiO_2 surface energy is much smaller than Ti surface energy (surface energy of Al, SiO_2, and Ti:866, 663, and 1650 erg/cm^2).[21] In addition, high Al atom migration prevents coalescence from occurring until thick Al deposition. This enhances the roughness of the deposited Al film surface, and the roughness of the deposited Al film does not depend on the roughness of the underlying insulator. In the case of reflow samples that are deposited at lower temperature and reflowed at high temperature, the surface roughness of the Al/Ti layered film is determined in the stage of Al deposition almost regardless of insulator surface roughness and stays the same even after reflow annealing, while the preferred orientation of the film is also determined at the stage of deposition, and it depends on the insulator surface roughness (see Figs. 5 and 6).

C. Electromigration performance in Al/Ti layered interconnects

In Table I, the samples that should be paid attention to here for electromigration tests are TEOS-O_3 SiO_2 films with and without CMP treatment. The Al alloy films deposited on these insulators without a Ti underlayer have a large difference in Al grain size (without CMP treatment: 2.57 μm; with CMP treatment: 9.59 μm). There is no such difference with a Ti underlayer (without CMP treatment: 1.17 μm; with CMP treatment: 1.34 μm). On the other hand, without a Ti underlayer, Al(111) peak intensity from XRD analysis shows only 1.5 times higher in the CMP treated sample compared with that in the non-CMP treated sample, however, with a Ti underlayer, two orders higher Al(111) peak intensity can be obtained in the CMP treated sample. These facts enable the Al(111) peak intensity to change without changing Al grain size when the Al alloy film is deposited on Ti on insulators of which surface roughness is controlled by CMP treatment. If the other factors such as Al–Ti reaction and Ti content in Al alloy film that will affect the electromigration performance of the metal lines are the same, the evaluation of electromigration performance as a function of the level in Al preferred orientation is considered to be possible by controlling insu-

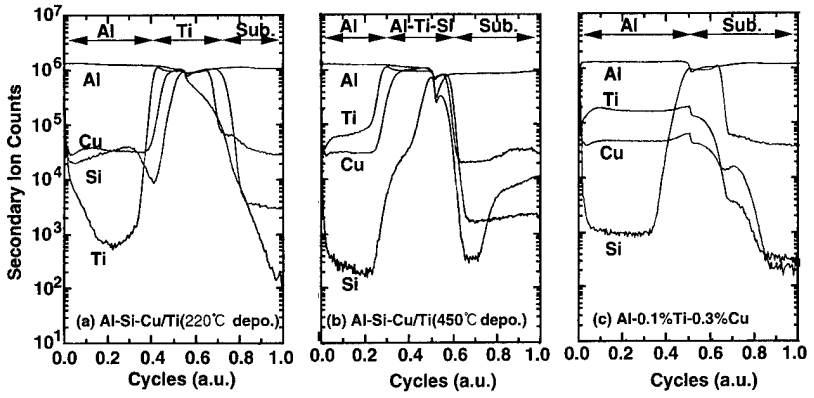


FIG. 13. Impurity concentration depth profiles measured by SIMS (a) 220 °C Al deposition, (b) 450 °C Al deposition, (c) control sample; Al-0.1% Ti-0.3% Cu).

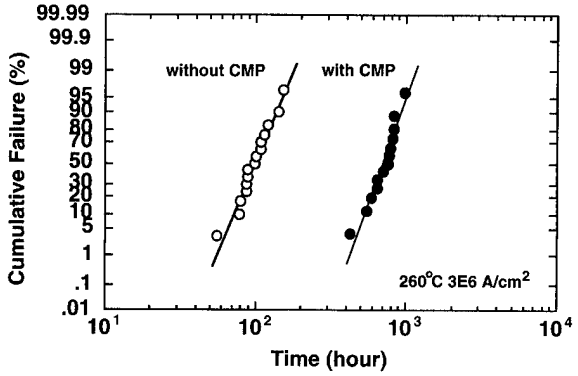


FIG. 14. One example of the electromigration test results. The samples are Al (500/nm)/Ti (50 nm) layered metal line formed on TEOS-O_3 SiO_2/plasma-TEOS SiO_2 double layer films with and without CMP treatment. The Al deposition temperature is 450 °C. The test condition is dc current density of 3×10^6 A/cm^2 at a temperature of 260 °C.

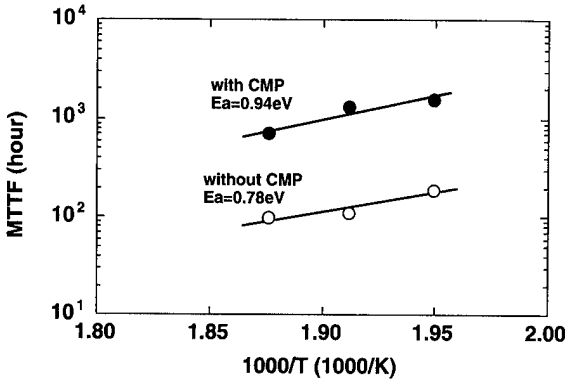


FIG. 15. The Arrhenius plot of the MTTFs in the temperature range of 240–260 °C. The current density is 3×10^6 A/cm^2.

lator surface roughness, assuming that there were no effects of insulator surface roughness on the electromigration performance. In that sense, the cross-sectional TEM observation of Al/Ti layered structures as a parameter of underlying insulator CMP treatment, as shown in Figs. 9(c) and 9(d), shows that there is no difference in the reaction product thickness and unreacted Ti thickness due to the CMP treatment. Ti concentration depth profile evaluated by SIMS shows that Ti content in Al alloy film increases with increasing deposition temperature, however it does not depend on the CMP treatment. Here, in Fig. 13, impurity concentration depth profiles measured by SIMS were presented at (a) 220 °C Al deposition, (b) 450 °C Al deposition, and (c) control sample: Al-0.1% Ti-0.3% Cu). On the assumption that the Ti concentration of the control sample (c) is correct, the Ti concentration in the 450 °C Al deposition sample has about 0.03% and that of 220 °C Al deposition sample has no Ti in Al alloy film.

The electromigration tests were carried out as a parameter of the CMP treatment of underlying insulator. The samples are Al (500 nm)/Ti(50 nm) layered metal lines formed on TEOS-O_3 SiO_2/plasma-TEOS SiO_2 double layer films with and without CMP treatment. The Al deposition temperature is 450 °C. Figure 14 shows one example of the electromigration test results. The test condition is dc current density of 3×10^6 A/cm^2 at a temperature of 260 °C. The MTTF of the sample with CMP treatment is about seven times longer than that without CMP treatment (702 and 98 h) in this test condition. The Arrhenius plot of the MTTF in the temperature range of 240–260 °C are shown in Fig. 15. The activation energies of the MTTFs for the samples with and without CMP treatment are 0.94 and 0.78 eV, respectively.

In order to consider the obtained electromigration results, the line resistance change during EM tests was monitored and the analysis of electromigration failure mode in the metal lines after EM tests was carried out. The resistance change of the metal lines during the electromigration test was monitored at the test condition of dc current density of 3×10^6 A/cm^2 and temperature of 250 °C. The results were shown in Fig. 16. Different resistance changes were obtained according to the CMP treatment. The samples without CMP treatment [Fig. 16(a)] show a monotonical resistance increase and result in line opens at the resistance increase of 2%–4% from the initial values. On the other hand, the samples with CMP treatment [Fig. 16(b)] show a different resistance change. They show a resistance increase during the first 100 h, however it then starts to decrease. The peak of the resistance change is not so high (less than 0.6%) compared to the resistance increase of Fig. 16(a) just before the lines open. The resistance continues to decrease lower than the initial resistance, then starts to increase again (the second

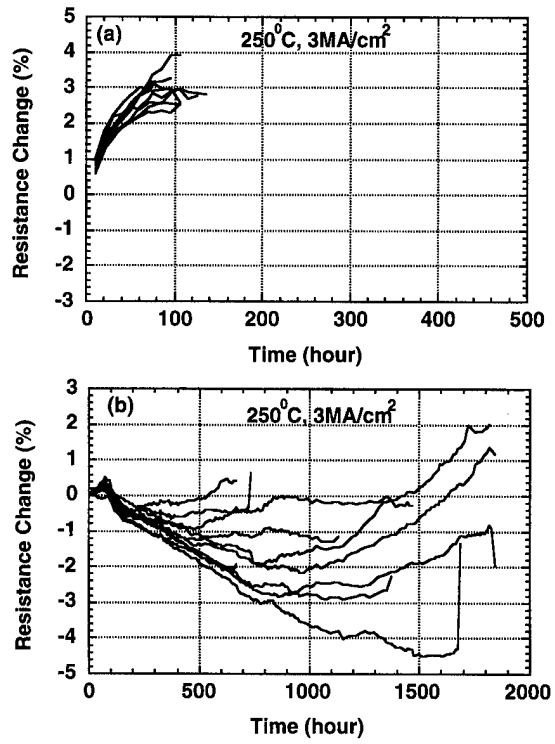


FIG. 16. The resistance change of the metal lines during the electromigration test monitored at the test condition of 3×10^6 A/cm^2 and 250 °C. The Al deposition temperature is 450 °C. (a) Non-CMP treated sample and (b) CMP treated sample.

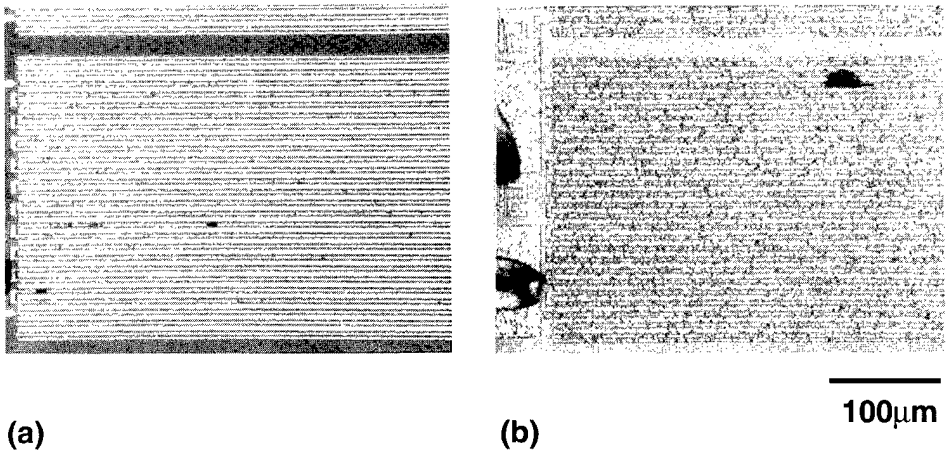


FIG. 17. Optical micrographs of the metal lines that show large parts of the serpentine metal line with dummy lines in-between after the electromigration test. (a) CMP treated sample and (b) non-CMP treated sample.

increase) and the metal lines become open. Some of the metal lines become open without the second resistance increase.

In Figs. 17 and 18, photographs of the metal lines after the electromigration test are shown. Figures 17(a) and 17(b) are the optical micrographs with low magnificent that show large parts of the serpentine metal line with wider dummy lines in-between (a) a CMP treated sample and (b) non-CMP treated sample). Figures 18(a) and 18(b) show the SEM photographs of the voids formed in CMP treated sample and the open part of the metal line in a non-CMP treated sample, respectively. In CMP treated samples, a large number of voids such as shown in Fig. 18(a) exist in the metal lines, and the open parts distribute randomly in the metal lines [Fig. 17(a)]. It is speculated here that the Al thickness variation in the non-CMP treated sample could have an effect on the void growth. On the other hand, in non-CMP treated samples, the open parts of the metal lines such as shown in Fig. 18(b) always exist at the first line from the negative pad, as shown in Fig. 17(b). Void distribution in the non-CMP treated samples cannot be discerned from the optical micrograph because of the random reflection from the rough Al surface [Fig. 17(b)].

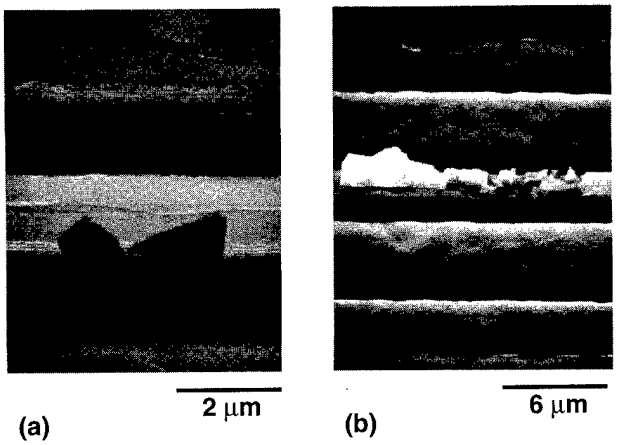


FIG. 18. SEM photographs of the voids formed in CMP treated sample (a) and the open part of the metal line in non-CMP treated sample (b).

From the results shown in Figs. 16, 17, and 18, the following speculations can be deduced. The monotonical resistance increase during electromigration test until the line open in the non-CMP treated samples is considered to be mainly caused by void formation and growth in the first line from the negative pad. Since the electromigration resistance of the highly oriented Al/Ti layered interconnects is extremely high, the electromigration tests were carried out at rather severe conditions, in particular at high current density to reduce the test time in this experiment. In the first line from the negative pad, there seems to be a temperature gradient because of the pad heat sink. The temperature difference between the ambient temperature and the metal line temperature estimated by the line resistance change is around 100 °C. The temperature gradient accelerates the void growth due to the relatively high Al atom mobility and Al flux divergence. On the other hand, the resistance increase during the first 100 h in the CMP treated samples is also caused by void formation and growth, however the place where the voiding occurs scatters all over the line including the first line from the negative pad. The voiding accelerated by the temperature gradient in the first line from the negative pad could also occur in this case, however the metal line has an endurance to keep line continuity and voiding is suppressed to limited size and number because the Al film in CMP treated samples has hyper textured grains and low Al flux divergence so the metal line is not susceptible to the influence of temperature gradient. The resistance decrease after the first resistance increase is considered to be the recovery of Al crystalline quality by long time annealing. This recovery is usually observed during a long storage time at relatively high temperature without current passage and is not special for EM testing. The second resistance increase is caused by void growth, and void length increase in some

places results in high current passage through the reaction product layer, accelerating void growth and line openings.

IV. CONCLUSION

Effects of insulator surface roughness on the overlying aluminum alloy film surface roughness and crystallographic orientation in Al alloy/Ti/insulator layered structures have been studied. The following conclusions have been obtained.

The underlying insulator surface roughness has a large effect on the surface roughness and crystallographic orientation of the Al alloy/Ti layered film deposited on the insulators. In particular, the effect is remarkable when the Al alloy film is deposited at high temperature. The Al surface smoothness and the Al(111) preferred orientation in Al/Ti layered film improves when the underlying insulator surface is smooth, and deteriorates when the underlying insulator surface is rough.

The underlying insulator smoothness enhances Ti preferred orientation first, and that enhances the Al layered growth on the Ti which has a large surface energy, even when the reaction between Al and Ti occurs. The lattice mismatch between Al, Ti, and main reaction product Al_3Ti explains well the enhancement in Al preferred orientation. Without Ti underlayer, such a smooth surface and a large crystallographic orientation change in Al alloy does not occur.

Al/Ti layered interconnects formed on insulators that have a very smooth surface by CMP treatment have an extremely high electromigration performance because of the high Al(111) oriented texture.

[1]M. Inoue, K. Hashizume, and H. Tsuchikawa, J. Vac. Sci. Technol. A **6**, 1636 (1988).

[2]Tsann Lin, K. Y. Ahn, J. M. E. Harper, and P. N. Chaloux, in *Proceedings of the Fifth International IEEE VLSI Multilevel Interconnection Conference* (IEEE, New York, 1988), pp. 76–84.

[3]F. S. Chen, Y. S. Lin, G. A. Dirit, R. Sundaresan, C. C. Wei, and F. T. Liou, in *IEEE-IEDM 1990, Technical Digest* (IEEE, San Francisco, 1990), pp. 51–54.

[4]H. Ono, Y. Ushiku, and T. Yoda, in the *1990 Proceedings of the Seventh International IEEE VLSI Multilevel Interconnection Conference* (IEEE, New York, 1990), pp. 76–82.

[5]H. Nishimura, T. Yamada, and S. Ogawa, in *Proceedings of the Eighth International IEEE VLSI Multilevel Interconnection Conference* (IEEE, New York, 1991), pp. 170–176.

[6]H. Nishimura, T. Yamada, R. Sinclair, and S. Ogawa, in *Technical Digest of 1992 Symposium on VLSI Technology* (IEEE, Seattle, 1992), pp. 74–76.

[7]K. Kikuta, T. Kikkawa, and H. Aoki, in *Technical Digest of 1991 Symposium on VLSI Technology* (IEEE, Oiso, 1991), pp. 35 and 36.

[8]K. Kikuta, T. Kikkawa, and M. Aoki, in Ref. 5, pp. 163–169.

[9]G. A. Dixit, M. F. Chisholm, M. K. Jain, T. Weaver, L. M. Ting, S. Poarch, K. Mizobuchi, R. H. Havemann, C. D. Dobson, A. I. Jeffryes, P. J. Holverson, P. Rich, D. C. Butler, and J. Hems, in Technical Digest, International Electron Devices Meeting, San Francisco, 1994 (unpublished), pp. 105–108.

[10]S. Vaidya and A. K. Sinha, Thin Solid Films **75**, 253 (1981).

[11]H. Kaneko, M. Hasunuma, A. Sawabe, T. Kawanoue, Y. Kohanawa, S. Komatsu, and M. Miyauchi, Proceedings of the 28th International Reliability Physics Symposium, 1990 (unpublished), p. 194.

[12]K. Hashimoto and H. Onoda, Appl. Phys. Lett. **10**, 120 (1989).

[13]H. Onoda, M. Kageyama, and K. Hashimoto, J. Appl. Phys. **77**, 885 (1995).

[14]H. Shibata, M. Murota, and K. Hashimoto, Jpn. J. Appl. Phys. **32**, 4479 (1993).

[15]T. Mitsuzuka, Jpn. J. Appl. Phys. **31**, L 1280 (1992).

[16]H. Toyoda, T. Kawanoue, M. Hasunuma, H. Kaneko, and M. Miyauchi, Proceedings of the 32nd International Reliability Physics Symposium, 1994 (unpublished), p. 178.

[17]N. Sato, T. Nakano, H. Yamamoto, and T. Ohta, Jpn. J. Appl. Phys. **32**, L 110 (1993).

[18]H. Onoda, T. Narita, and K. Hashimoto, Jpn. J. Appl. Phys. **34**, 4728 (1995).

[19]W. B. Pearson, *Handbook of Lattice Spacings and Structures of Metals* (Pergamon, Oxford, 1967), Vol. 2.

[20]F. A. Wells, *Structural Inorganic Chemistry*, 3rd ed. (Clarendon, Oxford, 1962).

[21]L. E. Murr, *Interfacial Phenomena in Metals and Alloys* (Addison-Wesley, Reading, MA, 1975).

Gap fill dependence of fluorinated polyimide films on solid content, adhesion promoter, spin dwell time, and solvent spray

Shi-Qing Wang[a] and Bin Zhao
Sematech, Austin, Texas 78 741

(Received 15 February 1996; accepted 10 May 1996)

Cross sectional scanning electron microscopy was used to study the dependence of gap fill capabilities in 0.25–2.0 μm wide Al–Cu metal gaps of DuPont fluorinated polyimide films on solid content, adhesion promoter, spin dwell time, and solvent spray. It was found that reduced solid content, application of adhesion promoter and solvent spray improve gap fill capabilities. However, these improvements are insignificant. Spin dwell time has little effect on gap fill. © *1996 American Vacuum Society.*

I. INTRODUCTION

To improve the performance in ultralarge scale integration (ULSI) devices with increased density requires insulating materials with low dielectric constant (low ϵ) as interlevel dielectrics (ILDs) (Ref. 1) so that device RC delay, interconnect cross talk, and power consumption can be reduced. DuPont fluorinated polyimide (FPI) A and B are promising candidates due to their low dielectric constants ($\epsilon_{\text{out-of-plane}}$=2.64 and 2.67, respectively, at 1 MHz, room temperature, and 40% relative humidity), high thermal stability (decrease in film thickness is 1.0% for FPI-A after 3 h at 500 °C, and 0% and 2.7% for FPI-A and FPI-B, respectively, after 8 h at 425 °C in vacuum), and low moisture absorption (1.50 and 1.45 wt %, respectively, after 1 h at 100% relative humidity and room temperature).

In the current standard multilevel interconnect process route using Al–Cu lines/chemical vapor disposition (CVD) W plugs, Al–Cu metal lines are patterned through reactive ion etching and then dielectric film is deposited over the lines, filling in the gaps between these lines (subtractive method). These gaps could have widths in the sub-0.5 μm range and aspect ratios greater than 4. Therefore, small gap fill capability is required for dielectrics. In this study, the dependencies of gap fill capability on solid content, adhesion promoter, spin dwell time, and solvent spray of spin-on FPI-A and FPI-B films are investigated in Al–Cu metal gaps with widths ranging from 0.25 to 2.0 μm.

II. EXPERIMENT

DuPont experimental fluorinated polyimide copolymers FPI-A and FPI-B were used for this study. They are solutions of polyamic acid precursors of an equivalent total amount of two dianhydrides: pyromellitic dianhydride (PMDA) and 4,4-(hexafluoroisopropylidene) diphthalic anhydride (6FDA) with an equivalent total amount of up to two diamines: 2,2-bis (trifluoromethoxy)benzidine (TFMOB) and *p*-phenylene diamine (PPD), dissolved in the solvent N-methyl-2-pyrrolidone (NMP). The molecular structure of this FPI family can be found in Refs. 2 and 3. Molar ratios of the dianhydride and diamines in FPI-A and FPI-B are listed in Table I.

The patterned wafers used in this gap fill study had a series of metal gaps ranging from 0.25 to 2.0 μm in width. The wafers were manufactured by depositing 600 nm Al-0.5 wt % Cu metal film on 200 mm unpatterned substrates of 600 nm thermally grows SiO_2 on Si(100), followed by metal gap pattern definition through a lithographic and etch process. Therefore, if no overetch into the Si/SiO_2 substrates occurred, the most aggressive aspect ratio for the gap (defined as the gap depth divided by the gap width) would be 2.4. In this study, we chose primarily the gaps which form a 90° angle relative to the radius. Therefore, the voids formed in unfilled gaps were observed dominantly on the gap sidewall oriented away from the wafer center.

The typical thickness for an ILD application is around 1 μm. The recipes to achieve 1-μm-thick final cured films on blanket wafers, when applicable, were used to spin FPI-A and FPI-B films on 200 mm patterned wafers. In some cases, DuPont VM-651 adhesion promoter was applied before the FPI film spin process. In addition, NMP solvent spray was used on one patterned wafer after the application of the VM-651 adhesion promoter and before the FPI film deposition. A two step spin process was used for coating 200 mm patterned wafers with either adhesion promoter or FPI films, where the first spin step spreads solution to cover the whole wafer surface followed by a second spin, which primarily determines the final film thickness. In some cases, a spin dwell time of 120 s was added between the two spins.

A Machine Technology Inc. FlexiFab was used for the spin and bake of adhesion promoter and FPI films. Cure of the FPI films was conducted in a Yield Engineering Systems

TABLE I. Molar ratios of the dianhydride and diamines in the FPI-A and FPI-B copolymers.

	Molar ratio of dianhydride		Molar ratio of diamines	
Copolymer	PMDA	6FDA	TFMOB	PPD
FPI-A	95	5	95	5
FPI-B	75	25	100	0

[a]On assignment from Fairchild Research Center, National Semiconductor Corp., Santa Clara, CA 95052.

TABLE II. Summary of properties and the spin parameters of the FPI films used in this study.

	FPI solution properties				Film preparation process				
							FPI spin		
Wafer ID No.	Copolymer	Molecular weight (m_w)	Solid content (wt %)	Viscosity (poise)	VM651 adhesion promoter	NMP solvent spray	First spin (s/rpm)	Dwell time (s)	Second spin (s/rpm)
1	FPI-B	350 K	3.0	0.5	Yes	No	10/500	0	15/1200
2	FPI-B	350 K	4.5	0.8	Yes	No	10/500	0	20/1300
3	FPI-B	350 K	6.0	4	Yes	No	15/500	0	30/1500
4	FPI-B	350 K	7.0	7	Yes	No	30/500	0	45/1750
5	FPI-B	350 K	8.0	12	Yes	No	30/500	0	60/2000
6	FPI-A	233 K	7.0	4.3	Yes	No	10/500	120	30/1500
7	FPI-A	233 K	7.0	4.3	Yes	No	10/500	0	30/1500
8	FPI-A	233 K	7.0	4.3	No	No	10/500	120	30/1500
9	FPI-A	233 K	7.0	4.3	No	No	10/500	0	30/1500
10	FPI-A	233 K	7.0	4.3	Yes	Yes	10/500	0	30/2000

Lamina 450PB polyimide bake oven. The process of applying the adhesion promoter, if required by the experiment, is listed as: (a) mix 0.05 wt % VM651 with 60 ml DI water well and add 40 ml methanol; (b) without aging, dispense 6 ml VM651 solution on 200 mm patterned wafers; (c) spin the wafers at a speed of 300 rpm for 15 s; (d) wait for 60 s while wafers stay still; (e) spin the wafers at a speed of 5000 rpm for 60 s; (f) bake the wafers at 110 °C for 60 s in air on the hot plate; (g) let the wafers cool down to room temperature.

One of the wafers received a NMP solvent spray at this stage. Six ml of NMP was dispensed on this wafer and the wafer was spun at 100 rpm for 5 s. Then all the wafers proceeded with the FPI film spin process. The spin time and speed for FPI-A and FPI-B films varied in different experiments. However, the film soft bake and cure recipes were the same. The sequence of this process was: (a) dispense various amount (depending on the viscosities, but generally from 6 to 10 ml of FPI-A and FPI-B solutions on 200 mm patterned wafers; (b) apply the first spin; (c) apply a spin dwell time of 120 s (wafer stays still) if required by experiment; (d) apply the second spin; (e) soft bake the films at 110 °C for 60 s in air on the hot plate; (f) cure the wafer in the oven in a heated N_2 atmosphere at a pressure of 4×10^4 Pa through a sequence of annealing at 150 °C for 30 min, ramping up from 150 to 230 °C, annealing at 230 °C for 30 min, ramping up from 230 to 300 °C, annealing at 300 °C for 30 min, ramping up from 300 to 400 °C, annealing at 400 °C for 60 min, cooling down from 400 to 150 °C. Both ramping and cooling rate are 2 °C min. A detailed summary of the FPI films prepared for this study and the spin parameters is given in Table II.

The scanning electron microscopy (SEM) used for all the cross sectional analysis was an Amray 1910FE field emission microscope with an accelerating voltage of 15 kV. Wafers deposited with FPI films were first capped with an Olin CONSI-4000™ conformal silane substitute (a Si–O–C mixture). Then a focused ion beam (FEI800) was used to prepare clean and smooth cross sectional samples in order to obtain clear and undistorted gap fill images of FPI films.

TABLE III. Experimental results of gap fill dependence on the solid content.

	Thickness (μm) over large metal pads	Gap fill capability		
Wafer ID No.		0.8 μm gap	1.0 μm gap	1.2 μm gap
1	0.33	medium void	filled	
2	0.28	medium void	filled	
3	0.84		small void	filled
4	0.90		large void	filled
5	1.05		large void	filled

III. RESULTS AND DISCUSSION

A. Solid content

For the study of gap fill dependence on solid contents, five different kinds of FPI-B solutions were used (wafer Nos. 1–5 in Table II). All the solution had an identical molecular weight (m_w) of 350 k. However, when diluted by different amounts of NMP solvent, the solid contents of these solutions varied from 3 to 8 wt %. As a result, their viscosities ranged from 0.5 to 12 poise. The VM651 adhesion promoter was used before the FPI film spin process but a spin dwell time was not employed. In order to prevent a high nonuniformity in film thickness from an excessively slow spin speed, 1200 rpm was set as the lower limit for the second spin.

Table III summarizes the gap fill results (SEM micrographs are not shown). Also listed in the table is the film thicknesses measured by SEM on top of large metal pads. From Table III, it can be concluded that the lower solid content and viscosity an FPI solution has with a fixed molecular weight, the smaller gap it can fill. However, the improvement is marginal.

Since the film thickness of FPI-B varies from 0.28 to 1.05 μm on top of the large metal pads, a question may arise whether gap fill capability is film thickness dependant. In this study, it was extremely difficult to achieve an identical

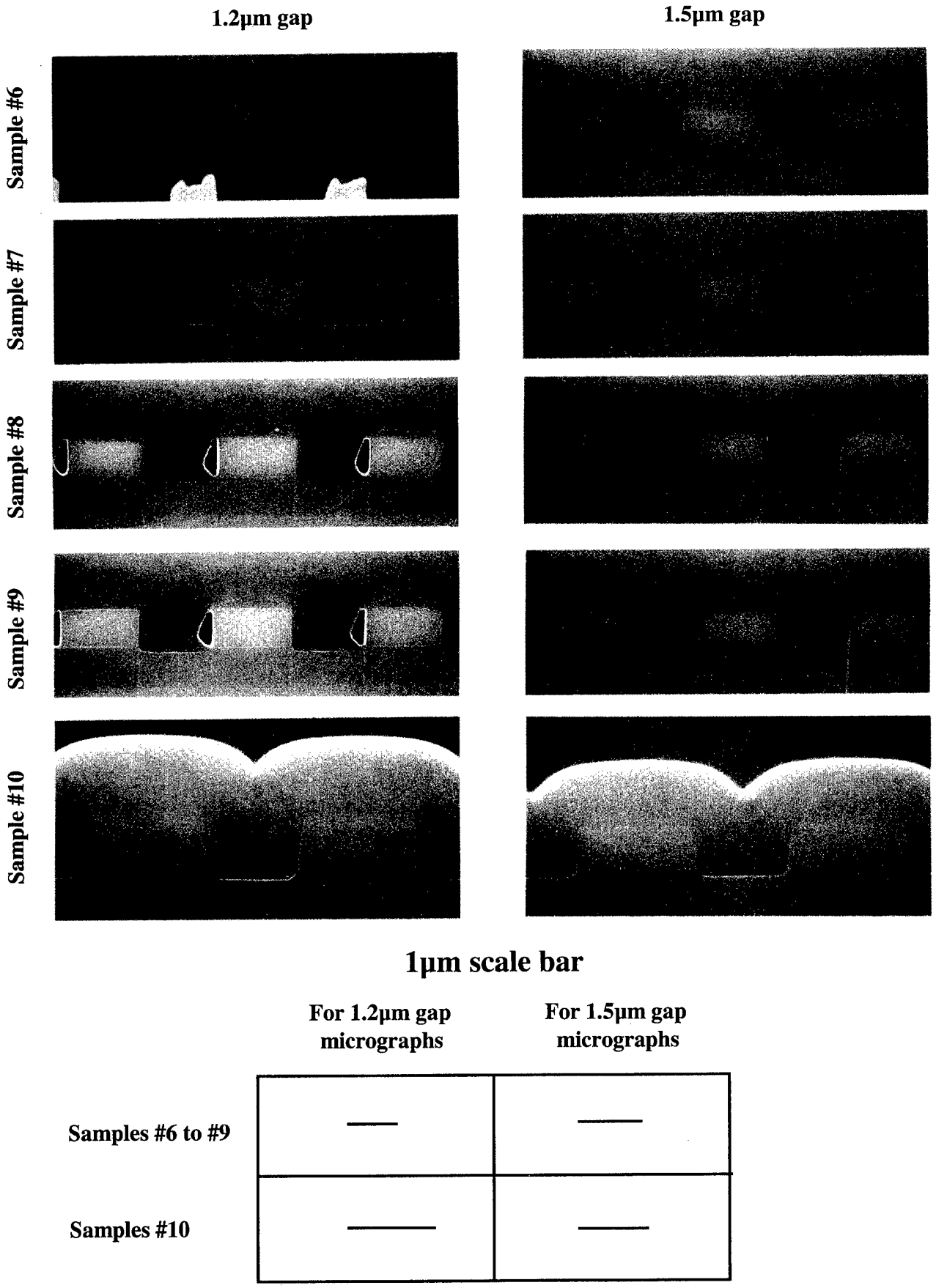


FIG. 1. Cross sectional SEM micrographs of cured FPI-A films in the 1.2 and 1.5 μm Al–Cu gaps on wafer Nos. 6, 7, 8, 9, and 10. The scales in these figures are different. In some micrographs, the top surface profile of the Olin CONSI-4000^{TM} conformal coating are visible while the FPI film and CONSI glass interfaces are not. Also visible in some micrographs are the residues and artifacts from sample preparation by focused ion beam.

final cured film thickness from the five FPI-B solutions because of their wide-spread viscosities and the limitation on how slow of a second spin speed can be used. A separate experiment using the same solution but a different second spin speed suggests that the dependence of the gap fill capability on the film thickness is minimal. It should be noted that the film on wafer No. 1 is thicker than that on wafer No. 2 although the solid content and the viscosity of the former sample is lower than those of the latter one. It was caused by the reduced second spin time/speed combination used for wafer No. 1 as compared to that of wafer No. 2.

B. Adhesion promoter

The effect of adhesion promoter on FPI-A gap fill capability can be investigated by comparing wafers Nos. 7 and 9 (without dwell time) and wafer Nos. 6 and 8 (with dwell time) in Table II. The NMP solvent spray was not used in this case. For gaps of identical width (1.2 or 1.5 μm), wafer Nos. 9 and 8 always have larger voids than wafer Nos. 7 and 6, respectively (Fig. 1). For example, medium size voids exist in the 1.5 μm gap on wafer No. 9 while the 1.5 μm gap on wafer No. 7 is completely filled. For the 1.2 μm gaps, the void size is larger on wafer No. 8 than that on wafer No. 6. It can therefore be concluded that application of the VM651 adhesion promoter slightly improves gap fill capability of the FPI-A film.

C. Spin dwell time

Further comparison between wafers Nos. 8 and 9 (without adhesion promoter) and between wafer Nos. 6 and 7 (with adhesion promoter) in Table III reveals that a spin dwell time of 120 s after the first slow spin and before the second fast spin of FPI-A solution, which is designed to provide additional time for FPI solution to flow into small features, shows no beneficial effect on gap fill (Fig. 1). Voids of similar size form in the gaps of identical widths in these two pairs of samples.

D. Solvent spray

A NMP solvent spray step was added after the application of the VM651 adhesion promoter and before the spin process of FPI-A films to investigate the effect of prewetting by solvent on the gap fill of the FPI-A films (wafer Nos. 7 and 10 in Table II). No spin dwell time was employed here.

Comparing wafer Nos. 7 and 10, it was found the 1.5 μm gaps were filled while voids formed in the bottom corners of the 1.2 μm gaps on both wafers (Fig. 1). The void size is smaller for wafer No. 10 than it is for wafer No. 7. Therefore, NMP solvent spray has a beneficial effect on gap fill. However, this effect is insignificant.

IV. CONCLUSIONS

Slight improvements were observed in Al–Cu gap fill of FPI-A and FPI-B films when the solid content is reduced, and adhesion promoter and solvent spray are applied. However, spin dwell time has little effect on gap fill.

This study is not targeted at filling the gaps as small as possible, but rather to investigate how reduced solid content, application of adhesion promoter and solvent spray, and use of spin dwell time affect gap fill property. For example, FPI used in this study have unnecessarily high molecular weights. In addition, optimization of spin/bake/cure recipes, modification of the solvent system by adding co-solvent and platicizer, and use of a more advanced spin coater equipped with a solvent saturation cup are just a few options that can improve gap fill capability. We have recently achieved excellent small gap fill with modified solutions containing the same FPI materials. The results will be published at an appropriate time.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contribution to this study by Dr. Brian Auman of DuPont (FPI solution preparation), by Felix Palacios of Sematech (FPI film processing), and by Danielle Roth of Sematech (SEM analysis).

[1] *The National Technology Roadmap for Semiconductors* (Semiconductor Industry Association, San Jose, CA, 1994).
[2] B. C. Auman, in *Low Dielectric Constant Materials-Synthesis and Applications in Microelectronics*, edited by T.-M. Lu, S. P. Murarka, T.-S. Kuan, and C. H. Ting, Mater. Res. Soc. Proc. No. 381 (Materials Research Society, Pittsburgh, PA, 1995), pp. 19–29.
[3] A. J. McKerrow, J. Leu, H-M. Ho, B. C. Auman, and P. S. Ho, in *Advanced Metallization and Interconnect Systems for ULSI Applications in 1995*, edited by R. Ellwanger, H. Komiyama, and S.-Q. Wang (Materials Research Society, Pittsburgh, PA, 1996), pp. 37–43.

Characterization of the Si/SiO_2 interface formed by remote plasma enhanced chemical vapor deposition from SiH_4/N_2O with or without chlorine addition

Young-Bae Park, Xiaodong Li,[a] and Shi-Woo Rhee[b]
Department of Chemical Engineering, Laboratory for Advanced Materials Processing (LAMP), Pohang University of Science and Technology (POSTECH), Pohang 790-784, Korea

(Received 13 February 1996; accepted 19 April 1996)

The Si/SiO_2 interface formed by remote plasma enhanced chemical vapor deposition (RPECVD) at low temperature with SiH_4/N_2O or $SiH_4/N_2O/Cl_2$ was studied and compared with thermal oxidation. The interface of the CVD SiO_2 without chlorine addition is rougher than that with chlorine addition. But the surface roughness of CVD SiO_2 films increases with chlorine addition. The thermal oxidation induces strong interface strains, and the strains generated by the CVD SiO_2 without chlorine addition are stronger and are distributed more nonuniformly than those by the chlorinated SiO_2. It is believed that chlorine addition during RPECVD affects the initial stages of deposition, and chlorine is combined with Si dangling bonds existing at the Si/SiO_2 interface through the formation of $Si-Cl_x$ bonds. It was also found that with chlorine addition during RPECVD, the strained layer thickness, interface trap density, and suboxide density could be lowered significantly.
© *1996 American Vacuum Society.*

I. INTRODUCTION

The Si/SiO_2 interface plays a crucial role in microelectronic devices and has been the subject of intense study for over 40 years. During the past decade, remote plasma enhanced chemical vapor deposition (RPECVD) has emerged as a viable technique for depositing SiO_2.[1,2] Many applications of SiO_2 films require low temperature processing. For example, in active matrix liquid crystal display (AMLCD) applications, low temperature deposition of SiO_2 films is required to utilize low cost glass substrates.

The incorporation of halogen compounds during thermal oxidation of Si results in a significant improvement in the electronic properties of the oxide and the interface formed with underlying Si.[3–10] However, for high halogen concentration and high temperature, a volatile phase begins to form at the interface, which may lift the oxide from the silicon, form bubbles, etch the silicon substrate, and increase the interface roughness.[11] Therefore, the determination of an appropriate amount of halogen compounds is important to optimize an oxidation window.

The effect of the addition of halogen compounds on the Si/SiO_2 interface and bulk properties in the PECVD process of SiO_2 film deposited at low temperature below 500 °C is still, to a large extent, unknown. Recently, Falcony and coworkers have reported on SiO_2 films deposited using a halogen chemistry, i.e., $SiF_4/H_2/N_2O$, $SiF_4/SiH_4/N_2O$, and $SiCl_4/O_2$, in order to reduce the incorporation of Si–H and Si–OH bonds and improve the breakdown field.[12,13] They could effectively decrease a significant amount of these bonds and improve the breakdown field up to 4–6 MV/cm.

In the present article, the interface between the Si substrate and the SiO_2 film obtained by RPECVD at low temperature with SiH_4/N_2O and $SiH_4/N_2O/Cl_2$ was studied using transmission electron microscopy (TEM), high frequency (1 MHz) capacitance–voltage ($C-V$) measurement, x-ray photoelectron spectroscopy (XPS), and electron spin resonance (ESR). The surface and interface roughness and morphology of deposited oxide films were observed by atomic force microscopy (AFM). For comparison, the Si/SiO_2 interface obtained by thermal oxidation was also studied.

II. EXPERIMENTAL PROCEDURE

A. Film preparation

The deposition of SiO_2 on Si was performed in a RPECVD reactor as shown in Fig. 1. N_2O (99.999% pure) was introduced into a quartz tube, excited by the induction with 13.56 MHz RF power to produce activated oxygen, and flowed into the deposition zone. SiH_4 (99.99% pure) mixed with Ar (99.9999% pure) diluent was introduced into the region 10 cm downstream of the plasma, through the gas dispersal ring. This then mixed with the excited oxygen species and flowed toward the substrate, 20 cm downstream of the plasma. Cl_2 (99.99% pure) in Ar diluent was also introduced into the region, 12 cm downstream of the plasma. Commercially supplied 10–20 Ω cm p-type Si wafers of (111) orientation were used as substrates. The modified RCA method, (i) dipping in $H_2SO_4:H_2O_2=3:1$ solution for 10 min and deionized (DI) water rinse, (ii) dipping in $HF:H_2O=1:7$ solution for 30 s and DI water rinse, was used for the predeposition cleaning. After the substrate was loaded in the reaction chamber, it was baked for 30 min to eliminate moisture from the wet cleaning. Thermal oxidation was carried out at 1000 °C in a furnace using ultrahigh-purity oxygen. Deposi-

[a]Present address: Department of Mechanical Engineering, Ohio State University, 206 West 18th Avenue, Columbus, Ohio 43210-1107; Electronic mail: ypark@vision.postech.ac.kr
[b]Author to whom all correspondence should be addressed; Electronic mail: srhee@vision.postech.ac.kr

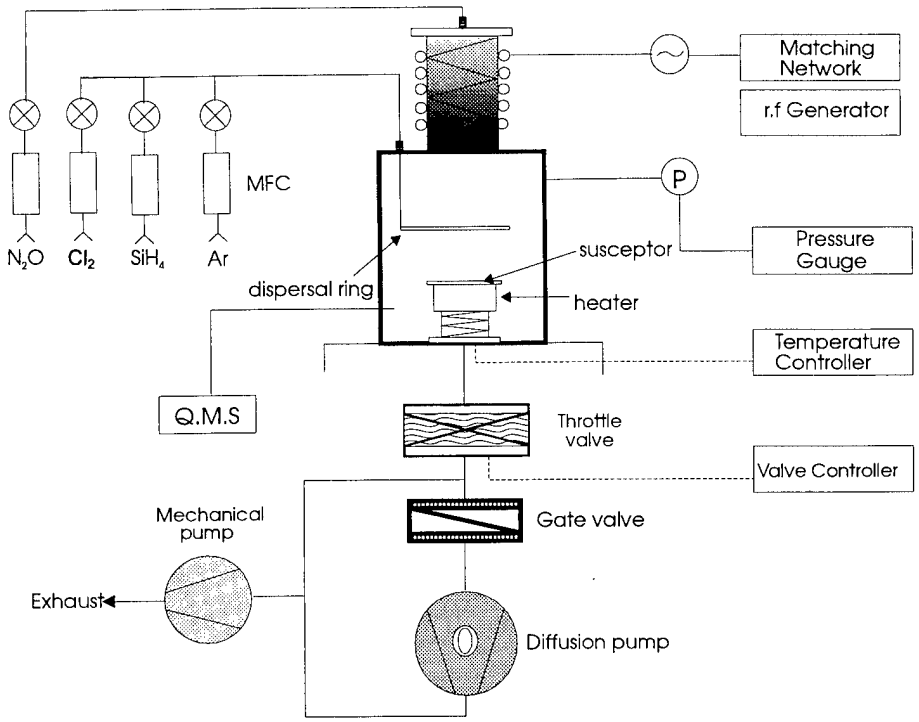


FIG. 1. Schematic diagram of the RPECVD system.

tion was performed at 250–300 °C. The amount of Cl_2 addition was 0 and 10 vol % of total gas flow. The total gas flow rate including Ar was 100 sccm, the N_2O flow rate was 12 sccm, and the SiH_4 flow rate was 3 sccm. The reactor pressure was 400 mTorr and the plasma power was 50 W.

B. Film characterization

The oxide thickness and refractive index were measured with an ellipsometer (NIIC EL-101A). As for the AFM observation of the Si/SiO_2 interface, optimization of the oxide stripping condition is essentially required. We used the HF-dipping method and showed the interfacial morphology in a three-dimensional real space. The chemical etch of the oxide film was performed in the buffered HF solution (49% $HF:H_2O:C_2H_5OH=1:10:110$) at room temperature.[14] The replica of the Si/SiO_2 interface may be preserved by such a treatment, because bare Si was not attacked by the diluted HF-based solution. After the HF dipping, the substrates were rinsed in DI water for 4 min and dried by blowing with dry N_2 gas. It is thought that the Si dangling bonds on the surface are terminated by hydrogen, which prevents surface reoxidation in air and damage of the original Si/SiO_2 interface morphology.

Cross-sectional specimens for TEM were first glued face to face, cut, then ground, dimpled, and thinned by ion milling. TEM observations were conducted using JEOL JEM-

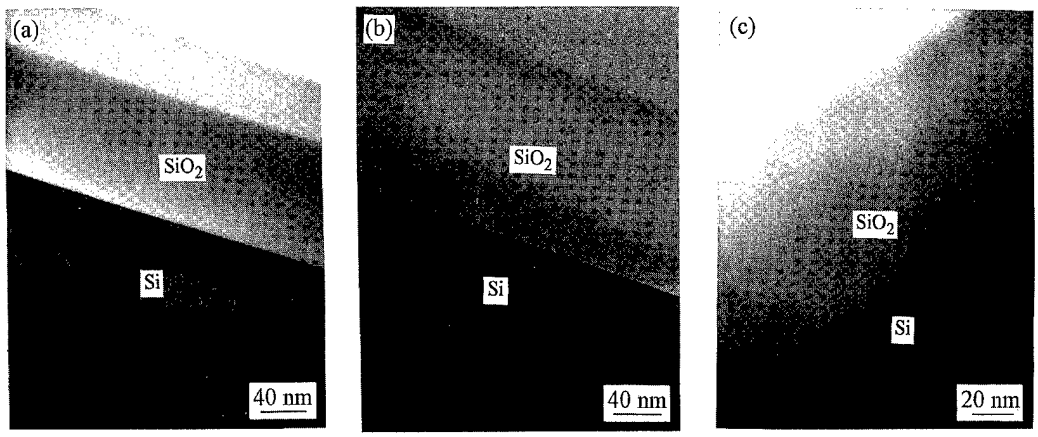


FIG. 2. TEM images of the Si/SiO_2 interface including the interface formed by (a) thermal oxidation, (b) CVD without chlorine, and (c) CVD with chlorine addition (6 vol %). PECVD oxides were deposited at the temperature of 300 °C with $SiH_4:N_2O=3:12$ sccm. The pressure was 0.4 Torr.

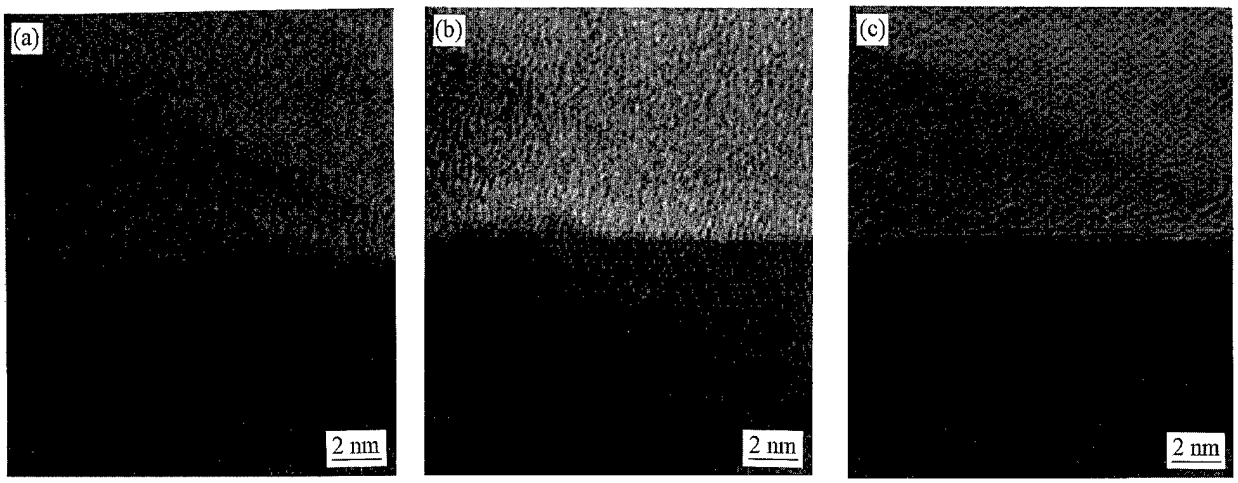


FIG. 3. High resolution TEM lattice images of the Si/SiO_2 interface formed by (a) thermal oxidation, (b) CVD without chlorine, and (c) CVD with chlorine addition (6 vol %). PECVD oxides were deposited at a temperature of 300 °C with SiH_4:N_2O=3:12 sccm. The pressure was 0.4 Torr.

4000FX transmission electron microscope operating at 400 kV. The surface roughness of the SiO_2 film was measured by AFM using a Park Science Instrument (PSI) Autoprobe-CP. The photoelectrons were detected using a Perkin Elmer PHI 5400 XPS with Al $K\alpha$ radiation. The photoelectrons were collected by hemispherical analyzer at a 45° take-off angle. Sputtering for the depth profile was performed at 3.5 kV with Ar ions. Si $2p$ core levels were measured and analyzed using a least-square curve-fitting method. The paramagnetic defects were examined by electron spin resonance using a Bruker ER 200D X-band (microwave frequency: 9.45 GHz) spectrometer at room temperature.

For electrical characterization, metal–oxide–semiconductor (MOS) capacitors were fabricated with deposited oxides. The aluminum dot electrode, 0.05 cm in diameter, was deposited by thermal evaporation through a metallic mask. The $C-V$ characteristics were analyzed at high frequency (1 MHz) using a Hewlett–Packard 4275 multifrequency LCR meter with a sweep voltage range of +10 to −10 V. The flat band voltage and the concentration of oxide fixed charges were obtained from the high frequency $C-V$ data. The interface trap density at the Si/SiO_2 interface was also obtained from the $C-V$ data using the Terman method.[15]

III. RESULTS AND DISCUSSION

Figure 2 shows the TEM images of the Si/SiO_2 interface formed by thermal oxidation and chemical vapor deposition (CVD) from SiH_4/N_2O and $SiH_4/N_2O/Cl_2$. Sections with homogenous oxide thickness were obtained with these three samples. The thickness of the SiO_2 sample was found to be about 96, 140, and 100 nm, respectively. The pattern of thickness fringes was observed in the Si substrate below the oxide in these three samples, indicating oxide-induced strains in the lattice.[16] For the thermally oxidized sample, three straight fringes with clear contrast exist in the depth range from the surface of the Si substrate to a depth about 12 nm, whereas only one fringe with poor contrast was observed for the SiO_2 samples by CVD with and without chlorine. This indicates that the thermal oxidation induces strong interface strains. The fringe in the SiO_2 sample with chlorine addition is thicker and more nonuniform than that in the

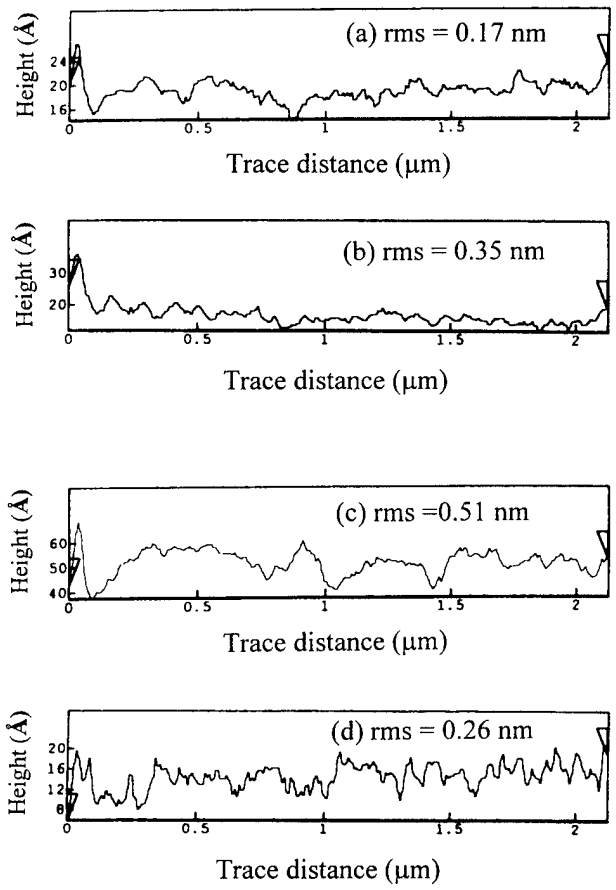


FIG. 4. Two-dimensional line scanning of AFM on the Si surface after stripping SiO_2: (a) bare Si, (b) thermal oxidation, (c) CVD without chlorine, and (d) CVD with chlorine addition (6 vol %). PECVD oxides were deposited at a temperature of 300 °C with SiH_4:N_2O=3:12 sccm. The pressure was 0.4 Torr.

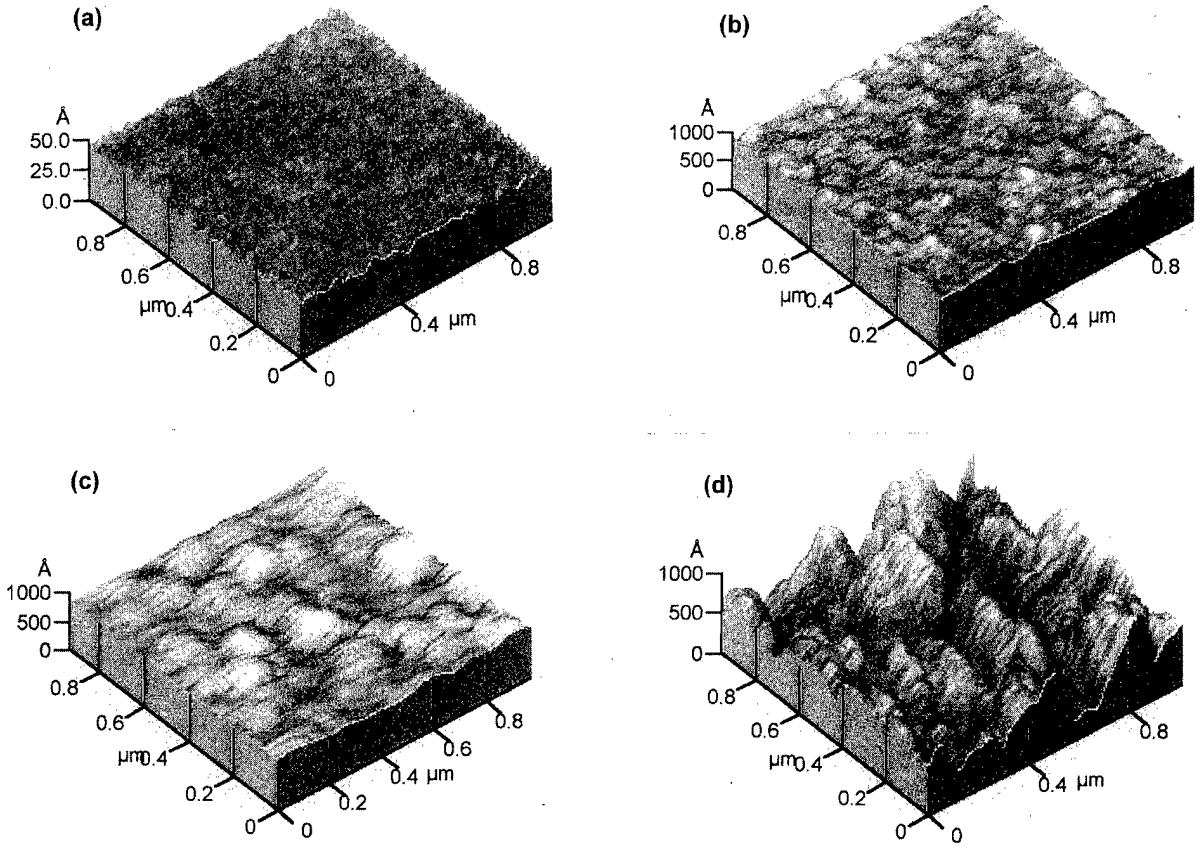


FIG. 5. Three-dimensional AFM for the deposited SiO_2 surfaces from (a) thermal oxidation, (b) CVD without chlorine, (c) CVD with chlorine addition (6 vol %), and (d) CVD with chlorine addition (10 vol %). PECVD oxides were deposited at a temperature of 250 °C with SiH_4:N_2O=3:12 sccm. The pressure was 0.4 Torr.

sample without chlorine addition, indicating that the strains generated by the chlorine-free CVD SiO_2 are stronger and are distributed more nonuniformly than those by the chlorine-added CVD SiO_2.

Figure 3 shows the high resolution TEM lattice images of the Si/SiO_2 interface formed by thermal oxidation and chemical vapor deposition from SiH_4/N_2O and $SiH_4/N_2O/Cl_2$. Taking the largest protrusion or step height observed in a sample as a measure of the Si/SiO_2 interface roughness, one sees that the interface roughnesses of the thermally oxidized, chlorine-free, and chlorine-added CVD SiO_2 samples are about 1, 0.67, and 0.33 nm, respectively. For the thermal oxidation, it has been reported that the interface roughness results from protrusions of Si into the SiO_2 rather than oxidized valleys in the silicon surface.[17] This is suggestive of locally retarded oxidation rather than a local acceleration of the reaction rate. The interface protrusions lead directly to a variation in oxide thickness. The local thickness variation is indicative of nonuniform oxidation.

Compared with thermal oxidation, RPECVD affects the underlying substrate much less.[18] It is obvious that Cl_2 addition decreases interface roughness [Fig. 3(c)]. Lattice strain fields in Si near the interface were observed in these three samples. The lattice strain field in the thermally oxidized sample is more irregular than those in the CVD SiO_2 samples. The strain field in the RPECVD sample with chlorine addition is weaker than that in the chlorine-free CVD SiO_2 sample. It seems that the Cl_2 addition can relax the interface strain by breaking the strained SiO_2 ring structure. Beyond the Si/SiO_2 interface roughness, there were no apparent structural defects or irregularities associated with the thin oxide films. We did not observe any evidence of Si inclusions in the SiO_2, amorphous oxide regions in the near interface Si, crystalline patches of oxide, metallic or other precipitates, or oxide pinholes. In thermal oxidation, Si is consumed to generate the oxide film, so that the Si/SiO_2 interface is located in the bulk of the Si, below the original Si surface. Similarly, the Si/SiO_2 interface deposited by remote PECVD is also formed by a subcutaneous oxidation of Si substrate that occurs at the initial stage of the film deposition.[19,20] This generates unwanted interfacial oxide thickness of 0.5–1 nm. At deposition temperatures above 300 °C and deposited oxide thicknesses above 25 nm, the amount of subcutaneous oxidation increased significantly, and the electronic quality of these interface rapidly degraded. The mechanism for defect generation may be related to a buildup of stress in the subcutaneous layer. This stress derives from a molar volume mismatch between Si and SiO_2

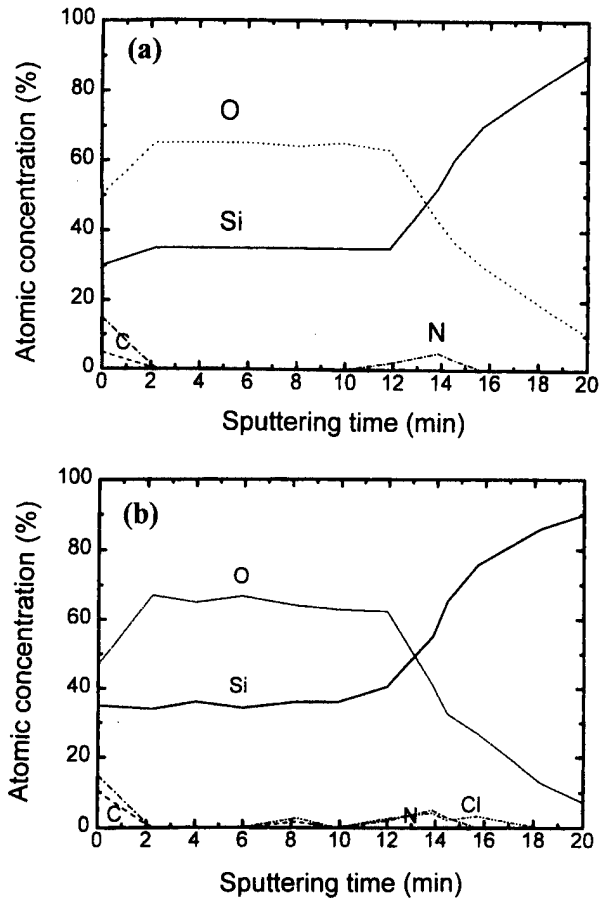


FIG. 6. XPS depth profile of chlorinated SiO_2 samples: (a) CVD without chlorine and (b) CVD with chlorine addition (6 vol %). PECVD oxides were deposited at a temperature of 300 °C with SiH_4:N_2O=3:12 sccm. The pressure was 0.4 Torr.

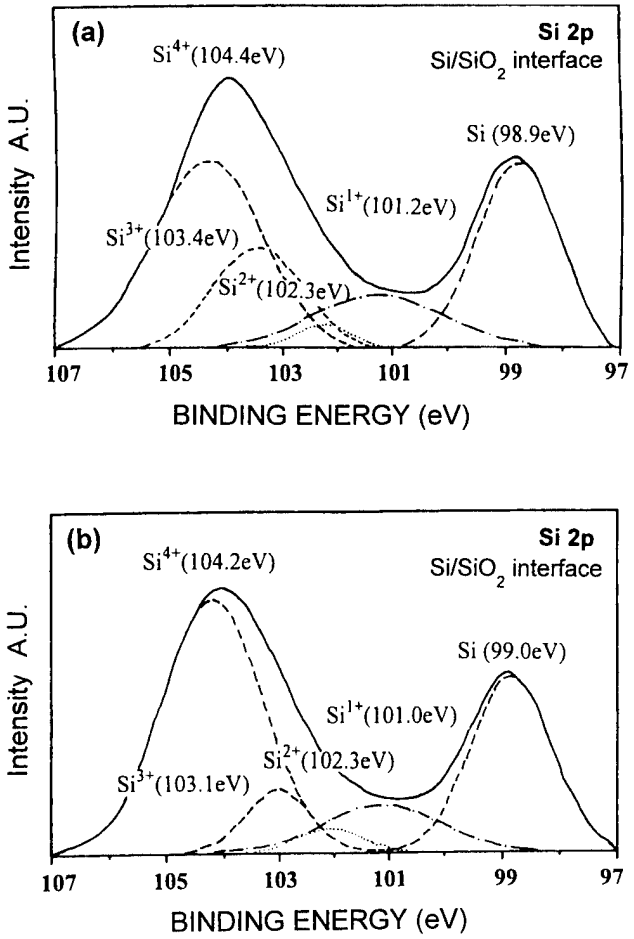


FIG. 7. XPS curve fitting for the Si $2p$ core level: (a) CVD without chlorine and (b) CVD with chlorine addition (2 vol %). PECVD oxides were deposited at a temperature of 300 °C with SiH_4:N_2O=3:12 sccm. The pressure was 0.4 Torr.

and cannot be relieved through a viscoelastic relaxation process at low temperatures in this RPECVD process. However, when chlorine was added into this RPECVD process gases, chlorine can strongly adsorb to an initial Si surface and prevent further subcutaneous oxidation from active oxygen species generated by remote plasma. It is believed that chlorine addition improves the surface states by reducing surface roughness and interface trap density.

Figure 4 shows two-dimensional line scanning for the Si/SiO_2 interface using AFM after the oxide film was removed. Compared with bare Si (rms=0.17 nm) after HF dipping, the rms values of surface roughness are 0.35 nm for thermal oxidation and 0.51 and 0.26 nm for CVD without and with chlorine addition, respectively. The AFM result for the Si/SiO_2 interface roughness gives us a better guide because high resolution TEM is focused only on a very small spot. CVD with chlorine addition has the lowest interfacial roughness. The roughness at the interface reduces the mobility of charge carriers in MOS-based devices, and a local field enhancement across the oxide lowers the oxide breakdown field strength. Interface states and also protrusions or hillocks can act as electron emitters, thus reducing the positive charge centers near the interface.[21]

Figure 5 shows three-dimensional (3D) AFM images of CVD oxide surface with chlorine addition. It can be seen that the oxide surface roughness was dependent on the amount of chlorine addition and the deposition temperature. The rms values of surface roughness for thermal oxide and CVD SiO_2 without chlorine addition were 0.2 and 1.6 nm, respectively. When the chlorine was added, the rms value was increased from 3.1 nm (2 vol % addition) to 4.2 nm (6 vol % addition) at a deposition temperature of 300 °C. The rms value was 15.2 nm with 10 vol % addition. Surface roughening was mainly due to the surface desorption of halide groups such as $Si–Cl_x$ and O–Cl.

Figure 6 shows the XPS depth profile with Ar ion sputtering. With chlorine addition, about 5 at. % chlorine atoms, irrespective of the amount of chlorine addition, piles up at the Si/SiO_2 interface. It is believed that chlorine addition affects the initial stages of deposition, i.e., neutralizes the Si dangling bonds through the formation of Si–Cl bonds. The substitution of the stronger Si–Cl bonds for the weak Si–O bonds can relax the strains at the Si/SiO_2 interface. The film stoichiometry, defined by the O/Si ratio, was still lower than 2, and there is no detectable chlorine in the bulk oxide. It

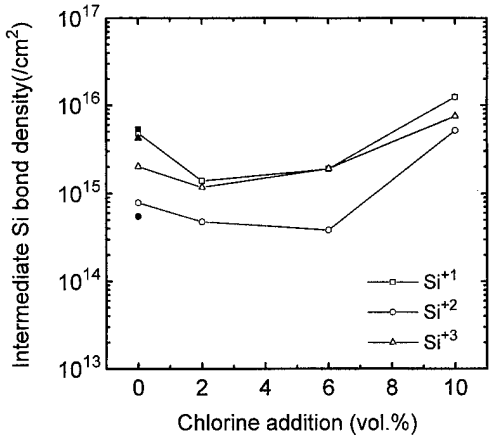


FIG. 8. The density of various intermediate Si states as a function of chlorine addition. The closed symbols represent the thermal oxide. PECVD oxides were deposited at a temperature of 300 °C with SiH_4:N_2O=3:12 sccm. The pressure was 0.4 Torr.

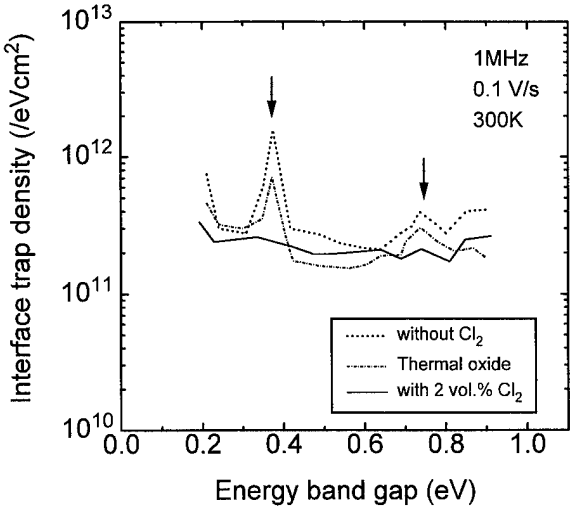


FIG. 9. Interface trap density distributions in the Si energy band gap: (·-·-) thermal oxidation, (···) CVD without chlorine, and (—) CVD with chlorine addition (2 vol %). PECVD oxides were deposited at a temperature of 300 °C with SiH_4:N_2O=3:12 sccm. The pressure was 0.4 Torr.

seems that the chlorine in the bulk oxide etches out the SiO_2 ring structure and desorbs out.

Figure 7 shows Si $2p$ spectra at the interface for CVD SiO_2 without and with chlorine addition. Si $2p$ peaks at the Si/SiO_2 interface show the chemical binding state of Si at the Si/SiO_2 interface at various deposition conditions. As mentioned in the TEM result, the transition (strained and mismatched) layer consisted of nonbonded silicon in a nonstoichiometric oxide and chloride with intermediate oxidation states, i.e., Si^{1+}, Si^{2+}, and Si^{3+}. The Si $2p$ peak was deconvoluted into intermediate oxidation states, which determine the microscopic oxide interface properties such as interface trap density.[22] The silicon–chlorine bond is located at 100.15 eV for Si–Cl, 101.22 eV for $Si–Cl_2$, and 102.34 eV for $Si–Cl_3$, respectively,[23] but no attempt is made to deconvolute silicon suboxides and subchlorides.

Figure 8 shows the intermediate Si bond density obtained by XPS based on the well established method of calculating the suboxide density, N_{SiO_x},[24]

$$N_{SiO_x}=n_{Si}\lambda_{Si}\frac{I^{+X}}{I^0}\frac{\sigma_{Si}}{\sigma_{SiO_x}}\sin(\theta), \quad (1)$$

where n_{Si} (5.0×10^{22} cm^{-3}) is the atomic density of silicon, λ_{Si} (0.33 nm) is the photoelectron mean free path in silicon,[25] I is the integrated peak intensity of the suboxide or the bulk silicon, θ (45°) is the photoelectron take-off angle, σ is the photoionization cross section ($\sigma_{SiO_x}/\sigma_{Si}$ = 1.0, 1.1, and 1.7 for Si^{1+}, Si^{2+}, and Si^{3+}, respectively).[25] The calculated values for the intermediate densities may not be absolute but are useful in showing relative changes as a function of deposition parameter. For the Si/SiO_2-based devices, properties of interfaces will certainly be affected by the microscopic structure at the Si/SiO_2 interface. With increasing chlorine addition, the intermediate silicon bond density is decreased compared with the thermal oxide and the CVD oxide without chlorine addition. The decrease of the intermediate silicon bond density means that chlorine at the interface reduces the mismatched suboxide density. At high chlorine partial pressure, however, the intermediate silicon bond density increased because the etching process dominantly occurs in competition with deposition as shown in the AFM image [Fig. 5(d)]. It shows that the deposited SiO_2 film surface is also roughened due to the desorption of halide group. Therefore it is possible that too much chlorine addition in the SiO_2 deposition process may cause the degradation of bulk properties as well as interface properties.

Figure 9 shows the interface trap density distributions of the interface formed by thermal oxidation and CVD without and with chlorine addition. The peaks at the valence band (E_v)+0.3–0.35 eV and 0.7–0.75 eV mainly appeared and they have been reported to be the trivalent Si dangling bond ($\cdot Si{\equiv}Si_3$, where $\cdot$ denotes an unpaired electron and $\equiv$ denotes three back bonds to the central atom's nearest neighbors) near the Si/SiO_2 interface.[26] These Si dangling bonds are called P_b centers. This center is a major source of interface traps in the 0.15–0.9 eV range of the Si energy band gap. The interface trap densities at E_v+0.3–0.35 eV for the thermal oxidation and CVD without and with chlorine addition are 1.2×10^{12}, 8×10^{12}, and 3.5×10^{11} eV^{-1} cm^{-2}, respectively. The CVD SiO_2 without chlorine addition has the highest interface trap density. It is clear that the Cl_2 addition during the RPECVD decreases the interface trap density greatly. The chlorine incorporation could reduce the interface trap density due to the passivation of positive (Si^{1+}) Si dangling bonds through the formation of $Si–Cl_x$ bonds and inhibit the incorporation of hydrogen and water through the formation of H–Cl or HO–Cl bonds.[10]

Poindexter *et al.*[26] have extensively studied the P_b center defect using ESR and found a very strong correlation between the P_b center spin density and the interface trap density in oxide on both Si(111) and (100) surfaces. Figure 10 shows ESR spectra for three differently prepared CVD

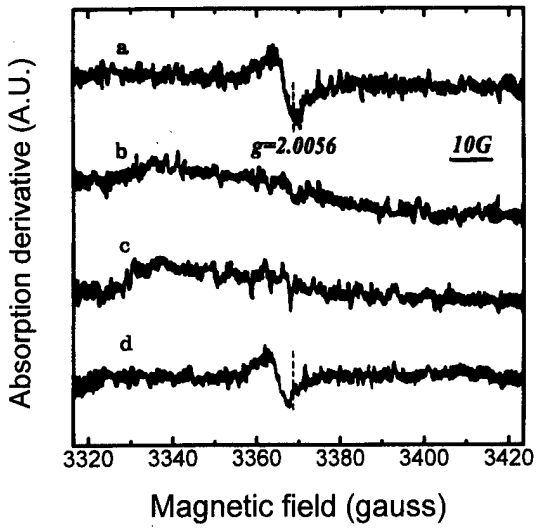


FIG. 10. ESR spectra of (a) CVD without chlorine, (b) and (c) CVD with chlorine (2 and 4 vol %) SiO_2. PECVD oxides were deposited at a temperature of 250 °C with SiH_4:N_2O=3:12 sccm and the pressure was 0.4 Torr. (d) The thermal oxide.

samples compared with the thermal oxide, and peak positions are around $g=2.0056$.

The correlation between the interface trap at $E_v+0.35$ and 0.75 eV and the peak amplitude of the P_b center at $g=2.0056$ in ESR measurements is shown in Fig. 11. We can conclude that interfacial states, resulting from lattice mismatch, suboxide states, interface roughness, and nonbonded defects are passivated by the addition of small amounts of chlorine during the low temperature CVD process.

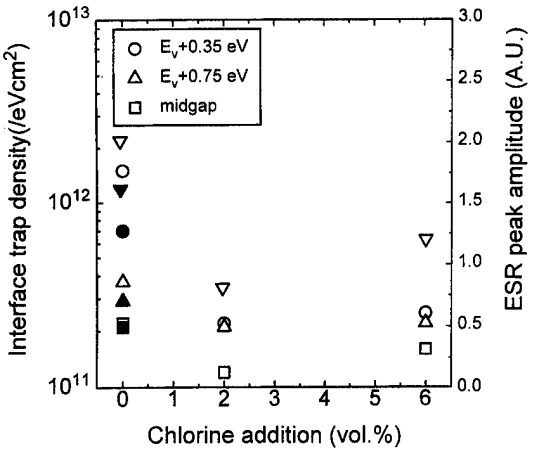


FIG. 11. Correlation between the interface trap density at $E_v+0.35$ and 0.75 eV and ESR amplitude at $g=2.0056$ for samples shown in Fig. 10. The interface trap density at $E_v+0.35$ eV (△) and +0.75 eV (○), at midgap (□), and at the peak amplitude of ESR (▽). The closed symbols represent the thermal oxide (▲, ●, ■, and ▼).

IV. CONCLUSIONS

We have studied and compared the Si/SiO_2 interfaces formed by thermal oxidation and RPECVD with or without chlorine addition. The thermal oxide has a rough interface and the interface of the CVD SiO_2 without chlorine addition is rougher than that with chlorine addition. The thermal oxidation induces strong interface strains. The strains generated by the CVD SiO_2 without chlorine addition are stronger and are distributed more nonuniformly than with chlorine addition. It was also found out that with chlorine addition during RPECVD, the occurrence of strained layers, the interface trap density, and the suboxide density could be lowered significantly. It is suggested that chlorine addition during RPECVD affects the initial stages of deposition, i.e., neutralizes the Si dangling bonds through the formation of Si–Cl bonds. Although the large amount of chlorine addition roughened the deposited SiO_2 surface, we have discovered the appropriate amount of chlorine addition in order to get device quality low temperature SiO_2 films.

ACKNOWLEDGMENTS

This research was supported by POSCO, and the support of The Engineering Research Center for Interface Science and Technology of Materials is also acknowledged.

[1]G. Lucovsky and D. V. Tsu, J. Vac. Sci. Technol. A **5**, 2231 (1987).
[2]L. Meiners, J. Vac. Sci. Technol. **21**, 655 (1982).
[3]R. J. Jaccodine, in *The Physics and Chemistry of SiO_2 and the $Si–SiO_2$ Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993), p. 45.
[4]H. Shiraki, Jpn. J. Appl. Phys. **15**, 83 (1976).
[5]K. Hirabayashi and J. Iwamura, J. Electrochem. Soc. **120**, 1595 (1973).
[6]D. R. Young and C. M. Osburn, J. Electrochem. Soc. **120**, 1578 (1973).
[7]R. J. Kiegler, Appl. Phys. Lett. **20**, 449 (1972).
[8]C. M. Osburn, J. Electrochem. Soc. **121**, 809 (1974).
[9]R. J. Kiegler, Y. C. Cheng, and D. R. Colton, J. Electrochem. Soc. **119**, 388 (1972).
[10]Y.-B. Park and S. Rhee, Appl. Phys. Lett. **66**, 3477 (1995).
[11]J. R. Monkowski, R. E. Tressler, and J. Stach, J. Electrochem. Soc. **125**, 1867 (1978).
[12]C. Falcony, J. C. Alonso, A. Ortiz, M. Garcia, E. P. Zironi, and J. Rickards, J. Vac. Sci. Technol. A **11**, 2945 (1993).
[13]A. Ortiz, S. Lopez, C. Falcony, M. Farias, L. Cota-Araiza, and G. Soto, J. Electron. Mater. **19**, 1411 (1990).
[14]M. Niwa, H. Iwasaki, and S. Hasegawa, J. Vac. Sci. Technol. A **8**, 266 (1990).
[15]L. M. Terman, Solid-State Electron. **5**, 285 (1962).
[16]F. Banhart, Ultramicroscopy **56**, 233 (1994).
[17]A. H. Carim and R. Sinclair, J. Electrochem. Soc. **134**, 741 (1987).
[18]G. G. Fountain, S. V. Hattangady, R. A. Rudder, R. J. Markunas, G. Lucovsky, S. S. Kim, and D. V. Tsu, J. Vac. Sci. Technol. A **7**, 576 (1989).
[19]G. Lucovsky, S. S. Kim, D. V. Tsu, G. G. Fountain, and R. J. Markunas, J. Vac. Sci. Technol. B **7**, 861 (1989).
[20]S. S. Kim, D. J. Stephens, G. Lucovsky, G. G. Fountain, and R. J. Markunas, J. Vac. Sci. Technol. A **8**, 2039 (1990).
[21]S. M. Goodnick, R. G. Gann, J. R. Sites, D. K. Ferry, C. W. Wilmsen, D. Fathy, and O. L. Krivanek, J. Vac. Sci. Technol. B **1**, 803 (1983).
[22]Z. H. Lu, M. J. Graham, D. T. Jiang, and K. H. Tan, Appl. Phys. Lett. **63**, 2941 (1993).
[23]C. C. Cheng, K. V. Guinn, V. M. Donnelly, and I. P. Herman, J. Vac. Sci. Technol. A **12**, 2634 (1994).
[24]P. J. Grunthaner, M. H. Hecht, F. J. Grunthaner, and N. M. Johnson, J. Appl. Phys. **61**, 629 (1987).
[25]F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Holliger, Phys. Rev. B **38**, 6084 (1988).
[26]E. H. Poindexter, G. J. Geradi, M. E. Rueckel, and P. J. Caplan, J. Appl. Phys. **56**, 2844 (1984).

Oxide growth on silicon (100) in the plasma phase of dry oxygen using an electron cyclotron resonance source

Keunjoo Kim, M. H. An, Y. G. Shin, M. S. Suh, C. J. Youn, Y. H. Lee, K. B. Lee, and H. J. Lee

Department of Physics and Semiconductor Physics Research Center, Jeonbuk National University, Jeonju 560-756, Korea

(Received 22 November 1995; accepted 1 May 1996)

Silicon dioxide films were grown using an oxygen plasma generated by an electron cyclotron resonance (ECR) source at several low temperatures. The plasma oxidation rate was investigated by varying the growth parameters. The oxide thickness parabolically increases with microwave power but decreases with increasing pressure or flow rate. A complementary model of the Deal–Grove oxidation theory is suggested for the plasma oxidation, and kinetic parameters are compared with the other plasma and thermal oxidation cases. The diffusion rate constant in O_2 plasma oxidation at room temperature is enhanced up to the level of the diffusion rate in thermal oxidation and the reaction rate constant is much larger than the thermal oxidation case. This may imply that, due to oxygen atoms dissociated by the ECR plasma, plasma oxidation is related to the atomic diffusion through oxide layer and the atomic chemical reaction at the Si–SiO_2 interface. The high quality of the ultrathin oxide film was characterized with a breakdown field of 14.8 MeV/cm and an interfacial state density of 1.2×10^{10} eV^{-1} cm^{-2}. © *1996 American Vacuum Society.*

I. INTRODUCTION

In silicon technology, the quality of ultrathin oxide films becomes a crucial factor for ultralarge scale integration (ULSI) devices because of decreasing sizes of integrated circuits. The conventional thermal oxidation of silicon gives rise to questions about applications in ULSI technology. At high temperatures, potential problems of the thermal oxidation such as dopant diffusion, impurity redistribution, defect formation, and thermally induced stress can seriously degrade the device performance.[1] Therefore, the low-temperature oxidation process is highly desirable for high quality ultrathin dielectric films. In order to reduce the processing temperature and to increase the growth rate, many studies on the sputtered gate oxide films and the films deposited by the plasma enhanced chemical vapor deposition (PECVD) were undertaken.[2] However, the deposited films show interfacial instability with large gap states generated by hydrogen atoms which are included in the deposition process, and so the degraded electrical property is not comparable with that in thermal oxides.

High quality films can be grown by low-temperature plasma oxidation. As one kind of plasma oxidation process, plasma anodization has improved the low oxidation rate of the plasma oxidation. Anodization with two electrodes that apply the external positive bias to the sample enhances the diffusion of oxygen ions into oxide layer. Unfortunately, the contamination caused by electrode sputtering is unavoidable.[3] O_2 plasma oxidation without an external bias shows an even larger oxidation rate than the thermal oxidation case. Plasma oxidation without external bias increases as the device geometry of the ULSI decreases. Recently, an electron cyclotron resonance (ECR) plasma source has been developed and has several advantages such as a low plasma potential, a high ionization rate (10%), dissociation of oxygen in low pressure (mTorr), and a long plasma lifetime. ECR plasma oxidation shows that the dielectric property of the film is comparable to that of thermal oxides.[4–6]

Usually, the driving force for both the chemical reaction and the diffusion process is the thermal activation energy in the thermal oxidation of silicon. Instead of using the thermal energy, plasma oxidation utilizes a highly activated oxygen plasma including excited neutral oxygen atoms and reactive radicals of oxygen ions. In an ECR plasma, there are optical emission spectra for neutral oxygen atoms (O I) and molecular oxygen ions (O_2^+).[4,5] Without the ECR condition, the oxygen plasma emits spectra not only from O I and O_2^+, but also from O^+.[7] However, Kimura *et al.*[4] suggested the migration of O^- ions into the oxide. Carl, Hess, and Lieberman[5] also suggested that the strong electric field induced by the plasma self-bias on the oxide layer can conduct O^- ions through the oxide. Vinckier *et al.*[8] showed that charged species generated from the microwave discharge in afterglows have no effect on the oxidation rate due to poor ionization efficiency. Joseph, Hu, and Irene[9] demonstrated the effect of applied bias, suggesting O^- as the main oxidizing species. These experimental findings contradict each other and provide little explanation for the mechanism of the plasma oxidation without an external bias. The plasma oxidation kinetics parameters derived from the Deal–Grove model[10] are still controversial. The reaction rates have been reported with a positive value by Carl *et al.* but with a negative value by Kimura *et al.*

In this work, we systematically analyze the kinetic parameters from the complete form of the oxidation model. Furthermore, the ECR plasma oxidation experiment without an applied bias on samples was carried out to confirm the plasma oxidation kinetics, and the oxidation rate was studied in terms of processing pressure, microwave power, substrate temperature, and oxidation time. The experimental results for plasma oxidation were fitted by the present formalism of oxidation theory. Furthermore, the electrical properties of the

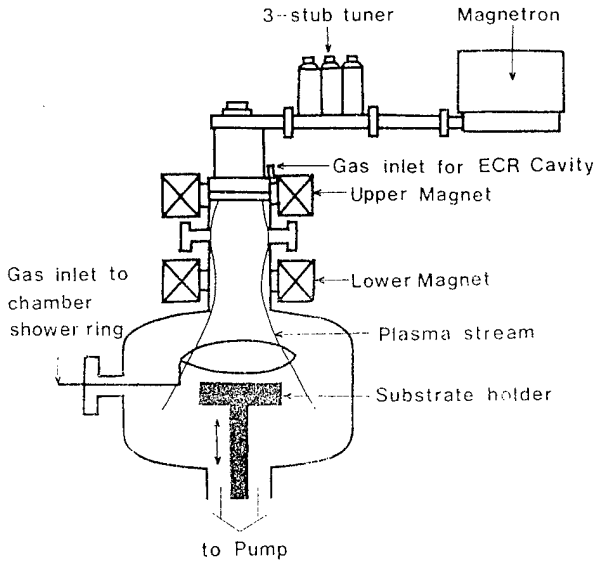


FIG. 1. Schematic of the divergent magnetic field-type ECR source.

breakdown field and the interfacial state density were characterized from metal–oxide–semiconductor (MOS) capacitor structures.

II. EXPERIMENTAL METHODS

A cross-sectional schematic view of the ECR system of Astex, Inc. is shown in Fig. 1. A microwave generated by a magnetron was guided through the waveguide to the ECR cavity. The waveguide was separated from the ECR cavity by a quartz disk. The ECR condition was adjusted by the three-stub tuner and accordingly the electromagnetic mode could be changed. In order to reproduce the process, it was important to choose a distinct electromagnetic mode that could provide a stable plasma. The ECR source was excited by a microwave power supply at 2.45 GHz with the range 50–1000 W. The ECR source was surrounded by magnetic solenoids to get the resonance condition. Sufficient electron kinetic energy could be obtained by introducing a steady magnetic field $B=875$ G for the electron cyclotron frequency $\omega_c=eB/mc$ in the resonance chamber, which could resonate with the applied microwave frequency $\omega=2.45$ GHz.

Pure oxygen gas (99.99%) entered the ECR source through a shower ring. The distribution of magnetic flux was of mirror type in the resonator and of divergent type in the discharge chamber. This divergent magnetic field could accelerate the activated oxygen radical stream to the sample by a magnetic longitudinal force regardless of the species charge. The samples were negatively self-biased with 12 V from the induced sheath potential of the ECR plasma without the applied bias. The distance between the shower ring to the sample stage was adjustable and set to 7.5 cm to optimize the uniformity of the plasma density. The sample stage was made of stainless steel and the temperature was measured using a resistance temperature detector placed directly underneath the stage. The temperature of the substrate holder could be controlled from room temperature up to 400 °C. The reflected power was kept below 10% of the forward microwave power. To maintain the low pressure, the system was evacuated by a 1000 ℓ/s turbomolecular pump. The chamber was pumped out to a base pressure of 2×10^{-6} Torr before each run. Plasma oxidation was processed in the pressure range of about $(1-2)\times10^{-4}$ Torr. The processing pressure is directly related to the oxygen flow rate.

The samples used for oxidation were boron doped Si(100) with a resistivity of 3–10 Ω cm and 7.5 cm in diameter. These samples were cleaned with a standard RCA-type wet chemical clean. The organic material on the surface could be dissolved in a boiling 4:1 mixture of H_2SO_4 and H_2O_2 at 110 °C for 10 min. The samples were dipped in a 20:1 diluted solution of H_2O_2:HF for 10 s to etch the native oxide layer. The wafers were also rinsed in deionized H_2O at room temperature after each of these steps. However, in spite of the wet chemical cleaning, native oxides with a thickness of 10 Å still existed. The oxide thickness was measured by a Rudolph ellipsometer II. A nominal index of refraction of 1.462 was used for oxide thickness measurements. The breakdown field and interface charge trap density were obtained from $I-V$ and $C-V$ measurements on MOS capacitors. These MOS capacitor structures were fabricated by depositing 1-μm-thick Al on both sides of wafers. The dot size of the capacitors was 7.85×10^{-3} cm^2 in area.

III. GROWTH PROPERTIES IN PLASMA OXIDATION

Instead of using thermal activation in conventional high-temperature oxidation, the ECR plasma enhances oxidation by increasing the activated oxygen species, and so plasma oxidation is useful for a low-temperature process in semiconductor technology. When a low-pressure gas is introduced, the gas breaks down and a discharge forms inside the chamber. The oxygen plasma diffuses along the expanding magnetic field lines into a process chamber toward a wafer holder. The oxidation rate asymptotically decreases with increasing pressure or O_2 flow rate as shown in Fig. 2. This is due to the higher plasma density at lower pressure as mentioned by other groups.[6,11] However, at very low pressure, the oxidation rate decreases with increasing intensity of the O_2^+ ion peak.[6]

The optical emission spectroscopy (OES) shows several peaks of excited neutral atomic oxygen (O I) and positively charged molecular oxygen (O_2^+). There are several spectrum lines associated with O I in plasma at 4368, 6158, 5329, 3947, 7775, and 8446 Å.[4,5] There are also a number of relatively broad bands associated with ionized molecular oxygen (O_2^+), the most prominent of which are at 5251, 5597, 5973, and 6351 Å.[4,12] In particular, Kimura *et al.* demonstrated the oxygen pressure dependence of the emission intensity from O I (436.8 nm) and O_2^+ (525.1 nm). The maximum intensities of two species are located at 0.1–0.2 mTorr, consistent with our pressure dependence of the oxide thickness as shown in Fig. 2. They showed that the dissociation of O_2 molecules into O I with an energy of 5 eV likely occurs in the high-pressure region due to low electron cyclotron energy, while

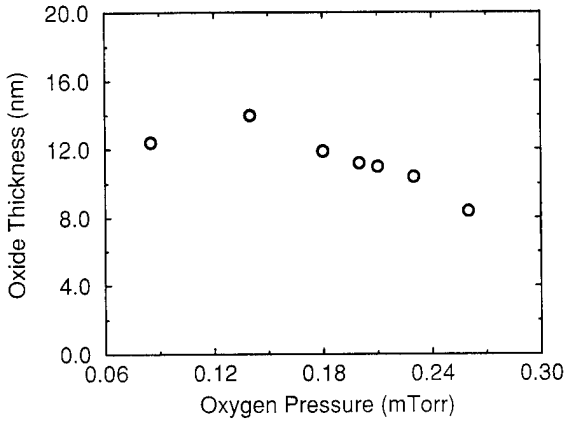


FIG. 2. Oxygen pressure dependence of oxide thicknesses with 1000 W at room temperature for 30 min. The thickness is increased by lowering the oxygen pressure, which is inversely related to the thermal oxidation case, and the pressure for the maximum oxidation rate is consistent with the maximum intensities of the oxidizing plasma species (Ref. 4).

the electron energy is high enough to give rise preferentially to ionization (13 eV) in the low-pressure region.

The effect of microwave power on the oxide thickness is shown in Fig. 3. The oxide films were grown under conditions such that the flow rate was 10 sccm with a pressure of 0.1 mTorr and a 7.5 cm distance to source for 30 min. The parabolic behavior of the oxidation rate to the incident microwave power P_W is observed and fitted with the curve of

$$x - x_0 = a_0 P_W^{1/2}, \tag{1}$$

where x, x_0, and a_0 are the oxide thickness grown for 30 min, the native oxide thickness, and the proportional constant, respectively. According to experimental data of present work, Sung and Pang[6] and Vinckier *et al.*,[8] the native oxide thicknesses are given by 10, 15, and 18.7 Å, respectively. The constant a_0 is fitted to 0.5 for both the present work and

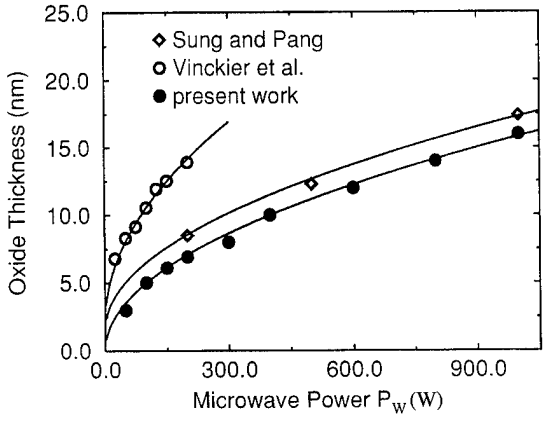


FIG. 3. Oxide thickness as a function of the microwave power. The experimental conditions are following: pressure=1 mTorr and oxidation time =100 min for Sung and Pang (Ref. 6); pressure=1.05 Torr and oxidation time=4 h for Vinckier *et al.* (Ref. 8); pressure=0.1 mTorr and oxidation time=30 min for the present work. As the power increases, the oxide thickness is parabolically increased, implying a direct contribution of the momentum of the O atoms.

the work of Sung and Pang and 0.86 for the work of Vinckier *et al.* The experiment on the dependence of the O-atom concentration on the applied microwave power was carried out by Vinckier *et al.* They showed that the atomic oxygen concentration linearly increases with microwave power and that the oxide thickness is proportional to the O-atom concentration C_O with the following relation:

$$x^{2.17} - x_0^{2.17} = A C_O, \tag{2}$$

where A is a constant. The oxide thickness and power relationship may support that the belief that the O atom is diffused into the interface as the oxidant. The oxidation rate may be directly related to the mass transport or the momentum of the O atom. Vinckier *et al.*[8] also showed that the voltage applied on samples has no effect on the oxidation rate. This implies that charged species generated from the oxygen plasma do not play a role in the oxidation process of silicon and that the major oxidizing species is the O atom. On the other hand, Joseph, Hu, and Irene[10] demonstrated the effect of applied bias on the oxidation rate. They suggested that the positive bias on samples provides significant enhancement of the oxidation process and that the most probable oxidizing species is O^-. The oxidation kinetics by charged species in the limit of a low electric field is compatible with the Cabrera–Mott theory, which purports oxidation occurs from the movement of both Si outward as a cation in the interlayer and the oxidant species O^- inward as an anion. However, there is still no evidence on optical emission spectra of O^- ions in the ECR oxygen plasma without an external bias.

Our concern is for the ECR system without the applied bias on samples and so the sheath potential of -12 V is self-biased between the plasma and the sample. This implies that the O^- ions are repelled by the negative potential on the silicon surface. Therefore, without bias on samples, the primary oxidizing species is concluded to be the neutral excited oxygen atoms generated from the dissociation of O_2 molecules by the ECR source. Some dissociated O atoms can be recombined with electrons on the silicon surface to reduce the large electronegativity of the oxygen atom. Even though growth parameters are different for available experimental data, the oxidation rate as the function of substrate temperature is shown in Fig. 4. Kimura *et al.*[4] and Carl, Hess, and Lieberman[5] demonstrated that the oxidation rate increases with increasing temperatures higher than 350 °C. Sung and Pang[6] showed the low rate at a room temperature of 16 °C and also Lee *et al.*[11] did for the data at 400 °C. Our measurement of the rates below 350 °C shows that the rate slightly decreases with increasing temperature. This may imply that the dissociated neutral oxygen atom is simultaneously diffused into the oxide layer below 350 °C and then thermalized above 350 °C, indicating the inclusion of the thermal diffusion effect on plasma oxidation at high temperatures.

IV. PLASMA OXIDATION KINETICS

Many authors introduced the Deal–Grove model in order to understand plasma oxidation kinetics.[4–6,11] Deal and

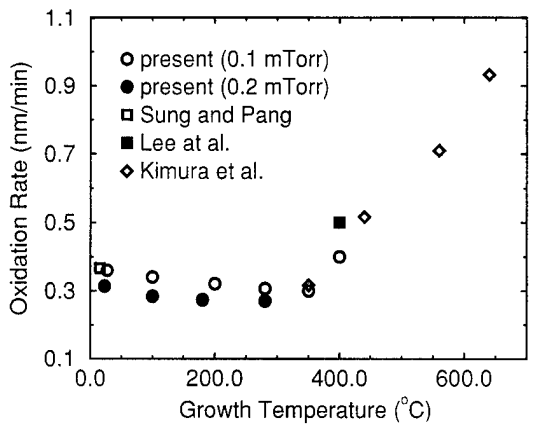


FIG. 4. Temperature dependence of oxidation rates for an oxidation time of 30 min. Despite the different experimental growth conditions, there is an obvious tendency of the growth rate. The rate gradually decreases below 350 °C, while it increases above that temperature.

Grove proposed that the oxidation rate was determined by a combination of two processes. One is the actual chemical reaction of oxygen with silicon at the interface, and the other is the diffusion of oxygen through the previously formed oxide film. The combination of these processes resulted in the linear–parabolic relationship between the oxide thickness x and the oxidation time t,

$$x=\frac{A}{2}\left[\left(1+\frac{4B(t+\tau)}{A^2}\right)^{1/2}-1\right], \tag{3}$$

where

$$A=\frac{2D}{k}, \tag{4}$$

$$B=2D\frac{C_O}{C_{ox}}, \tag{5}$$

and

$$\tau=\frac{x_0^2+Ax_0}{B}, \tag{6}$$

where k is the first-order reaction rate constant for the oxidation, C_O the concentration of oxidant at the oxide surface, C_{ox} the concentration of oxygen molecules incorporated the oxide layer, and D the effective diffusion coefficient of oxygen in the silicon dioxide. x_0 is the native oxide thickness and τ is the corresponding oxidation time. In Eq. (5), B is the parabolic rate constant while the ratio of B/A is the linear rate constant. The rate constants obtained from several groups are summarized in Table I. Since the experimental conditions of the references given in Table I are different from each other, the comparison of the rate constants in terms of the substrate temperature can be arguable. However, the negative value of the reaction rate of B/A is strange even for plasma oxidation, because the negative rate is unphysical. Otherwise, the negativeness of the reaction rate is perhaps a unique characteristic of plasma oxidation, which may differ from thermal oxidation. This may imply that the Deal–Grove theory is not valid for plasma oxidation kinetics, which is involved with the atomic diffusion of O atoms.

According to a modified Deal–Grove model proposed by Tiller,[13] it is shown that B/A can be negative if migration of oxygen ions is considered. In Tiller's model, the effect of field-aided ion migration is added into B/A, B, and τ terms. However, Kimura *et al.* argued that the index of the power law relation exceeds 2, in contrast to Tiller's results of the Deal–Grove model. Note that the index of the exponent in the Deal–Grove model is varied from 1 to 2 according to a change of the mechanism from the linear region in a thin oxide to the parabolic region in a thick oxide. Vinckier *et al.* used an exponent index of 2.17 for their model calculations of the silicon dioxide growth rate[8] from the following relation:

$$x^n-x_0^n=bt, \tag{7}$$

where x_0 is the initial thickness of the native oxide layer and b a constant independent of the oxidation time t. This indicates that before analyzing the plasma oxidation kinetics, the validity of the Deal–Grove model has to be examined fundamentally. We consider again the analysis of the silicon oxidation kinetics of the Deal–Grove model in order to apply it to the plasma oxidation mechanism.

For the diffusion process of O atoms in the oxide layer, Fick's first law is written by

$$F_D=-D\frac{dC(x')}{dx'}, \tag{8}$$

where D is the effective diffusivity of the O atoms in the oxide layer and $C(x')$ is the oxidant concentration at any

TABLE I. Comparison of rate constants for several plasma oxides and thermal oxides. The experimental data were fitted by the Deal–Grove model for 16, 400, and 640 °C (Refs. 4, 6, and 11) and by the power law for 350 °C (Ref. 5). The present experimental result at 27 °C was fitted by the present theory.

Oxidation	T (°C)	B (μm^2/h)	B/A (μm/h)	τ (h)	References
Plasma	27	1.94×10^{-3}	0.388	3.16×10^{-3}	Present work
	16	8.01×10^{-5}	-4.47×10^{-3}	−1.57	6
	350	1.26×10^{-2}	0.21	0.03	5
	400	9.80×10^{-4}	−0.14	−0.34	11
	640	4.92×10^{-4}	-1.23×10^{-2}	−1.44	4
Thermal	920	4.90×10^{-3}	2.08×10^{-2}	1.40	10
	1000	1.20×10^{-2}	0.07	0.37	10

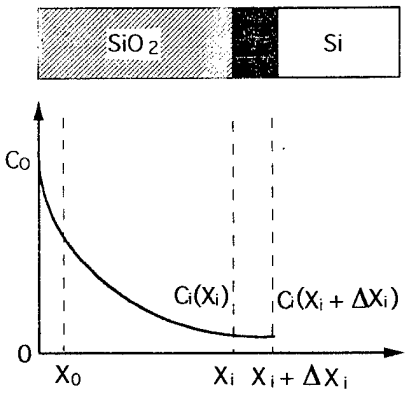


FIG. 5. Schematic diagram for the oxidation of silicon in the O_2 plasma phase. The native oxide is denoted as x_0 and the interface as x_i. In steady state, the whole diffused flux at the Si/SiO_2 interface participates in the chemical reaction.

position x' of the diffused species. It is valuable to note that the above diffusion equation is fundamentally valid for a steady state process.[14,15] The variable of integration x' is any arbitrary distance from the gas/oxide interface in the position $0 \leqslant x' \leqslant x_i$. In this work, however, our concern is limited only to the position of the interface x_i because oxidation is assumed to occur at the interface. As shown in Fig. 5, the diffused flux at the interface can be represented from the gradient of the concentration of oxygen for an infinitesimal element of the interfacial position x_i,

$$F_D = -D \lim_{\Delta x_i \to 0} \frac{C_i(x_i+\Delta x_i) - C_i(x_i)}{(x_i+\Delta x_i) - x_i}, \tag{9}$$

$$= -D \frac{dC_i(x_i)}{dx_i}, \tag{10}$$

where the x_i denotes the position of the interface and C_i is the concentration of oxygen species at the interface. We again remind ourselves that this time-invariant flux at the interface is in a steady state. It is noted that Deal and Grove approximated Fick's law to the linear form of the concentration gradient from the assumption of steady state oxidation. This led to the difference form of the diffusion flux at any point of the interface with the surface concentration C_O,

$$F_D = -D \frac{C_i(x_i) - C_O}{x_i}. \tag{11}$$

The flux involving the oxidation reaction at the interface is related to the interfacial oxidant concentration with the linear reaction coefficient k,

$$F_R = kC_i(x_i). \tag{12}$$

Under the steady state oxidation of the O-atom diffusion in the equation of continuity,

$$\frac{\partial C_i}{\partial t} = -\nabla F = 0, \tag{13}$$

the invariant diffusion flux comes into the bulk oxide and finally reaches the Si/SiO_2 interface. Therefore, the continuity equation is valid for the interface as well as in the bulk oxide. Every diffused flux participates in and is completely being used for the chemical reaction that results in growth of the oxide layer at the interface, implying the relationship of $F_D = F_R = \text{const}$. By equating the diffusion flux at the interface in Eq. (10) to the flux required for the chemical reaction at the interface in Eq. (12), the differential equation provides the oxidant concentration and the flux at the interface from the initial boundary condition of the surface oxidant concentration, C_O,

$$-D \frac{dC_i(x_i)}{dx_i} = kC_i(x_i), \tag{14}$$

$$C_i(x_i) = C_O \exp\left(-\frac{x_i}{L_D}\right), \tag{15}$$

and

$$F_R(x_i) = kC_O \exp\left(-\frac{x_i}{L_D}\right), \tag{16}$$

where $L_D = D/k$ is represented for the characteristic diffusion length in the exponentially decaying incident flux at the interface.

In addition, the exponential term of the flux equation can be approximated in the form of $(1 + x_i/L_D)^{-1}$ for both limits of thin and thick oxide films x_i. This approximation corresponds to the flux form derived by Deal and Grove,

$$F_R = \frac{kC_O}{1 + kx_i/D}. \tag{17}$$

The steady state oxidation theory requires constant flux at the interface for both processes of diffusion and reaction at any instant. When the oxygen concentration at each point in the oxide no longer changes with time, the quantity of oxygen passing through the oxide per unit time is a constant. However, since the flux is the flow per unit area, it is a function of the oxide thickness and again it cannot be simply approximated to linear form. The difference between the present theory and the Deal–Grove theory is the interpretation of the steady state diffusion flux at the interface. Deal and Grove thought that Eq. (8) is not a steady state form and needed revision due to the relationship of $-\nabla F_D = 0$. However, this claim may be challenged because the steady state condition applies only to the time variable and the continuity equation provides information on the time-independent flux of $F_D = -DdC/dx' = \text{const}$, indicating Eq. (8) in a steady state form.[14,15] From now on, let us denote the position of the interface x_i to the oxide thickness x.

The oxidant flux results in oxidation with the growth rate of the oxide layer x through the relationship

$$F_R = C_{ox} \frac{dx}{dt}, \tag{18}$$

where C_{ox} is the concentration of the oxygen incorporated in the oxide. By equating Eqs. (16) and (18), the following

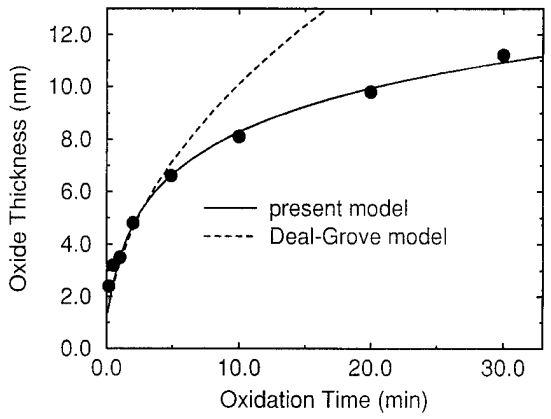


FIG. 6. Time dependence of the oxide growth at room temperature. The oxide films were grown under the following experimental conditions: oxygen flow rate=10 sccm, pressure=0.2 mTorr, microwave power=1000 W, and source to sample distance=7.5 cm. The agreement of the experimental fit for the ultrathin film from the complete form of the silicon oxidation model is remarkable and gives positive kinetic parameters.

differential rate equation and the relationship between the oxide thickness and the oxidation time can be obtained by taking into account the initial oxidation conditions,

$$\frac{dx}{dt}=\frac{kC_{\mathrm{O}}}{C_{\mathrm{ox}}}\exp\left(-\frac{x}{L_D}\right) \tag{19}$$

and

$$x=\frac{A}{2}\ln\left(\frac{2B(t+\tau)}{A^2}+1\right), \tag{20}$$

with

$$\tau=\frac{A^2}{2B}\left[\exp\left(\frac{2x_0}{A}\right)-1\right], \tag{21}$$

where A and B are defined to be the same as the Deal–Grove case in Eqs. (4) and (5). The present oxidation kinetics shows that the oxide layer grows logarithmically with increasing the oxidation time. Therefore, this logarithmic limit shows faster saturation of the diffusion-limited process compared with the parabolic limit from the Deal–Grove model.

The experimental fit from this oxidation theory for the ultrathin oxide film is shown in Fig. 6. The agreement of the experimental fit is remarkable. The oxidation kinetic parameters are compared with others in Table I. The diffusion rate constant (B) of the O-atom species was found to be 1.94×10^{-3} μm^2/h and the reaction rate constant (B/A) at the Si/SiO$_2$ interface was found to be 0.388 μm/h. The time shift constant τ, which is related to the presence of the initial oxide, was found to be 3.16×10^{-3} h. The diffusion length is characterized as 2.5 nm and this means that during a plasma oxidation of 30 min, the oxidation is well into the region where the diffusion-limited process of O atoms into the oxide layer dominates. In comparison, the rate constants for the thermal oxidation of Si in dry oxygen at 920 and 1000 °C were 4.9×10^{-3} and 1.2×10^{-2} μm^2/h for B, 2.08×10^{-2} and 0.07 μm/h for B/A, and 1.4 and 0.37 h for τ, respectively. The results of Carl, Hess, and Lieberman are similar to the thermal oxidation case at 1000 °C for B, but much smaller for B/A. Our result for the diffusion rate of B shows a value similar to that in the thermal oxidation case at 920 °C. However, the reaction rate of B/A is much larger compared to other thermal and plasma oxidations.

The Deal–Grove model could not overcome the negativeness of the reaction rate of B/A (Refs. 4, 6, and 11) in the plasma oxidation kinetics. Although a spontaneous reaction can occur with a negative reaction energy, the negativeness of the reaction rate is, however, too complicated to understand the oxidation phenomenon, because the reaction rate must be increased without increasing the concentration of the oxidant species at the interface, indicating evidence of the additional process. It is also valuable to note that Deal–Grove estimates fail in predicting even thermal oxidation rates in dry oxygen within the thin film regime (<20 nm). However, for the ultrathin oxide film, our theory clarifies the problem of the negative reaction rate and provides a comparable result for kinetic parameters for thermal oxidation, implying the plasma enhancement of the low-temperature oxidation.

V. ELECTRICAL PROPERTIES

The samples used in this study were p-type Si(100) with a boron doping concentration of 2.17×10^{16} cm^{-3}. Silicon wafers were treated with the standard cleaning process before plasma oxidation. After oxidation, aluminum was deposited as the gate metal. The area of contact was 7.85×10^{-3} cm^2. Postoxidation and premetallization annealing was performed in a rapid thermal process in vacuum at 400 °C for 2 min. The MOS capacitor structure with the thin oxide layer of 77 Å was characterized using $I-V$ and $C-V$ measurements at room temperature.

The breakdown field is defined as the electric field that will introduce 1 μA leakage current across the MOS capacitor. The present average breakdown field of plasma oxide films was found to be 14.8 MeV/cm. This is higher than 7.8 and 12.39 MeV/cm from Salbert, Reinhard, and Amussen[16] and Sung and Pang,[6] respectively. The measured 1 MHz capacitance–voltage curve with (real line) and without (dashed line) the low-temperature anneal at 400 °C for 2 min is shown in Fig. 7. The annealing effect lowers the oxide charge from 1.87×10^{-5} to 2.54×10^{-6} C/cm^2 and shifts the flat band voltage from -2.67 to -0.38 V. The work function difference for the annealed case is -0.638 eV. The Si/SiO$_2$ interface state density D_{it} is determined from the slope of the $C-V$ curve by deep level transient spectroscopy (DLTS) measurements. The trap density of D_{it} is found to have a minimum of 1.2×10^{10} eV^{-1} cm^{-2}, as shown in Fig. 8. This is much lower than 10^{11} eV^{-1} cm^{-2} from Sung and Pang[6] and 5×10^{11} eV^{-1} cm^{-2} from Salbert, Reinhard, and Amussen.[16] The low density of interface trapping states is comparable to the thermal oxide MOS capacitors. This indicates that the plasma damage on the oxide layer is not significant and the atomic diffusion of neutral O atoms into the Si/SiO$_2$ interface leads to the layer-by-layer oxidation. Addi-

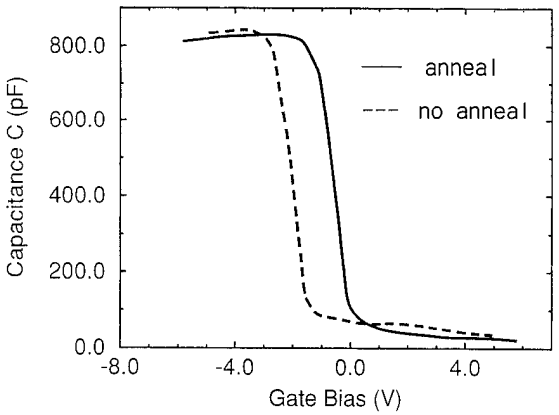


FIG. 7. MOS $C-V$ characteristics at 1 MHz for a 77-Å-thick oxide with (real line) and without (dashed line) a low-temperature anneal at 400 °C for 2 min. The gate oxide was grown with ECR oxygen plasma under the following conditions: substrate temperature=300 °C, O_2 flow rate=20 sccm, pressure=0.2 mTorr, microwave power=400 W, source to sample distance=7.5 cm, and oxidation time=30 min.

tionally, the large reduction in interface states is attributed to the removal of traps or defects after annealing. We strongly argue that in the ECR plasma oxidation, the boron dopant can be diffused out to the interface from the silicon bulk due to the plasma oxidation environment. The sheath potential attracts the effectively positively charged boron in silicon. The interfacial boron segregation results in the reduction of Si dangling bonds and reflects the asymmetric distribution of the density of the interfacial trapping state just below the conduction band edge as shown in Fig. 8. More study is needed about the dopant behavior at the interface. Note that the ultrathin gate oxide for submicron devices (0.25 μm channel length) is required to be 70 Å in order to reduce the fringe effect. The device quality of the electrical property of

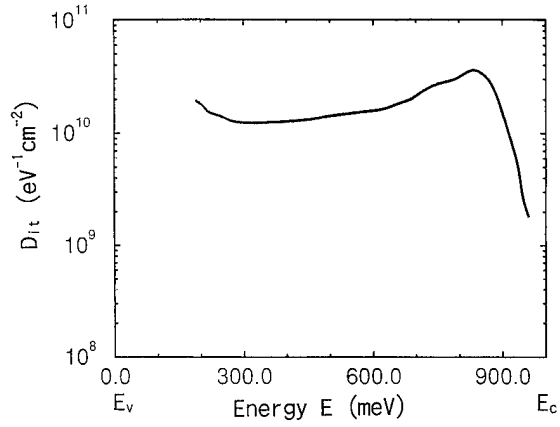


FIG. 8. Si/SiO_2 interface state density D_{it} as determined from the slope of the $C-V$ curve by the DLTS measurement. The high quality of the interface is comparable to the case of thermal oxide films of $D_{it}=10^{10}$ eV^{-1} cm^{-2}.

the ultrathin gate oxide grown by the ECR oxygen plasma seems to be applicable for ULSI semiconductor technology.

VI. SUMMARY

Plasma oxidation of silicon has been performed by using O_2 gas generated by an ECR source at low temperatures. The oxidation rate was studied in terms of processing pressure, microwave power, substrate temperature, and oxidation time. The oxide thickness was found to increase with microwave power parabolically. The pressure dependence of the oxidation rate is consistent with the maximum intensities of the optical emission spectra of the O atom and O_2^+ species. Below a 350 °C substrate temperature the rate slightly decreases, and it rapidly increases above that temperature. For ultrathin film growth, the plasma oxidation kinetics provides a diffusion rate and chemical reaction rate comparable to the thermal oxidation cases. The proposed model shows a more general formulation of the oxidation kinetics than the Deal–Grove model, especially for plasma oxidation, and gives remarkable agreement with the experimental fit. As a conclusion, in spite of the uncertainty of the contribution of charged atomic species, plasma oxidation is associated with the atomic diffusion process rather than the molecular species related process. The high quality of the MOS capacitor using plasma oxidation films suggests that these films may be suitable for gate dielectrics in ULSI devices.

ACKNOWLEDGMENT

This work is supported by the Korean Science and Engineering Foundation (KOSEF) through the Semiconductor Physics Research Center (SPRC) at Jeonbuk National University.

[1]S. M. Hu, J. Appl. Phys. **70**, R53 (1991).
[2]J. Batey and E. Tierney, J. Appl. Phys. **60**, 3136 (1986).
[3]A. K. Ray and A. Reisman, J. Electrochem. Soc. **128**, 2460 (1981).
[4]S. Kimura, E. Murakami, K. Miyake, T. Warabisaka, H. Sunami, and T. Tokuyama, J. Electrochem. Soc. **132**, 1460 (1985).
[5]D. A. Carl, D. W. Hess, and M. A. Lieberman, J. Vac. Sci. Technol. A **8**, 2924 (1990).
[6]K. T. Sung and S. W. Pang, J. Vac. Sci. Technol. B **10**, 2211 (1992).
[7]S. Dzioba, G. Este, and H. M. Naguib, J. Electrochem. Soc. **129**, 2537 (1982).
[8]C. Vinckier, P. Coeckelberghs, G. Stevens, M. Heyns, and S. De Jaegere, J. Appl. Phys. **62**, 1450 (1987).
[9]J. Joseph, Y. Z. Hu, and E. A. Irene, J. Vac. Sci. Technol. B **10**, 611 (1992).
[10]B. E. Deal and A. S. Grove, J. Appl. Phys. **36**, 3770 (1965).
[11]J. Y. Lee, S. W. Kang, J. W. Lee, C. H. Han, and C. K. Kim, J. Korean Electron. Eng. Soc. A **32**, 461 (1995).
[12]R. W. B. Pearse and A. G. Gaydon, *The Identification of Molecular Spectra*, 4th ed. (Chapman and Hall, London, 1976).
[13]W. A. Tiller, J. Electrochem. Soc. **127**, 625 (1984).
[14]P. G. Shewmon, *Diffusion in Solids* (McGraw-Hill, New York, 1963), p. 5.
[15]J. Philibert, S. J. Rothman, and D. Lazarus, *Atom Movements: Diffusion and Mass Transport in Solids* (Les Editions de Physique, Paris, 1991), p. 2.
[16]G. T. Salbert, D. K. Reinhard, and J. Asmussen, J. Vac. Sci. Technol. A **8**, 2919 (1990).

Gate quality Si_3N_4 prepared by low temperature remote plasma enhanced chemical vapor deposition for III–V semiconductor-based metal–insulator–semiconductor devices

D. G. Park,[a] M. Tao, D. Li, A. E. Botchkarev, Z. Fan, Z. Wang, S. N. Mohammad, A. Rockett, J. R. Abelson, and H. Morkoç[b]
Materials Research Laboratory and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

A. R. Heyd and S. A. Alterovitz
NASA Lewis Research Center, Cleveland, Ohio 44135

(Received 13 December 1995; accepted 19 April 1996)

We report the properties of silicon nitride films deposited by the electron cyclotron resonance remote plasma enhanced chemical vapor deposition method on Si substrates using SiH_4 and N_2. The effects of nitrogen/silane gas ratio ($R=N_2/SiH_4$), electron cyclotron resonance power, substrate temperature, and H on growth, refractive index, chemical composition, and etch rate were investigated. Nominally stoichiometric Si_3N_4 films were obtained with a refractive index of 1.9~2.0 at a wavelength of 632.8 nm. The etch rate of the films in a buffered HF solution (7:1) was low (~0.7 nm/min) and increased with increasing H_2 gas flow rate and decreasing substrate temperature during deposition. Fourier transform infrared spectroscopy and high temperature thermal evolution experiments showed very small amounts of H in the films. A leakage current less than 100 pA/cm^2 at a field of 2 MV/cm, a resistivity of $>4\times10^{17}$ Ω cm, and breakdown strengths of 6–11 MV/cm at a current density of 1 μA/cm^2 were observed. These properties are comparable to those of Si_3N_4 prepared by conventional high temperature (700 °C) chemical vapor deposition. The performance of GaAs-based field-effect-transistors in switching and power applications can be enhanced substantially by employing a metal-insulator-semiconductor structure. By taking advantage of an *in situ* process approach, insulator-GaAs structures were successfully gated with excellent interfacial properties. © *1996 American Vacuum Society.*

I. INTRODUCTION

The nearly ideal SiO_2/Si interface has made the integration of large numbers of silicon transistors on a single chip possible. Analogous devices in compound semiconductor structures (which have more favorable electrical characteristics than Si) have not come to fruition because of inferior dielectrics and dielectric/semiconductor interfaces. Since O induces a large density of trap states, the adsorption of less than a monolayer of O on GaAs causes Fermi level pinning.[1] Deposited insulator/compound-semiconductor interfaces have been reported to exhibit quality superior to those formed exclusively with native oxides.[2] Si_3N_4 appears to be a favorable gate dielectric for III–V compound semiconductor-based metal–insulator–semiconductor (MIS) devices.

Silicon nitride (Si_xN) layers have been extensively used as gate dielectrics in thin film transistors (TFTs),[3,4] as interlayer dielectrics in metal–nitride–oxide–Si (MNOS) devices,[5] and as final passivation layers in integrated circuits.[6] Attributes of high quality Si_xN films used for gate dielectrics are high dielectric constant, high breakdown field, low leakage current, and low interface trap density.[7] Compared to SiO_2, Si_3N_4 is known to be a good impurity barrier against sodium ions and water.[8] Nominally stoichiometric Si_3N_4 has a dielectric constant about 75% larger than that of SiO_2. Such a large dielectric constant results in a larger transconductance, a larger gate insulator capacitance, and hence a smaller threshold voltage shift for a given defect charge concentration. On the contrary, nonstoichiometric films have both lower dielectric breakdown field and higher leakage current as compared to those of Si_3N_4 films.[9]

Si_3N_4 films are typically prepared either by plasma enhanced chemical vapor deposition (PECVD) or low-pressure chemical vapor deposition (LPCVD) method. In the current work, remote plasma enhanced chemical vapor deposition (RPECVD) was used. In RPECVD the plasma is separated from the surface which reduces damage caused by plasma field and energetic particles. PECVD Si_xN films are nonstoichiometric and contain large concentration[10] of H which begins to desorb at ~350 °C.[11] This may lead to instabilities in MIS device performance. To minimize the H incorporation during growth, N_2 has been used in place of NH_3. However, PECVD has the advantage of being successful at a substrate temperature around 300 °C which is much lower than about 700 °C used in thermal chemical vapor deposition (CVD). Such a lower temperature prevents evaporation of the volatile group V species from the III–V compound semiconductor surface.

Recent strides for the realization of amenable GaAs metal–insulator–semiconductor field-effect-transistors (MISFETs) utilize epitaxial Si or Ge interlayers grown on

[a]Electronic mail: dpark@uiuc.edu
[b]On sabbatical at Wright Laboratory, Wright Patterson Air Force, under a University Resident Program; Eletronic mail: Morkoc@uiuc.edu

GaAs in order to avoid the formation of an intervening native oxide prior to the insulator deposition and the outdiffusion of the volatile group V element.[12–14] Pseudomorphic silicon cap layer deposited onto GaAs was reported to unpin the surface Fermi energy[13,15] and reduce the density of tunneling-related traps.[14] Given that Si on GaAs can be grown epitaxially and that high quality dielectrics can be deposited on Si, a practical approach to achieving high quality GaAs MIS structures would be to introduce a pseudomorphic Si layer between GaAs and a suitable dielectric.[14] However, because of the thinness of Si and the weak nature of the III–V compound semiconductor surface involved, the dielectric deposition must be performed with no exposure to atmospheric conditions at any time during the entire process. To insure this, we constructed a deposition apparatus which allows this very approach and utilized it to attain GaAs-based MIS structures with properties enhanced to a point at which meaningful analyses of the interface and field effect transistor (FET) can be undertaken.

In this article we discuss the effect of RPECVD deposition parameters on the physical and electrical characteristics of the Si_3N_4 thin films. We implemented these Si_3N_4 films, prepared at low temperature (<300 °C), as gate dielectrics for the GaAs-based MIS devices. It is noted that the properties of these Si_3N_4 films are comparable to those of thermal CVD Si_3N_4 films, and hence these films are promising candidates as gate dielectrics for III–V MIS devices.

II. EXPERIMENT

The ultrahigh vacuum (UHV) RPECVD system used for this study and described in detail previously,[16] consists of an electron cyclotron resonance (ECR) source coupled to an ion and turbomolecular pumped cold-wall stainless-steel chamber with a base pressure of 1×10^{-10} Torr. He and N_2 are introduced at the top of the ECR chamber where the plasma is generated. SiH_4 is introduced into the CVD chamber through a "showerhead" located halfway between the aperture of the ECR chamber and the substrate.

The Si substrates were P-doped n-type (100), had a diameter of 76 mm, and were misoriented 4° toward (110). The substrates were etched in 1:10 HF:H_2O solution followed by 3:1:1 HCl:H_2O_2:H_2O solution to remove the native oxide and to oxidize the Si surface, respectively. This procedure was repeated three times to insure the removal of the contaminants on the Si surface as well as to form a high quality native oxide. The substrates were then etched in the same HF solution for 1 min, and loaded into the vacuum system. Without further treatment they showed a (1×1) reflection high-electron energy diffraction (RHEED) pattern, indicating dihydride surfaces.[17] When the Si substrate temperature was increased to about 550 °C, the surface structure changed from (1×1) to the (2×2) (two orthogonal 2×1's) structure due to hydrogen desorption. It has been reported that hydrogen desorbs from the Si surface in a temperature range of 450–510 °C.[17] ECR PECVD Si_xN was deposited at substrate temperatures T_s ranging between 100 and 650 °C under the application of an ECR power (P_μ) of 100–250 W at 2.45 GHz. The film composition varied as a function of the SiH_4, N_2, H_2, and He flow rates. During the growth the N_2/SiH_4 ratio ranged from 1–50 as determined from relative partial pressures at fixed N_2 flow rate of 30 sccm. H_2 and He were added to the gas phase at a rate up to about 10 sccm.

An ellipsometry apparatus was used to determine the film thickness (t) and refractive index (n). The L116 Gaertner ellipsometer uses a rotating analyzer photometric design and operates at a wavelength of 632.8 nm. Values of n and t of the films obtained by this ellipsometer were in excellent agreement ($\Delta n\sim\pm0.02$, $\Delta t\sim\pm0.2$ nm) with the corresponding data obtained from a spectroscopic ellipsometer. A Perkin–Elmer 5300 x-ray photoelectron spectroscopy (XPS) facility permitted *in situ* analysis of the Si/N ratio in the films and testing for incorrect Si bonding. Mg $K\alpha$ (1253.6 eV) radiation was used in all studies reported here. Near surface depth profile data was obtained by varying the photoelectron take-off angle from 20°–70°. Mixed Gaussian–Lorentzian curves were fitted to the spectra using the χ^2 method with the fraction Gaussian character as a fitting parameter. A quantitative analysis of the H content and bonding species of the films was performed by fourier transform infrared (FTIR) spectroscopy using Si substrates polished on both sides. Thermal H evolution was also used.[18] Thermal stability of the film was estimated by measuring the film thickness change with high temperature annealing at 850 °C in N_2 for 30 min. Etch rates of the layers in 300 K NH_4F buffered-HF (BHF) solutions (NH_4F:HF=7:1) were measured as a function of the substrate growth temperature and the various flow rates of H_2. For comparison, the etch rates of Si_3N_4 grown by thermal CVD at 800 °C and by radio frequency (RF) PECVD at 400 °C from SiH_4-NH_3 were also evaluated.

The deposition of GaAs capacitors was initiated by the growth of a 0.5 μm, 10^{16} cm^{-3} Si-doped GaAs buffer layer by molecular beam epitaxy (MBE) on a (100) n^+ GaAs:Si substrate. The sample was then transferred via an UHV transfer tube (low 10^{-9} Torr) to an ultrahigh vacuum chemical vapor deposition (UHVCVD) chamber where Si and Si_3N_4 were deposited sequentially on the pristine, *as grown* GaAs surface. After deposition of 10 Å of Si on the *as grown* GaAs surface at 300 °C the substrate was rapidly heated to 600–700 °C with a ramp rate of 50 °C/min and kept at that temperature for a few minutes to crystallize the *as deposited* Si. The substrate was then cooled to 300 °C and the deposition of a nominally 150-Å-thick Si_3N_4 layer ensued. Depletion-mode transistors were grown in a similar fashion with the exception that the GaAs was grown on a semi-insulated substrate and consisted of 0.5 μm of undoped GaAs and active regions. GaAs channel region was capped by a 1000-Å-thick, 2×10^{17} cm^{-3} Si-doped GaAs.

Electrical properties were obtained from the current–voltage ($I-V$) characteristics of Al/Si_xN/n^+-Si MIS capacitors and from capacitance–voltage ($C-V$) curves of Al/Si_xN/epi-Si/n^+-Si and Al/Si_3N_4/Si/n-GaAs MIS capacitors. Nominally undoped Si epitaxial-layers (carrier concentrations $<5\times10^{16}$ cm^{-3}) were deposited using 60 W of microwave power. This was followed by deposition of the Si_xN

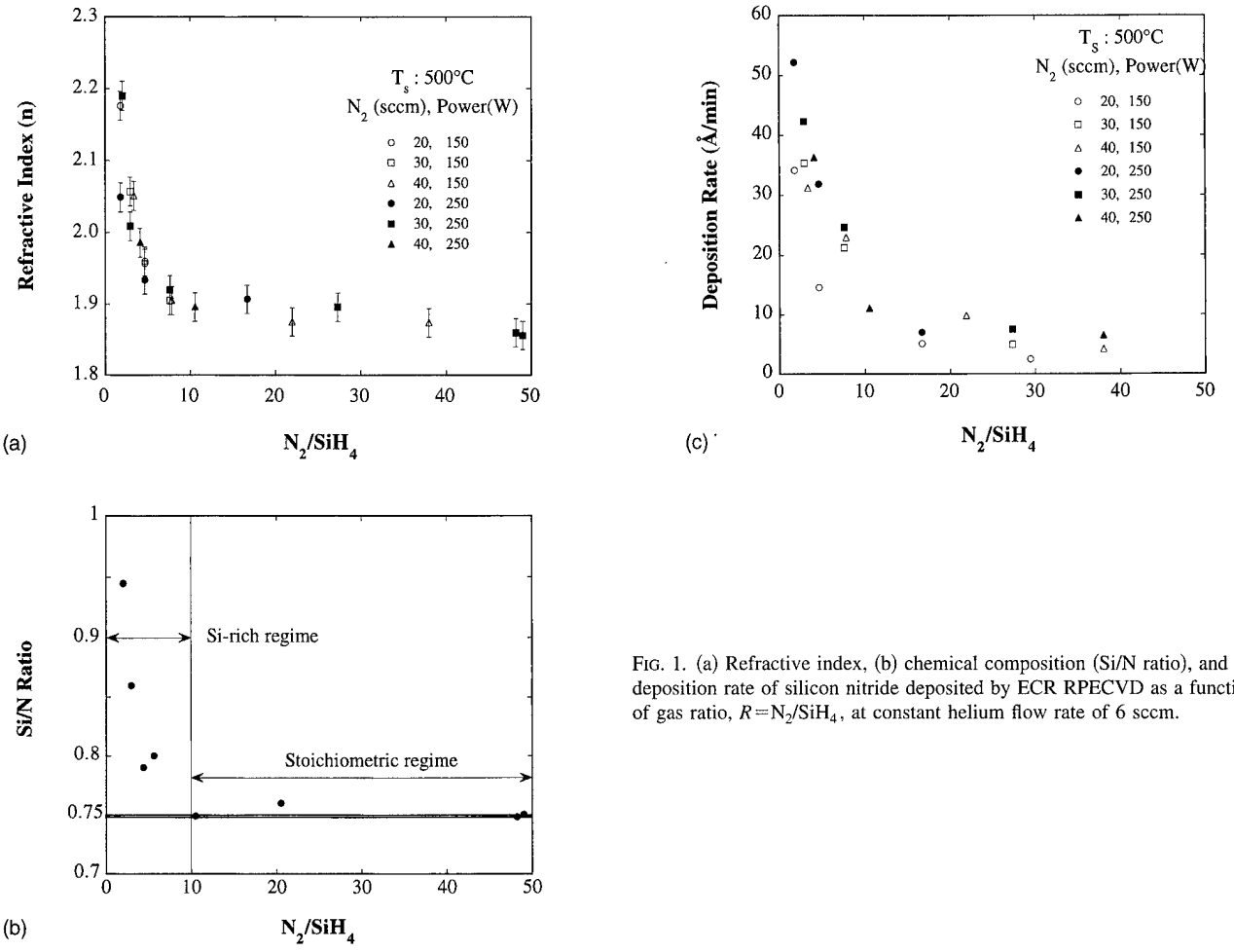


FIG. 1. (a) Refractive index, (b) chemical composition (Si/N ratio), and (c) deposition rate of silicon nitride deposited by ECR RPECVD as a function of gas ratio, $R=N_2/SiH_4$, at constant helium flow rate of 6 sccm.

without air exposure. After removal from the growth chamber, the samples were heat treated by rapid thermal annealing (RTA) at 450–500 °C for 30 s. An Al layer, about 1000 Å thick, was then evaporated on the top and patterned photolithographically into 300-μm dots to serve as the gate electrodes of MIS capacitors. The depletion-mode transistors were fabricated using a self-aligned technique in which Al was first evaporated over the entire wafer. The step was followed by a mesa etch to define the individual devices. Next, the Al was etched away from the source/drain regions, followed by etching of the Si_3N_4 using a reactive ion etching step which employs CF_4 as the etchant gas. Finally, AuGe/Ni/Au was evaporated to form the ohmic source/drain contacts.

In order to probe the $C-V$ properties of the MIS junctions, small-signal admittance measurements at frequencies ranging from 20 Hz to 1 MHz were carried out using an HP4284A LCR meter. $I-V$ measurements were made by using a HP4140B picoampmeter/voltage source. The leakage current was determined at a field of 2 MV/cm and the resistivity was measured from the slope of the $I-V$ curves. We estimated the breakdown strength as the breakdown field at a current density of 1 $\mu A/cm^2$. The dielectric constant of the Si_xN layer was measured by $C-V$ technique at various frequencies. The interface characterization of the GaAs MIS structure was undertaken by $C-V$ and conductance measurement. Direct current transfer characteristics measurements for the MISFETs were performed using an HP4145.

III. RESULTS AND DISCUSSION

A. Growth and related physical properties

Figures 1(a)–1(c) show the refractive index, the Si/N ratio, and the deposition rate as functions of gas flow ratio. For all measurements of these data the He flow rate was kept constant at 6 sccm. From Fig. 1(a) it may be noted that, for $R<2$, the layers were silicon-rich with $n\sim2.16$. However, for 300 °C$<T_s<$650 °C, ECR power of 150–250 W, and $R>10$, stoichiometric Si_3N_4 was obtained with $n=1.93$–1.99. Si has a much higher refractive index than Si_xN. Because of this, a higher refractive index corresponds to a Si-rich sample, which is possible when, as compared to SiH_4 flux, the N_2 flux into the deposition chamber is relatively small. Unlike Si-rich sample, the stoichiometric film is made of comparable amounts of Si and N. Higher composition of N thus causes the refractive index of the sample to be lower. Our experiments indicate that, for $R>10$, the refractive index of the sample becomes independent of both the flow ratio and the ECR power. This occurs because, although higher flux of N_2 leads to a higher generation of reactive nitrogen

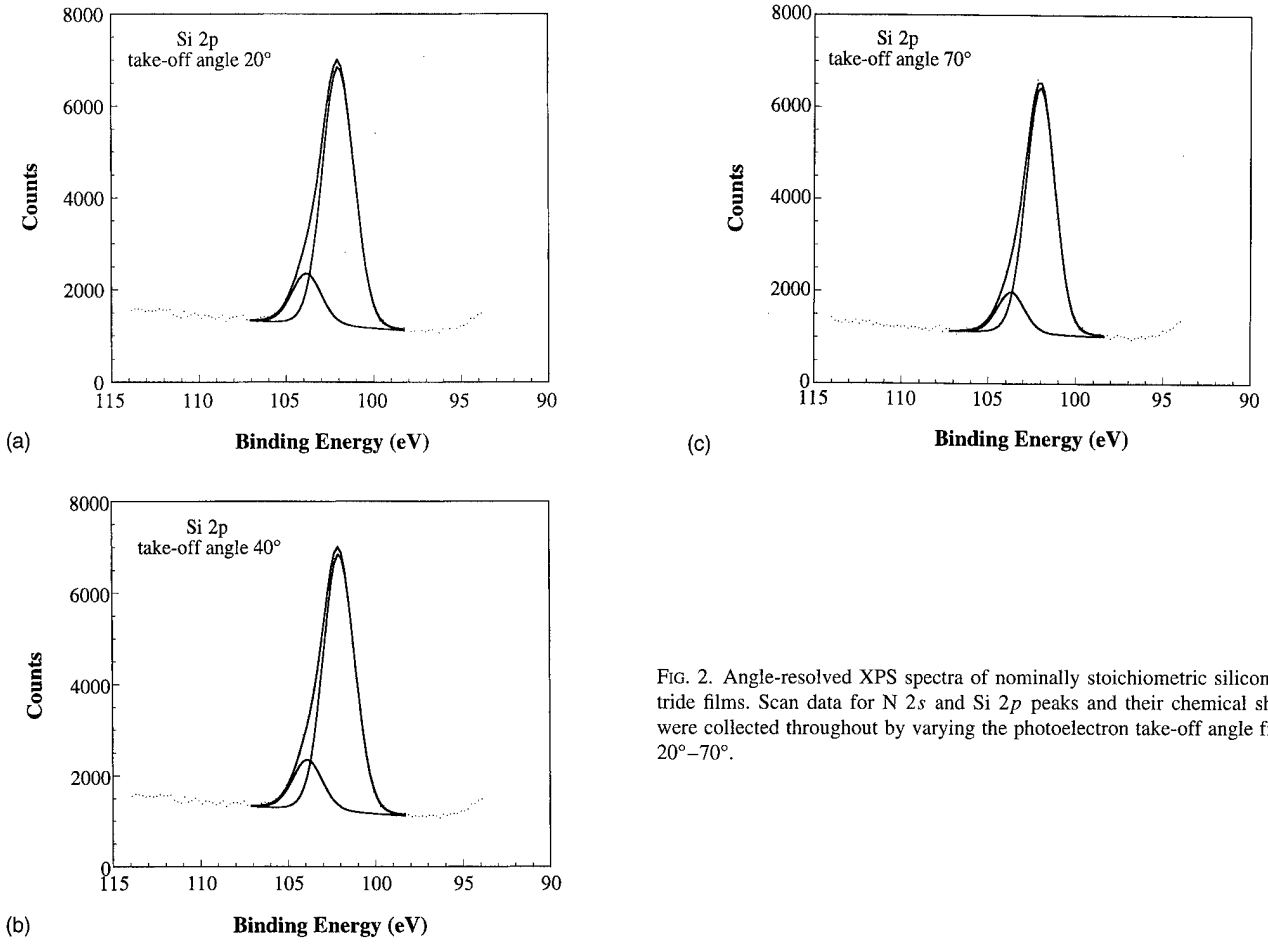


FIG. 2. Angle-resolved XPS spectra of nominally stoichiometric silicon nitride films. Scan data for N 2*s* and Si 2*p* peaks and their chemical shifts were collected throughout by varying the photoelectron take-off angle from 20°–70°.

species, most of them do not find reactive Si species to react and form silicon nitride. So, eventually these nitrogen species form back N_2 molecules and leave the system.

We note that the n values in Fig. 1(a), measured at a constant photon energy ~1.96 eV and a constant extinction coefficient $k=-0.020$, were normally 0.02–0.04 lower than those measured in representative films by spectroscopic ellipsometry. The extinction coefficients measured for absorption were an order of magnitude lower than −0.020, suggesting that the actual n values for our nominally stoichiometric samples are 1.95–2.02. That the silicon nitride formed in the regime $10 \leqslant R \leqslant 50$ is stoichiometric silicon nitride is apparent from Figs. 1(b) and 1(c). From these two figures one can see that Si/N ratio, and hence the film deposition rate in this regime, are about 0.75 and 10 Å/min, respectively. This indicates that the composition of the film is Si_3N_4 which is formed due to steady state incorporation of Si and N into the deposited layer. From these two figures one can also see that in the regime $0 \leqslant R \leqslant 10$, Si/N ratio and the film deposition rate decrease rapidly from their maximum values, 0.95 and 52 Å/min, respectively, with increasing R. This is a Si-rich regime with the deposition rate determined primarily by SiH_4 flow rate in agreement with previous results,[19,20] and with Si content of the silicon nitride films decreasing with increasing N_2 flux. The thickness uniformity of Si_3N_4 is within 4% (2σ).

Figures 2(a)–2(c) show Si 2*p* core level spectra as a function of electron take-off angles. These spectra have been shifted as appropriate to account for sample charging. From these figures one may visualize that the peak counts at 101.8 ±0.2 eV correspond to Si–N bonds, while the peak counts at 103.6±0.2 eV correspond to Si–O bonds. Any Si peak from Si–Si bonds (~99.15 eV) was not discernible, suggesting that excess Si was indeed insignificant. Similar results on the absence of the excess Si from Si LVV Auger electron spectra of Si_3N_4 also classified the layer to be "quasi-nominally stoichiometric Si_xN." [3] The constancy of the Si 2*p* component peak intensities for various take-off angles indicates no noticeable depth dependence. The film investigated here was found to be rather uniform since the relative magnitude of peaks of Si for Si–O and Si–N bonds does not change with take-off angle. The possible explanation for the presence of SiO_2 is that a uniform but a minute amount of SiO_2 had been present through the depth of the film due to the chamber contamination, which led to be the frequent deposition of SiO_2 within the same system. We have observed almost negligible oxygen contamination on Si interlayer grown on GaAs buffer layer prior to the silicon nitride deposition.

Among the gas ratios employed, a setting of $R=27$ led to stoichiometric films with optimal film properties. The dependence of the deposition rate at the substrate temperature of 300 °C on ECR power at this value of R and with 6 sccm He flow rate is shown in Fig. 3. It was observed that the depo-

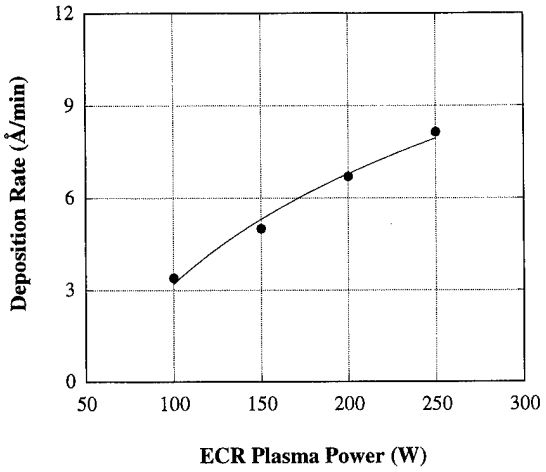


FIG. 3. Deposition rate of silicon nitride at the substrate temperature of 300 °C deposited by ECR RPECVD as a function of ECR power at $R=27$ and He flow rate of 6 sccm.

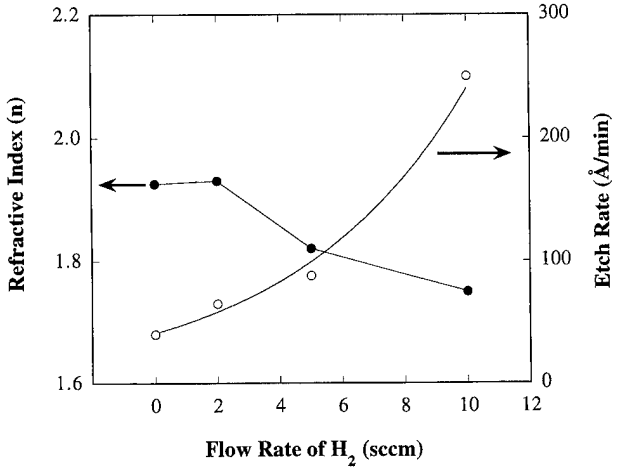


FIG. 5. Refractive index and etch rate as a function of flow rate of H_2 gas at $R=27$ and He flow rate of 6 sccm.

sition rate increased with increasing ECR power, and that the increase was about ~60% at P_μ between 150 and 250 W. This suggests that the plasma power influences the amount of N_2 and SiH_4 dissociation and enhances the formation of reactive Si and N species. The increase of deposition rate, however, gradually slowed down as the ECR plasma power approached 300 W. An exactly similar trend of variation of deposition rate with the ECR power was observed earlier by Matsuo and Kiuchi,[21] who used 10 sccm of SiH_4 and 10 sccm of N_2 to deposit Si_3N_4 films as confirmed by Si–N bond peak at the wave number of 845 cm^{-1}. As indicated earlier [see Figs. 1(a)–1(c)], film composition and n of Si_3N_4 were, however, unaffected by power under these conditions.

Figure 4 shows the growth and etch rates of stoichiometric silicon nitride as functions of ECR power and the substrate temperature at $R=27$. It was noted that, for

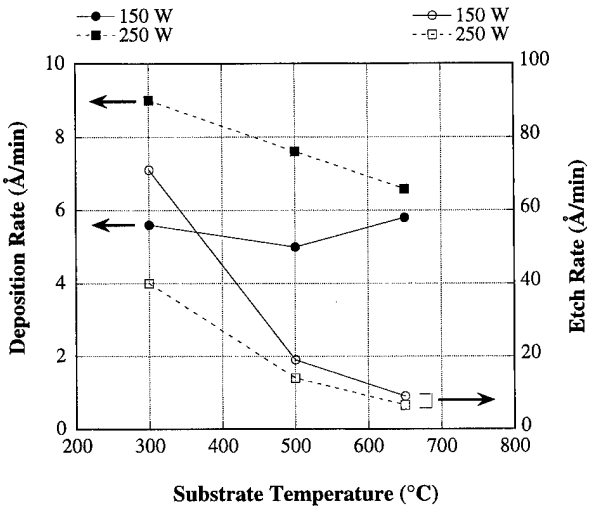


FIG. 4. Deposition rate and etch rate as functions of substrate temperature and plasma power. Etch rate was measured in NH_4F buffered-HF (BHF) solution (NH_4F:HF=7:1) at room temperature.

300 °C$<T_s<$500 °C the refractive index n is ~1.93–1.95, and at $T_s=650$ °C, n~1.99; both were observed to be consistent with nominally stoichiometric Si_3N_4. The deposition rate was independent of T_s at $P_\mu=150$ W, but increased with decreasing substrate temperature at $P_\mu=250$ W. This behavior is uncommon, although Piccirillo *et al.*[20] observed similar trends using a rf plasma source and SiH_4+NH_3 gas. At high power levels, there occurs an efficient dissociation of SiH_4 and a generation of more excited nitrogen atoms and/or radicals, which lead to an increase in the deposition rate of the films. Increasing substrate temperature decreases the amount of excited nitrogen by allowing surface reaction of nitrogen with hydrogen to form volatile NH_x species.[19] This results in a decrease in the growth rate of the films. The deposition rate was found to increase by about 30% as the flow rate of He increased from 6–10 sccm. Excited He is known to supply energy to surface species thus enhancing surface diffusion, generating reactive radicals, and increasing the deposition rate.

The etch rate in BHF is generally employed to gauge hydrogen and hydroxyl ion concentration in nitride and oxide films. Slower etch rates indicate smaller H and OH contents. The etch rate in BHF solution decreased as the substrate temperature and the plasma power increased. It may be noted from Fig. 4 that, at $T_s=650$ °C and $P_\mu=250$ W, the etch rate was ~0.7 nm/min. This is comparable to that of thermal CVD Si_3N_4 deposited at 800 °C. At lower substrate temperatures, for example, at $T_s=300$ °C and $P_\mu=250$ W the etch rate was five times higher than that for layers prepared at $T_s=650$ °C and $P_\mu=250$ W, but five times lower than that for rf plasma assisted Si_xN deposited at 400 °C.

The effects of H_2 flow rate on the refractive index and etch rate of hydrogenated Si_3N_4 are shown in Fig. 5. Deliberate introduction of H_2 at a rate up to 10 sccm, while keeping $R=27$ and He flow rate constant at 6 sccm, did not change the deposition rate significantly at $T_s=300$ °C and $P_\mu=250$ W. However, n decreased from 1.97 to 1.75 for H_2 flow rate increasing from 2 to 10 sccm. Decreasing n with

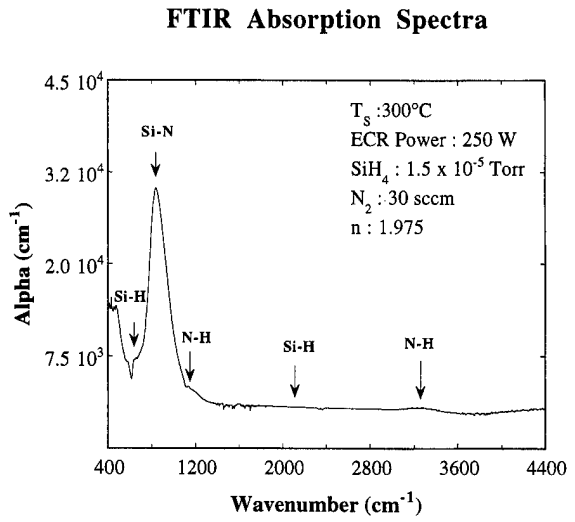


FIG. 6. FTIR absorption spectrum of an optimized nominally stoichiometric Si_3N_4 layer ($n\sim1.975$) prepared at T_s=300 °C, P=250 W, and $R\sim27$.

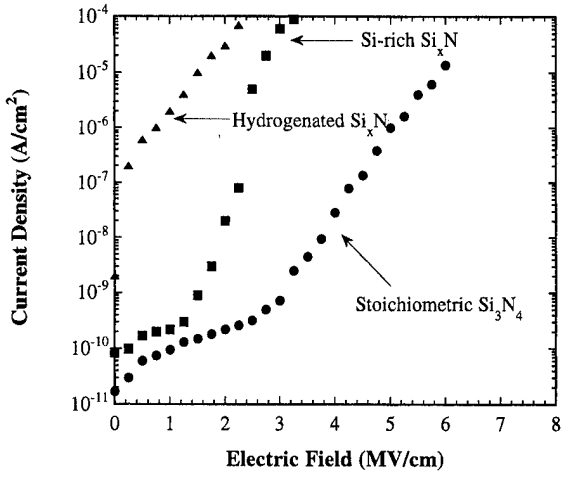


FIG. 7. Current density (J) vs electric field (E) characteristic for the nominally stoichiometric, Si-rich, and hydrogenated Si_xN prepared by ECR RPECVD.

increasing H incorporation has also been reported in amorphous silicon.[22] Hydrogen replaces Si–Si and Si–N bonds presumably with Si–H and N–H bonds which has a strong effect on n.[23] Bonds to H can very well be nonbridging (e.g., non-network forming). Because of this, a greater number of N–H bonds than Si–H bonds can make the material nitrogen rich. The bond lengths for both Si–H and N–H bonds are smaller than the Si–Si and Si–N bonds, which cause an increase in the number density of atoms in the medium. The angle of refraction increases with increase in density of the medium. If the angle of incidence is unchanged, the refractive index n consequently decreases with increase in the H_2 flow rate into the chamber.

Unlike the refractive index, the etch rate increased almost exponentially to 25 nm/min as a function of H_2 flow rate of 10 sccm. This value is similar to etch rates for rf plasma enhanced Si_xN. The etch rate of the Si_xN is reported[24,25] to be strongly affected by the H content in the film. The low etch rate of the ECR-deposited Si_3N_4 films confirms the low H content of the films. The constancy of the n value and insignificant film shrinkage (<2%) were observed after furnace annealing at 850 °C for 30 min for the Si_xN layer deposited at 300 °C and under 250 W. This shows that the silicon nitride films were thermally stable as well.

Figure 6 depicts the FTIR absorption spectrum of an optimized nominally stoichiometric Si_3N_4 layer ($n\sim1.975$) prepared at T_s=300 °C, and $R\sim27$ under the influence of P_μ=250 W. From Fig. 6 it is apparent that the dominant signal of the FTIR spectrum is the Si–N stretching band occurring at 870 cm^{-1}. No Si–H stretching mode at 2200 cm^{-1} was observable and the N–H stretching at 3360 cm^{-1} was negligible. Both stretching bands were considered to be the main peaks of the H-related bonds in Si_xN layers. Very small Si–H wagging and N–H bending bonds were detected at 550 and 1150 cm^{-1}, respectively.

The H concentration ([H]) determined by thermal evolution[15] during heating to 1200 °C was $\sim3\times10^{21}$ cm^3, or 3 at. %. This value is similar to the H content observed in thermal LPCVD Si_3N_4.[10] On the other hand, [H] deposited at 100 °C increased to 1.1×10^{22} H atom/cm^3, suggesting that SiH_4 cannot be easily decomposed into SiH, and that the H desorption from the reacting surface may not be possible at this temperature. The hydrogenated Si_xN layer (grown with a 10-sccm H_2 flow rate) was found to contain about 1.5×10^{22} H atom/cm^3, consistent with the higher etch rate. It is interesting to note that the H concentration reported in the present investigation, which uses N_2 as a nitrogen source gas, is much lower than that reported for films deposited with NH_3 as a nitrogen source.[18]

B. *I-V* and *C-V* characteristics

Figure 7 shows the current density (J) versus electric field (E) plots of the nominally stoichiometric Si_3N_4, Si-rich, and hydrogenated films. From this figure it may be seen that nominally stoichiometric Si_3N_4 exhibited a leakage current of ~100 pA cm^{-2} at an electric field of 2 MV/cm. The J vs E curves for the Si-rich Si_xN and hydrogenated Si_xN films showed orders of magnitude higher leakage currents and premature failures, respectively. This occurs probably because Si-rich nitrides contain relatively high concentrations of Si dangling bonds and distorted excess Si–Si bonds resulting in midgap defect states or bulk traps. These traps create trap states or potential wells in the conduction and valence bands inducing electrical conduction[3,26] by Frenkel–Poole emission. In the case of stoichiometric sample, there are essentially three regimes in the J versus E curve: The regime at lower electric field corresponds to electrical conduction primarily by hopping. The regime at moderately high electric field corresponds, on the other hand, to Frenkel–Poole emission. Fowler–Nordheim tunneling takes place almost always at higher electric fields due to reduction of potential barrier to the mobile carriers. Threshold voltage drift is also thought to be due to the loss of H when the device is operated at elevated temperatures.[10] A resistivity $\rho_s>4\times10^{17}$

Ω cm at a field lower than 1 MV/cm was measured for several optimal samples. Typical values of ρ_s for nominally stoichiometric Si_3N_4 were found to be 10^{15}–10^{17} Ω cm.[7]

As determined from our J versus E plots, the breakdown field in the case of Si-rich sample is quite low, about 2–4 MV/cm. The breakdown field for nominally stoichiometric Si_3N_4 samples is much higher, about 6–11 MV/cm. This is comparable to that of Si_3N_4 prepared by thermal CVD.[26] I-V measurements revealed that, for the stoichiometric Si_3N_4 films at high electric field (5.6 MV/cm), the current is temperature independent. This indicates that the electrical transport in stoichiometric Si_3N_4 films at high electric field is indeed a thermally nonactivated process, which is Fowler–Nordheim tunneling. Based on this observation it may be safely concluded that the breakdown was caused by the Fowler–Nordheim Tunneling of electrons through the conduction band potential barrier, and not by other breakdown mechanisms. This tunneling is identified to be the main current conduction mechanism of the thermally grown silicon dioxide.[26] Fowler–Nordheim tunneling in Si_3N_4, studied in a number of other reports[27–29] attests to the fact that the films prepared by ECR RPECVD are indeed of high quality.

Hydrogenation is a common problem with deposited silicon nitride samples.[30–32] To determine precisely the nature of conduction in the Si-rich and hydrogenated samples we plotted $\log(J/E)$ vs $E^{1/2}$ for E up to its breakdown limit. In both cases we found that the resulting curves were straight lines and dependent on temperature, although the slopes for the curves corresponding to Si-rich sample was much higher than that corresponding to hydrogenated sample. For these two samples with film thickness about 150 Å the leakage current was very high at relatively low electric fields. This is probably because in both cases the conduction was dictated by Frenkel–Poole emission and that the difference between the two resulted plausibly from different dielectric constants and densities of traps. The $\log(J/E)$ versus $E^{1/2}$ plot (not shown here) for stoichiometric film corresponding to Fowler–Nordheim tunneling was also a straight line. However, it was observed only at very high field at $E>3$ MV/cm. For Si-rich and hydrogenated samples, there were breakdowns at $E>3$ MV/cm.

The static dielectric constant was determined by using the capacitance value of MIS capacitors in strong accumulation. The thickness employed for this was determined by the ellipsometer. Si_3N_4 with $n=1.93$–1.99 had relative dielectric constants of 6.8–7.1. For the Si-rich Si_xN ($n\sim2.2$) the dielectric constant was ~7.5 in agreement with values for rf PECVD materials.[30]

The typical C-V characteristics of the Si_3N_4/Si/GaAs capacitors are shown in Fig. 8(a). The thickness of Si interlayer in Si_3N_4/Si/GaAs capacitors with static *in situ* annealing of Si was 10 Å, at which the capacitor exhibits the best $C-V$ characteristics.[33] The interface trap densities (D_{it}) were determined using the conductance loss [Fig. 8(b)] due to the interface traps from the measured admittance of the device when biased in depletion. The series resistance and insulator capacitance effects were removed from the measured conductance peaks to avoid artificially low interface trap densities.[34] By assuming that the defects obey the single-time constant model,[34] the interface trap densities were found to be around 8×10^{10} eV^{-1} cm^{-2}. Previous $C-V$ measurements of Si_3N_4/Si MIS structure indicated[16] that D_{it} is about 2×10^{10} eV^{-1} cm^{-2}. The discrepancy between the two results suggests that the main contribution to D_{it} in Si_3N_4/Si/GaAs MIS structures comes from Si/GaAs interface states.

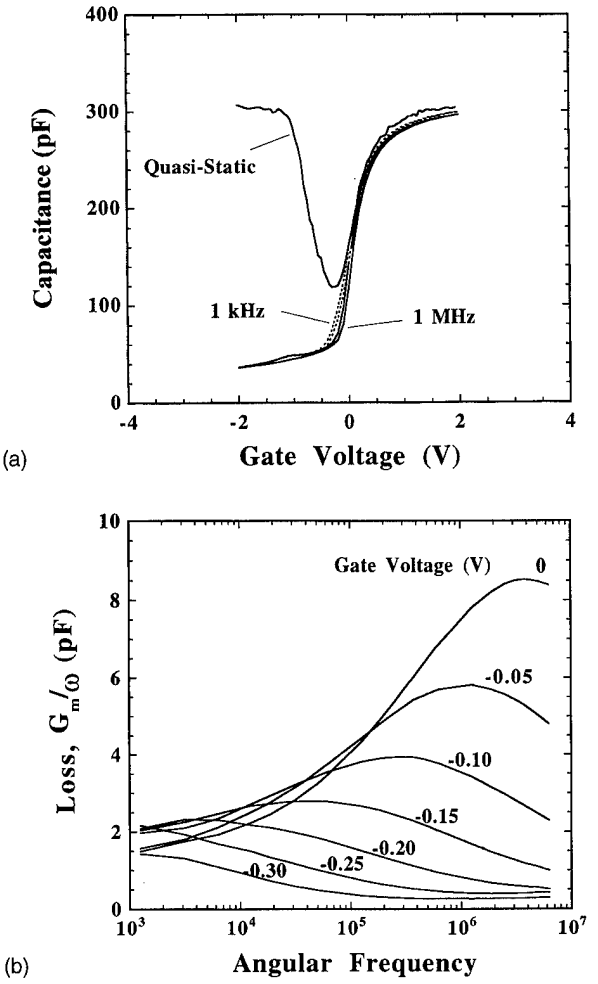


FIG. 8. (a) Measured 1 kHz, 1 MHz HF $C-V$ curves and QSCV curve of a Si_3N_4/Si/n-GaAs MIS capacitor. (b) Loss-angular frequency of the Si_3N_4/Si/n-GaAs capacitors. The applied gate voltages are shown.

Figures 8(a) and 8(b) depict the $C-V$ and loss-angular frequency curves for a specific Si_3N_4/Si/GaAs capacitor. As is apparent from Fig. 8(a), the flatband voltage shift was about −10 mV, which demonstrates a very small amount of fixed charge in the interfacial region. A typical hysteresis value of about 50 mV was obtained from retraced capacitance measurements at 1 kHz, which resulted from subjecting the MIS capacitor to the electric fields at 1.3 MV/cm (2.0 V). The hysteresis for a field swing of +4 to −4 MV/cm was about 150 mV. Frequency dispersion was found to be around 100 mV from the deviation of the 1 MHz and 1 kHz C-V curves. Both hysteresis and frequency dispersion degraded as we went down into the GaAs bandgap. As shown in Fig.

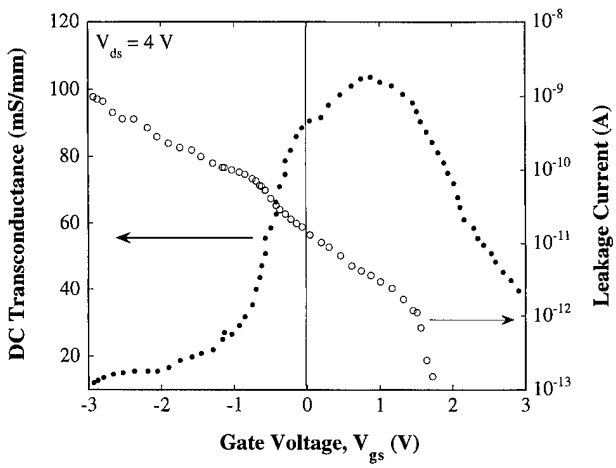


FIG. 9. Direct current transconductance and gate leakage current of a nominally 4.5-μm-gate length Si_3N_4/Si/*n*-GaAs depletion mode MISFET as a function of gate voltage V_{gs}.

8(b), the minimum conductance peak was obtained at a gate bias of -0.3 V, corresponding to about 0.25 eV above the midgap. Beyond this bias the conductance peak occurred at frequencies less than 1 kHz. An increase in trap density in the upper half of the bandgap was also observed in Fig. 8(b).

Superimposed on the high frequency curves in Fig. 8(a) is the quasistatic curve. The nature of this curve demonstrates that the excellent insulator quality of the nitride has also allowed the realization of quasistatic characteristics. It is noteworthy to mention that the quasi-static $C-V$ (QSCV) with big dip can only be obtained when the leakage current level of our capacitator is lower than 1 pA under the applied field of ~2 MV/cm, as determined by the $I-V$ characteristics. Notably, the trap density determined by low frequency (LF)/high frequency (HF) method using QSCV and 1-MHz signal frequency was slightly higher, $D_{it}\sim 3\times 10^{11}$ eV^{-1} cm^{-2}. Since the interface trap capacitance (C_{it}) must be extracted from the measured capacitance which consists of insulator capacitance, depletion layer capacitance, and interface trap capacitance, higher inaccuracies are likely to arise in extracting the D_{it} from the measured LF/HF capacitance rather than from the conductance measurement.[26] In other studies the $C-V$ characteristics of GaAs MIS capacitor with similar device structures[12,13] suffered from large hysteresis and frequency dispersion, and quasistatic $C-V$ curves for them were characteristically absent.

C. GaAs-based MISFET

Recent interests in GaAs-based MISFETs[35,36] stem from promising features of GaAs based electronic circuits for high-speed low-power applications. In an attempt to determine the practical applicability of our silicon nitride films, we measured various transfer characteristics of GaAs-based depletion-mode MISFETs. These typical dc transfer characteristics, for example, for MISFETs with channel length $L=4.5$ μm and channel width $W=120$ μm are shown in Fig. 9. These transfer characteristics involve the variations of the transconductance (g_m) and gate leakage current (I_g) as a function of gate-source bias V_{gs}. As may be evident from this figure, the variation of the dc transconductance with applied gate-source bias obtained from a MISFET with channel length $L=4.5$ μm shows a peak at about 1 V, and that the maximum transconductance varies between 90–110 mS/mm. Although the leakage current of this MISFET at a voltage of -3 V is of the order of 10^{-9} A, it dies down rapidly being reduced to 10^{-13} A at voltage of 3 V. It might be advantageous to use the dc transconductance of the depletion mode MISFETs for analyzing the as-measured results. This transconductance $g_m=(dI_d/dV_{gs})$ may be given by[37]

$$g_{m,dc}=KC_dC_i/(C_i+C_d+C_{inv}+C_{it}), \quad (1)$$

where $K=\mu WV_{ds}/L$, C_i is the gate insulator capacitance (6.5×10^{-7} F/cm^2), C_d is the depletion capacitance, C_{inv} is the inversion layer capacitance, C_{it} is the static capacitance of interface traps ($C_{it}=qN_{it}$, N_{it} is the interface trap density), μ is the carrier mobility, and W is the channel width.

Let us consider the region where the gate-source voltage is higher than the flatband voltage, viz., $V_{gs}>V_{fb}$ (~0.1 V). As may be apparent from Fig. 9, this region corresponds to the active channel in accumulation with the largest channel electron concentration and exhibits the transconductance at the maximum and the drain current saturated condition. The flatband condition occurs at about 0.7 V before the transconductance ($g_{m,dc}$) reaches its maximum value, resulting in $g_m=110$ mS/mm at $V_{gs}=0.8$ V and $V_{ds}=4.0$ V. However, sweeping V_{gs} toward depletion causes the channel to be partly depleted. The rate of change of channel conductance with V_{gs} thus becomes smaller, resulting in a decrease of the transconductance g_m. During transition from depletion to weak inversion, C_d decreases progressively, leading to a reduction of g_m. At the onset of strong inversion, the maximum depletion is reached and C_d saturates. On the other hand, C_{inv} increases exponentially leading to a rapid fall of $g_{m,dc}$.

It should be noted that the decrease of $g_{m,dc}$ in strong accumulation is due to the reduction of the surface mobility with V_{gs}. Indeed, the surface conduction becomes dominant in strong accumulation, and g_m is obtained by replacing C_d with C_{acc} and μ with μ_s (electron mobility at the surface). Due to high carrier scattering resulting from the presence of dangling bonds at the interface, the surface electron mobility μ_s is quite low, which causes the transconductance to be reduced.

Using this transconductance and other relevant parameters, namely, the gate dielectric capacitance C_0 per unit area, the drain voltage V_{ds}, the channel length L, and the channel width Z, the *field-effect mobility* μ_{FE}[36] may be calculated from

$$\mu_{FE}=Lg_{mi}/ZC_0V_{ds}. \quad (2)$$

The field-effect electron mobility determined in this way is found to be around 1000 cm^2/V s for the Si_3N_4/Si/GaAs structure at room temperature. For bulk GaAs doped with Si to 2×10^{17} cm^{-3}, it should be about 4000 cm^2/V s. Such low values of electron mobility may be ascribed to (1) the inter-

TABLE I. The peak values of the extrinsic transconductances and gate leakage of GaAs-based MISFETs as a function of channel length. The gate leakage was measured at $V_{dg}=5$ V.

Channel length (μm)	Extrinsic transconductance (mS/mm)	Gate leakage (nA)
3	100–140	<0.01
4.5	90–110	0.1
10	60–70	1
20	35–40	1

face trap density at the Si_3N_4/Si/GaAs interfaces which can act as recombination or trap centers, (2) surface roughness scattering of the GaAs active layer, and (3) Coulomb scattering from fixed charges and interface states.[38] That it is dominantly affected by the surface trap density is apparent also from an earlier investigation by Reed[39] for Al/Si_3N_4/Si/GaAs MISFETs in which transconductance is found to decrease with an increase in interface trap density, when the latter exceeds a level of about 10^{11} eV^{-1} cm^{-2}.

Table I summarizes the peak values of the extrinsic transconductances and gate leakage of GaAs-based MISFETs as a function of channel lengths. The transconductances shown in Table I are extrinsic, which include the effect of parasitic resistance values (e.g., the source contact resistance and the ungated channel resistance). If the parasitic resistance effect is not taken into account, the intrinsic trancondutance (g_{mi}) thus obtained for Si_3N_4/Si/GaAs devices exhibits a value of about 165 mS/mm for $L=4.5$ μm, 88 mS/mm for $L=9.5$ μm, and 45 mS/mS for $L=19.5$ μm. These results are superior to the reported GaAs metal–oxide–semiconductor field-effect-transistor (MOSFET) result which is 60 mS/mm for a 8-μm gate length device.[13] This superiority may be ascribed to the improved interface quality provided by the Si_3N_4/Si/GaAs structure, and larger dielectric constant of silicon nitride.

Important issues which deserve immediate attention are improved leakage current of GaAs MISFETs with respect to GaAs metal–semiconductor field-effect-transistors (MESFETs). As mentioned earlier, one of the motivations driving the development of GaAs MISFETs is to realize a MIS device with low gate leakage current. It was noted that, at the same gate bias (−6.0 V), while the gate leakage for the MESFET is several microamperes,[40] the gate leakage for the MISFET varies between several tens of picoamperes and several nanoamperes. This improvement of more than three orders of magnitude in the gate leakage of the MISFET over that of the MESFET may be considered quite remarkable, being beneficial to the development of high-speed, low-power GaAs integrated circuits (ICs) used for wireless communication systems.

IV. CONCLUSION

We have described the deposition characteristics and related properties of nominally stoichiometric Si_3N_4 prepared by the ECR RPECVD method. This description can provide valuable direction to the optimization of high quality silicon nitride films for practical applications. Our study convincingly demonstrates that the films deposited from SiH_4 and N_2 can be relatively free from H contents, which endow them with low etch rates and excellent dielectric strength. The properties of these stoichiometric silicon nitride films are comparable to those of thermal CVD Si_3N_4, and are found to be promising candidates as gate dielectrics for III-V MIS devices. Nominally stoichiometric Si_3N_4 films are found to exhibit electrical characteristics superior to those of Si-rich and hydrogenated Si_xN, suggesting that all MIS structures should, for example, prefer insulator gates from stoichiometric silicon nitrides. On the other hand, the memory devices should be made from Si-rich silicon nitrides. High/low frequency $C-V$, conductance versus frequency, and MISFETs were used for a comprehensive characterization of the n-type MIS structures. From the stringent conductance measurements, interface state densities in the high 10^{10} eV^{-1} cm^{-2} have been obtained. Self-aligned gate depletion mode MISFETs, which utilized stoichiometric silicon nitride as a gate dielectric and which had gate lengths of 3 μm exhibited transconductances of about 100–140 mS/mm for GaAs channels. Noting that the development of GaAs-based MISFET technology is at its embriyonic stage, it should be considered quite remarkable.

ACKNOWLEDGMENTS

This work is supported by the Department of Energy under Contract No. DEFGO2-96-ER45439, Air Force Office of Scientific Research under Contract No. F49620-95-1-0298, and National Science Foundation under Contract No. DMR 93-12422. The authors wish to thank two referees of the article for very constructive comments. The assistance of Doug Jeffers in the Epicenter is acknowledged. One of us (D.G.P.) gratefully appreciates support of the Korean Government through the Ministry of Education under File No. 91-0052. D.G.P. also acknowledges the assistance of Dr. H. Liang and C. S. McCormick during FTIR and TE measurement.

[1]W. E. Spicer, P. W. Chye, P. R. Streath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. **16**, 1422 (1979).
[2]P. D. Gardner, S. Y. Narayan, and Y. H. Yun, Thin Solid Films **117**, 173 (1984).
[3]S. S. He and G. Lucovsky, Mater. Res. Soc. Symp. Proc. **336**, 787 (1994).
[4]J. Kanicki, F. R. Libsch, J. Griffith, and R. Polastre, J. Appl. Phys. **69**, 2339 (1991).
[5]S. Fujita, M. Nishihara, W. L. Hoi, and A. Sasaki, Jpn. J. Appl. Phys. **20**, 917 (1981).
[6]R. C. G. Swan, R. R. Mehta, and T. P. Cage, J. Electrochem. Soc. **114**, 713 (1971).
[7]Y. Manabe and T. Mitsuyu, J. Appl. Phys. **66**, 2475 (1989).
[8]A. C. Adams, in *VLSI Technology*, edited by S. M. Sze (McGraw-Hill, New York, 1988).
[9]A. J. Lowe, M. J. Powell, and S. R. Elliot, J. Appl. Phys. **59**, 1251 (1986).
[10]J. Z. Xie, S. P. Muraka, X. S. Guo, and W. A. Lanford, J. Vac. Sci. Technol. B **7**, 150 (1989).
[11]J. V. Nguyen, J. Electron. Mater. **16**, 275 (1987).
[12]H. Hasegawa, L. He, H. Ohno, T. Sawada, and T. Haga, J. Vac. Sci. Technol. B **5**, 1097 (1987).
[13]S. Tiwari, S. L. Wright, and J. Batey, IEEE Electron Devices Lett. **9**, 488 (1988).
[14]D. S. L. Mui, Z. Wang, and H. Morkoç, Thin Solid Films **231**, 107 (1993).

[15]A. J. Sambell and J. Wood, IEEE Trans. Electron Devices **37**, 88 (1990).
[16]D. S. L. Mui, S. F. Fang, and H. Morkoç, Appl. Phys. Lett. **58**, 70 (1991).
[17]B. S. Meyerson, F. J. Himpsel, and K. J. Uram, Appl. Phys. Lett. **57**, 1034 (1990); see also, A. B. Muller, F. Reinhardt, U. Resch, W. Richter, and K. C. Rose, Thin Solid Films **233**, 19 (1993); J. J. Boland, Phys. Rev. Lett. **65**, 3325 (1990).
[18]N. Maley, A. Meyer, M. Pinabasi, D. Leet, J. R. Abelson, and J. A. Thornton, J. Vac. Sci. Technol. A **7**, 1267 (1989).
[19]J. R. Flemish, R. Pfeffer, W. Buchwald, and K. A. Jones, Mater. Res. Soc. Symp. Proc. **284**, 15 (1994).
[20]A. Piccirillo and A. L. Gobbi, J. Electrochem. Soc. **137**, 3910 (1990).
[21]S. Matsuo and M. Kiuchi, Jpn. J. Appl. Phys. **22**, L210 (1983).
[22]G. F. Feng, M. Katiyar, J. R. Abelson, and N. Maley, Phys. Rev. B **45**, 9103 (1992).
[23]C.-H. Lin, J. B. Wachtman, G. H. Sigel, R. L. Pfeffer, T. P. Monahan, and R. T. Lareau, Mater. Res. Soc. Symp. Proc. **284**, 21 (1993).
[24]R. Chow, W. A. Landford, W. Ke-Ming, and R. S. Rosler, J. Appl. Phys. **53**, 5630 (1982).
[25]W. A. Lanford and M. J. Rand, J. Appl. Phys. **49**, 2473 (1978).
[26]S. M. Sze, J. Appl. Phys. **38**, 2915 (1967); see also, *Physics of Semiconductor Devices* (Wiley, New York, 1981).
[27]M. Tao, D. G. Park, S. N. Mohammad, D. Li, A. E. Botchkarev, and H. Morkoç, Philos. Mag. B (in press).
[28]D. S. L. Mui, H. Liaw, A. L. Demirel, S. Strite, and H. Morkoç, Appl. Phys. Lett. **59**, 2847 (1991).
[29]D. Wang, T. P. Ma, J. W. Golz, B. L. Halpern, and J. Schmitt, IEEE Electron Devices Lett. **13**, 69 (1994).
[30]M. Maeda and Y. Arita, J. Appl. Phys. **53**, 6852 (1982).
[31]J. Robertson, Philos. Mag. **69**, 307 (1994).
[32]C. Kaya, T. P. Ma, T-C. Chen, and R. C. Barker, J. Appl. Phys. **64**, 3949 (1988).
[33]J. Reed, M. Tao, D. G. Park, K. Suzue, A. E. Botchkarev, Z. Fan, D. Li, S. J. Chey, J. V. Nostrand, D. G. Cahill, and H. Morkoç, Solid State Electron. **38**, 1351 (1995).
[34]E. H. Nicollian and J. R. Brews, *MOS Physics and Technology* (Wiley, New York, 1982).
[35]C. L. Chang, L. J. Mahoney, K. B. Nicholes, M. J. Manfra, B. M. Gramstroff, K. M. Molver, R. A. Murphy, and E. R. Brown, IEEE Electron Devices Lett. **16**, 199 (1995).
[36]K. Kaviani, A. Madhukar, J. J. Brown, and L. E. Larson, Electron. Lett. **30**, 669 (1994).
[37]H. Haddara, T. Elewa, and S. Cristoloveanu, IEEE Electron Devices Lett. **9**, 35 (1988).
[38]D. K. Schroder, *Semiconductor Material and Device Characterization* (Wiley, New York, 1990).
[39]J. C. Reed, Ph.D. thesis, University of Illinois at Urbana-Champaign (1994).
[40]D. G. Park, G. B. Gao, Z. Fan, and H. Morkoç (unpublished).

B/Si(100) surface: Atomic structure and epitaxial Si overgrowth

Z. Zhang, M. A. Kulakov,[a] B. Bullemer, and I. Eisele
Universität der Bundeswehr München, Fakultät für Elektrotechnik, Institut für Physik, D-85577 Neubiberg, Germany

A. V. Zotov
Universität der Bundeswehr München, Institut für Physik, D-85577 Neubiberg, Germany and Institute of Automation and Control Processes, Russian Academy of Sciences, 690041 Vladivostok, Russian Federation

(Received 23 February 1996; accepted 24 May 1996)

A boron-accumulated Si(100) surface with boron coverage up to 0.1 monolayer has been prepared by high-temperature annealing of B-doped Si samples. Scanning tunneling microscopy has been used to monitor the transformation in surface morphology and surface atomic structure induced by boron surface accumulation. The specific boron-induced surface features have been elucidated and a model of their atomic structure has been proposed. On the boron-accumulated surface epitaxial Si films with a thickness of 0.1–3.0 monolayers have been grown using solid phase epitaxy and molecular beam epitaxy. The results show that under appropriate growth conditions B surface segregation can be suppressed even on the atomic scale. The effect of boron induced features on the growth process has been discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

The formation of delta-function-like spikes of dopants in Si (*Si delta doping*) has recently attracted considerable interest both from a scientific viewpoint and because of its technological applications. Such delta-doped layers have been produced successfully in Si using molecular beam epitaxy (MBE) and MBE-related solid phase epitaxy (SPE).[1–7] A new challenge for the delta-doping technique is the fabrication of delta-doped layers with ordered two-dimensional structures of the dopants (*ordered delta doping*).[8,9] Dopant ordering offers two prospective advantages for device applications: (i) The lowering of the statistical limit for the device size in deep submicron VLSI and (ii) the possibility of reducing the ionized impurity scattering which one can anticipate from an ordered array of dopants.[9] The most appropriate approach to ordered delta doping requires the formation of an ordered submonolayer surface phase of adsorbate on an atomically clean Si surface[10] followed by low-temperature epitaxial Si overgrowth by means of MBE or SPE. Achievement of this goal requires sufficient stability of the ordered surface phase with respect to Si overgrowth. Unfortunately, most of the dopant surface phases lose their ordered structure already upon room-temperature (RT) deposition of Si submonolayers.[11] Of the possible dopants, only boron shows promise for preservation of the ordered reconstruction at the buried interface. The Si(111) $\sqrt{3} \times \sqrt{3} - B$ surface phase is known to persist after a RT deposition of an amorphous Si cap.[8] However, at the relatively high temperatures required for the epitaxial Si(111) overgrowth (400 °C for MBE and 600 °C for SPE), ordered delta doping cannot be achieved. The buried Si(111) $\sqrt{3} \times \sqrt{3} - B$ surface phase suffers from both boron disordering[12] and outsmearing of the boron profile towards the surface.[13]

Compared to the Si(111) surface, the temperature required for epitaxial Si growth on the Si(100) surface is much lower. As a result, preservation of the 2×1 ordered boron submonolayer beneath the crystalline Si cap grown at about 200 °C has been confirmed both by x-ray diffraction[8] and transmission electron microscopy.[9] Electrical characterization of these structures showed that the embedded ordered boron yields the highest electrically active concentration of boron in silicon ever reported.[7,8] However, the expected large mobility enhancements have not been observed. For optimization of the ordered boron delta-doped structures detailed information concerning the atomic structure of the B-adsorbed Si(100) surface and of the main regularities of Si epitaxial overgrowth are highly desirable. Despite the widespread use of boron as a *p*-type dopant, the investigation of the interaction of B with the Si(100) surface has been started just recently. Until now there have been only a few publications devoted to the atomic structure of the B/Si(100) surface and there is no atomic-scale information about Si growth on the B/Si(100) surface.

In the present work, the MBE and SPE overgrowth of a few Si monolayers on B/Si(100) are studied by scanning tunneling microscopy (STM). The B/Si(100) surface is prepared by high-temperature annealing of the B-doped Si(100) samples and the atomic structure of the surface formed is discussed.

II. EXPERIMENT

The samples used in this study were Si bars, 14×4.5×0.3 mm^3 in size, cut from 0.01 Ω cm B-doped Si(100) wafers (Wacker–Chemitronic). Before loading into the UHV-STM system the samples were cleaned ultrasonically with acetone and methyl alcohol. Then they were fixed by molybdenum fasteners and transferred into the ultrahigh vacuum (UHV) chamber. After degassing for about 5 h at 550 °C the samples

[a]Corresponding author; Electronic mail: e91bzeus@rz.unibw-muenchen.de

 0734-211X/96/14(4)/2684/6/$10.00 ©1996 American Vacuum Society

were heated by direct current to 1230 °C to induce boron accumulation at the sample surface. The pressure during annealing did not exceed the value of 5×10^{-10} Torr.

Si cap layers were deposited from a resistively heated Si stripe at a rate of about 0.3 monolayer (ML)/min [1 ML = 6.8×10^{14} cm^{-2}, the site density of the Si(100) plane]. During deposition the pressure was below 4×10^{-10} Torr. The deposition rate was determined in calibration experiments from the shape, size, height, and number density of Si islands displayed by STM after deposition for a fixed time. The temperatures of the sample above 350 °C were monitored by an infrared pyrometer with an accuracy of ±15 °C. The temperatures from 20 to 350 °C were estimated by extrapolating the "heating current-temperature" dependence to lower currents. The accuracy in the latter case was about ±50 °C.

The STM observations were performed using Omicron UHV STM apparatus driven by custom-made electronics.[14] All the STM images presented here were obtained at a constant current of 0.2 nA and at sample biases in the range from −3.0 to +3.0 V. Images are shown with a conventional gray scale keyed to the surface height.

III. RESULTS AND DISCUSSION

A. B/Si(100) surface

High-temperature annealing of B-doped Si samples is known to result in boron accumulation at the sample surface.[15–18] The saturating surface density of B atoms achieved upon Si sublimation at a given temperature is the result of a competition between the boron enrichment at the surface, due to (i) the much lower evaporation rate of B atoms from Si surface compared to the one of Si atoms, and (ii) the boron diffusion into the bulk driven by the B concentration gradient forming in the near-surface region.[16] Auger electron spectroscopy measurements showed that maximum boron coverage achieved for the present experimental conditions, namely, after annealing of 0.01 Ω cm Si(100):B samples at about 1230 °C for 2 h was about 0.1 ML.

STM observations have revealed that B accumulation causes transformations both in surface morphology and surface atomic structure. The typical large-scale transformations are illustrated by Fig. 1. After moderate annealing (1215 °C for 10 min) boron surface concentration is low (<0.03 ML) and the surface displays a STM image [Fig. 1(a)] much like the one known for an atomically clean Si(100) surface. One can see the regular distribution of the monolayer-high steps reflecting the sample miscut. The step-down direction is from the upper right to the lower left. The downward steps running parallel to the dimer rows (S_A steps, according to Chadi's notation[19]) are straight whereas those running perpendicular to the dimer rows (S_B steps) are rather rough. The B accumulation changes the surface morphology to the one shown in Fig. 1(b). This surface corresponds to the annealing for about 1 h at 1230 °C and B coverage of 0.06–0.08 ML. The apparent attribute of the surface is enhanced roughness of the S_B steps. Such morphology implies the formation of certain surface features that pin some dimer rows during sublimation.

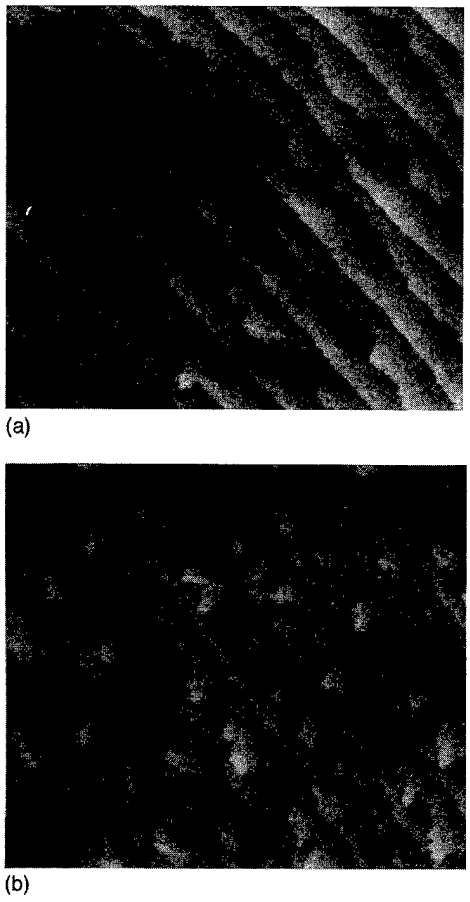


FIG. 1. Large-scale transformations induced by boron accumulation on the Si(100) surface. STM images of a 0.01 Ω cm Si(100):B sample surface after annealing in UHV. (a) at 1215 °C for 10 min, (b) at 1230 °C for 1 h. Scale: 1870×1560 Å^2.

The STM images with atomic resolution show specific features which are believed to be associated with boron dopant atoms.[20] As one can see in Fig. 2, these are nearly round protrusions which occur mostly in pairs and rarely as individuals or as chains containing three and more elements. In

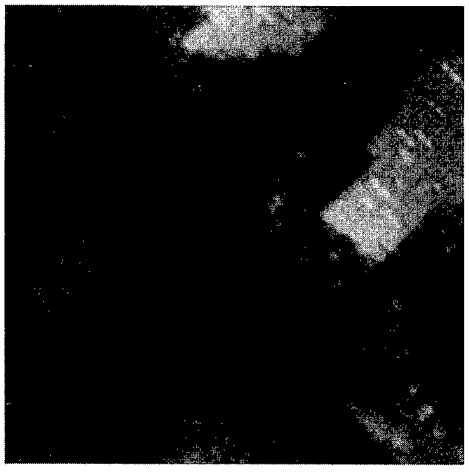


FIG. 2. STM image of the B/Si(100) surface. Scale: 230×230 Å^2.

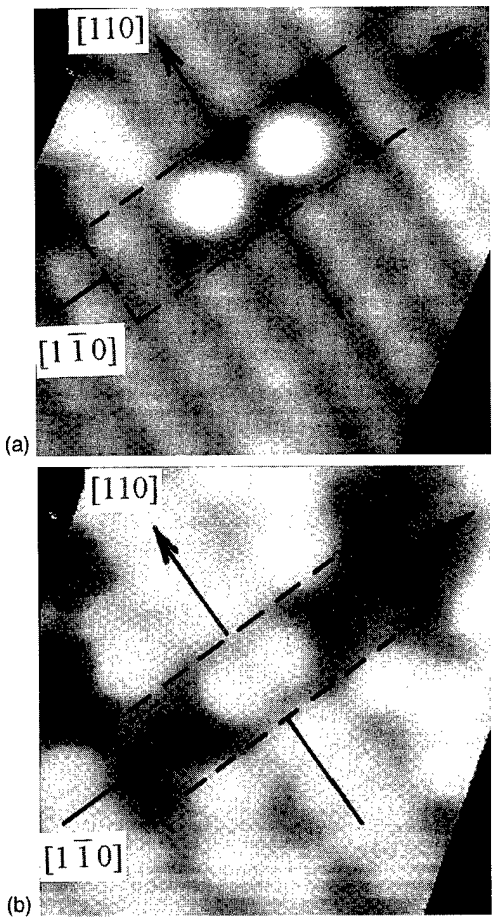


FIG. 3. STM images of the boron-induced double-protrusion feature (marked by a dashed frame). (a) Empty states (sample bias is +2.0 V), (b) filled states (sample bias is −3.0 V). The axes are drawn for convenience of determining the local symmetry and the position of the feature with respect to the Si(100)2×1 dimer rows. Scale: 28×28 Å^2.

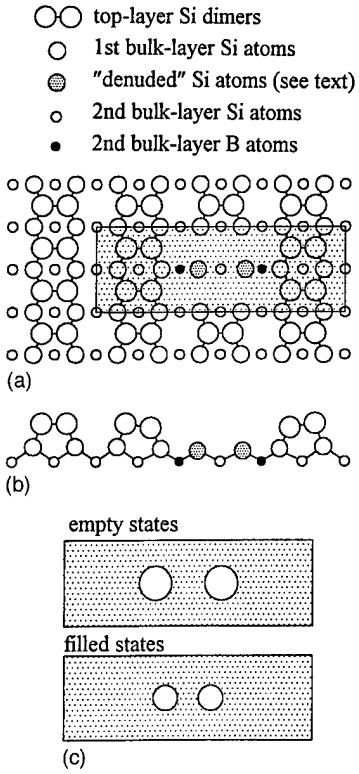


FIG. 4. Proposed model for the boron-induced double-protrusion feature. (a) top view, (b) side view, (c) schematic diagram showing the spatial location of the boron-induced protrusions within the area of the feature (hatched region) in empty- and filled-state STM images.

particular, one can see that many of them are located adjacent to the edges of S_B-steps terminating dimer rows. Thus, it can be assumed that the boron-induced features pin the dimer rows during sublimation leading to the specific form of the S_B-step edges seen in the large-scale STM images [Fig. 1(b)]. The reason for such pinning might be related in some way to changes of the local stress around a boron-induced feature due to the fact that Si–B bond is shorter than the Si–Si one.

To elucidate the atomic structure of the boron-induced features, high-resolution STM images have been obtained. As an example, Fig. 3 shows a small fragment of the surface with a double-protrusion feature recorded at positive [Fig. 3(a)] and negative [Fig. 3(b)] bias polarities. One can see that the feature replaces two dimers in a central dimer row. The double-protrusion feature shares the symmetry axis of the central dimer row and lies on the ⟨110⟩ line drawn between two neighboring dimers as shown in Figs. 3(a) and 3(b).

Figure 4 shows a structural model of the double-protrusion feature proposed on the basis of the atomic-resolution STM measurements, on the data available in the literature concerning the B/Si(100) surface, namely, that this surface is not a dangling bond-free surface[21] and on the presumption that boron atoms occupy substitutional sites in the Si lattice.[8] In the model two boron atoms (shown in Fig. 4 as small closed circles) occupy substitutional sites in the second bulklike layer. These boron atoms are separated by $2a$, where a is the unit length of the Si(100) plane. Due to the shorter Si–B bond as compared to the Si–Si bond, the B-bonded Si atoms of the first bulklike layer are shifted from their normal positions towards the B atoms. The stresses originating from these shifts are believed to be the reason of the bond disruption and the elimination of two Si surface dimers. Thus, two Si atoms of the first bulklike layer (shown in Fig. 4 as shaded circles) become "denuded." Each of these atoms has two dangling bonds and we assume that these electron densities are responsible for the double protrusions observed in the STM images. The images are obviously affected by the boron atoms in the substitutional positions, which act as negatively charged ions simultaneously repelling the electron density and causing a partial depletion of the electron states of the neighboring Si atoms. As a result, the protrusions appear to be more closely spaced in the filled-state image [Fig. 3(b)] compared to the empty-state image [Fig. 3(a)]. Basically similar structural models are applicable for the description of features containing three or more protrusions. The periodic occupation of $2a$ spaced substitutional sites in the second bulklike layer by B atoms, in the manner described above, produces a linear chain with the corresponding number of protrusions in the STM image. More detailed analysis of these structures is presented elsewhere.[22]

Comparison of the present results with recent STM data by Wang *et al.*[23,24] for the B/Si(100) surface prepared by decomposition of boron hydrides implies that the surface structure formed might be noticeably different for the different preparation procedures. Wang *et al.* reported the B/Si(100) surface to be a mixture of regions of several reconstructions consisting of silicon dimers and dimer vacancies. In addition to double-protrusion features reported here and labeled as "A features" in Refs.23 and 24 regions of $c(4\times4)$ reconstruction were detected in Refs. 23 and 24 but have never been observed on our samples. The STM images of $c(4\times4)$ presented by Wang *et al.*[23,24] are identical in both polarities to the ones reported for metastable Si(100)$c(4\times4)$ structure formed under hydrogen exposure followed by hydrogen desorbing sample annealing.[25] This identity makes us suspect that the reconstruction observed in Refs. 23 and 24 might be induced by the interaction of the Si(100) surface with the hydrogen released during decomposition of B_2H_6 and $B_{10}H_{14}$ precursors rather than by the presence of boron at the surface. The procedure used in the present work (i.e., boron extraction from the bulk) is free from these side effects and $c(4\times4)$ does not form. Unfortunately, our study is restricted to a maximum B coverage of about 0.1 ML and thus the question of what the completed (stoichiometric) B/Si(100) surface looks like still remains open. The additional studies (e.g., STM observation of the B/Si(100) surface prepared by deposition of elemental B) are required to clarify the intrinsic structure of the "ideal" Si(100)-B surface phase.

B. Si overgrowth on B/Si(100) surface

Using the B/Si(100) surface as a substrate, Si overgrowth has been carried out both in SPE and MBE regimes, i.e., by RT deposition of amorphous Si followed by annealing and by Si deposition onto the heated substrate, respectively. We should admit that the experimental conditions of the present study do not coincide completely with the ones required for the delta-doping technique as (i) the surface is not a completed Si(100):B surface phase and (ii) boron is accumulated in a rather thick near-surface layer and not only at the surface. However, the main regularities of Si overgrowth on such "nonideal" B/Si(100) substrate are believed to remain valid for the fabrication of Si:B delta-doped layers.

1. Si-SPE growth

Figure 5(a) shows a STM image of a 3-ML-thick Si film deposited onto a B/Si(100) surface held at RT. One can see that the film is basically amorphous, although the presence of a few short pieces of dimer rows can be noticed. Annealing at temperatures of about 350 °C induces the SPE crystallization as evidenced by the appearance of well-defined dimer rows [Fig. 5(b)]. The observation of the SPE growth at such low temperatures indicates a facilitation of epitaxial crystallization in Si films of a few monolayers thickness, as compared to Si films of the more conventional thickness range of

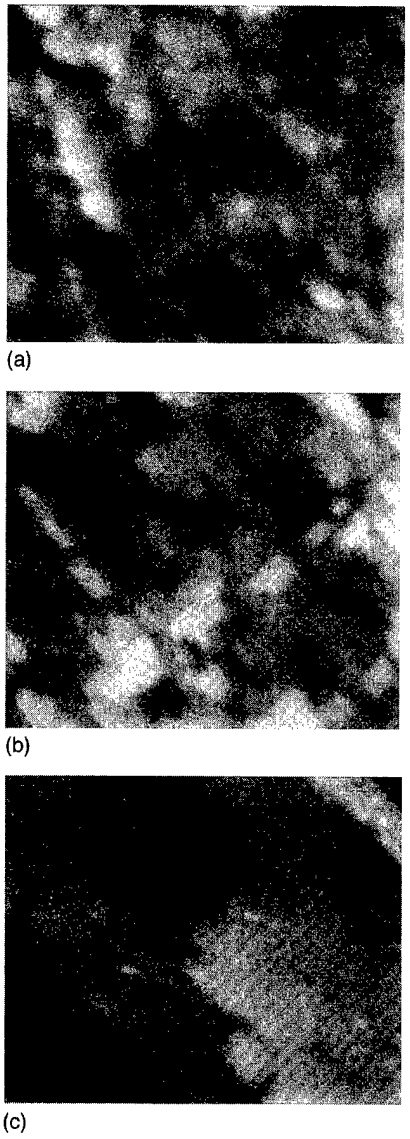
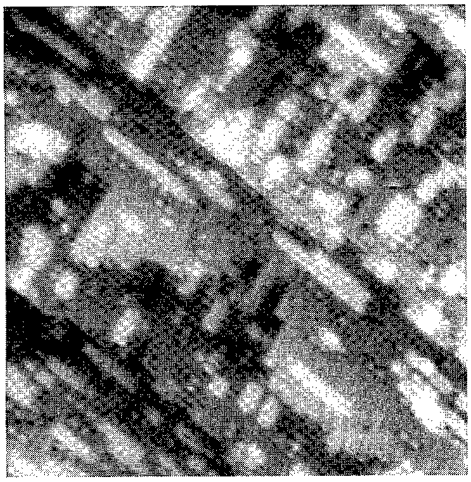


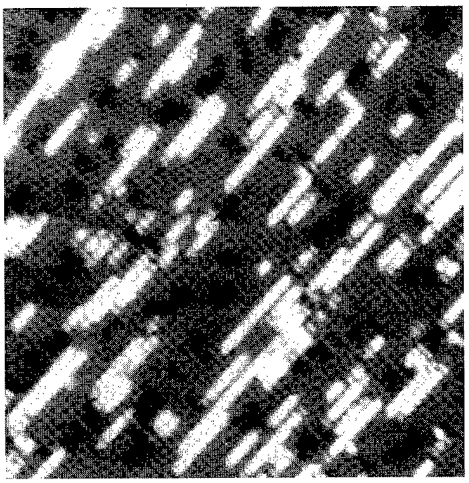
FIG. 5. Si-SPE growth on B/Si(100) surface. (a) 3 ML Si film as deposited at room temperature, (b) annealed at 350 °C for 5 min, (c) annealed at 500 °C for 7 min. Scale: 265×220 Å^2.

10–100 nm. The extrapolation to 350 °C from the known data for "thick" Si films[26] yields the negligible SPE crystallization rate of about 10^{-5} nm/min.

A detailed STM inspection of the annealed film has not revealed the presence of B-induced features on the surface. Selected protrusions initially suspected to be boron-induced features appear, at a greater magnification, to be isolated Si dimers or intrinsic Si(100) surface defects. The crystallized film is rough and up to four different atomic layers can be distinguished in Fig. 5(b). After annealing at higher temperatures [e.g., at 500 °C as illustrated by Fig. 5(c)] the film surface becomes perfect and flat and only very few B-induced features can be found on it. Thus, one can see that under appropriate heat treatment epitaxial Si overgrowth can be achieved by SPE without noticeable boron segregation, even on the atomic scale. It should be noted, however, that sufficient thickness (at least 3 ML) of the deposited Si film



(a)



(b)

FIG. 6. Surface after deposition of 0.2 ML Si at about 200 °C. (a) B/Si(100) surface and (b) clean Si(100). Scale: 230×230 Å^2.

appears to be a necessary requirement for the suppression of boron segregation from the doped substrate (ideally from the delta-doped layer) during Si-SPE regrowth. In fact, upon crystallization of amorphous 1–2 ML thick Si films, the B-induced features already reappear at the film surface at 350 °C. The existence of such a critical thickness needed to avoid surface segregation seems to be an inherent feature of Si-SPE and has also been reported for other dopants (e.g., for Sb[27]).

2. Si-MBE growth

Epitaxial growth of a Si cap layer on the B/Si(100) surface starts at a temperature of about 200 °C. The initial stage of epitaxial growth is illustrated by Fig. 6(a) which shows a B/Si(100) surface after deposition of about 0.2 ML Si. For comparison, Fig. 6(b) shows a clean Si(100)2×1 surface after deposition of the same amount of Si. One can see that in both cases the formation of long islands occurs with a preferential growth direction perpendicular to the dimer rows of the original substrate. However, in the case of the B/Si(100) substrate this growth anisotropy is less pronounced. This

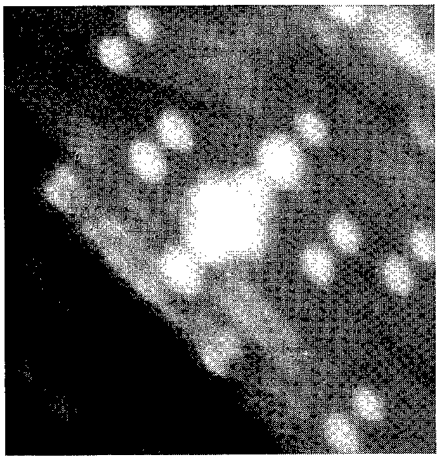


FIG. 7. Top Si dimer trapped between two boron-induced features.

specific feature of Si epitaxy on B/Si(100) surface seems to be caused by the enhanced topographical roughness of the original surface which does not provide enough area for continuous growth of long dimer strings [see Fig. 1(b)] and by the presence of B-induced features. The effect of these features may be twofold. First, a boron-induced feature may serve as a preferential site for island nucleation and, second, the propagation of the growing dimer string can be blocked on reaching a boron-induced feature. Analysis of many STM images shows that boron-induced protrusions are located mainly at the ends of newly grown dimer strings. The opposite arrangement is rare, i.e., there are only very few dimer strings or islands that have a boron-induced feature in the center while keeping their edges free. This may be an indication that boron-induced features are potential barriers for adatom diffusion. Another argument in favor of this assumption is the sporadic presence on the annealed B/Si(100) surface of features like the one shown in Fig. 7. It presents a remainder silicon dimer of a completely sublimated upper layer that has just been trapped between two boron-induced features.

Epitaxial growth of thicker films is presented in Fig. 8. Figure 8(a) shows a surface after deposition of 3 ML Si at about 200 °C. The film exhibits a well-defined epitaxial structure though it is rough on an atomic scale, and up to 5 different atomic layers are seen much like the case of SPE growth at 350 °C [Fig. 5(b)]. No apparent indication of any boron segregation can be found at this growth temperature. However, an increase in temperature leads to pronounced surface segregation of boron. As an example, Fig. 8(b) shows the surface of the film grown by MBE at about 500 °C. One can see the presence of numerous boron-induced double-protrusion features in each atomic layer. The threshold temperature for the onset of boron segregation could not be determined accurately in the present study due to limited accuracy of the temperature measurements. Our estimation suggests that for a growth rate of about 0.3 ML/min it should be situated between 230 and 300 °C.

The existence of a temperature range in which Si(100) epitaxial overgrowth takes place, but in which boron segre-

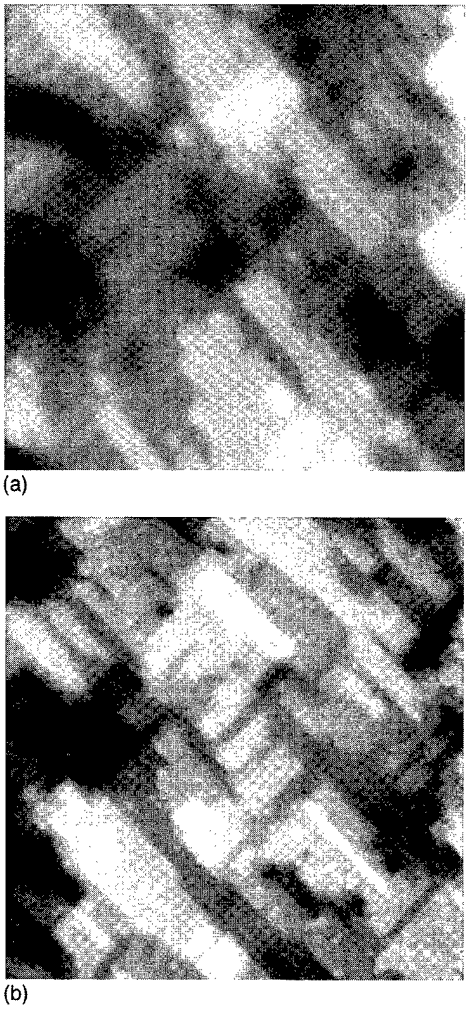


FIG. 8. Si-MBE growth on B/Si(100) surface: (a) 3 ML Si deposited at about 200 °C, (b) 1 ML Si deposited at about 500 °C. Scale: (a) 175×175 Å^2 and (b) 350×350 Å^2.

gation is still suppressed even on the atomic scale, is very promising for Si:B ordered delta doping. In contrast, such a temperature interval does not exist for Si growth on the B/Si(111) surface where B surface segregation already occurs at the minimal epitaxial temperature of about 400 °C.[13]

IV. CONCLUSION

Scanning tunneling microscopy has been used to study (i) the structure of the B/Si(100) surface prepared by annealing of B-doped Si(100) samples and (ii) regularities of epitaxial Si overgrowth on this surface. The appearance of specific boron-induced features has been detected and a model of their atomic structure has been proposed on the basis of STM observations and known literature data. In our model B atoms are suggested to reside in the second bulklike layer, occupying substitutional sites spaced by two periods of Si(100) plane while Si top dimers are eliminated denuding the underlying atoms. Epitaxial growth of a few monolayers has been monitored by STM for both MBE and SPE regimes. The growth conditions under which the suppression of B segregation occurs have been determined, namely, for MBE deposition temperatures should be below 230 °C, while for SPE annealing temperatures up to 500 °C can be used provided the Si film thickness is greater than 3 ML.

ACKNOWLEDGMENTS

One of the authors (A. V. Z.) acknowledges the Alexander von Humboldt Foundation for a research fellowship. Dr. P. Wagner (Wacker–Chemitronic) is also acknowledged for providing silicon wafers.

[1]H. P. Zeindl, T. Wegehaupt, I. Eisele, H. Oppolzer, H. Reisinger, G. Tempel, and F. Koch, Appl. Phys. Lett. **50**, 1164 (1987).
[2]A. A. van Gorkum, K. Nakagawa, and Y. Shiraki, J. Appl. Phys. **65**, 2485 (1989).
[3]N. L. Mattey, M. Hopkinson, R. F. Houghton, M. G. Dowsett, D. S. McPhail, T. E. Whall, E. H. C. Parker, G. R. Booker, and J. Whitehurst, Thin Solid Films **184**, 15 (1990).
[4]A. V. Zotov, A. A. Saranin, V. G. Lifshits, and E. A. Khramtsova, Surf. Sci. **230**, L147 (1990).
[5]I. Eisele, Superlattices Microstruct. **6**, 123 (1989).
[6]H. -J. Gossmann and F. C. Unterwald, Phys. Rev. B **47**, 12618 (1993).
[7]B. E. Weir, L. C. Feldman, D. Monroe, H. -J. Gossmann, R. L. Headrick, and T. R. Hart, Appl. Phys. Lett. **65**, 737 (1994).
[8]R. L. Headrick, B. E. Weir, A. F. J. Levi, D. J. Eaglesham, and L. C. Feldman, Appl. Phys. Lett. **57**, 2779 (1990).
[9]B. E. Weir, D. J. Eaglesham, L. C. Feldman, H. S. Luftman, and R. L. Headrick, Appl. Surf. Sci. **84**, 413 (1995).
[10]V. G. Lifshits, A. A. Saranin, and A. V. Zotov, *Surface Phases on Silicon. Preparation, Structure and Properties* (Wiley, Chichester, 1994).
[11]A. V. Zotov, S. V. Ryzhkov, and V. G. Lifshits, Surf. Sci. **328**, 95 (1995).
[12]R. L. Headrick, B. E. Weir, A. F. J. Levi, D. J. Eaglesham, and L. C. Feldman, J. Cryst. Growth **111**, 838 (1991).
[13]A. V. Zotov, M. A. Kulakov, S. V. Ryzhkov, V. G. Lifshits, B. Bullemer, and I. Eisele, Surf. Sci. **352–354**, 358 (1996); A. V. Zotov, M. A. Kulakov, B. Bullemer, and I. Eisele, Phys. Rev. B **53**, 12 902 (1996).
[14]P. Heuell, S. Cuzdi, M. A. Kulakov, and B. Bullemer, Thin Solid Films **264**, 217 (1995).
[15]V. V. Korobtsov, V. G. Lifshits, and A. V. Zotov, Surf. Sci. **195**, 466 (1988).
[16]A. V. Zotov, S. V. Ryzhkov, and V. G. Lifshits, Surf. Sci. **295**, L1005 (1993).
[17]S. Bensalah, J. -P. Lacharme, and C. A. Sebenne, Surf. Sci. **211/212**, 586 (1989).
[18]T.-C. Shen, C. Wang, J. W. Lyding, and J. R. Tucker, Phys. Rev. B **50**, 7453 (1994).
[19]D. J. Chadi, Phys. Rev. Lett. **59**, 1691 (1987).
[20]The number density of features grows with annealing following the boron surface concentration monitored by AES. These features have been reproducibly observed for all Si(100):B samples tested but never after the same heat treatment of Si(100) samples doped with other impurities (e.g., P and As).
[21]R. Cao, X. Yang, and P. Pianetta, J. Vac. Sci. Technol. B **11**, 1455 (1993).
[22]M. A. Kulakov, Z. Zhang, A. V. Zotov, B. Bullemer, and I. Eisele, Appl. Surf. Sci. (in press).
[23]Y. Wang, R. J. Hamers, and E. Kaxiras, Phys. Rev. Lett. **74**, 403 (1995).
[24]Y. Wang and R. J. Hamers, J. Vac. Sci. Technol. A **13**, 1431 (1995).
[25]R. I. G. Uhrberg, J. E. Northrup, D. K. Biegelsen, R. D. Bringans, and L. -E. Swartz, Phys. Rev. B **46**, 10251 (1992).
[26]A. V. Zotov and V. V. Korobtsov, J. Cryst. Growth **98**, 519 (1989).
[27]S. Kono, T. Goto, Y. Ogura, and T. Abukawa, in Abstracts of the Second Russia–Japan Seminar on Semiconductor Surfaces, Osaka, Nov. 13–17, 1995, p. 28.

High phosphorus doping of epitaxial silicon at low temperature and very low pressure

X. D. Huang, P. Han, H. Chen, Y. D. Zheng, L. Q. Hu, R. H. Wang, S. M. Zhu, and D. Feng
Department of Physics, Nanjing University, Nanjing 210008, China and Institute of Solid State Physics, Nanjing 210008, China

(Received 26 September 1995; accepted 3 May 1996)

In situ phosphorus doped Si epitaxial layers have been grown at 600 °C in a very low pressure chemical vapor deposition system, using SiH_4 and PH_3 diluted in H_2. The presence of H_2 was found to decelerate the epitaxial growth rates. Secondary-ion mass spectrometry measurement shows that the constant phosphorus concentration with depth for a steady flow of PH_3 was achieved. Namely, dopant concentration is a function of gas phase dopant concentration. Chemical concentration as high as 2.5×10^{20} P/cm^3 was obtained in Si epitaxial layers though with very low growth rate. Epilayers with constant doping levels from 1.5×10^{18} P/cm^3 to 1.2×10^{20} can readily be grown. Despite the relatively low growth rate, there is no evidence of a time-dependent accumulation of phosphorus on the growth surface or the reactor wall. The reasons accounting for this phenomenon are discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

Remarkable progress in silicon-based device technology has been made by taking advantage of the low-temperature preparation of epitaxial films in the Si/Ge materials system. A variety of unique devices[1–3] with high performance have been fabricated. Lately, the world's first highly integrated SiGe IC, a 1.0 GHz 12-bit digital-to-analog converter, was developed by IBM and Analog Devices.[4] Low-temperature growth of *n*-type films will be necessary in future integrated circuit fabrication. However, *n*-type doping to high levels in low-temperature (450–750 °C) Si deposition processes has been difficult due to the lack of control of phosphorus content and the heavy reduction of the growth rate.[5,6] In addition, when attempting to incorporate phosphorus concentrations in excess of 1×10^{18} P/cm^3, a steady flow of PH_3 produces an monotonically increasing or unstable level of phosphorus doping.[6,7] In this case, it is believed that phosphorus accumulates on and eventually passivates the growth surface. To achieve high levels of *n*-type doping at low temperature, study using $SiCl_2H_2$ as Si gaseous source has been carried out under atmospheric pressure.[8] Study of high level ($>5\times10^{19}$ cm^{-3}) *n*-type doping using SiH_4 as gaseous source has not been reported in the literature. In this article, we report an investigation of *in situ* phosphorus doping of Si epilayers at high levels using SiH_4 and 1800 ppm PH_3 diluted in H_2.

II. EXPERIMENT

Phosphorus-doped layers in this study were prepared in rapid thermal process/very low pressure chemical vapor deposition (RTP/VLP-CVD) reactor which has been previously described elsewhere.[9] The main feature of RTP/VLP-CVD is the use of tungsten-halogen lamps around the quartz tube reactor as energy source to activate deposition reaction. In this article, all samples were grown at 600 °C. Epitaxial layers were deposited on (100)-oriented Si substrates. Overall initial substrate preparation involved a standard RCA cleaning process followed by a 10:1 de-ionized (DI) water/HF bath for 15 s. This preclean process produces an oxide-free hydrogen passivated silicon surface.[10] Before deposition, the wafers were prebaked in 14 sccm H_2 flow at 600 °C for 15 min,[11] then the major residual species in the reactor was hydrogen. The main growth gaseous source was SiH_4. The gaseous source employed for the purpose of doping was 1800 ppm PH_3 in H_2. The partial pressure of SiH_4 was maintained at 10 mTorr during deposition. The ratio of diluted PH_3 flow rate to SiH_4 flow rate varied from 0.1 to 1.4. The total growth pressure varied correspondingly from 10 mTorr to 25 mTorr. Growth duration was between 45 and 150 min producing 80–350-nm-thick epitaxial layers. Epilayer thickness and phosphorus incorporation was determined from spreading resistance profiling (SRP), secondary-ion mass spectroscopy (SIMS) profiling.

III. RESULTS AND DISCUSSION

Figure 1 shows a typical SIMS profile of phosphorus doped Si epilayer in which chemical stability and abrupt dopant transition have been achieved. The flat depth profile shows that the epitaxial layer has nearly constant phosphorus level. The incorporation efficiency of phosphorus from the gas phase is stable. A steady flow of PH_3 produces a constant level of incorporated phosphorus rather than a monotonically increasing one.[7] Phosphorus content of epitaxial layer is determined by the phosphine atomic fraction present in the growth source. In our case, a surface poisoning effect[12] is not observed. There is no dopant buildup on the growth surface or the reactor wall. Figure 1 also shows that the dopant concentration turn-on slope is about 18 nm/decade. In addition, Fig. 2 shows the phosphorus doping concentration as a function of the ratio of diluted PH_3 flow rate to SiH_4 flow rate. The data suggest that the phosphorus concentration is proportional to diluted PH_3 flow rate and phosphorus does not

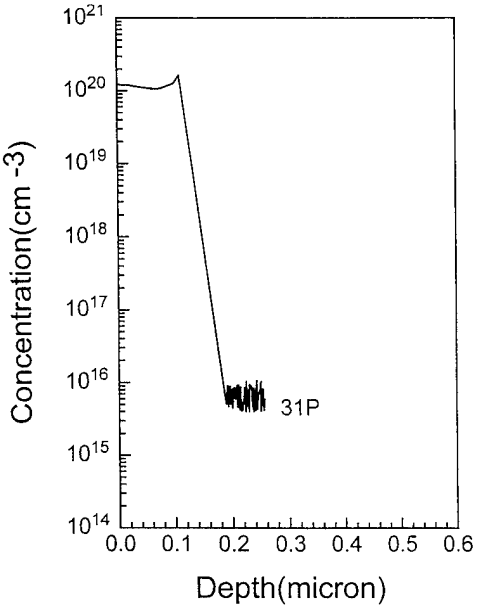


FIG. 1. A typical SIMS profile of phosphorus in epitaxial Si deposited at 600 °C from SiH_4 and 1800 ppm PH_3 in H_2.

block surface sites. Many phosphorus doped samples show that the reproducible range of phosphorus concentration is from 1.5×10^{18} to 1.2×10^{20} P/cm^3. We obtained growth rates from a series of SIMS profiles related to samples grown under different PH_3 flow rates, as plotted in Fig. 3. The growth rate decreases as the ratio of diluted PH_3 flow rate to SiH_4 flow rate increases. When the flow rate ratio varies from 0.1 to 1.4, the growth rate decreases from 3 to 0.4 nm/min.

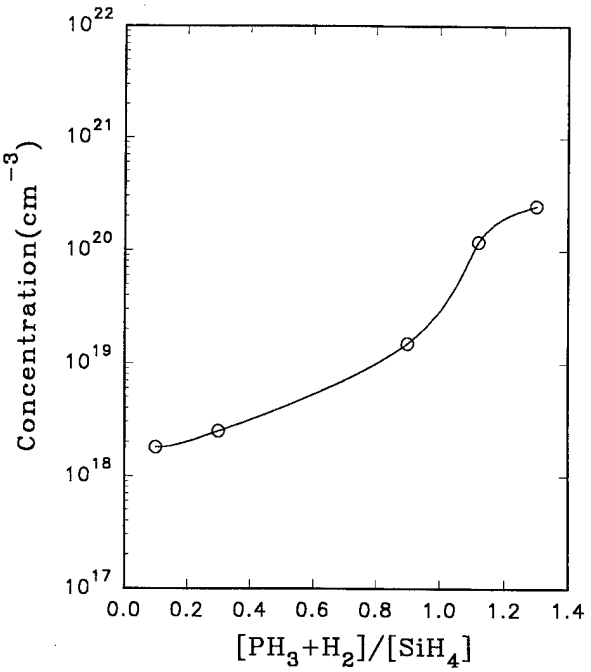


FIG. 2. Phosphorus incorporation in Si as a function of the ratio of diluted PH_3 flow rate to SiH_4 flow rate.

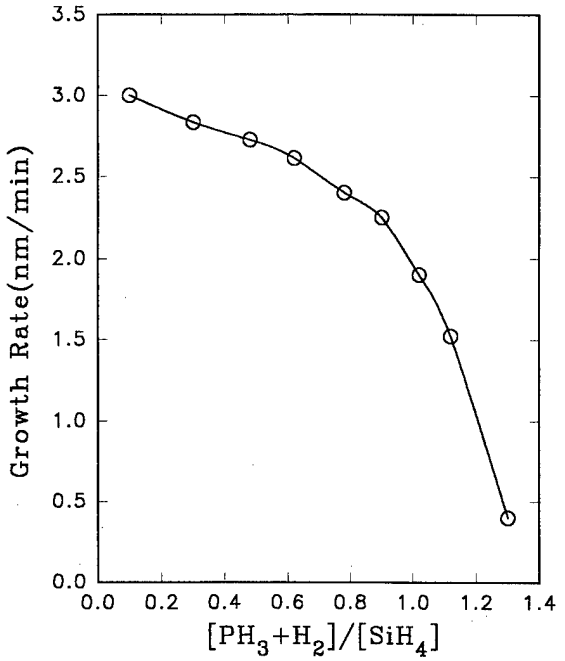


FIG. 3. Phosphorus doped epitaxial Si growth rate dependence on the ratio of diluted PH_3 flow rate to SiH_4 flow rate.

The reason, why a constant level of incorporated phosphorus is achieved, is attributed to the introduction of a large amount of diluted gas H_2 and the features of RTP/VLP-CVD. As reported in previous literature,[12] the study of the interaction of SiH_4 and PH_3 with Si(100) surface demonstrates that a stable layer of adsorbed phosphorus will form during the phosphorus-doping deposition process. This high phosphorus coverage leads to accumulation of phosphorus on Si surface, which accounts for the rapid reduction of the Si deposition rate and the lack of control of phosphorus content. With presence of a large amount of dilute gas H_2, the desorption of the reaction by-product, H_2 from surface becomes less efficient, so the surface reactive sites are mostly occupied by surface species, H, SiH, SiH_2, SiH_3, rather than phosphorus. The formation of the overlayer of adsorbed phosphorus is hindered by the presence of dilute gas H_2. In this instance, there is no accumulation of phosphorus on Si surface. This results in a constant level of incorporated phosphorus. On the other hand, the dilute gas H_2 will decrease the growth rates, since the desorption rate of hydrogen which is the reaction-limited process will decrease as a function of H_2 overpressure. Although, the incorporation of phosphorus can cause growth rate reduction, we believe the reductions is mostly a result of the introduction of dilute gas H_2 under lower incorporation of phosphorus from 1.5×10^{18} P/cm^3 to 1.2×10^{20}. When the content of phosphorus is larger than 1.2×10^{20} say 2.5×10^{20}, both introduction of dilute gas H_2 and incorporation of phosphorus decrease the reduction of growth rate equally.

IV. SUMMARY

By using 1800 ppm PH_3 diluted in H_2 as dopant in RTP/VLP-CVD, we obtained samples in which phosphorus dop-

ing concentrations are constant. Chemical concentration as high as 2.5×10^{20} P/cm^3 was achieved. No evidence shows time-dependent accumulation of phosphorus on the growth surface or the reactor wall. In addition, the growth rates decrease as the flow rate of diluted phosphine increases. We believe that this reduction of growth rate is mostly a result of the introduction of dilute gas H_2 under lower content of phosphorus, and that the dilute gas H_2 has important effect on the phosphorus doping process by reducing the desorption rate of hydrogen. As a result, the surface reactive sites are mostly occupied by surface species, H, SiH, SiH_2, SiH_3, rather than phosphorus. This leads to the phosphorus doping process in which there is no accumulation of phosphorus and the growth rates are relatively low.

ACKNOWLEDGMENT

The authors would like to thank Jilan Chen at Charles Evans and Associates for SIMS measurements.

[1]H. Temkin, J. C. Bean, A. Antreasyan, and R. Leibenguth, Appl. Phys. Lett. **52**, 1089 (1988).
[2]H. Daemgkes, H. Herzog, H. Jorke, H. Kibbel, and E. Kaspar, IEEE Trans. Electron Devices **ED-33**, 633 (1986).
[3]K. Ismail, B. S. Meyerson, and P. J. Wang, Appl. Phys. Lett. **59**, 973 (1991).
[4]A. Mcdonald, III–Vs Rev. **7**, 38 (1994).
[5]B. S. Meyerson and W. Olbricht, J. Electrochem. Soc. **131**, 2361 (1984).
[6]Syun-Ming Jang, Kenneth Liao, and Rafael Reif, Appl. Phys. Lett. **63**, 1675 (1993).
[7]B. S. Meyerson, Proc. IEEE **80**, 1592 (1992).
[8]T. O. Sedgwick, P. D. Agnello, D. Nguyen Ngoc, T. S. Kuan, and G. Scilla, Appl. Phys. Lett. **58**, 1896 (1991).
[9]Y. Zheng, R. Zhang, L. Hu, S. Gu, R. Wang, P. Han, and R. Jiang, Mater. Res. Soc. Symp. Proc. **263**, 227 (1992).
[10]B. S. Meyerson, F. J. Jimpsel, and K. J. Uram, Appl. Phys. Lett. **57**, 1034 (1990).
[11]J. O. Borland, *Symposium on Advanced Science and Technology of Silicon Materials*, edited by the 145th Committee of the Japan Society for the Promotion of Science, November 1991 (unpublished).
[12]B. S. Meyerson and M. L. Yu, J. Electrochem. Soc. **131**, 2366 (1984).

Characterization by x-ray photoelectron spectroscopy of the chemical structure of semi-insulating polycrystalline silicon thin films

Fabio Iacona and Salvatore Lombardo
Istituto Nazionale di Metodologie e Technologie per la Microelettronica (IMETEM), Consiglio Nazionale delle Richerche, Stradale Primosole 50, I-95121 Catania, Italy

Salvatore U. Campisano
Dipartimento di Fisica dell'Università, Corso Italia 57, I-95129 Catania, Italy

(Received 13 September 1995; accepted 11 April 1996)

The x-ray photoelectron spectroscopy (XPS) technique has been used to investigate the composition of semi-insulating polycrystalline silicon (SIPOS) films having oxygen contents of 10 and 35 at. % prepared by low pressure chemical vapor deposition. XPS analysis has demonstrated that the film compositions can be qualitatively described by means of the five Si-Si_xO_{4-x} tetrahedra (with $0 \leqslant x \leqslant 4$ and integer) predicted by the statistical random bonding model (RBM). However, the quantitative analysis of the XPS spectra has demonstrated that the concentrations of the various tetrahedra found in the SIPOS films are remarkably different from those predicted by a statistical approach, i.e., by assuming that each Si atom forms with equal probability bonds with either Si or O. We have also found that the composition of high temperature (up to 1000 °C) annealed films further departs from that predicted by the RBM model; indeed, the anneal promotes the decomposition of partially oxidized Si-Si_xO_{4-x} tetrahedra in Si-Si_4 tetrahedra (that form Si nanocrystals) and Si-O_4 tetrahedra (that enrich the oxygen content of the amorphous phase).
© *1996 American Vacuum Society.*

I. INTRODUCTION

Semi-insulating polycrystalline silicon (SIPOS) is a generic acronym used to indicate non-stoichiometric silicon oxide films, generally prepared by low pressure chemical vapor deposition (LPCVD) or by atmospheric pressure chemical vapor deposition (APCVD). SIPOS composition is usually expressed as SiO_x, with x that can be varied from 0 (polycrystalline or amorphous silicon) to 2 (stoichiometric silicon dioxide); the various compositions are obtained by changing the flow ratio between the silicon-bearing gas (generally SiH_4) and the oxygen-bearing one (generally N_2O).

SIPOS thin films are widely used as passivation layers and as resistive field shield in high voltage Si devices,[1,2] as emitters in heterojunction bipolar transistors[3,4] and for solar cell applications.[5,6] More recently, Er-doped SIPOS films have been proposed for optoelectronic application.[7] Because the physical properties of SIPOS films are strictly related to their oxygen content, a large number of papers have been published in the last few years dealing with the chemical and structural characterization of these films. It is generally recognized that as-deposited SIPOS films are amorphous mixtures of silicon and oxygen, while SIPOS films annealed at temperatures of 800–900 °C or higher are multiphasic materials, in which silicon nanocrystals are dispersed in an amorphous matrix.[8–10]

The microstructure of the amorphous phase can be described by means of two different models. The first one is the random bonding model[11] (RBM); according to this model SIPOS is formed by five different tetrahedral units, namely Si-Si_4, Si-Si_3O, Si-Si_2O_2, Si-SiO_3 and Si-O_4, corresponding to the five possible oxidation states of silicon (from 0 to +4). At each composition the tetrahedra distribution is determined by statistical laws, without the influence of any chemical driving force. So, for example, a SIPOS film having an oxygen content of 50 at. % (a "silicon monoxide" film) should be characterized by a tetrahedra distribution consisting in 6.25% of Si-Si_4, 25% of Si-Si_3O, 37.5% of Si-Si_2O_2, 25% of Si-SiO_3 and 6.25% of Si-O_4. Obviously, this symmetric distribution mirrors the fact that in a $SiO_{1.00}$ film each Si atom has the same probability to be surrounded by a Si or by an O atom.

The second model, the random mixture model (RMM), neglects the existence of the intermediate Si oxidation states, suggesting that all the possible SIPOS compositions are simply obtained by the suitable combination of Si-Si_4 and Si-O_4 tetrahedra.[12] According to this model, a $SiO_{1.00}$ film should be composed half of Si-Si_4 tetrahedra and half of Si-O_4 tetrahedra.

In this work we have investigated by means of the x-ray photoelectron spectroscopy (XPS) technique the chemical structure of two different SIPOS films prepared by low pressure chemical vapor deposition, having nominal oxygen contents of 10 and 35 at. %. These chemical structures have been compared with those predicted by the two above discussed models. We have also investigated the effect of high temperature annealing processes (up to 1000 °C) on the chemical structure of SIPOS.

II. EXPERIMENT

SIPOS films were deposited on ⟨100⟩ silicon substrates by using the LPCVD technique at a temperature of 620 °C. The gaseous precursors were SiH_4 and N_2O; the N_2O/SiH_4 flow ratio was set to values of 0.40 and 0.12 in order to obtain SIPOS films having nominal oxygen contents of 35 and 10

at. %, respectively. The deposition rates were about 20 Å/min for the 35 at. % O SIPOS film and about 40 Å/min for the 10 at. % O. The 35 at. % O films were about 3700 Å thick, while the 10 at. % O films were about 6000 Å thick. The bulk composition and the thickness of the films were determined by Rutherford backscattering spectrometry (RBS) of 1.7 MeV ^{4}He ions. Electron diffraction analysis of the SIPOS films was made on plan view transmission electron microscopy samples, prepared by mechanical thinning and backside chemical etching. The analysis was made in a JEOL 2010 transmission electron microscope (TEM), operating at 200 kV accelerating voltage. SIPOS films were annealed in a rapid thermal annealing apparatus, at temperatures ranging from 700 to 1000 °C, in nitrogen atmosphere. HF etched samples were prepared by dipping small fragments of the wafers for 15 s in a diluted HF solution (5% in H_2O) immediately before the introduction in the XPS vacuum chamber.

XPS analysis was performed using a Kratos ES 300 electron spectrometer, operating at a base pressure of 1×10^{-9} Torr. All spectra were obtained by using the unmonochromatized Al $K\alpha$ radiation. Angular resolved XPS (ARXPS) measurements were performed at two different take-off angles (measured with respect to the sample normal), i.e., 10°, roughly corresponding to the maximum sampling depth, and 60°, corresponding to a sampling depth reduced of approximately a factor of two.

The energy scale of the XPS spectra was calibrated assuming at 285.0 eV the binding energy (B.E.) of the C 1*s* peak due to sample surface contamination; under this assumption, the binding energy of the elemental Si component in the Si 2*p* peaks was 99.5 ± 0.15 eV for all samples. The Si 2*p* XPS peaks were analyzed by means of a computer program of gaussian peak fitting, using an inelastic Shirley-type background. The results of the peak fitting procedure were used to evaluate the oxygen content of the films, by considering the concentration of each component and the related oxygen content. The oxygen content was also evaluated from the XPS (Si 2*p*)/(O 1*s*) ratio, by introducing the appropriate corrections for the photoionization cross sections[13] and the experimentally determined transmission function of the spectrometer. The latter method generally gave higher oxygen contents (+10–20%) with respect to the former, indicating the presence of excess oxygen on the SIPOS surface, probably as C-O bonds in contamination layers, detectable also as a shoulder in the C 1*s* peaks. Consequently, the oxygen contents reported in the article are in all cases those obtained by the fitting procedure of the Si 2*p* peaks.

III. RESULTS AND DISCUSSION

A. Characterization of the as-deposited SIPOS films

In order to determine the average oxygen content of our films we have used the RBS technique. Figure 1 shows a typical RBS spectrum for a SIPOS film having a nominal oxygen content of 35 at. % (solid line). In order to better distinguish the backscattering yields of O and Si in SIPOS,

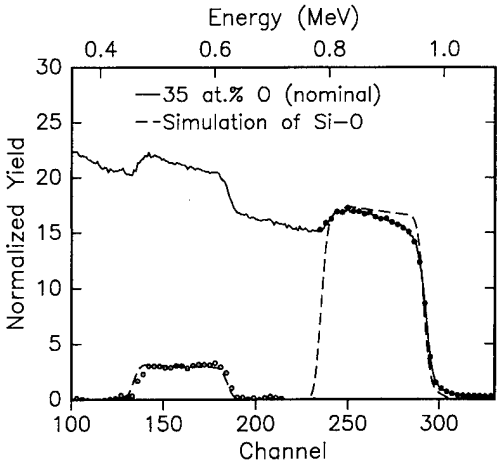


FIG. 1. Experimental (solid line) and background subtracted (open circles) RBS spectra for a SIPOS film having a nominal oxygen content of 35 at. %, compared with the simulation of a RBS spectrum for a Si-O mixture containing 35 at. % of oxygen (dashed line).

the sample was oriented with the ⟨100⟩ axis of the Si substrate aligned with the He beam (channelling conditions), thus reducing the backscattering yield due to the substrate. The figure shows also the spectrum after the subtraction of the background due to backscattering from the substrate in the 100–220 channel region (open circles), thus leaving only the signal due to the O of SIPOS. This background subtracted spectrum fits quite well with a simulation relative to a Si-O mixture containing 35 at. % of oxygen (dashed line), even if a small O enrichment is observed in the surface region. Similarly, the RBS analysis has demonstrated that the SIPOS films having a nominal oxygen content of 10 at. % are really characterized by this atomic concentration of oxygen.

According to the above described RBM model, the oxygen atoms contained in SIPOS can exist under the form of $Si\text{-}Si_3O$, $Si\text{-}Si_2O_2$, $Si\text{-}SiO_3$ and $Si\text{-}O_4$ tetrahedra.[11] On the other hand, the RMM model predicts only the existence of $Si\text{-}O_4$ tetrahedra.[12] In this work we have employed the XPS technique in order to investigate if the compositions of our SIPOS films agree with one of the proposed models.

To extract detailed chemical information from the XPS Si 2*p* peaks it is necessary to establish a suitable peak fitting procedure. Our procedure was based on the use of five gaussian-shaped peaks to simulate the five silicon oxidation states; the full widths at half maximum (FWHM) used were 1.6 eV for the $Si\text{-}Si_4$ tetrahedra and 1.8 eV for the oxidized species. We have found that for all our SIPOS samples the binding energies of the elemental Si components fall in the range 99.5 ± 0.15 eV, in good agreement with literature[14]; the binding energies of the various Si-O species were obtained from the empirical relationship $\Delta E = 2.2x$, which correlates the binding energy chemical shift (ΔE) of a SiO_x species with respect to elemental Si with its oxygen content x.[15]

An example of this approach is shown in Fig. 2 for a SIPOS film containing 35 at. % of oxygen. The figure shows

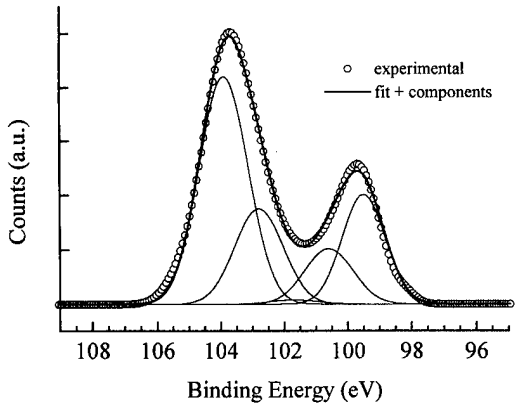


FIG. 2. Background subtracted Si 2*p* XPS peak for a 35 at. % O SIPOS film, illustrating the peak fitting procedure. The five components correspond to the five Si oxidation states, i.e., to Si-Si_4, Si-Si_3O, Si-Si_2O_2, Si-SiO_3 and Si-O_4, tetrahedra (in order of increasing binding energies).

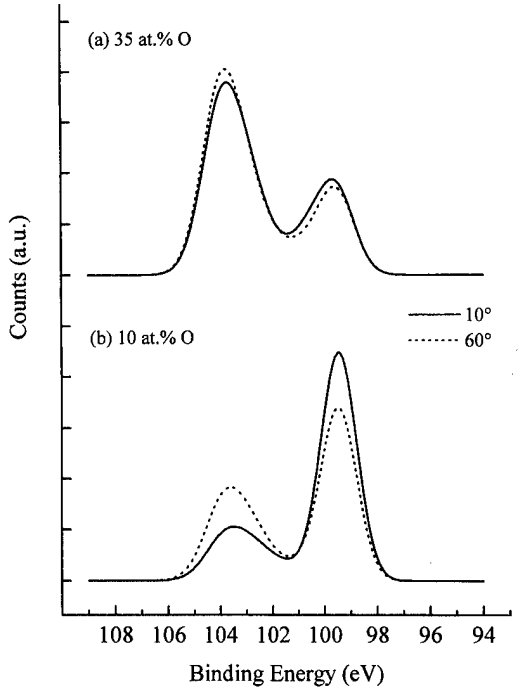


FIG. 3. Background subtracted Si 2*p* XPS peaks for SIPOS films containing (a) 35 at. % and (b)10 at. % of oxygen, obtained at two different take-off angles, 10° (maximum sampling depth) and 60° (enhanced surface sensitivity).

that all the five different Si species predicted by the RBM model are present. The main species are Si-Si_4 (99.5 eV B.E.) and Si-O_4 (103.9 eV B.E.), but relevant contributions of Si-Si_3O (100.6 eV B.E.) and Si-SiO_3 (102.8 eV B.E.) are present, while Si-Si_2O_2 (101.7 eV B.E.) is the only kind of tetrahedron having an almost negligible intensity. The quantitative analysis of the peak fitting gives an oxygen content of about 57 at. %, a value much greater than that expected on the basis of the RBS measurements (35 at. %). We explain this result with an oxidation phenomenon of the SIPOS surface due to the exposure to the atmosphere, similar to that observed in crystalline silicon, where a thin (~10–15 Å) SiO_2 layer is formed. In order to characterize the oxidized overlayer of SIPOS, we have used the ARXPS technique.

Figure 3 shows the background subtracted XPS Si 2*p* peaks obtained at two different take-off angles (10°, roughly corresponding to the maximum sampling depth, valuable in about 90 Å on the basis of the values of inelastic mean free path of Si and SiO_2,[16] and 60°, corresponding to a sampling depth reduced of about a factor of two) for both the studied SIPOS compositions. In particular, Fig. 3a refers to the 35 at. % O SIPOS film. The absence of any large angular dependence indicates for this sample the presence of a partially oxidized overlayer, whose thickness is comparable to the maximum XPS sampling depth or larger.

A surface oxygen enrichment is also observed for the 10 at. % O SIPOS films, as shown in Fig. 3b. The peak fitting procedure performed on the Si 2*p* peak obtained at a take-off angle of 10° indicates a very high oxygen content (about 32 at. %) with respect to the value given by the RBS measurements (10 at. %). Differently than the previous case, the ARXPS analysis gives evidence that the Si 2*p* peak obtained in a situation of enhanced surface sensitivity (60°) exhibits a more intense oxide component; the quantitative analysis of these ARXPS data indicates the presence of a 15 Å thick overlayer, mainly composed of Si-O_4 and Si-SiO_3 tetrahedra, on the 10 at. % O SIPOS surface.

The conclusion is therefore that both the SIPOS surfaces are not representative of the bulk of the films, due to an oxidation phenomenon induced by the exposure to the atmosphere. Furthermore, these data indicate that the SIPOS reactivity towards atmospheric oxygen increases with the O concentration of the film, contrary to the expected situation.

The surface oxidation of SIPOS is a general problem, encountered by all the researchers dealing with the XPS analysis of this material.[14,17–20] There are two possible ways to remove the surface oxidized layer and expose to the XPS analysis a "true" SIPOS surface, i.e., sputtering and wet chemical cleaning, by using HF solutions. We did not use sputtering, because it is well known that a secondary effect of the ion bombardment is to induce chemical modifications in the bombarded surface. It has been found that also the very stable SiO_2 surface can be affected by inert gas bombardment, undergoing loss of oxygen.[21] Therefore, we expect greater effects for the less stable, partially oxidized Si-Si_xO_{4-x} tetrahedra contained in SIPOS, such as rearrangements of the Si-O bonds, and preferential sputtering of oxygen.[17,18] Other reported deleterious effects induced by the sputtering are reactions with carbonaceous overlayers, leading to the formation of Si-C bonds,[20] and the oxygen uptake due to the exposure of the very reactive sputtered surface to residual H_2O vapor during the pumping down subsequent to the sputtering treatment.[14]

Due to the above discussed problems, we have preferred to use the HF wet etching method to remove the oxidized layer from the surface of our films. This method is widely used in literature,[14,18–20] even if some caution must be taken

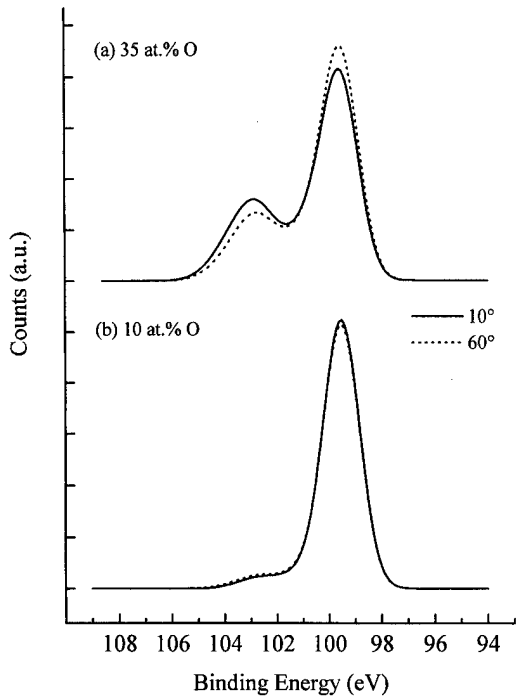


FIG. 4. Background subtracted Si $2p$ XPS peaks for HF etched SIPOS films containing (a) 35 at. % and (b) 10 at. % of oxygen, obtained at two different take-off angles, 10° (maximum sampling depth) and 60° (enhanced surface sensitivity).

with regard to the possibility of etching the SIPOS film also.[17] To avoid this problem we have used very diluted HF solutions (5% in H_2O), characterized by very low etch rates, and very low dipping times (15 s). Moreover, we do not expect to have etching effects, since the O content of our films is relatively low. Indeed, it has been reported that SIPOS etching by concentrated HF solutions occurs only for films having an oxygen content greater than 34 at. %.[20] As a further test, we have verified that, for both the studied SIPOS compositions, the XPS spectra do not appreciably change their shapes even if the HF dipping time is doubled (30 s). The etched samples have been introduced in the XPS analysis chamber immediately after the HF dip (the total time spent between the HF dip and introduction in the UHV chamber is less than 1 min).

The background subtracted Si $2p$ XPS spectra obtained after HF etching for both SIPOS samples are shown in Fig. 4. Figure 4a in particular shows the Si $2p$ XPS peaks for the 35 at. % O SIPOS film, obtained at take-off angles of 10° and 60°. The spectrum obtained at a take-off angle of 10° shows a marked decrease of the contribution of the oxidized species with respect to the situation found for the unetched film. The peak fitting analysis indicates an oxygen content of 37 at. %, in excellent agreement with the value obtained by RBS (35 at. %). The absence of any relevant dependence on the take-off angle for the two peaks shown in figure demonstrates the complete elimination of the oxidized overlayer.

Figure 4b shows the Si $2p$ XPS peaks for the 10 at. % O SIPOS film, obtained at take-off angles of 10° and 60°, after the HF dip. The spectra show that also in this case the HF dip produces a marked decrease of the oxidized species concentration. The peak fitting procedure indicates an oxygen content of 10 at. %, in agreement with the RBS data. Also in this case the Si $2p$ band shape does not depend on the take-off angle, demonstrating the absence of oxidized overlayers.

TABLE I. Experimental and calculated percent concentrations of the $Si\text{-}Si_xO_{4-x}$ tetrahedra, for SIPOS films containing 10 and 37 at. % of oxygen. The experimental data refer to the as-deposited, HF etched samples.

	Percent concentrations					
	10 at. % O			37 at. % O		
Tetrahedron	Expt.	RMM	RBM	Expt.	RMM	RBM
$Si\text{-}Si_4$	87.7	94.5	79.7	52.0	70.7	24.9
$Si\text{-}Si_3O$	6.7	···	18.6	15.8	···	41.4
$Si\text{-}Si_2O_2$	1.9	···	1.6	4.1	···	25.8
$Si\text{-}SiO_3$	3.6	···	0.1	20.5	···	7.1
$Si\text{-}O_4$	0.1	5.5	···	7.6	29.3	0.8

The absence of angular dependence shown by the Si $2p$ peaks, as well as the excellent agreement between the total oxygen concentrations obtained by XPS and those obtained by RBS, demonstrates the capability of the HF etching method in producing clean, homogeneous SIPOS surfaces, representative of the bulk of the films. Therefore, in the rest of the article the oxygen concentrations given by the XPS analysis of the HF etched surfaces (10 and 37 at. %) will be used to identify the two studied SIPOS films.

The oxygen distributions for the two HF etched SIPOS films, as obtained by the peak fitting procedure of the relative Si $2p$ peaks, are reported in columns 2 and 5 of Table I, in terms of percent concentrations of the various $Si\text{-}Si_xO_{4-x}$ tetrahedra. These data demonstrate that in as-deposited SIPOS oxygen is present under the form of all the $Si\text{-}Si_xO_{4-x}$ tetrahedra predicted by the RBM model; for both films the main species is $Si\text{-}Si_4$, but, in the film containing 10 at. % O there is also a strong component due to the $Si\text{-}Si_3O$ tetrahedra, whilst in the 37 at. % O there is a large contribution due to $Si\text{-}SiO_3$. Therefore, the tetrahedra distribution in SIPOS seems to depend on the total oxygen content of the film. These differences in the tetrahedra concentrations probably determine also the different reactivity toward atmospheric oxygen shown by the XPS spectra of the two unetched SIPOS films. Indeed, the high reactivity of SIPOS toward oxygen is due to the presence of partially oxidized tetrahedra; Table I shows that the 37 at. % O SIPOS contains about 48% partially oxidized tetrahedra, and therefore is more reactive than the 10 at. % O, which contains only about 12% of this kind of tetrahedra.

B. Comparison with the RBM and RMM models

Our XPS data have demonstrated that SIPOS is composed by all the tetrahedra predicted by the RBM model. This means that the RMM model should be totally unsatisfactory to describe the composition of our films because it does not predict the existence of any partially oxidized tetrahedron.

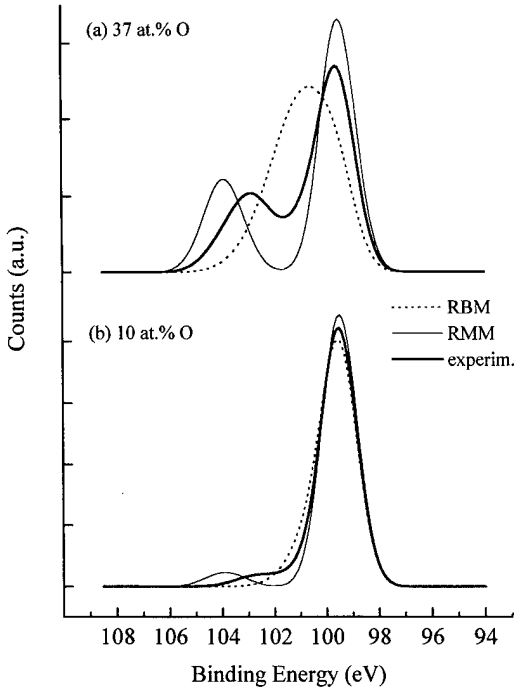


FIG. 5. Comparison among the experimental Si 2*p* XPS peak, and those built according to the RBM and RMM models, for SIPOS films containing (a) 37 at. % and (b)10 at. % of oxygen.

However, in order to compare in a more careful way our experimental data with the two models, we have calculated the distributions of the Si-Si_xO_{4-x} tetrahedra predicted by the two models for the two studied SIPOS compositions (10 and 37 at. % O). The results of the calculation are shown in columns 3, 4, 6 and 7 of Table I. We have used these theoretical distributions to build the relative Si 2*p* XPS peaks, by applying the same B.E. and FWHM used to fit the experimental spectra. Figure 5 shows the comparison between these calculated spectra and the experimental ones.

Figure 5a compares the two calculated Si 2*p* peaks with the experimental one for the 37 at. % O SIPOS film. The figure demonstrates that the experimental peak significantly differs from those predicted by both models. In particular, though the RMM model is clearly inadequate to describe the 37 at. % O SIPOS, which is characterized by a relevant content of intermediate Si-O species, the overall aspect of the experimental spectrum is closer to this model than to the RBM one. Indeed, the RBM spectrum is a wide, unresolved band, peaked in correspondence of the B.E. of the Si-Si_3O tetrahedra. On the other hand, the RMM spectrum exhibits a double-peaked structure, like the experimental one, even if the oxidized peak of the experimental spectrum has its maximum at lower B.E. with respect to the calculated one.

Figure 5b compares the two calculated Si 2*p* peaks with the experimental one for the 10 at. % O SIPOS film. The figure shows also in this case marked differences between experimental and calculated peaks. In particular, the experimental peak clearly differs from the one predicted by the RMM model, because the experimental data show a relevant content of partially oxidized Si-Si_xO_{4-x} tetrahedra, whose existence is neglected by this model; on the other hand it appreciably differs also from the one predicted by the RBM model, because, according to this model, almost all its oxidized components should exist as Si-Si_3O tetrahedra.

Therefore, the chemical structures of our SIPOS films do not satisfactorily agree with those predicted by the RBM and RMM models. The RMM model fails by neglecting the intermediate Si-O species existence, which has been clearly demonstrated by XPS, not only in SIPOS films, but also in evaporated or sputter deposited substoichiometric SiO_x films,[22–24] and at the interface between single crystal Si and native SiO_2.[25–27] On the other hand, even if the approach of the RBM model, by taking into account the existence of the intermediate oxidation states of silicon, is more correct, the deviation of the experimental spectra from this model is huge, demonstrating the failure of this pure statistical model. This means that the formation of the various tetrahedra during the film deposition cannot be treated as a pure statistical problem, but chemical factors must be taken into account. Even if, during deposition, the arrival of the O and Si atoms on the SIPOS surface can be described in terms of statistical laws, according to the RBM model, the subsequent step, i.e., the formation of the Si-O bonds, must be clearly influenced by chemical factors. The relative weights of the chemistry and the statistics in determining the film structure are determined by the deposition rate coupled with the surface diffusion rate of the reacting species. Indeed, it is known that the Si-O bond, due to the greater electronegativity of O, has an ionic character that determines a bond polarization which can be described as $Si^{\delta+}$-$O^{\delta-}$. For low deposition rates and high surface diffusivities, an O atom arriving at the SIPOS surface prefers to bind to a polarized Si atom, thus forming Si-O_4 tetrahedra, according to the RMM model. On the other hand, for high deposition rates and low surface diffusivities, the atoms have no time to be influenced by this factor, and therefore form tetrahedra according to the statistical laws of the RBM model, because they are quenched in their positions by the fast deposition of new layers.[28]

According to the above discussion, it is therefore not surprising that our LPCVD SIPOS films contain less partially oxidized Si-Si_xO_{4-x} tetrahedra than predicted by the RBM model. Indeed, in our case the deposition rate is sufficiently low (20–40 Å/min) and the surface diffusion rate sufficiently high (due to the high deposition temperature of 620 °C) to allow a relevant number of oxygen atoms to bond to already polarized Si atoms, thus forming a higher number of Si-SiO_3 and Si-O_4 tetrahedra than predicted by the RBM model. Also in other cases reported in literature[14,17–20] the shape of the XPS spectra of LPCVD SIPOS films seems to confirm this model, even if a detailed analysis of the Si 2*p* peaks in terms of the five silicon oxidation states is generally missing.

The influence of the deposition rate in determining the tetrahedra composition of SIPOS has been already demonstrated on the basis of infrared data coupled with data about stress and reactivity to HF.[28] In this work, LPCVD SIPOS films (typical deposition rates of 30 Å/min, roughly corre-

sponding to the deposition of a monolayer each 6 s) have been found to exhibit a marked SiO_2 character with respect to APCVD SIPOS films (typical deposition rates of 500 Å/min, corresponding to about 3 monolayers/s).

This analysis can be extended also to the case of SiO_x films prepared by evaporation. It has been shown that the chemical structure of these films is characterized by high concentrations of partially oxidized $Si\text{-}Si_xO_{4-x}$ tetrahedra.[22,23] In this case the very high deposition rate (about 1000 Å/min) and the low surface diffusion rate (due to the low deposition temperature) minimize the influence of the chemical factor, leading to the deposition of films having compositions in good agreement with those predicted by the RBM model.

C. Effect of high temperature annealing

The chemical characterization of high temperature annealed SIPOS films must take into account the deep structural rearrangements undergone by these films. Indeed, by annealing amorphous SIPOS films at temperatures of about 800–900 °C, the nucleation of Si nanocrystals occurs.[8–10] We have analyzed by XPS the same SIPOS films containing 10 and 37 at. % of oxygen already characterized in the as-deposited state, after annealing at temperatures ranging from 700 to 1000 °C. Because we have found that also the surfaces of the annealed films contain an excess of oxygen, also in this case we have used a HF dip to remove this oxidized overlayer before the XPS analysis.

Figure 6 reports the composition, obtained by XPS, as a function of the annealing temperature, for the 37 at. % O SIPOS. In particular, Fig. 6a shows the atomic percent composition as a function of the annealing temperature. Note that the oxygen concentration is constant up to a temperature of 900 °C and decreases from 37 to 28 at. % at 1000 °C. Figure 6b reports the atomic percent concentrations of the different Si species, derived from the peak fitting analysis of the Si 2*p* XPS peaks, as a function of the annealing temperature. The figure shows that the tetrahedra concentrations are strongly dependent on temperature. In order to discuss this figure it is useful to divide it into two parts: a first one, up to 900 °C, corresponding to the temperature range in which the film composition, according to Fig. 6a, is constant, and a second part referring to the 1000 °C annealing. The figure shows that up to 900 °C the concentrations of the $Si\text{-}Si_4$ and SiO_4 tetrahedra increase, while all the other contributions decrease. This means that the anneal promotes the transformation of partially oxidized tetrahedra in those having extreme compositions. As a result, the composition of SIPOS films annealed up to a temperature of 900 °C further departs from the RBM model, approaching the RMM model.

By extrapolating from the data up to 900 °C the SIPOS behavior at higher temperatures, one should expect a further increase of the $Si\text{-}Si_4$ and $Si\text{-}O_4$ concentrations, with a corresponding decrease of the other species. Figure 6b instead shows that at 1000 °C all the concentrations of the oxidized tetrahedra decrease, while only the $Si\text{-}Si_4$ contribution largely increases.

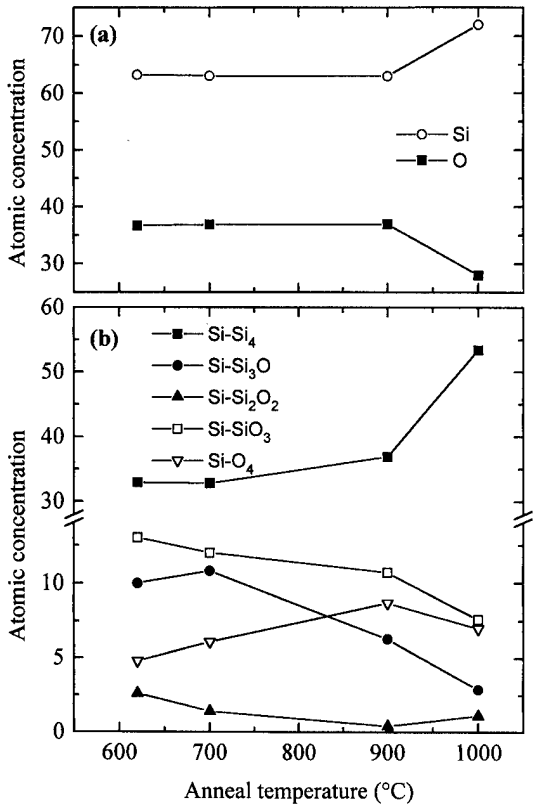


FIG. 6. (a) Atomic percent composition, and (b) $Si\text{-}Si_xO_{4-x}$ tetrahedra percent concentration as a function of the annealing temperature, for a SIPOS film containing 37 at. % of oxygen.

This discontinuity must be related to the decrease of the total oxygen concentration occurring at this temperature shown in Fig. 6a. According to the low temperature trends, we think that at 1000 °C the conversion of the $Si\text{-}Si_3O$, $Si\text{-}Si_2O_2$ and $Si\text{-}SiO_3$ tetrahedra into $Si\text{-}Si_4$ and $Si\text{-}O_4$ becomes very efficient. This effect is not shown by the XPS spectra, because, contrarily to the as-deposited and low temperature annealed SIPOS films, the films annealed at 1000 °C are not stable towards the HF dip used to remove the surface oxidized layer, but undergo a selective etching the $Si\text{-}O_4$ tetrahedra; this explains both the decrease in the total amount of oxygen (see Fig. 6a) and of $Si\text{-}O_4$. This interpretation of the XPS data is also confirmed by RBS. Indeed, for all the anneal temperatures, the SIPOS compositions measured by RBS remain constant to the 35 at. % value, demonstrating that the loss of oxygen occurring at 1000 °C is not a bulk phenomenon due to the anneal, but an artifact due to the HF dip. Therefore, the stability of SIPOS films with respect to the HF etching seems to depend not only on the total oxygen content of the film,[20] but also on the kind of tetrahedra composing the film.[28]

All these findings are confirmed also by the behavior of the 10 at. % O SIPOS. Figure 7a shows the atomic percent composition of these films, derived by the XPS analysis, as a function of the annealing temperature. Also in this case we observe two distinct regions; a first one, up to 800 °C, in which the SIPOS composition remains almost constant, and

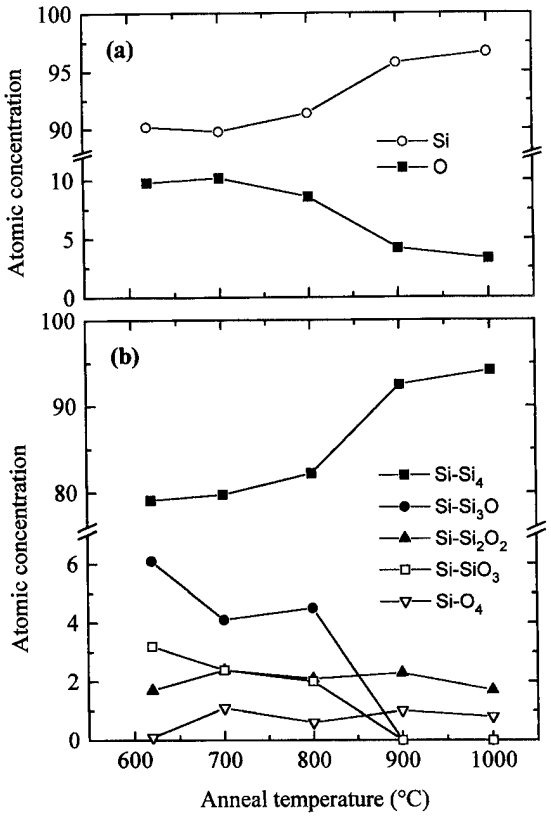


FIG. 7. (a) Atomic percent composition, and (b) $Si\text{-}Si_xO_{4-x}$ tetrahedra percent concentration as a function of the annealing temperature, for a SIPOS film containing 10 at. % of oxygen.

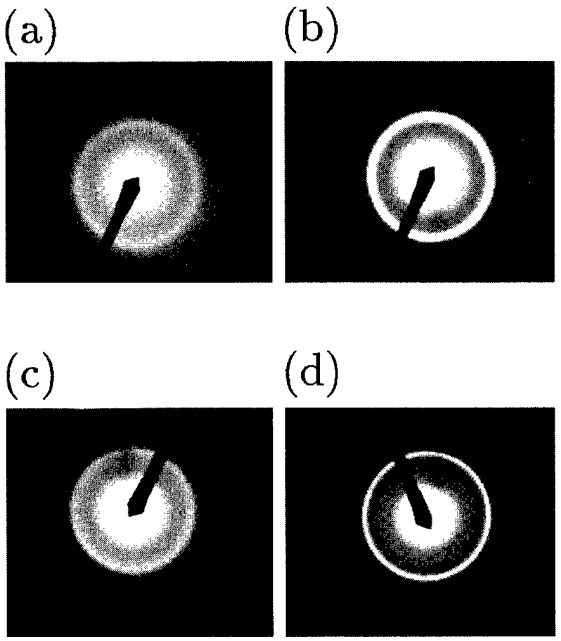


FIG. 8. Electron diffraction patterns for a 10 at. % O SIPOS film after annealing at (a) 800 and (b) 900 °C, and for a 37 at. % O SIPOS film after annealing at (c) 900 and (d) 1000 °C.

a second one, at higher temperatures, in which the oxygen content decreases (4 at. % at 900 °C, 3 at. % at 1000 °C). Figure 7b reports the atomic percent concentrations of the different tetrahedra, derived from the peak fitting analysis of the Si 2*p* XPS peaks, as a function of the annealing temperature. The figure shows that also in this case the various contributions depend on the temperature. In particular, up to 800 °C a slow decrease of the main oxidized species concentrations is observed, while at 900 °C all the contributions of the oxidized tetrahedra drop to 0, or become only slightly greater than 0. Correspondingly, the contribution of the $Si\text{-}Si_4$ tetrahedron increases up to a value of about 94%.

Therefore, even if the low oxygen content makes difficult a clear evaluation of the phenomena occurring in the low temperature range (620–800 °C), by applying the same considerations reported above for the 37 at. % O SIPOS, we can conclude that also for the 10 at. % O SIPOS films the effect of the annealing mainly consists in a conversion of the partially oxidized tetrahedra into $Si\text{-}Si_4$ and $Si\text{-}O_4$; this effect becomes relevant at 900 °C, and at this temperature the high concentration of $Si\text{-}O_4$ tetrahedra formed by the annealing leads to a material that is unstable towards the HF dip we used to remove the surface oxidized layer, so that a preferential etching of SiO_2 occurs. Also in this case these conclusions are supported by RBS data, showing that the bulk oxygen content of the film does not change due to the anneal.

The observation that also the 10 at. % O SIPOS can be affected by the HF dip is particularly remarkable. Indeed, according to literature data,[20] the 37 at. % O SIPOS films are in the boundary region between stability and instability towards the HF etching. Therefore, we would expect the 10 at. % O samples to be totally stable, due to the very low oxygen content. The observed instability of these films allows to us to affirm that the main factor determining the stability towards the HF etching of SIPOS is not the total oxygen content, but the way in which this oxygen is bonded to the Si atoms.

A last point to be discussed is the different temperatures at which the prominent formation of SiO_2 starts (and consequently also the preferential etching due to HF) for the two studied SIPOS compositions. We have already shown that the 37 at. % O SIPOS is stable up to 900 °C, while the 10 at. % O one is stable up to 800 °C. It is remarkable that these temperatures are exactly the same at which the crystallization of these materials starts, as shown in Fig. 8. This figure shows the transmission electron diffraction patterns obtained for the 10 at. % O SIPOS after annealing at 800 and 900 °C (a, b), and for the 37 at. % O after annealing at 900 and 1000 °C (c, d). Note that the 10 at. % O sample is amorphous at 800 °C, as demonstrated by the presence of broad and diffuse rings in the diffraction pattern; at 900 °C the rings become well defined structures, indicating the presence of the nanocrystalline Si phase. Similarly, for the 37 at. % O sample, the electron diffraction patterns demonstrate the presence of an amorphous phase up to 900 °C, while the crystallization starts at 1000 °C.

Therefore, the XPS analysis has allowed to us to demonstrate a correlation between structural changes (formation of

Si nanocrystals) and chemical transformations (formation of $Si-Si_4$ and $Si-O_4$ tetrahedra) induced by high temperature annealing in SIPOS. In particular, the nucleation of silicon nanocrystals occurring at high temperatures can be related to the formation of $Si-Si_4$ tetrahedra, due to the decomposition of the partially oxidized Si tetrahedra shown in Figs. 6 and 7. These nanocrystals are surrounded by an amorphous phase that is more oxygen-rich compared to that present in the as-deposited film, since the former also contains the $Si-O_4$ tetrahedra formed due to the temperature-induced decomposition of the partially oxidized tetrahedra. This conversion of the $Si-Si_3O$, $Si-Si_2O_2$ and $Si-SiO_3$ tetrahedra into $Si-Si_4$ and $Si-O_4$ shifts the compositions of the annealed films towards those predicted by the RMM model.

IV. CONCLUSIONS

We have described in this article the results of the XPS analysis of LPCVD SIPOS films containing 10 and 37 at. % of oxygen. We have found that the as-deposited samples are mixtures of all the possible $Si-Si_xO_{4-x}$ tetrahedra predicted by the RBM model and that each SIPOS composition exhibits a peculiar tetrahedra distribution. However, these distributions do not satisfactorily agree with those predicted by this model, because the detected concentrations of partially oxidized $Si-Si_xO_{4-x}$ tetrahedra are too low.

We have explained this deviation from the RBM model in terms of deposition rate and surface diffusivity. In general, high deposition rate and low surface diffusivity involve the prevailing of the statistical factor, and therefore the formation of films characterized by high concentrations of intermediate Si oxidation states (RBM model), while a low deposition rate and a high surface diffusivity allow the deposition of films with a greater SiO_2 character (RMM model). Our LPCVD films are characterized by a sufficiently low deposition rate and by a sufficiently high surface diffusivity to have a composition remarkably different from the one predicted by the RBM model.

We have also observed that the SIPOS films containing 37 at. % of oxygen are more reactive than the film with 10 at. % of oxygen with respect to the exposure to atmospheric oxygen, since they tend to oxidize for quite large thicknesses (larger than 90 Å). We have explained this effect with the higher concentration of partially oxidized tetrahedra observed in the 37 at. % O SIPOS. This oxidized layer can be selectively removed from the surface of the as-deposited samples by using diluted HF solutions without any modification of the SIPOS film.

We have also demonstrated that the nucleation of Si nanocrystals induced in SIPOS by a high temperature anneal is related to changes in the $Si-Si_xO_{4-x}$ tetrahedra distribution. Indeed, the effect of the anneal mainly consists in the decomposition of partially oxidized tetrahedra in $Si-Si_4$ and $Si-O_4$ tetrahedra. The $Si-Si_4$ tetrahedra form nanocrystalline Si grains, while the $Si-O_4$ tetrahedra enrich the oxygen concentration of the amorphous phase. We have shown that a high concentration of $Si-O_4$ tetrahedra makes annealed SIPOS films unstable towards the HF solution used to prepare a clean surface before XPS measurements. The high temperature induced formation of $Si-Si_4$ and $Si-O_4$ tetrahedra further makes more distinct the compositions of the two studied SIPOS films than those predicted by the RBM model.

ACKNOWLEDGMENTS

This work was partially supported by Consorzio per la Ricerca in Microelettronica nel Mezzogiorno (CORIMME), Catania, Italy. S. Pannitteri is gratefully acknowledged for the preparation of the samples for the TEM analysis.

[1]T. Matsushita, T. Aoki, T. Otsu, H. Yamoto, H. Hayashi, M. Okayama, and Y. Kawana, Jpn. J. Appl. Phys. Suppl. **15**, 35 (1976).
[2]H. Mochizuki, T. Aoki, H. Yamoto, M. Okayama, M. Abe, and T. Ando, Jpn. J. Appl. Phys. Suppl. **15**, 41 (1976).
[3]T. Matsushita, H. Hayashi, N. Oh-Uchi, and H. Yamoto, Jpn. J. Appl. Phys. Suppl. **20**, 75 (1980).
[4]T. Matsushita, N. Oh-Uchi, H. Hayashi, and H. Yamoto, Appl. Phys. Lett. **35**, 549 (1979).
[5]E. Yablonovitch, T. Gmitter, R. M. Swanson, and Y. H. Kwark, Appl. Phys. Lett. **47**, 1211 (1985).
[6]R. Hezel and K. Jaeger, J. Electrochem. Soc. **136**, 518 (1989).
[7]S. Lombardo, S. U. Campisano, G. N. van den Hoven, A. Cacciato, and A. Polman, Appl. Phys. Lett. **63**, 1942 (1993).
[8]M. Catalano, M. J. Kim, R. W. Carpenter, K. Das Chowdhury, and J. Wong, J. Mater. Res. **8**, 2893 (1993).
[9]J. Wong, D. A. Jefferson, T. G. Sparrow, J. M. Thomas, R. H. Milne, A. Howie, and E. F. Koch, Appl. Phys. Lett. **48**, 65 (1986).
[10]S. Lombardo, S. U. Campisano, and F. Baroetto, Phys. Rev. B **47**, 13561 (1993).
[11]H. R. Philipp, J. Non-Cryst. Solids **8–10**, 627 (1972).
[12]R. J. Temkin, J. Non-Cryst. Solids **17**, 215 (1975).
[13]J. H. Scofield, J. Electron Spectrosc. Relat. Phenom. **8**, 129 (1976).
[14]J. H. Thomas III and A. M. Goodman, J. Electrochem. Soc. **126**, 1766 (1979).
[15]S. I. Raider and R. Flitsch, IBM J. Res. Dev. **22**, 294 (1975).
[16]S. Tanuma, C. J. Powell, and D. R. Penn, Surf. Interface Anal. **20**, 77 (1993).
[17]P. Brüesch, Th. Stockmeier, F. Stucki, and P. A. Buffat, J. Appl. Phys. **73**, 7677 (1993).
[18]E. A. Irene, N. J. Chou, D. W. Dong, and E. Tierney, J. Electrochem. Soc. **127**, 2518 (1980).
[19]M. Hamasaki, T. Adachi, S. Wakayama, and K. Kikuchi, J. Appl. Phys. **49**, 3987 (1978).
[20]K. M. Brunson, D. Sands, C. B. Thomas, C. Jeynes, and J. F. Watts, Philos. Mag. B **61**, 361 (1990).
[21]I. Bertoti, R. Kelly, M. Mohai, and A. Toth, Surf. Interface Anal. **19**, 291 (1992).
[22]T. P. Nguyen and S. Lefrant, J. Phys. Condensed Matter **1**, 5197 (1989).
[23]R. Alfonsetti, L. Lozzi, M. Passacantando, P. Picozzi, and S. Santucci, Appl. Surf. Sci. **70/71**, 222 (1993).
[24]J. Finster, D. Schulze, and A. Meisel, Surf. Sci. **162**, 671 (1985).
[25]P. J. Grunthaner, M. H. Hecht, F. J. Grunthaner, and N. M. Johnson, J. Appl. Phys. **61**, 629 (1986).
[26]F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B **38**, 6084 (1988).
[27]T. Hattori, J. Vac. Sci. Technol. B **11**, 1528 (1993).
[28]W. R. Knolle and H. R. Maxwell, Jr., J. Electrochem. Soc. **127**, 2254 (1980).

Inherent possibilities and restrictions of plasma immersion ion implantation systems

S. Mändl,[a] J. Brutscher, R. Günzel, and W. Möller
Research Center Rossendorf, Inc., Institute for Ion Beam Physics and Materials Research, P.O. Box 51 01 19, D-01314 Dresden, Germany

(Received 19 June 1995; accepted 1 May 1996)

Design specifications for plasma immersion ion implantation systems are discussed. An analytical model of the plasma sheath evolution is used to deduce restrictions for the dimension of the vacuum chamber and the pressure range. The avoidance of arcing implies an upper limit for the maximum electric field in the sheath and hence for the plasma density. The total fluence per pulse, which determines the implantation duration and the power requirement, is calculated. Furthermore, safety considerations, especially the shielding of x-rays generated by secondary electrons, are discussed.
© *1996 American Vacuum Society.*

I. INTRODUCTION

Ion implantation is a versatile conventional method for injecting atoms into a substrate, thereby changing the composition and the properties of the surface. It is widely used in semiconductor device manufacturing. Additionally, implantation into metals, ceramics and polymers is an emerging technology by which new surface alloys are created to improve the tribological properties.[1] However, conventional ion implantation has its limitations: low beam currents lead to long implantation times; beam scanning and target manipulation are necessary for getting uniform implantation profiles in non-flat items; and ion sources are rather complex.

Plasma immersion ion implantation (PIII) has been developed as a technique without these limitations to allow for cost-effective implantation.[2] By applying negative high voltages pulses to the substrate, which is immersed in a plasma, the positive ions are extracted from the plasma and accelerated to the surface. PIII processes have been demonstrated for metallurgical[3,4] and semiconductor applications.[5,6] Recently, cost estimates for commercial PIII were published.[7]

For reaching widespread attention and industrial significance, design parameters for PIII systems must be specified. In this article, specifications for the vacuum system, different possibilities of plasma generation, high voltage pulse generation, and safety aspects in regard to commercial applications are discussed. In Sec. II, an analytical model[8] is used for approximating the evolution of the positively charged ion sheath and the ion flux during the high voltage pulse. The range for different system parameters and subsequent design restrictions is explored, and recommendations for PIII system specifications are given in Sec. III. Finally, a summary is given in Sec. IV.

II. ANALYTICAL MODEL

A high voltage pulse of amplitude $-V_o$ and duration t_p is applied to a target in a homogeneous plasma of density n_o. The square pulse is repeated after an off time t_{off}, corresponding to a repetition frequency $f_r=1/(t_p+t_{\text{off}})$ and a duty cycle $R=t_p/(t_p+t_{\text{off}})$. Fur sufficiently low pressure the mean free path is larger than the sheath width. Hence collisions in the plasma sheath can be neglected. The rise time t_r of the pulse is assumed to be shorter than the inverse ion plasma frequency $1/\omega_{pi}$ and long compared to the inverse electron plasma frequency $1/\omega_{pe}$. Therefore the electrons can be considered inertialess and neglected. An exact calculation should be performed in spherical coordinates.[9] Albeit, for our present purposes a planar calculation is sufficient.

A. Sheath evolution

Initially, for $t<1/\omega_{pi}$ a positively charged ion sheath forms. Its width x^{ini} can be calculated from Poisson's equation to be

$$x^{\text{ini}}=\sqrt{\frac{2\varepsilon_o V_o}{en_o}}, \tag{1}$$

where ϵ_o is the free-space permittivity and e is the ion charge. After the formation of the initial sheath, a quasi-static Child-Langmuir sheath is formed on a timescale of $1/\omega_{pi}$. The ions with the mass M are accelerated in the pre-sheath to the Bohm (ion sound) speed $u_B=(kT_e/M)^{1/2}$, where T_e is the electron temperature and k the Boltzmann constant. For this acceleration, a potential drop of $kT_e/2$ in the pre-sheath has to occur, corresponding to a decrease in the ion density $n_s=n_o e^{-1/2}\approx 0.6n_o$. The total ion current density J_c is supplied by the ion drift current and by the uncovering of ions due to the moving sheath edge[8]

$$j_c=0.6en_o\left(u_B+\frac{dx}{dt}\right), \tag{2}$$

which across a sheath of the thickness x is equal to[10]

$$j_c=\frac{4}{9}\varepsilon_o\left(\frac{2e}{M}\right)^{1/2}\frac{V_o^{3/2}}{x^2}. \tag{3}$$

Therefore, the evolution of the sheath edge is characterized by the following equation:

[a]Electronic mail: maendl@fz-rossendorf.de

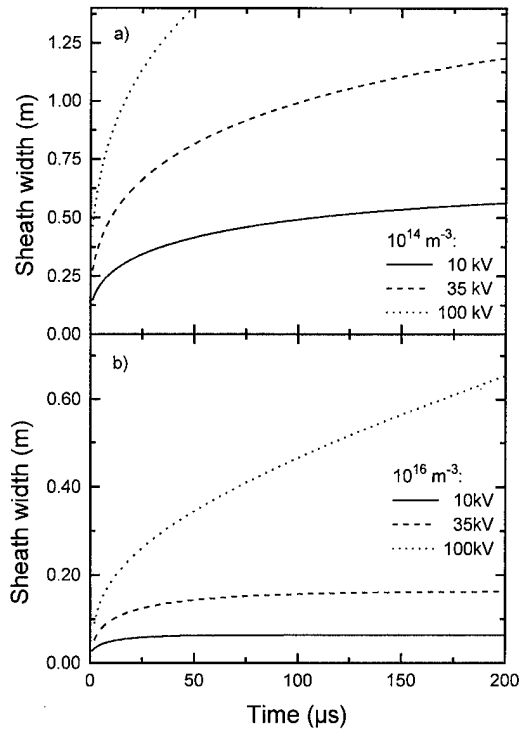


FIG. 1. Sheath edge evolution in planar geometry for different values of plasma parameters: N_2^+ plasma, T_e=2 eV, V_o=10, 35, 100 kV; (a) $n_o=10^{14}$ m^{-3} and (b) $n_0=10^{16}$ m^{-3}.

$$\frac{dx}{dt}=\frac{4\varepsilon_o V_o^{3/2}}{0.6\times 9x^2 n_o}\sqrt{\frac{2}{eM}}-u_B. \tag{4}$$

The solution of this differential equation yields the sheath evolution for a square pulse and different plasma densities, which is shown in Fig. 1. The sheath is expanding monotonically during the pulse, and the maximum sheath width $x^{\max}$, corresponding to a stationary sheath edge at $t_p\rightarrow\infty$ is

$$x^{\max}\approx\frac{V_o^{3/4}}{(ekT_e)^{1/4}}\sqrt{\frac{\varepsilon_o}{n_o}}. \tag{5}$$

The sheath width can reach several meters for low plasma densities of 10^{14} m^{-3} and voltages higher than 35 kV. However, for a density of about 10^{16} m^{-3}, the sheath can be as small as 15 cm at 35 kV. It can expand for 200 μs and longer in the former case, whereas the stationary width is reached in less than 50 μs in the later case.

B. Ion fluence

The ion fluence density F per pulse of length t_p is composed of the ions uncovered by the receding sheath edge and the ions crossing the sheath edge with the Bohm speed[8]:

$$F=0.6n_o\left(x(t_p)+t_p\sqrt{\frac{kT_e}{M}}\right). \tag{6}$$

The ions impinging from the Child-Langmuir sheath are implanted with the maximum energy $W_{\max}=eV_o$, whereas the ions from the initial matrix sheath as given by Equation (1) arrive at the surface of the target with an energy continuously distributed between $W_{\min}$=0 eV and $W_{\max}$.

III. DESIGN CONSIDERATIONS

For PIII applications, the control of the energy of the implanted ions is desired, i.e., monoenergetic ions should reach the target during the application of the pulse. Therefore the collision rate during the transit through the sheath should be negligible, and the fraction of ions implanted from the initial matrix sheath small compared to the total fluence.

A. Vacuum equipment and plasma generation

The pressure range that can be used by commercial PIII is limited by two factors: In batch processing, the time needed for pumping down the system to the base pressure must be comparable to the total implantation time.[7] The working pressure in the plasma source must be higher than the base pressure by some orders of magnitude to reduce the implantation of residual impurity ions during PIII.

The second factor is the requirement of a collisionless flight through the sheath, which is only possible if the mean free path is larger than the sheath width. It has been shown that collisions are important for sheath width greater than 0.5 times the mean free path Λ,[11] which is related to the pressure p by the expression $p\times\Lambda$=constant. The constant is 6.1×10^{-3} Pa·m for nitrogen, and of the same order of magnitude for other gases.[12] For small targets, the planar approximation is no longer valid and the sheath widths should be calculated within spherical coordinates, leading to a decrease by a factor of 2–4 for pulses longer than some tens of microseconds. In PIII applications, the lower limit of the mean free path for short pulses and long pulses is the initial sheath width x^{ini} and the steady state sheath width $x^{\max}$, respectively. This condition provides an upper limit for the working pressure.

In Fig. 2(a) the initial x^{ini} and steady state $x^{\max}$ sheath widths calculated from Equations (1) and (5) are plotted as functions of ion density for three different pulse voltages V_o. It can be seen that the sheath widths increase for higher voltages and during the voltage pulse. A lower limit for the sheath width, corresponding to an upper limit for the plasma density, is given by the maximum electrical field in the initial sheath for which no dielectric breakdown with subsequent arcing can occur. For pressures lower than 1 Pa, this field strength is approximately 10 kV/mm,[13] depending among others on the specific geometry, i.e., field enhancement at edges or protrusions from the workpiece, and surface properties. This limiting field strength occurs at the beginning of the pulse because the field decreases during the extension of the sheath. For an initial matrix sheath the maximum field $E^{\max}$ can be calculated to be

$$E^{\max}=\left|\left(\frac{dV}{dx}\right)_{x=0}\right|=\sqrt{\frac{2V_o e n_o}{\varepsilon_o}}. \tag{7}$$

The widths and densities for an electric field of 10 kV/mm are indicated with crosses for the different voltages V_o. This

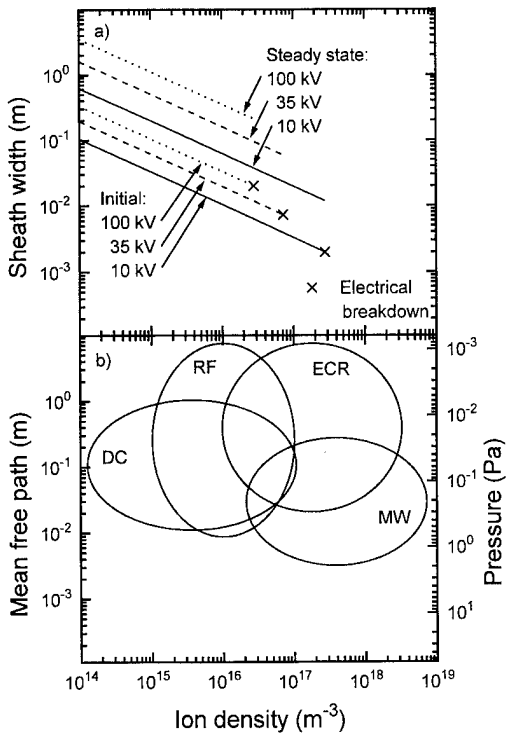


FIG. 2. (a) Sheath width depending on ion density in the initial and steady state case [see Equations (1) and (5)] for three different pulse voltages (10, 35, and 100 kV). The crosses indicate the sheath widths where arcing can occur [Equation (7)]. (b) Parameter range of ion density and gas pressure of different plasma sources (dc filament discharge,[14] RF radio-frequency discharge,[14] ECR electron cyclotron resonance,[14] and MW microwave source[15]).

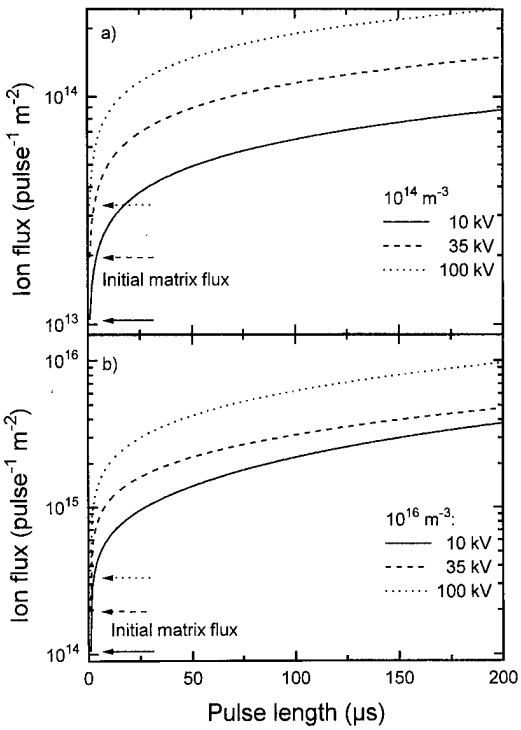


FIG. 3. Ion fluence as function of pulse length calculated from Equation (6) for different plasma parameter: N_2^+ plasma, T_e=2 eV, V_o=10, 35, 100 kV; (a) $n_o=10^{14}$ m^{-3} and (b) $n_o=10^{16}$ m^{-3}. The arrows indicate the ion flux from the initial matrix sheath with a continuously distributed energy.

limitation can be partly alleviated with longer rise times of the pulse or smooth workpieces. The minimum sheath width in the order of some millimeters to decimeters, depending on the voltage, is also a lower limit for the radius of holes and trenches which can be homogeneously implanted. The implantation of a smaller opening is accompanied by the displacement of the plasma out of it, leading to a simultaneous decrease of the flux into it.

At the same time, the dimension of the plasma chamber must be larger than the maximum sheath width $x(t_p)$ for the applied pulse length t_p and the size of the workpiece taken together. Otherwise the implantation of ions would be limited by the depletion and the subsequent extinction of the plasma. On the other hand, a plasma chamber that is too large wastes electrical power in plasma generation, so the vacuum chamber should be as small as possible. When implanting several workpieces in one run, care must be taken that the distance between the targets is clearly larger than twice the maximum sheath width. A treatment chamber with a diameter of several meters implies a mean free path of the same order of magnitude. This results in a operating pressure of the plasma source of 10^{-3} Pa, or less.

In Figure 2(b) typical pressure and ion density ranges for different plasma sources are shown.[14,15] A suitable plasma source for a given application is found by the following procedure. Proceeding from the required magnitude of the high voltage, the possible ranges for the sheath extension and plasma density are determined in Fig. 2(a). Then a plasma source can be selected from Fig. 2(b) using the restraint of a sheath extension smaller than the mean free path λ. It can be seen that microwave and dc discharges cannot be used for applications requiring an implantation energy of more than 40 keV. Using hot filaments can also result in additional contamination of the plasma and therewith the implanted workpiece.

B. High voltage equipment and pulse generation

Square pulse generators are widely used and available,[16–18] so that the model presented in Section II has practical relevance. However, it is only valid for short rise times. The energy distribution, which is deduced from this model, presents a reasonable approximation for the fraction of ions that are implanted with the energy eV_o for pulse rise times up to a few microseconds.[16] A model for finite pulse rise times t_r in the order of or larger than $1/\omega_{pi}$ is available.[19] In that case, the fraction of ions with an energy lower than eV_o is increased, so that the model used in this paper yields an upper limit for the fraction of ions impinging with the maximum energy.

In Fig. 3 the ion fluence depending on the pulse length is plotted for the plasma same parameters as used in Fig. 1. In the beginning, the ions uncovered by the moving sheath edge and the ions traversing the sheath contribute to the flux. For longer pulses, the sheath becomes stationary and the flux

is due to the drift current, linearly increasing with time. It can be seen that the fraction of ions with a kinetic energy below eV_o—the arrows in Fig. 3—decreases with increasing ion density n_o corresponding to a decreasing initial sheath width x^{ini}. For an ion density of 10^{16} m^{-3} pulse lengths of 30 μs are sufficient to reduce the ion fraction from the initial sheath below 10% of the total fluence. When the density decreases by a factor of 100, the length must be 200 μs or greater to achieve the same effect.

The total current, which must be supplied by the pulse generator at the constant voltage V_o is composed of the following parts: the ion current impinging on the target [Equation (3)] times the area of the workpiece; the displacement current due to the capacitance of the system; and the electron current generated at the surface by the ions with a secondary electron coefficient of 5 to 10.[20] This current can reach tens of amperes for large workpieces and determines the dimensions of the high voltage feed. The time integrated current and the implantation voltage V_o yield the energy which must be stored and recharged during the pulses.

The depleted plasma near the workpiece is subsequently refilled between the voltage pulses by ambipolar diffusion of ions and electrons from regions of higher density. The time evolution of the density can be calculated to be[21]

$$n(t)=n_o\cdot\left[1-\left(1-\frac{n_d}{n_o}\right)e^{-\gamma t}\right], \tag{8}$$

with a time constant

$$\gamma=\frac{1}{x(t_p)}\frac{T_e}{T_i}\left(\frac{kT_i}{2\pi M}\right)^{1/2}. \tag{9}$$

n_d is the density at the end of the pulse of width t_p. For sheath widths of 0.5 m or larger, the time required to fill in the depleted region up to 95% of the undisturbed plasma density $t_{95\%}=3/\gamma$ can be 0.5 ms or longer. At high repetition rates the initial density of the ion matrix in the beginning of each voltage pulse will decrease with each pulse. An equilibrium will be reached in which the sheath width during the pulse will be larger than predicted in Section II A, corresponding to a lower ion flux per pulse and eventually the occurrence of collisions during the flight through the sheath. Therefore, the repetition rate should be of the order of 1 kHz or less.

In a pulse with a duration of 100 μs, the fluence density varies between 5×10^{13} and 10^{16} ions per m^2, for plasma densities between 10^{14} and 10^{16} m^{-3} and voltages between 10 and 100 kV. An implantation dose of 10^{21} m^{-2} is reached after 10^5 to 2×10^7 pulses, corresponding to 100 s to 5 1/2 h at a repetition rate of 1 kHz. The energy deposited by the implanted ions is 100 mJ–200 J per pulse and per m^2 for voltages between 10 and 100 kV. Due to the higher current in the beginning of a pulse, average power densities of 500 W/cm^2 can be reached in short pulses. For commercial applications, a substrate area of 1 m^2 is reasonable. Therefore, the average power that must be dissipated by cooling—unless implantation at elevated temperatures is desired—can reach a few kilowatts.

The increase in the target temperature during a pulse can be estimated. Typical ion ranges are in the order of 100 nm. However, the ion energy during a pulse of length $_p$ is dissipated over a conduction length[22]

$$l=2\sqrt{\kappa t_p}=2\sqrt{\frac{\lambda t_p}{\rho c_p}}, \tag{10}$$

where κ, λ, ρ, and c_p are, respectively, the thermal diffusivity, the thermal conductivity, the density, and the specific heat capacity under constant pressure. This equation yields conduction lengths of some micrometers for pulse lengths of some tens of microseconds. Therefore, the heating of the workpiece can be approximated by a semi-infinite plane with a heat flux F_o impinging on the surface during the time t_p. The solution of the thermal conduction equation, neglecting loss due to radiation or conduction, yields the temperature evolution[22]:

$$T(x,t)=\frac{2F_o}{\lambda}\left[\sqrt{\frac{\kappa t_p}{\pi}}e^{-x^2/4\kappa t_p}\frac{x}{2}\mathrm{erfc}\left(\frac{x}{2\sqrt{kt_p}}\right)\right], \tag{11}$$

$$T(0,t)=\frac{2F_o}{\lambda}\sqrt{\frac{\kappa t_p}{\pi}}, \tag{12}$$

with erfc denoting the complementary error function. A flux $F_o=500$ W/cm^2 with a pulse length $t_p=100$ μs yields a temperature increase at the surface of 5.4 K per pulse for a 304 stainless steel target [$\kappa=0.081$ cm^2/s, $\lambda=0.30$ W/cm^2 (Ref. 23)], which is not enough for partially melting the target during a pulse.

For determining the equilibrium temperature during an implantation, the total heat balance must be obtained, including all fluxes to and from the workpiece:

$$\rho c_p V\frac{dT}{dt}=\dot{Q}=(A_hF_h-A_rF_r-\dot{Q}_c), \tag{13}$$

where V is the target volume, A_h the heated area, F_h the incident heat flux (including ion flux times duty cycle $F_o \times R$, plasma heating, radiation from hot filaments, microwave heating and x-ray absorption, where applicable), A_r the radiative area, F_r the radiative flux, and $\dot{Q}_c$ the loss due to conduction. The radiative heat flux is given by the Stefan–Boltzmann law:

$$F_r=\varepsilon\sigma(T^4-T_b^4), \tag{14}$$

with ε the emissivity, σ the Stefan–Boltzmann constant, and T_b the chamber temperature. Assuming an ion energy flux $F_o=500$ W/cm^2 and a duty cycle $R=10^{-3}$, which yields an average flux of 0.5 W/cm^2, and an additional heating from the plasma source of the same size, an incoming flux $F_h=1$ W/cm^2 is reasonable. With an emissivity $\varepsilon=0.4$, a radiative area twice the size of the heated area, i.e., both sides of the implanted sample, and a chamber at room temperature, the equilibrium temperature of the target is 405 °C in the worst case of no thermal contact between the workpiece and the holder. This temperature can be decreased by good thermal contact between the sample and the mounting, and by additional cooling. In the limiting case of an ideal

contact to a heat sink at 20 °C, the temperature at the surface of a 10 cm thick stainless steel target will be 53 °C.

For high voltage (75 kV and higher) and a given size of the workpiece, either the repetition frequency f_r and/or the pulse length t_p must be reduced. Otherwise the power consumption and the cooling system become prohibitively large and expensive. The pulse length and therewith the sheath extension is correlated with the size of the vacuum chamber, and the pulse length is correlated with the energy distribution of the implanted ions. A compromise must be found between a small system and a monoenergetic implantation. Nevertheless, implantation times of 20 min for doses of 10^{21} m^{-2} with 100 keV for areas larger than 0.1 m^2—three 8-in. wafers—should be possible.

Up to now, it was implicitly assumed that the applied voltage is present at the surface of the target, corresponding to a very low specific resistivity of the substrate. However, it is possible that isolating substrates, or conducting substrates into which an isolating layer is deposited during implantation, are used. A model was developed by Emmert[24] for this case. The applied high voltage drops across the insulator and the plasma sheath, depending on the respective impedances. During the pulse, the target becomes more positively charged due to the ion current, amplified by the emission of secondary electrons. Hence, the average ion energy will decrease during the pulse. Between the pulses, the positive surface attracts electrons, so that charge neutrality on the surface is rapidly restored. The energy distribution is broadened and shifted to lower energies, depending on the pulse length and the resistivity of the substrate. Care must be taken for inhomogeneous substrates and surfaces where insulator-metal junctions are exposed to the plasma. Due to the high electric field arcing is possible and must be avoided.

C. Safety considerations

In a PIII system, there are three inherent sources of hazards that can lead to serious injuries or death. All possible precautions must be taken to protect the operators of commercial systems.

First, there is the high voltage generator that supplies the current for the pulse generator. As shown above, the average current can be as high as 1 A at 100 kV with a peak current of up to 50 A. It is imperative that all high voltage components of the system are screened to prevent incidental touching.

A second health hazard is the intensive bremsstrahlung. It is caused by secondary electrons emitted from the substrate which are accelerated through the plasma to the walls of vacuum chamber. Therefore, all windows of the chamber should be made of lead-glass. For voltages higher than 50 kV, additional lead shielding is necessary, because the steel walls normally used for vacuum chambers become partially transparent for x-rays at this energy.

Using the following assumptions, the thickness of the shielding can be estimated: secondary electrons with a kinetic energy eV_o are generated with an average power density of 5 W/cm^2, corresponding to an ion flux $F_o=500$ W/cm^2, a duty cycle $R=10^{-3}$, and a secondary electron coefficient of 10. For a distribution of the electrons to all chamber walls and an isotropic x-ray emission the power density and hence the dose will be smaller at the surface of the chamber and decreasing with the distance from it. The generated bremsstrahlung spectrum $I_\gamma(E)$ can be approximated by the following formula[25]:

$$I_r(E)\propto\frac{I_eZ}{e^2}(eV_o-E), \tag{15}$$

where I_e is the secondary electron current and Z the atomic number of the decelerating material. This spectral distribution of x-rays is attenuated during the transit through the shielding and the maximum intensity is shifted to higher energies due to the stronger absorption for lower energies.[26] The shielding thickness in Fig. 4 was calculated as a function of the applied high voltage V_o to yield a radiation dose of 7 μS/h, which is, for instance, below the maximum level of 7.5 μS/h allowed in Germany for systems that do not request special authorization.[27] As can be seen, a reduction of the intensity by a factor of 10 will only reduce the required shielding by 33%, whereas a decrease of the applied voltage by 50% will reduce the thickness by a factor of 8. Therefore, the thickness is mostly determined by the applied voltage and not by the consumed power. Above the Pb K-edge at an energy of 87.9 keV, the rapidly rising absorption coefficient results in a shielding thickness nearly independent of the voltage up to 150 kV.

The third source of danger is the microwave-source that can be used as a plasma generator in PIII systems. Microwaves endanger health, especially at the power densities used in plasma sources. All non-metallic parts of the system should therefore be covered with a metallic mesh to absorb the emitted radiation.

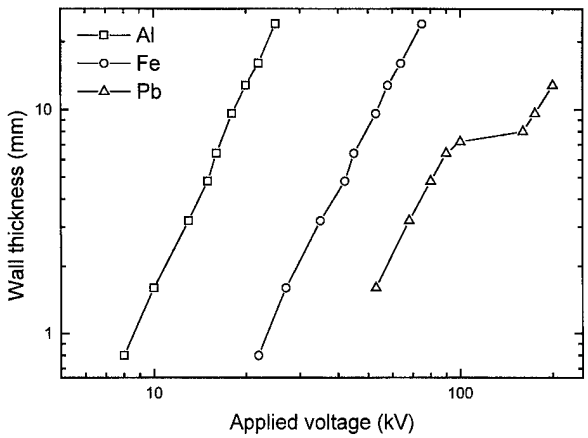


FIG. 4. Wall thickness necessary to reduce the radiation dose at the outer surface of the chamber to 7 μS h^{-1} for Al, Fe, and Pb coatings. The rising absorption above the Pb K-edge is responsible for the slight increase in the thickness for Pb shielding between 90 and 150 kV.

IV. SUMMARY

PIII is a method for fast and cost effective implantation of large non-planar areas. Essential requirements for monoenergetic implantations are a collisionless transit through the positively charged ion sheath and a square voltage pulse. A mean free path larger than twice the sheath width limits the available pressure range in the vacuum system. To avoid arcing, the ion density must be below 4×10^{16} m^{-3} for implantation energies of 100 keV. At the same energy, a dose of 10^{21} m^{-2} in areas of 0.1 m^2 can be reached in 20 min. The repetition rate and therefore the duration of one implantation run is restricted by the effectivity of the cooling system. Additional lead shielding is necessary for ion energies above 50 keV to screen the x-ray radiation emitted by the generated secondary electrons.

ACKNOWLEDGMENT

This research is supported by Bundesministerium für Forschung und Technologie Research Contract No. 13N6443.

[1]K. G. Budinski, *Surface Engineering for Wear Resistance* (Prentice-Hall, Englewood Cliffs, 1988), p. 163.
[2]J. R. Conrad, J. L. Radtke, R. A. Dodd, and F. J. Worzala, J. Appl. Phys. **62**, 4591 (1987).
[3]R. Günzel, E. Wieser, E. Richter, and J. Steffen, J. Vac. Sci. Technol. B **12**, 927 (1994).
[4]M. Samandi, B. A. Shedden, T. Bell, G.A. Collins, R. Hutchings, and J. Tendys, J. Vac. Sci. Technol. B **12**, 935 (1994).
[5]N. W. Cheung, W. En, E. C. Jones, and C. Yu, Mater. Res. Soc. Symp. Proc. **279**, 297 (1993).
[6]E. C. Jones and N. W. Cheung, IEEE Electron Device Lett. **14**, 444 (1993).
[7]D. J. Rej and R. B. Alexander, J. Vac. Sci. Technol. B **12**, 2380 (1994).
[8]W. En and N. W. Cheung, J. Vac. Sci. Technol. B **12**, 833 (1994).
[9]J. R. Conrad, J. Appl. Phys. **62**, 777 (1987).
[10]C. D. Child, Phys. Rev. **32**, 492 (1911).
[11]T. E. Sheridan and J. Goree, Phys. Fluids B **3**, 2796 (1991).
[12]G. Janzen, *Plasmatechnik* (Hüthig, Heidelberg, 1992), p. 50.
[13]J. Brutscher, Ph.D. thesis, Johann Wolfgang Goethe-Universität Frankfurt/M., 1993.
[14]G. D. Alton, Nucl. Instrum. Methods B **73**, 221 (1993).
[15]L. A. Berry and S. M. Gorbatkin, Nucl. Instrum. Methods B **56/57**, 1133 (1991).
[16]J. J. Vajo, J. D. Williams, R. Wei, R. G. Wilson, and J. N. Matossian, J. Appl. Phys. **76**, 5666 (1994).
[17]B. P. Wood, I. Henins, R. J. Gribble, W. A. Reass, R. J. Faehl, M. A. Nastasi, and D. R. Rej, J. Vac. Sci. Technol. B **12**, 870 (1994).
[18]M. J. Goeckner, S. M. Malik, J. R. Conrad, and R. A. Breun, Phys. Plasmas **1**, 1064 (1994).
[19]R. A. Stewart and M. A. Lieberman, J. Appl. Phys. **70**, 3481 (1991).
[20]M. Shamim, J. T. Scheuer, R. P. Fetherston, and J. R. Conrad, J. Appl. Phys. **70**, 4756 (1991).
[21]B. P. Wood, J. Appl. Phys. **73**, 4770 (1993).
[22]M. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon, Oxford, 1959), p. 75.
[23]J. P. Blanchard, J. Vac. Sci. Technol. B **12**, 910 (1994).
[24]G. A. Emmert, J. Vac. Sci. Technol. B **12**, 880 (1994).
[25]H. Kuhlenkampf and L. Schmidt, Ann. Phys. (Leipzig) **43**, 494 (1943).
[26]J. Brutscher, Master thesis, Johann Wolfgang Goethe-Universität Frankfurt/M., 1988.
[27]Verordnung über den Schutz vor Schäden durch Röntgenstrahlung (RöV), Bundesgesetzblatt, 1987.

X-ray photoelectron spectroscopy study on native oxidation of As-implanted Si (100)

Fumiko Yano,[a] Akiko Hiraoka, Toshihiko Itoga, Atsuko Matsubara, Hisao Kojima, Keiichi Kanehori, and Yasuhiro Mitsui[b]
Central Research Laboratory, Hitachi, Ltd., Kokubunji, Tokyo 185, Japan

(Received November 11 1994; accepted April 11 1996)

The native oxidation of As-implanted Si surfaces used in actual ultralarge scale integrated processes are investigated. Quantitative analysis of oxidation is made possible by x-ray photoelectron spectroscopy spectral decomposition of Si $2p$ into Si^{4+}, Si^{x+}, and Si^{0+}, and by calculation of SiO_2 and SiO_x thicknesses using the decomposition results. Here, the sensitivity is such that less than 1 Å change is detectable. The native oxidation of As-implanted (1×10^{15}–1×10^{16}/cm^2, at 25 kV) Si(100) is compared to that of Si(100) without implantation. The results show that the oxidation rate of As-implanted Si is faster than that of Si without implantation, that the native oxide on As-implanted Si includes more suboxide than that on Si without implantation, and that As is oxidized in deeper regions than Si. These results indicate that As implantation changes the Si native oxidation mechanism itself. We propose an oxidation model of As-implanted Si to explain our observations, and discuss the validity of this model. © *1996 American Vacuum Society.*

I. INTRODUCTION

Native oxidation on Si surfaces causes problems in ultralarge scale integrated (ULSI) processes and properties including the metal/Si ohmic contact, low-temperature epitaxial growth of silicide, and dielectric breakdown of thin SiO_2. HF treatment is a very effective way to suppress the native oxidation of Si; it removes surface oxide, makes the Si surface flat, and forms a stable surface with hydrogen atoms terminating dangling bonds.[1–3] Therefore, HF-treated Si surfaces are widely investigated by x-ray photoelectron spectroscopy (XPS), Fourier-transform infrared spectroscopy (FTIR), and other analytical methods.[1–7] However, most of these studies are directed toward Si surfaces that are ideally flat, ideally crystallized, and without implanted impurities. There have been far fewer investigations on realistic surfaces formed in ULSI fabrication. These include ion-implanted Si surfaces and polycrystallized Si surfaces.

The present work investigates the native oxidation of Si(100) surfaces having implanted As ions (1×10^{15}–1×10^{16}/cm^2) in anticipation of a future necessity for controlling native oxidation during semiconductor device fabrication. The oxide thickness is quantitatively estimated by XPS with a sub-1 Å sensitivity.[7] Comparing the oxidation of As-implanted Si with that of Si without implantation demonstrates that the implanted As atoms change the oxidation mechanism of Si. This article proposes a model of the native oxidation mechanism of As-implanted Si and discusses the validity of this model.

II. EXPERIMENT

Table I shows the sample preparation parameters. As^+ ions were implanted into p-type 4° off Si (100) wafers at 25 keV (1×10^{15}–1×10^{16}/cm^2). These wafers were annealed at 850 °C for 30 min and cut into small pieces for convenience in XPS measurement. They were dipped in 5 mol % HF solution for 10 min, blown dry in N_2 gas, exposed in a clean-room atmosphere (40% relative humidity) for 0 to 48 h, and analyzed by XPS. For comparison, a Si wafer without implantation was prepared in the same way. Ar^+-sputtered clean Si and thermally oxidized Si were also analyzed as the reference.

The photoelectron spectra were measured with the Shimazu-KRATOS XSAM800 system, using a Mg $K\alpha$ x ray (12 kV, 10 mA) as the x-ray source. The energy resolution of the measured spectrum was 0.96 eV, which was determined as the FWHM (full width at half-maximum) of the Ag $3d$ 5/2 peak. The base pressure in the analysis chamber was below 7×10^{-8} Pa.

For quantitative analysis, measured Si $2p$ spectra were decomposed into Si^{0+} (Si $2p$ from Si), Si^{x+} ($x=1$, 2, and 3, Si $2p$ from Si bonding with x oxygen atoms and $4-x$ silicon atoms), and Si^{4+} (Si $2p$ from SiO_2). Then, the thickness of silicon dioxide (d_o) and that of silicon suboxide (d_x) are calculated.[7] The decomposition process and the calculation are described briefly.

The decomposition process is shown in Fig. 1. We subtracted Si^{0+} and Si^{4+} from the Si $2p$ spectra and the remainder was treated as the summation of all Si^{x+} components. For subtraction of Si^{4+}, the binding energy of Si^{4+} [$E_b(Si^{4+})$] was calculated under the assumption of a constant energy difference between the binding energies of O $1s$ [E_b(O $1s$)] and $E_b(Si^{4+})$:

$$E_b(Si^{4+}) = E_b(O\ 1s) - \Delta E_b(\mathrm{ref}), \quad (1)$$

$$\Delta E_b(\mathrm{ref}) = 429.25\ \ (\mathrm{eV}). \quad (2)$$

Here, ΔE_b(ref) denotes the energy difference between

[a]Electronic mail: fumiko@crl.hitachi.co.jp
[b]Present address: Semiconductor and Integrated Circuits Division, Hitachi, Ltd., Kodaira, Tokyo 187, Japan.

TABLE I. Sample preparation parameters.

	Wafer	As implantation	Annealing	Treatment	Exposure time in clean-room air (h)
	Si(100) *p* type	...	...		0–96
Measured samples	Si(100)	25 kV, 1×10^{15}/cm^2		10 min dipping in 5% HF solution	5
	p type 4° off	25 kV, 2×10^{15}/cm^2	850 °C, 30 min		0–24
		25 kV, 1×10^{16}/cm^2			5
Reference samples	Si(100) *p* type	... ...	... ...	Ar sputtering for Si^{0+} Thermal oxidation for Si^{4+}	... ...

$E_b(\mathrm{O}\,1s)$ and $E_b(\mathrm{Si}^{4+})$ and is experimentally determined from an XPS spectrum of a reference sample having a 100-nm-thick thermal oxide. The FWHMs and peak shapes of Si^{0+} and Si^{4+} were also experimentally determined from XPS spectra of reference samples, clean Si for Si^{0+} and 100-nm-thick thermal Si oxide for Si^{4+}. The FWHMs are 1.24 and 1.64 eV, respectively.

Layer thicknesses of the multilayered structure can be calculated when the sensitivity, the escape depth, and the XPS peak intensity for the typical element in each layer are given.[8,9] We assume the native oxide structure of $SiO_2/SiO_x/Si$. Then, thickness of SiO_2 (d_o) and thickness of SiO_x (d_x) are given in Eqs. (3) and (4), using the Si^{4+}, Si^{x+}, and Si^{0+} peak intensities (I_o, I_x, and I_s) and sensitivity factors Q (sensitivity of Si^{4+} to Si^{0+}) and Q' (Si^{x+} to Si^{0+}), which are defined in Eqs. (5) and (6):

$$d_x=\lambda_{\mathrm{Si}}^{\mathrm{SiO}_x}\sin\theta\ln\left(\frac{1}{Q'}\frac{I_x}{I_s}+1\right), \tag{3}$$

$$d_0=\lambda_{\mathrm{Si}}^{\mathrm{SiO}_2}\sin\theta\ln\left(\frac{1}{Q}\frac{I_0}{I_s/\exp(-d_x/\lambda_{\mathrm{Si}}^{\mathrm{SiO}_x}\sin\theta)}+1\right), \tag{4}$$

$$Q=\frac{R^{\mathrm{SiO}_2}f_{\mathrm{Si}}^{\mathrm{SiO}_2}n_{\mathrm{Si}}^{\mathrm{SiO}_2}\sigma_{\mathrm{Si}}^{\mathrm{SiO}_2}\lambda_{\mathrm{Si}}^{\mathrm{SiO}_2}}{R^{\mathrm{Si}}f_{\mathrm{Si}}^{\mathrm{Si}}n_{\mathrm{Si}}^{\mathrm{Si}}\sigma_{\mathrm{Si}}^{\mathrm{Si}}\lambda_{\mathrm{Si}}^{\mathrm{Si}}}, \tag{5}$$

$$Q'=\frac{R^{\mathrm{SiO}_x}f_{\mathrm{Si}}^{\mathrm{SiO}_x}n_{\mathrm{Si}}^{\mathrm{SiO}_x}\sigma_{\mathrm{Si}}^{\mathrm{SiO}_x}\lambda_{\mathrm{Si}}^{\mathrm{SiO}_x}}{R^{\mathrm{Si}}f_{\mathrm{Si}}^{\mathrm{Si}}n_{\mathrm{Si}}^{\mathrm{Si}}\sigma_{\mathrm{Si}}^{\mathrm{Si}}\lambda_{\mathrm{Si}}^{\mathrm{Si}}}. \tag{6}$$

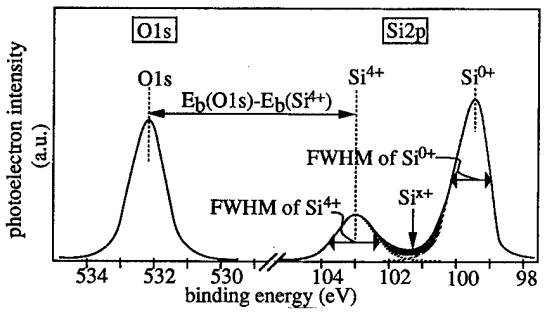


FIG. 1. Schematic diagram of the decomposition process.

Here, R is the surface roughness, f is the photoelectron transmission function, n is the density of Si atoms, σ is the atomic photoionization cross section, λ is the escape depth, and θ is the photoelectron take-off angle. Q is determined to be 0.67 based on XPS measurements of the reference samples, clean Si, and thermal Si oxide. Q' cannot be determined experimentally because we have no standard samples of SiO_x. Therefore, we use Q' of 0.84, which is interpolated between that of Si and SiO_2. This may cause error in calculated results; however, the error in the total thickness is very small because the amount of SiO_x is relatively small. We have also experimentally determined the Si escape depth λ_{Si} in SiO_2 (2.7 nm) using the XPS results and a cross-sectional TEM observation of thin thermal SiO_2 on Si.[10–12]

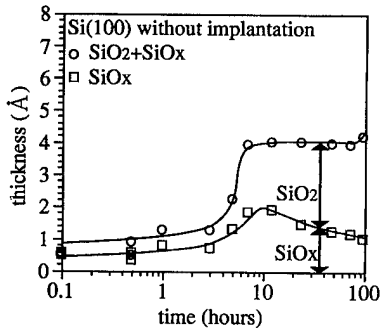
III. RESULTS

Figures 2(a) and 2(b) show the native oxidation of Si(100) without implantation and that of Si(100) with As-implantation, respectively. The open and filled circles denote the total oxide thicknesses (SiO_2+SiO_x) and the open and filled squares denote SiO_x thickness on the Si surface.

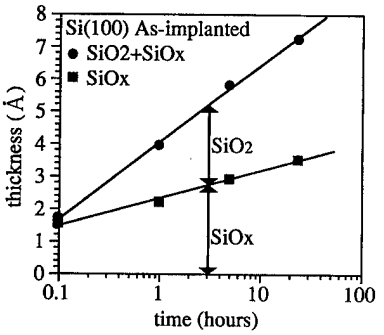
In Fig. 2(a), the total oxide thickness increases with the exposure time in the clean-room atmosphere. It increases slowly up to 3 h, increases abruptly after 3 h, and saturates after 10 h with oxide thickness of about 4 Å. The SiO_x thickness increases slightly up to 10 h but decreases after that. This decrease shows that some of the SiO_x turns into SiO_2 with increasing exposure time. Based on the amount of SiO_x, SiO_x is thought to be located only at the interface between Si and SiO_2.[7] These figures show layer-by-layer oxidation, a well known phenomenon.[13] The saturation after 10 h indicates a similar oxidation rate to that of the previous work[4] and it can be explained with the structural model of SiO_2/Si by Herman *et al.*[14] if Si atoms from two surface layers are oxidized.

In Fig. 2(b), the total oxide thickness increases with the exposure time. Here, there is no step, which indicates layer-by-layer oxidation. The SiO_x thickness also increases with the exposure time. This behavior is completely different from that seen for Si(100) without implantation.

Figure 3 shows the dependency of the total oxide thickness and the SiO_x thickness on the As dose. The



(a)



(b)

FIG. 2. Native-oxide thickness on HF-treated Si as a function of exposure time. (a) The results for Si (100) without implantation. ○ SiO_2+SiO_x thickness; □ SiO_x thickness, showing layer-by-layer oxidation. (b) The results for As-implanted Si (100) (2×10^{15}/cm^2). ● SiO_2+SiO_x thickness; ■ SiO_x thickness. This behavior is completely different from that seen in (a).

samples had been exposed in a clean-room atmosphere for 5 h. With an increasing As dose, in the range between 1×10^{15}–1×10^{16}/cm^2, both thicknesses increase similarly.

Summarizing these results, it becomes clear that the HF-treated, As-implanted Si is oxidized faster than the Si without implantation, and it does not exhibit layer-by-layer native oxidation. Furthermore, the native oxide on the As-implanted Si includes a thicker suboxide layer than the native oxide on Si without implantation and both the total-oxide thickness and the SiO_x thickness increase as the As dose increases.

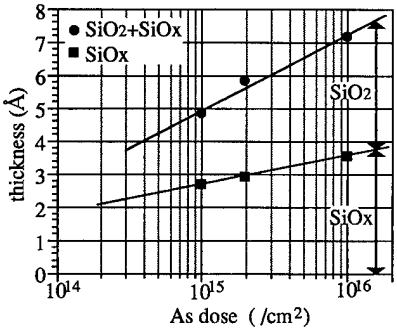


FIG. 3. Oxide thickness on Si (100) as a function of the As dose. The As-implanted Si was dipped in a 5 mol % HF solution and exposed in a clean-room atmosphere for 5 h. ● SiO_2+SiO_x thickness; ■ SiO_x thickness.

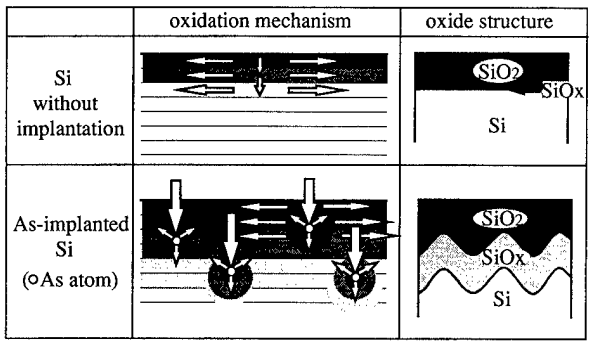


FIG. 4. Models of the native oxidation process on Si without implantation and on As-implanted Si. In the case of Si without implantation, Si atoms are oxidized layer by layer. The suboxide is located only at the interface between Si and SiO_2. In the case of As-implanted Si, the As atoms are preferentially oxidized first, then the Si atoms near As are oxidized. Because oxidation spreads from the oxidized As, more suboxide is created in the oxide layer.

IV. DISCUSSION

A. Model of the native oxidation of As-implanted Si(100)

Figure 2(a) shows that the total-oxide thickness and the SiO_x thickness of Si without implantation exposed for 5 hs in a clean-room atmosphere are 2.3 and 1.3 Å, respectively. Comparing these thicknesses and Fig. 3, it becomes clear that As-implantation affects the native oxidation of Si when the dose is 10^{13}/cm^2 or greater. It is well known that the surface of ion-implanted Si is rougher than that of Si without implantation. In addition, wafers misorientated 4° from (100), which were used for As-implanted samples, have more steps than perfectly oriented (100) wafers. The roughness and steps may accelerate oxidation, however, they are not supposed to change the mechanism of layer-by-layer oxidation. The change of the native oxidation mechanism of Si is caused by the high concentration of implanted As atoms. To explain the oxidation on As-implanted Si surfaces, we have proposed a model shown in Fig. 4.

In the case of Si without implantation, Si atoms are oxidized layer by layer. Although the microscopic mechanism of layer-by-layer oxidation is not firmly established, it is quite natural to assume that the vertical oxide-growth rate is much slower than the lateral oxide-growth rate in layer-by-layer oxidation process. As explained in Sec. III, SiO_x is regarded as being localized at the interface between Si and SiO_2, as depicted in the oxide structure shown in the upper right corner of Fig. 4.

On the other hand, layer-by-layer oxidation is blurred in the native oxidation of As-implanted Si. Also, it is natural to assume that this blurring is the result of the faster vertical oxide-growth rate than the lateral oxide-growth rate. From this point of view, we assumed a model of the native oxidation of As-implanted Si in which implanted As atoms work

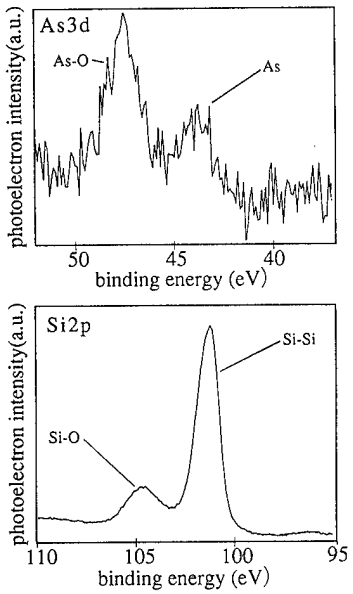


FIG. 5. As $3d$ and Si $2p$ spectra of As-implanted Si (2×10^{15}/cm^2), dipped in a 5 mol % HF solution and exposed in the clean-room atmosphere for 5 h.

as oxidation initiators. This model includes the following processes, i.e., As atoms are preferentially oxidized in Si. After As oxidation, the Si atoms near the As atoms are also easily oxidized. This is because Si bonds are known to be easily oxidized after one of the bonds combines with an oxygen atom. This makes the vertical oxide-growth rate faster and blurs the layer-by-layer oxidation. As a consequence of this model, more suboxide is produced around the oxidized As atoms, contributing to the interfacial region between SiO_2 and Si. The structure of this oxide is shown in the lower right corner of Fig. 4, in which there is a layer of suboxide between Si and SiO_2.

B. Examining the model

In order to verify this oxidation model, it is important to show that the As atoms are preferentially oxidized in As-implanted Si. For this purpose, As $3d$ and Si $2p$ XPS spectra are investigated in more detail.

Figure 5 shows the As $3d$ and Si $2p$ XPS spectra of the As-implanted Si (100) surface after 5 h of exposure in the clean-room atmosphere. There are two well-resolved peaks in both the As $3d$ and Si $2p$ spectra. In the As $3d$ region, the peak at the higher binding energy reflects oxidized As. The lower energy peak reflects unoxidized As. However, it is unclear whether this unoxidized As peak corresponds to As bonding with Si atoms or As binding with As atoms (As clusters in Si). In the Si $2p$ region, the higher energy peak corresponds to SiO_2 and the lower one corresponds to Si in the Si substrate. Roughly 65% of the As atoms and 30% of the Si atoms are oxidized in the region probed by XPS. The depth of As oxidation can be calculated when the As depth profile in Si and the escape depth of As $3d$ photoelectrons in Si and SiO_2 are known. In calculating this As oxidation depth, we make two assumptions.

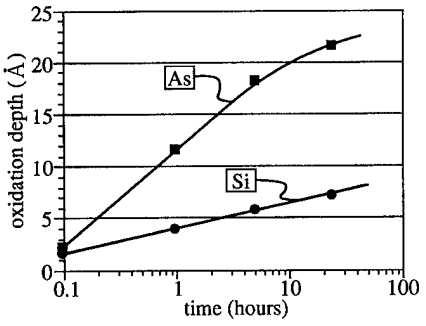


FIG. 6. Si native-oxide thickness on As-implanted Si (2×10^{15}/cm^2) and the oxidation depth of As. As is oxidized in a deeper region than Si, confirming that As is preferentially oxidized in Si.

First, we assume that the As concentration is constant in the depth probed by XPS (it is shallower than 50 Å when the take-off angle is 30°). This is based on the As depth profile measured by HG-ICPS (hydride generation inductively coupled plasma spectrometry).[15] In this profile, the topmost Si layer contains 1.1×10^{21}/cm^3 of As. However, this layer is removed during HF treatment. Then, the As concentration either remains constant or increases in the range between 3.5×10^{20}/cm^3 and 7.5×10^{20}/cm^3, with increasing depth in the depth sampled by XPS. So as not to overestimate the oxidation depth of As, we assume a constant concentration of As in the region probed by XPS.

Second, we assume that the escape depths of As $3d$ photoelectrons in Si and SiO_2 are approximately the same as those of Si $2p$. This assumption is based on the escape depth dependence on the kinetic energy of the photoelectron. The kinetic energy of As $3d$ is about 1210 eV and differs by only 5% from that of Si $2p$, about 1150 eV.

Based on these assumptions, the As oxidation depth is calculated in exactly the same way as the SiO_2 thickness is calculated.

Figure 6 shows the Si native oxide thickness (total thickness of SiO_2 and SiO_x) and the As oxidation depth in As-implanted Si (100) (2×10^{15}/cm^2) as a function of the exposure time in the clean-room atmosphere. As is oxidized in a deeper region than Si, showing that the As atoms are preferentially oxidized. At an exposure time of 24 h, As atoms located 15 Å below the Si/SiO_2 interface are oxidized. This is a reasonable result for this oxidation model because it is estimated that there is one As atom per a 15 Å cube when the As concentration is 5×10^{20}/cm^3.

There are several factors that could influence native oxidation of As-implanted Si; damage and surface roughness caused by the As-implantation process, a chemical effect of implanted As, or an electrical effect.[16] The experimental result of preferential As oxidation suggests that As-implantation affects the native oxidation of Si via its chemical effect.

V. CONCLUSION

The authors have investigated the native oxidation process of As-implanted (1×10^{15}–1×10^{16}/cm^2) Si(100) by

quantitative XPS analysis in which the thicknesses of SiO_2 and SiO_x are calculated with a sensitivity of about $\frac{1}{10}$ of a monolayer. The results show that the oxidation of As-implanted Si is faster than that of Si without implantation, that the native oxide on As-implanted Si contains more suboxide than that on Si without implantation, and that As is oxidized in a deeper region than Si is. This means high-concentration As implantation changes the native oxidation process of Si itself.

From these results, we developed an oxidation model for As-implanted Si in which As atoms are preferentially oxidized in Si, acting as oxidation initiators, followed by subsequent oxidation of the surrounding Si. The oxidation depth of As-implanted samples was determined, confirming the validity of this model.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Masanobu Miyao and Dr. Keizou Suzuki for their encouragement throughout this study.

[1]V. A. Burrows, Y. J. Chabel, G. S. Higashi, K. Raghavachari, and S. B. Christman, Appl. Phys. Lett. **53**, 998 (1988).
[2]T. Takahagi, I. Nagai, A. Ishitani, H. Kuroda, and Y. Nagasawa, J. Appl. Phys. **64**, 3516 (1988).
[3]G. W. Trucks, K, Raghavachari, G. S. Higashi, and Y. J. Chabel, Phys. Rev. Lett. **65**, 504 (1990).
[4]M. Morita, T. Ohmi, E. Hasegawa, M. Kawakami, and M. Ohwada, J. Appl. Phys. **68**, 1272 (1990).
[5]T. Hattori, Ext. Abstr. Int. Conf. Solid State Devices Mater. Yokohama **410** (1994).
[6]Y. J. Chabel, G. S. Higashi, K. Raghavachari, and V. A. Burrows, J. Vac. Sci. Technol. A **7**, 2104 (1989).
[7]F. Yano, A. Hiraoka, T. Itoga, H. Kojima, K. Kanehori, and Y. Mitsui, J. Vac. Sci. Technol. A **13**, 2671 (1995).
[8]A. Ishizaka and S. Iwata, Appl. Phys. Lett. **36**, 71 (1980).
[9]C. S. Fadley, J. Electron Spectrosc. **5**, 725 (1974).
[10]R. Flitsch and S. I. Raider, J. Vac. Sci. Technol. **12**, 305 (1975).
[11]F. Bechstedt and K. Hubner, Phys. Status Solidi A **67**, 517 (1981).
[12]S. Tanuma, C. J. Powell, and D. R. Penn, Surf. Interface Anal. **17**, 927 (1991).
[13]E. A. Taft, J. Electrochem. Soc. **135**, 1022 (1987).
[14]F. Herman, J. P. Batra, and V. Kasowski, *The Physics of SiO_2 and Its Interface*, edited by S. Pantelides (Pergamon, New York, 1978).
[15]A. Matsubara, H. Kojima, T. Itoga, and K. Kanehori, Jpn. J. Appl. Phys. **34**, 3965 (1995).
[16]I. W. Lyo, E. Kaxiras, and Ph. Avouris, Phys. Rev. Lett. **63**, 1261 (1989).

Quantification of metal trace contaminants on Si wafer surfaces by Laser-SNMS and TOF-SIMS using sputter deposited submonolayer standards

A. Schnieders,[a] R. Möllers, M. Terhorst,[b] H.-G. Cramer,[b] E. Niehuis, and A. Benninghoven
Physikalisches Institut der Universität Münster, Wilhelm-Klemm-Strasse 10, D-48149 Münster, Germany

(Received 10 July 1995; accepted 1 May 1996)

We have shown that quantitative determination of metal trace impurities in the range between 10^9 cm^{-2} and 10^{12} cm^{-2} on top of the native oxide of a Si wafer by laser postionization of sputtered neutrals in combination with time-of-flight mass spectrometry (Laser-SNMS) and time-of-flight secondary ion mass spectrometry (TOF-SIMS) is possible. Trace metal standards with concentrations in the range between 10^9 cm^{-2} and 10^{12} cm^{-2} were used for calibration. These standards were prepared by sputter deposition and were independently controlled by Total Reflection X-Ray Fluorescence Analysis. Relative sensitivity factors for 12 metals on Si for Laser-SNMS were determined. Additionally, we compared postionization with different wavelengths (193 nm and 248 nm). With Laser-SNMS it was then possible to determine the influence of UV/ozone treatment on the measured surface concentration of metal species. The UV/ozone treatment is necessary to achieve high and reproducible useful yields for the metal species in TOF-SIMS. With this knowledge, we were able to determine relative sensitivity factors for the metals on UV/ozone-treated Si wafer surfaces measured by TOF-SIMS. Detection limits down to 10^8 cm^{-2} and 10^{12} cm^{-2} for sample surface areas of 100 μm in diameter and 1 μm in diameter, respectively, were found for both Laser-SNMS and TOF-SIMS. © *1996 American Vacuum Society.*

I. INTRODUCTION

The detection and quantitative determination of metal trace contaminants is an important issue in wafer processing and integrated circuit (IC) fabrication. Detection limits well below 1 ppm in a single monolayer are required. From the results of shallow depth profiling of oxide layers on Si, it is known that these surface contaminants are preferentially located on top of the oxide layer, i.e., the uppermost monolayer.[1] Standard techniques for surface analysis like Auger electron spectroscopy and x-ray photoelectron spectroscopy cannot be applied because of their limitations in sensitivity. The same holds true for the electron microprobe. The standard technique for the quantitative determination of metals on wafer surfaces is total reflection x-ray fluorescence spectroscopy (TXRF),[2] widely applied in the semiconductor industry for research and development as well as production control. The most serious limitations of this technique are the detection limits (in the range of 10^{10} cm^{-2}), the lack of lateral resolution and a larger information depth (about 10 monolayers).

We have demonstrated the usefulness of TOF-SIMS for metal trace detection on GaAs, down to the ppm-range with the areas examined as small as 50 μm in diameter.[3] More recently these investigations were extended to Si wafer surfaces.[4–6] In this article we report on a combined Laser-SNMS/TOF-SIMS study of metal contaminants on Si wafer surfaces. For calibration, standards were produced by controlled sputter deposition of metal submonolayers on Si wafer surfaces under TXRF control. Quantitation was achieved for both Laser-SNMS and TOF-SIMS with comparable detection limits down to 10^8 cm^{-2} and 10^{10} cm^{-2} for sample surface areas of 100 μm in diameter and 1 μm in diameter, respectively.

II. PROCEDURE

In previous work[7,8] it has been shown that quantitative, high sensitivity bulk analysis of metal alloys is possible with Laser-SNMS by applying nonresonant multiphoton postionization. The important quantity used in this context is the relative sensitivity factor (RSF) S(A, B) of element A with respect to element B. It is defined by the following equation:

$$S(\mathrm{A,B}) = \frac{I(\mathrm{A})}{I(\mathrm{B})} \cdot \frac{c(\mathrm{B})}{c(\mathrm{A})}, \tag{1}$$

$I(\mathrm{X}) \triangleq$ detected photoion signal intensity of element X,

$c(\mathrm{X}) \triangleq$ bulk concentration of element X.

RSFs are determined by the sputtering and the ionization process. They are in particular influenced by:

(i) Preferential sputtering. This means that from a multicomponent sample, some species are dominantly sputtered, whereas other species are enriched on the surface. Only when sputter equilibrium is reached (after removing many, or, in some cases, several hundred monolayers) the flux of sputtered particles does represent the composition of the sample material.

(ii) Different angular and energy distributions for various sputtered species. These distributions determine the

[a]Electronic mail: schnial@uni-muenster.de
[b]Present address: Ion-TOF GmbH, Mendelstrasse 11, D-48149 Münster, Germany.

fraction of sputtered particles passing the ionization volume. The ionization volume is the intersection volume of the acceptance of the mass analyzer and the portion of the laser focus in which the intensity is high enough for ionization.

(iii) Different ionization cross sections of various species at a given laser wavelength and laser power density.

For Laser-SNMS with nonresonant multiphoton ionization, the experimentally determined RSFs of different elements are within approximately one decade.[8,9]

Variations of the RSF for a given pair of elements in different samples can be partially explained by different sputter yields (the average number of sputtered particles per primary ion) due to preferential sputtering, and changes in the angular and energy distributions of the sputtered species. In the case of sputter equilibrium, it has been shown in Laser-SNMS that the RSFs remain constant within 20%, in most cases, even for strong matrix variations.[8] In addition to sputter induced variations, changes in the ionization probability are possible. Switching from a metal surface to an oxidized surface for example, the RSFs of most elements are preserved within a factor of 1.5.[8,10] An exception are elements with a relatively low multiphoton ionization cross section. Their RSFs can be enhanced by as much as a factor of 5 for the most dramatic case (e.g., by changing from Si to PtSi[8]). This is primarily caused by large amounts of major molecular sputtering and dissociative ionization of sputtered molecules or clusters, featuring a higher ionization cross section than the corresponding atomic species. Consideration of all these effects makes standard calibration of Laser-SNMS necessary.

These results for Laser-SNMS bulk analysis were the starting point for our attempt to quantify trace metals on Si wafer surfaces by Laser-SNMS. Since the general features of our samples [low concentrations (10^9 cm^{-2} up to 10^{12} cm^{-2}) of metals on top of the native oxide of a Si wafer] did not change, variations of RSFs for different samples are not expected. As long as we remain within static sputtering conditions (primary ion dose $<4\times10^{13}$ cm^{-2}), the sputter induced changes in the surface structure and composition can be neglected. Consequently, the influence of preferential sputtering and changes in the angular and energy distributions of the sputtered particles on the RSFs are insignificant. Therefore, good reproducibility of our measurements for various samples was expected, once a standard calibration for the laser parameters has been carried out.

A technique was necessary for the preparation of standards which provides homogeneous surface coverages by different metals in various combinations in a concentration range between 10^9 cm^{-2} and 10^{12} cm^{-2}. Additionally, the amount of deposited material should be controllable and reproducible. Sputter deposition meets these requirements. In contrast to solution or evaporation preparation, this technique supplies very homogeneous coverages. Furthermore, it is possible to sputter virtually all metals and alloys. The amount of deposited material can be controlled and calculated by the total primary ion dose applied to the sputtered target.

As an additional control, the standard wafers were analyzed after metal deposition by TXRF to determine the resulting surface concentrations of the metals. TXRF is the standard surface analytical tool for research, development and control in IC manufacturing. It is the only reference technique which is comparable to Laser-SNMS and TOF-SIMS, in terms of detection limits (between 10^{10} cm^{-2} and 10^{11} cm^{-2} for the transition metals), and information depth ($\approx$10 monolayers, Laser-SNMS and TOF-SIMS $<$3 monolayers). The different lateral resolution of the techniques is no problem for calibration because our samples are homogeneously covered. The homogeneity was shown by Laser-SNMS analysis of different spots on one sputter deposited wafer.

With these types of standards, RSFs and detection limits were determined for Laser-SNMS. The RSFs allow correction for the effects mentioned above arising from the particular choice of samples. Prior to these studies, the sample preparation itself has been investigated in some detail. A comparison of postionization with different wavelengths (193 nm and 248 nm) follows.

In TOF-SIMS, a greater effort to achieve quantification is necessary because of the more pronounced matrix effect. Here, the RSFs can vary by several decades, e.g., when comparing a metal surface to an oxidized one. In these investigations, the oxidation of the surface is essential to achieve high ionization yields for the metal contaminants and a well-defined chemical environment for these species. With the Laser-SNMS results, it was possible to determine the influence of UV/ozone treatment on the surface concentrations of metal species. Based on these results, RSFs and detection limits for various metals on Si surfaces with TOF-SIMS have been determined.

III. INSTRUMENTATION

A. Time-of-flight mass spectrometer

All measurements presented here were performed in a reflectron type time-of-flight mass spectrometer developed at the University of Münster (TOF III type).[11] Both techniques, TOF-SIMS and Laser-SNMS, are implemented in this type of spectrometer. A schematic drawing of the instrument is shown in Fig. 1. The instrument is equipped with a 10 keV Ar^+ electron impact (EI) primary ion source, which is coupled to a 90°-deflection pulsing unit.

In the Laser-SNMS mode, primary ion pulses with a pulse width of 1 μs and a pulse current of 200 nA are used. The beam diameter is about 100 μm. In the Laser-SNMS mode, it is important to suppress the sputtered secondary ions. This is achieved for the positive secondary ions by a repulsive potential, which is applied to the extractor during the primary ion pulse of length t_p, and the time delay t_d for the laser pulse, according to the timing diagram of Fig. 2. The generated negative secondary ions are suppressed by the analyzer, which is held at a negative potential for the analysis of the positive photoions. For postionization, a standard excimer

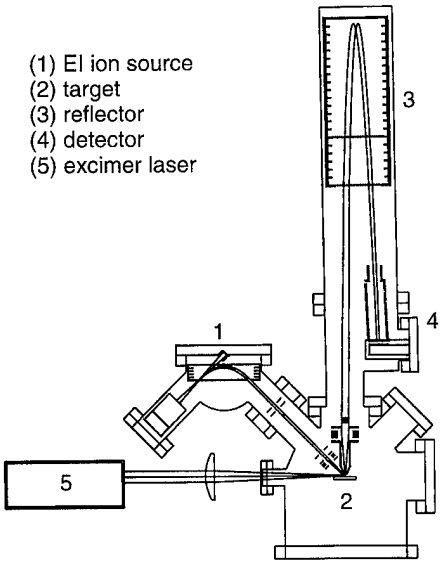


FIG. 1. Scheme of the time-of-flight mass spectrometer.

laser (Lambda EMG 103 MSC, pulse width 15 ns, maximal pulse energy 120 mJ) is used at different wavelengths (193 nm (ArF) and 248 nm (KrF)). The laser beam is focused by a single quartz lens with $f=250$ mm. The resulting ionization volume has a diameter of about 200 μm, which is comparable to the acceptance of the analyzer ($\varnothing \approx 500$ μm). The laser focus is located as close to the target surface as possible, in order to maximize the portion of sputtered particles in the ionization volume. Nearly simultaneously with the laser pulse, the extraction voltage of 2 kV is switched on to accelerate the generated photoions over the extraction distance of 3 mm into the analyzer. The effective flight path of the ions is 3 m. The energy spread is focused in second order by a double stage reflectron. The transmitted ions are detected by a dual channelplate detector (chevron arrangement). In Laser-SNMS, there are up to several hundred photoions of one mass generated in a single shot. Therefore, the photoions must be detected in analog registration mode rather than by single ion counting. The signals are recorded by an 8 bit-transient digitizer (Tektronix RTD 720) with a maximum sampling rate of 2 GHz. For the results reported here, the data transfer to the PC limited the repetition rate to 8 or 16 Hz, depending on the recorded mass range.

For the TOF-SIMS measurements we have used a similar spectrometer (TOF III type) which has been optimized for TOF-SIMS applications. This instrument has an improved primary ion gun (11 keV Ar^+) whose 90°-deflection pulsing unit produces, in combination with an electrodynamic buncher, short mass separated ion pulses up to 3000 primary ions per pulse with a minimal pulse width of 600 ps or a minimal spot diameter of 5 μm. In the TOF-SIMS mode, secondary ions are extracted into the mass analyzer by a static extraction field in front of the target. Because of the smaller energy spread of secondary ions compared with that of photoions, which are generated in an extended volume in front of the target, a single stage reflectron for time focusing

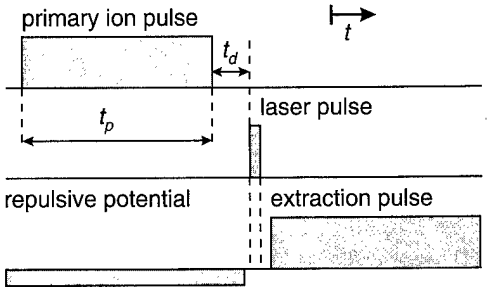


FIG. 2. Timing diagram of one acquisition cycle in Laser-SNMS.

is sufficient. This has the advantage of a higher transmission (factor 2) than a double stage reflectron. TOF-SIMS spectra are recorded in single ion counting mode by a multistop time-to-digital converter (TDC) with ultimate time resolution (50 ps, 512 stops, 2 M channels) and a maximal repetition rate of 20 kHz.

The mass resolution R of a time-of-flight mass spectrometer is given by

$$R=\frac{m}{\Delta m}=\frac{1}{2}\cdot\frac{t}{\Delta t}=\frac{1}{2}\cdot\frac{t}{\sqrt{\Delta t_{ana}^2+\Delta t_{\text{start}}^2+\Delta t_{\text{stop}}^2}}. \quad (2)$$

In the following, the 50% definition of the mass resolution is used. This means that Δm is the full width at half maximum (FWHM) of the respective peak at mass m. In this equation, Δt_{start} and Δt_{stop} are the time durations of the start event and the stop event, respectively, and Δt_{ana} is the time dispersion of the particles during their flight in the mass analyzer. Whereas Δt_{start} and Δt_{stop} are independent of the observed mass, Δt_{ana} is proportional to the total flight time of the ion.

If a static ion extraction is used, Δt_{start} is determined by the duration of the ion formation and therefore in the case of Laser-SNMS is comparable with the laser pulse width. However, this duration can be eliminated by the use of a pulsed and delayed extraction (Fig. 2), where the start moment of the ions is given by the rising edge of the extraction pulse. The duration Δt_{stop} is given by the pulse width of the detector (800 ps) and the time resolution of the transient digitizer (channel width 500 ps). The dispersion Δt_{ana} is caused by the energy distribution of the photoions. This energy distribution has two components: The first one is due to the initial energy distribution of the sputtered particles which is of the order of a few eV. The second one originates from the spatial extension of the ionization volume in the extraction field. Depending on the extraction field strength and the ionization volume, this second energy spread can be 50 times larger than the initial energy spread of the sputtered particles. Unfortunately, these two energy distributions cannot be focused in time simultaneously,[12] so that, in general, the double stage reflectron is used for second order focusing of the energy spread caused by the axial space distribution of the photoions in the ionization volume.

In TOF-SIMS, higher mass resolution is achieved than in Laser-SNMS. The short time interval in which secondary ions are generated (primary ion pulse width of 600 ps) does

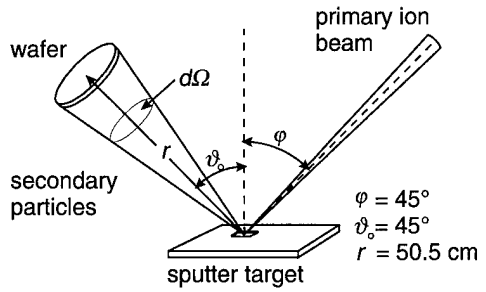


FIG. 3. Scheme of the sputter device. The wafer surface is positioned at a distance r to the sputter target and an angle ϑ_0 relative to the normal of the target.

not require a pulsed and delayed extraction. In addition, all secondary ions start at the same potential. This means the time-of-flight analyzer can be used to focus the initial energy spread of the sputtered secondary ions. Secondary ion spectra are recorded in single ion counting mode by a TDC with an ultimate time resolution corresponding to the channel width of 50 ps. This means that the time error of the stop event Δt_{stop} is considerably smaller than that for Laser-SNMS analysis. In addition, the determination of the stop event with a constant fraction discriminator is very precise and independent of the pulse width of the detector response.

B. Preparation of surface standards by sputter deposition

Standards have been prepared by sputtering corresponding metals or alloys by 10 keV Ar^+ ions (primary ion current 200 nA dc) with an angle of incidence φ=45°. Details of the arrangement are given in Fig. 3. The primary ion beam was scanned over an area of 0.5 mm^2 on the sputter target to compensate for possible inhomogeneities of the alloys. Sputter equilibrium was ensured by presputtering in the case of alloys. The amount of deposited material was controlled by the total primary ion dose N_{PI} applied to the sputter target. Assuming sputter equilibrium, a cosine distribution $dN/d\Omega$ of the sputtered particles, and a sticking probability of 1 for the sputtered metal atoms on the Si-surface, one can calculate the surface coverages Θ on the wafer surface by the following equation:

$$\Theta(\mathrm{A}) = a(\mathrm{A})c(\mathrm{A})N_{\text{PI}}\ Y_s(\varphi)\cdot\frac{\iint_{\Omega}\frac{dN}{d\Omega}d\Omega}{F}$$

$$\approx a(\mathrm{A})c(\mathrm{A})N_{\text{PI}}\cos^{-f}(\varphi)Y_s(\varphi=0)\cdot\frac{1}{\pi}\cdot\frac{\cos(\vartheta_o)}{r^2}, \tag{3}$$

where $c(\mathrm{A})$ is the relative concentration of metal A in the sputtered target, and $Y_s(\varphi)$ the total sputter yield for bombardment at angle φ. The values for $Y_s(\varphi = 0)$ are taken from Matsunami *et al.*[13] Ω is the solid angle and F the area of the wafer. The exponent f has following value[14]:

$$f=\left\{\begin{matrix}\approx\frac{5}{3}\\ \to 1\end{matrix}\right\}\quad \text{for}\quad \frac{M_{\text{target}}}{M_{\text{projectile}}}\left\{\begin{matrix}\leqslant 3\\ \geqslant 3\end{matrix}\right., \tag{4}$$

M is the mass of a projectile particle and a target particle, respectively. Factor a in Equation (3) is a correction factor which considers deviations from the assumptions made above. Its value must be determined experimentally.

IV. RESULTS

A. Preparation and calibration of standards

For the preparation of standards, 4 in.-Si wafers (orientation (100), thickness (525±15) μm, resistivity 16–24 Ω cm, B-doped) were used without any pretreatment from the standard production of Wacker Chemitronic GmbH, Burghausen, FRG.

Qualitative Laser-SNMS investigations show that the intensities of hydrocarbon peaks in a mass spectrum of a wafer after metal deposition by sputtering are higher than those from the original wafer. The surface coverages by the sputter deposited metals are not stable over time or at higher temperatures without any further treatment. This effect for wafers after metal deposition was investigated for an extended time period (six months) and by heating a metal covered wafer for 10 min to a temperature of about 60 °C under vacuum. In both cases, the majority of the deposited metals disappeared from the surface. Hydrocarbon adsorption on the freshly deposited metals which then form volatile metallo-organic compounds may explain these observations. A UV/ozone treatment (30 min in a commercial device, UV-1, SAMCO, Japan) of the wafers immediately after sputter deposition was found to stabilize the metal concentrations. We assume that by this treatment, organic contaminants are fragmented and oxidized to small volatile molecules like CO_2, H_2O, or N_2. A qualitative investigation of UV/ozone treated wafers by Laser-SNMS has shown that the majority of organic contaminants disappear, whereas the surface coverages by the sputter deposited metals change only slightly but become stable even over long times.

For calibration, one wafer with 12 different metals on its surface was prepared (see Table I). In addition, a series of six wafers was prepared with the four metals V, Ni, Co, and Cu in different surface coverages in the range between some 10^9 cm^{-2} and 10^{12} cm^{-2} for testing the functional dependence on the surface coverage of the measured peak intensities in Laser-SNMS and TOF-SIMS. Since the isotopes of these four metals have no mass interferences with other elements and, additionally, these elements are either monoisotopic or their main isotope accounts for over 65% of the element, the uncertainty caused by background in the spectrum is small, even for low surface coverages. Absolute coverages of the standards were determined by TXRF with an ATOMIKA 8010 (excitation with Mo-$K\alpha$). They were found to be in good agreement with the values calculated by Equation (3) after setting the factor a=2 (see also Table I). This factor may be explained in the following manner. As the collision cascade, in the case of nonperpendicular incidence

TABLE I. Standard for metal trace determination on Si wafer surfaces.

Element	Surface coverage/cm^{-2}	
	Calculated by Equation (3) (by setting the factor $a=2$)	Determined by TXRF
Al	5.0×10^{11}	...
Ti	7.0×10^{11}	...
V	1.0×10^{12}	1.0×10^{12}
Cr	8.0×10^{11}	8.3×10^{11}
Fe	1.6×10^{12}	1.1×10^{12}
Co	9.0×10^{11}	1.0×10^{12}
Ni	2.6×10^{12}	2.1×10^{12}
Cu	3.0×10^{12}	2.2×10^{12}
Ga	1.4×10^{12}	1.5×10^{12}
As	1.4×10^{12}	8.0×10^{11}
Mo	9.2×10^{11}	...
W	9.8×10^{11}	4.0×10^{11}

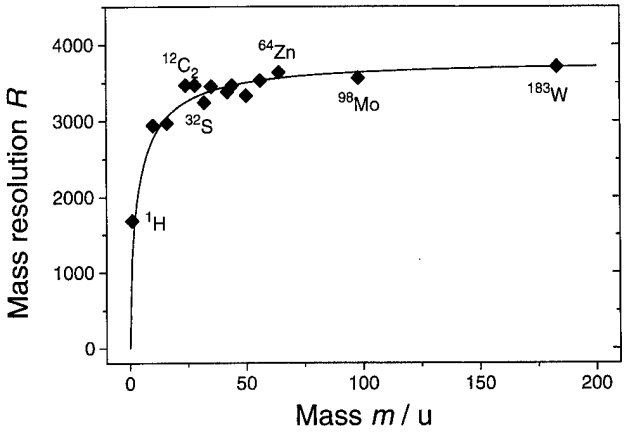


FIG. 5. Mass resolution of the mass spectrometer in the Laser-SNMS mode. The solid curve is fitted to the experimental data.

of the primary particles, reaches the surface sooner than in case of normal incidence, the excitation direction cannot be neglected. The angular distribution of sputtered particles is more pronounced in this direction.[15] Therefore the assumption of a cosine distribution in Equation (3) results in too small values of the calculated surface concentrations. For the following investigations, the surface coverages determined by TXRF were used. For Al, Ti, and Mo, the values calculated by Equation (3) were used (factor $a=2$), because they could not be detected by TXRF in our investigations: the surface coverages by Al and Ti on the samples were below the detection limits of TXRF due to overlap of Al and Ti peaks with the intense Si peak. Mo could not be detected in TXRF because the samples were excited by Mo-$K\alpha$ radiation.

B. Laser-SNMS results

The typical mass resolution achieved in Laser-SNMS is demonstrated in Fig. 4. Since pulsed and delayed extraction of the photoions is performed, the width of the H^+ peak (1.5 ns) is caused by the mass independent duration of the detector pulse (800 ps) and the channel width of the transient digitizer (500 ps). This leads to a time error $\Delta t_{stop} \approx 940$ ps. The time dispersion Δt_{ana} can be assumed to be about 670 ps since the overall flight time is known to be about 5.1 μs. Therefore, the duration of the starting process Δt_{start} can be estimated as about 950 ps [Equation (2)]. The pulsed and delayed extraction minimizes the duration of the ionizing event (laser pulse width 15 ns). In contrast to H, the time width for masses above 12 u is mainly determined by the time dispersion Δt_{ana}. In this mass range, the mass resolution approaches the limit given by the time-of-flight analyzer ($R\approx3800$). In Fig. 5 the experimentally determined mass resolution of Laser-SNMS is plotted as a function of mass. The solid curve represents the functional form of the Equation (2) fitted to the experimental values by following parameters: $\sqrt{\Delta t_{start}^2+\Delta t_{stop}^2}=1.4$ ns, $t/2\Delta t_{ana}=3800$.

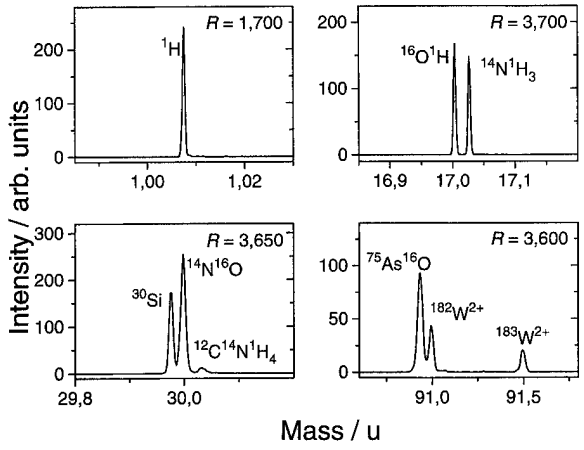


FIG. 4. Details of a Laser-SNMS spectrum demonstrating the high mass resolution of the double reflectron TOF mass spectrometer in the Laser-SNMS mode.

In Fig. 6 a logarithmically scaled spectrum of a Si wafer covered with metals (Ni, Cu, Zn) at a total surface concentration of about 10^{12} cm^{-2} is shown. It was acquired under low dose sputtering conditions (primary ion fluence $<4\times10^{13}$ cm^{-2}, equivalent to sputtering of less than 10% of a monolayer) over 10 000 cycles (acquisition time 21 min). The sputtered particles were postionized with a laser power density of about 10^{10} W cm^{-2} (wavelength $\lambda=248$ nm, laser pulse width $\tau=15$ ns). The most important features of this spectrum may be summarized as follows:

(i) The dynamic range in the spectrum is about six orders of magnitude.

(ii) There is almost no background over the entire mass range. Obviously, the generated secondary ions are effectively suppressed by the combination of the repelling field during sputtering and the delayed extraction after postionization. Since the majority of sputtered molecular species are photodissociated by interaction with the laser beam, metastable decay in the analyzer does not contribute to the background in the spectrum.

(iii) The achieved mass resolution is sufficient to separate the elemental peaks from molecular peaks at the same nominal mass. Only some H-free molecules, mainly

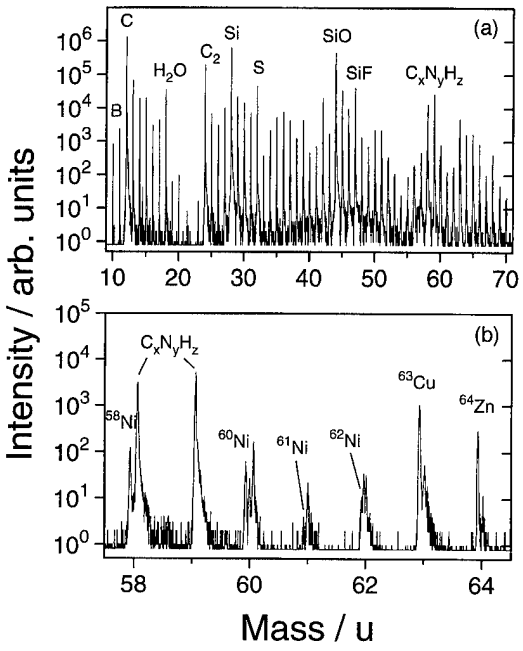


FIG. 6. Logarithmically scaled Laser-SNMS spectrum (primary ion fluence 4×10^{13} cm^{-2}, postionization: $\lambda=248$ nm, $I=1\times10^{10}$W cm^{-2}) of a Si wafer covered by various metals (Ni, Cu, Zn). (a) Overview and (b) detail.

monoxides like SO or VO, show interferences with elemental signals and require higher mass resolution.

(iv) The base peak in the spectrum, $^{12}C^+$, originates primarily from photodissociation of sputtered hydrocarbons. The contribution to this peak of gas phase ionization from the residual gas (total pressure $<5\times10^{-9}$ mbar) or thermally desorbed organic molecules was found to be below 1%.

(v) The matrix-signals Si^+, SiO^+ are relatively low. We found the ionization cross section of SiO at 248 nm to be low compared to those of metals like V, Cr, etc. and that of Si to be even smaller. The Si^+-signal is mostly formed from SiO photofragmentation.

(vi) The intense S^+ peak indicates a high concentration of S on the wafer surface. S has not been deposited by our preparation but was on the surface before, as indicated by spectra of untreated wafers.

(vii) Most of the detected molecular photoions are of the general structures $C_xO_yH_z$ or $C_xN_yH_z$. Due to the strong dissociation of sputtered molecules, the interpretation of this part of the spectrum is difficult.

(viii) Some doubly charged ions are generated for several metals like V, Cr, Mn ($\approx10\%$ depending on the laser power density). These species must be considered for accurate quantification of the corresponding elements. This is performed by summing up the intensities of the singly charged and doubly charged photoions, because the doubly charged ions are mainly generated by successive absorption of additional photons by previously formed singly charged ions.[16]

The accuracy and reproducibility of Laser-SNMS results have been demonstrated by the analysis of a series of Si

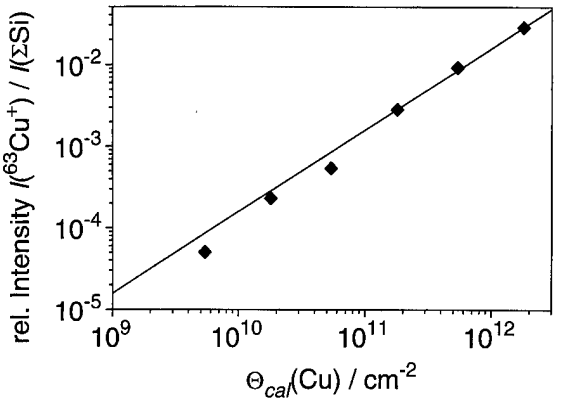


FIG. 7. Linear correlation between the normalized Cu^+ Laser-SNMS signal and the calculated Cu surface coverage Θ_{cal}. The solid line is the linear regression fit to the experimental data.

standards covered by metals in the range of some 10^9 cm^{-2} to 10^{12} cm^{-2}. In Fig. 7 the linear relation between the Cu surface coverage and the $^{63}Cu^+$ intensity (intensity corresponds to the integrated peak area) normalized to the sum of the Si^+ and the SiO^+ intensities is shown. The excellent correlation observed (correlation coefficient $R=0.997$) is also confirmed by corresponding results for V, Co, and Ni.

In Table II the RSFs of several metals deposited on Si wafers and the detection limits are listed for the two laser wavelengths (193 nm and 248 nm) applied. The RSFs are defined by Equation (1) wherein the bulk concentration c(A) is replaced by the surface coverage Θ(A). The normalization of the metal ion intensities to the sum of Si^+ and SiO^+ takes into account small variations in experimental conditions (primary ion current, transmission, position of the laser focus) during data acquisition. Additionally, the effects of various oxidation states of the surface are considered. Since the sputtered SiO species are either photoionized to SiO^+ or photofragmented to atomic Si^+, the sum of the Si^+ and the SiO^+ intensities is an approximate measure for the total number of sputtered Si (elemental or oxidized).

At 193 nm the ion yield for Si^+ is enhanced by a factor of 20 compared to 248 nm, because of an accidental resonance. Therefore the combined ion yield of Si^+ and SiO^+ is of the same order as the ion yield of metals like Cr and Fe. Al and Ga are ionized in a single photon process at 193 nm, which leads to an enhancement of the ionization probability by one decade for Al and two decades for Ga. This improves the detection limits for these two metals accordingly. A slight enhancement of the ion yield is observed for the other metals, too, in agreement with the investigations of Hayashi *et al.*[17] All these facts results in an accidental improvement of uniformity of the RSFs at 193 nm in contrast to 248 nm. A disadvantage of the 193 nm wavelength is the higher ionization probability for hydrocarbons. For this wavelength, even in the higher mass range ($m>70$ u), organic molecules are detected in the spectrum with high intensities, and the intense $^{12}C^+$ peak can lead to detector saturation, making quantification impossible.

TABLE II. Laser-SNMS: Relative sensitivity factors and detection limits for metals on Si wafer surfaces. Detection limits may originate from different factors (see Sec. V B 1).

	Relative sensitivity factor $S(\mathrm{Me},\Sigma\mathrm{Si})$		Detection limits in cm^{-2}		
Element	193 nm	248 nm	193 nm	248 nm	limited by
Al	0.9	0.9	8×10^{9}	3×10^{10}	useful yield
Ti	1.9	9	5×10^{9}	5×10^{9}	mass interference
V	2.1	42	3×10^{8}	3×10^{8}	data rates
Cr	2.9	28	3×10^{9}	3×10^{9}	mass interference
Fe	1.6	31	5×10^{9}	5×10^{9}	mass interference
Co	1.1	6.9	5×10^{8}	5×10^{8}	data rates
Ni	0.9	14	1×10^{9}	1×10^{9}	isotopy, data rates
Cu	1.2	24	5×10^{9}	5×10^{9}	mass interference
Ga	4.4	0.4	5×10^{8}	8×10^{9}	isotopy, data rates
As	1.9	10	3×10^{8}	3×10^{8}	data rates
Mo	1.8	32	3×10^{8}	3×10^{8}	isotopy, data rates
W	5.1	39	5×10^{8}	5×10^{8}	isotopy, data rates

The detection limits for the respective metals are also given in Table II. For a detailed discussion see Sec. V B1.

C. TOF-SIMS results

In Fig. 8 the mass resolution achieved in the TOF-SIMS mode is plotted as a function of mass. The scatter in the data is caused by large differences in the energy distributions of elemental and various molecular secondary ion species. The resolution of the mass analyzer is about 12 000 and the time error of the start and stop events can be estimated as $\sqrt{\Delta t_{\mathrm{start}}^2+\Delta t_{\mathrm{stop}}^2}\approx 700$ ps. In Fig. 9 a detail of a positive secondary ion spectrum from a Si wafer is shown demonstrating the high mass resolution achieved.[4] Eleven peaks at the nominal mass 56 u can be identified. The mass difference between $^{56}Fe^+$ and $^{28}Si_2^+$ is only 0.0189 u. A mass resolution of more than 3000 is required to separate the corresponding peaks, provided they have equal intensities.

For quantitative determination of metal contaminants on Si wafers with TOF-SIMS, a reproducible oxidation state of the metal atoms at the surfaces is required in order to control the matrix effect. This reproducible oxidation state has been achieved by a 30 min UV/ozone treatment of the wafer surface. This oxidation procedure results in a high and reproducible probability that sputtered metal atoms are emitted as positive ions. In addition, surface coverage by hydrocarbons is considerably reduced.

The effect of the UV/ozone treatment is demonstrated in Fig. 10. Molecular surface contaminations of a Co-covered ($\Theta=3\times10^{10}$ cm^{-2}) wafer results in a significant background in the mass spectrum [Fig. 10(a)]. Co^+ ions are not detectable. Figure 10(b) shows the spectrum of the same wafer after a 30 min UV/ozone treatment, resulting in a well separated small Co^+ peak. The remaining background originates from surface adsorption of hydrocarbons from the ambient air during transfer of the wafer from the UV/ozone chamber to the mass spectrometer. In principle, this re-coverage could be avoided by a UV/ozone treatment directly in the vacuum chamber of the spectrometer.

We have investigated the influence of the UV/ozone treatment on the Laser-SNMS spectra, since this method is much less sensitive to changes in the surface composition than TOF-SIMS. There is only a change in the relative metal ion intensities $I(\mathrm{Me}^+)/I(\Sigma\mathrm{Si}^+)$ by a factor of 2 in the worst case, and a considerable reduction of the hydrocarbon ion intensities. These results justify the assumption that the trace metal concentration is not affected by the UV/ozone pretreatment.

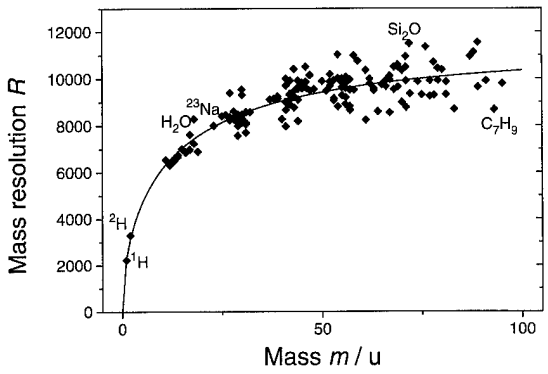
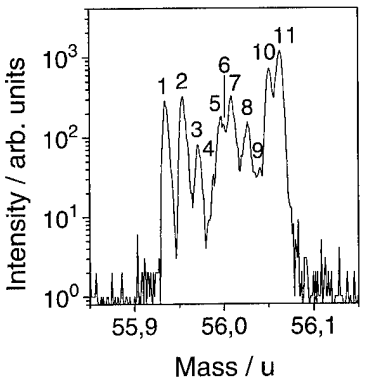


FIG. 8. Mass resolution of the mass spectrometer in the TOF-SIMS mode. The solid curve is fitted to the experimental data. The scattering of the values is caused by large differences in the energy distributions of elemental and various molecular secondary ion species.



Nr.	peak	mass / u
1	^{56}Fe	55.9343
2	$^{28}Si_2$	55.9536
3	$^{28}SiCO$	55.9706
4	$^{29}SiCNH$	55.9869
5	$^{28}SiCNH_2$	55.9956
6	$^{29}SiC_2H_3$	55.9999
7	$^{28}SiC_2H_4$	56.0079
8	C_3OH_4	56.0256
9	$C_2N_2H_4$	56.0385
10	C_3NH_6	56.0500
11	C_4H_8	56.0618

FIG. 9. Detail of a secondary ion spectrum demonstrating the high mass resolution of the mass spectrometer in the TOF-SIMS mode.

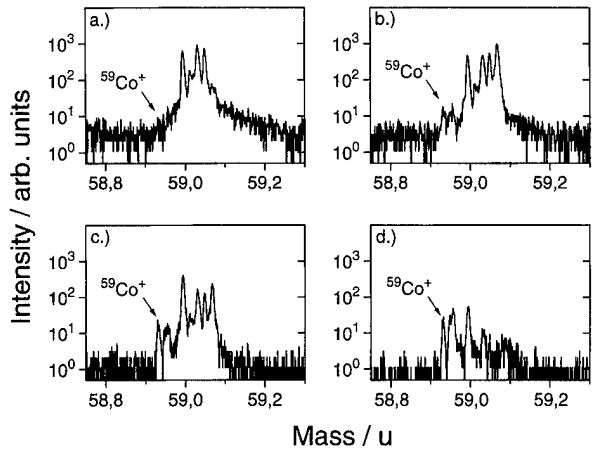


FIG. 10. Details of secondary ion peaks at 59 u [Θ(Co)=3×10^{10} cm^{-2}] (primary ion dose 1.25×10^9, analyzed area 700 μm^2). (a) Untreated; (b) 30 min UV/ozone treatment; (c) 30 min UV/ozone treatment, presputtered (primary ion fluence 2×10^{13} cm^{-2}); and (d) 30 min UV/ozone treatment, presputtered (primary ion fluence 8×10^{13} cm^{-2}).

To reduce the remaining background in the secondary ion spectrum after UV/ozone treatment, moderate presputtering (primary ion fluence $<8\times10^{13}$ cm^{-2}) is helpful. This can be seen from Figs. 10(c) and 10(d). Presputtering results in a significant decrease of molecular ions with only negligible sputtering of metal atoms. One exception is Cu: a possible explanation is the assumption of preferential sputtering of Cu from a Si surface or a strong decrease of the ionization probability of Cu after presputtering. Unfortunately, the determination of low Cu surface coverages is not possible without presputtering.

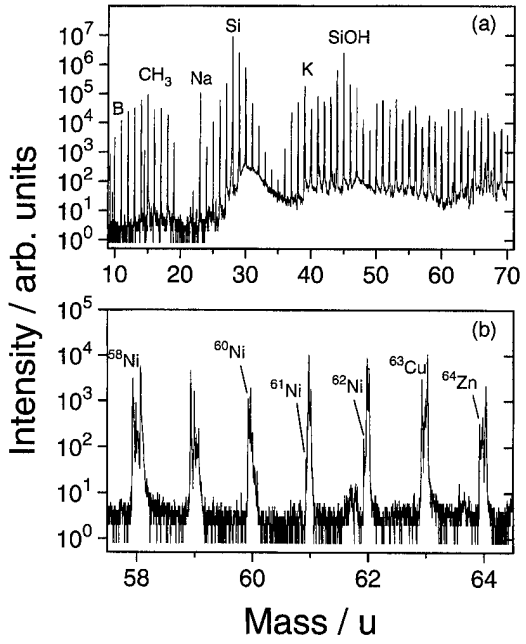


FIG. 11. Logarithmically scaled secondary ion spectrum (primary ion dose 1.25×10^9, analyzed area 700 μm^2) of a Si wafer covered by various metals (Ni, Cu, Zn). (a) Overview and (b) detail.

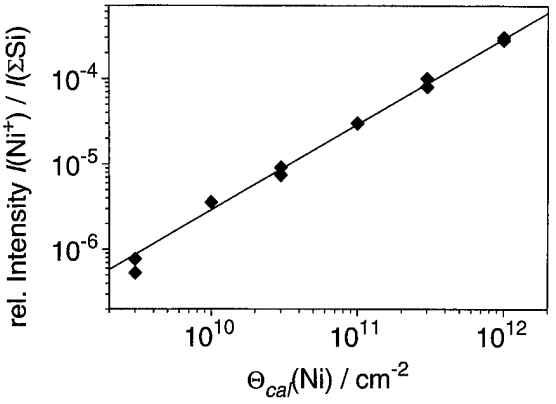


FIG. 12. Linear correlation between the normalized Ni^+ secondary ion signal and the calculated Ni surface coverage Θ_{cal}. The solid line is the linear regression fit to the experimental data.

Figure 11 presents a logarithmically scaled TOF-SIMS spectrum from the same wafer as Fig. 6. The wafer was presputtered with a primary ion fluence of 5×10^{13} cm^{-2}. For spectral acquisition, the surface was bombarded with a total dose of 1.25×10^9 primary ions in a spot size of 700 μm^2 in 15 min. This is equivalent to an erosion of about 40% of the outer monolayer assuming a sputter yield of about 2. The most important features of this spectrum may be summarized as follows:

(i) The dynamic range in the spectrum is about six orders of magnitude.
(ii) The background in the spectrum originates from metastable decay of highly excited molecular secondary ions in the analyzer.
(iii) The high mass resolution allows separation of nearly all peaks at each nominal mass.
(iv) The base peak in the spectrum is the matrix signal Si^+. There are also intense peaks of SiH^+ and $SiOH^+$.
(v) The Na^+ and K^+ peaks are relatively high, because of the high ionization probabilities for alkali atoms in SIMS.
(vi) Most of the molecular ions can be identified as ions of the general compositions C_xNH_y, SiC_xH_y, C_xOH_y and $Si_xO_yH_z$. They can supply important information on the molecular composition at the wafer surface.[18]

The quantification of metal contaminants is based upon the assumption of a functional dependence of the observed integrated peak intensities on the corresponding surface coverage. We have investigated by TOF-SIMS the same metal covered Si wafers as by Laser-SNMS. Since the transformation probability $P(\text{Me}\rightarrow\text{Me}^+)$ is higher than $P(\text{Me}\rightarrow\text{Me}^-)$ for almost all metals, our investigations were restricted to positive secondary ion spectra. The metal ion intensities were normalized to the matrix-signals $^{28}Si^+$ plus $^{28}Si^{16}O^+$ ($=\Sigma$Si) in order to minimize small variations in the experimental conditions. Figure 12 presents as an example the results for Ni. A good linear correlation (correlation coefficient R=0.999) was found in the concentration range between some 10^9 cm^{-2} and 10^{12} cm^{-2}. This was confirmed also for

TABLE III. TOF-SIMS: Relative sensitivity factors and detection limits for metals on Si wafer surfaces (30 min UV/ozone treated). Detection limits may originate from different factors (see Sec. V B 1).

Element	Relative sensitivity factor $S(\mathrm{Me},\Sigma\mathrm{Si})$	Detection limits in cm^{-2}	Limited by
Al	1.6	5×10^{8}	background
Ti	0.9	3×10^{9}	background
V	1.1	2×10^{9}	useful yield
Cr	1.5	1×10^{9}	useful yield
Fe	1	3×10^{9}	useful yield
Co	0.5	2×10^{9}	useful yield
Ni	0.4	4×10^{9}	useful yield, isotopy
Cu	0.5	5×10^{9}	useful yield, isotopy
Ga	1.8	2×10^{9}	useful yield, isotopy
As	0.1	2×10^{10}	useful yield
Mo	0.2	3×10^{10}	useful yield, isotopy
W	...	$>1\times10^{12}$	useful yield, isotopy

V, Co, and Cu. Assuming the same linear behavior for the other investigated metals one can define RSFs $S(\mathrm{Me},\Sigma\mathrm{Si})$ as in Laser-SNMS. Table III presents the values for $S(\mathrm{Me},\Sigma\mathrm{Si})$ and the corresponding detection limits.

The concentration of W on the reference wafer is below the TOF-SIMS detection limit for this metal. A higher secondary ion yield is expected for the oxidized W atoms (WO_n^-) in the negative spectrum. Similar considerations hold for As.

The RSFs presented in Table III are valid only for the chemically well defined oxidized Si surface with metal coverages in the submonolayer range. For higher surface concentrations of the metals ($>1\times10^{13}$ cm^{-2}), sputter yields as well as ionization probabilities may be affected.

V. DISCUSSION

A. Fundamental aspects

1. Sample preparation

Controlled ion beam sputter deposition was applied for preparation of all investigated standards. The surface concentration of metals prepared by such a sputter deposition is determined by:

(i) The flux of sputtered particles.
(ii) The sticking probability of the metal atoms arriving at the surface.
(iii) The behavior of the metal atoms on the surface after their deposition (surface diffusion, cluster formation, diffusion into the surface oxide, desorption, etc.).

For describing the coverage of the wafer surface by sputtered metal atoms [Equation (3)] the following assumptions were made:

(i) The flux of sputtered metal atoms is determined by the composition of the target, the sputter yield of the respective species, and their angular distribution. Sputter yields have been applied according to Matsunami *et al.*[13] These yields are calculated by an empirical formula with best fit of the parameters to the available experimental data. They are assumed to be accurate within 20%. The calculated total sputter yields are valid for normal incidence of the projectile particles. The dependence of these yields on the angle of incidence of the projectiles was considered as well as the dependence of the angular distribution of sputtered particles by the factor a in Equation (3). The value of a has been found to be about 2 for all sputtered targets by comparing the calculated surface coverages and those determined by TXRF. The good agreement between the calculated values for the surface coverages according to Equation (3) and those determined by TXRF confirms an adequate modelling of sputtering by the equations used.

(ii) The assumption of sticking probabilities near 1 for all sputter deposited metals on a Si wafer surface is a priori not justified. However, the good agreement between the calculated surface coverages and those determined by TXRF supports this assumption. A possible sticking probability of less than 1 is considered by the correction factor a in Equation (3). This factor has been found to be nearly the same for all deposited metals indicating that the sticking probability does not depend on the metal species. This all indicates a value of 1 for the sticking probability.

(iii) Thermal desorption of sputter deposited submonolayers of metal atoms on the Si wafer surface was observed. To understand this, the structure of the wafer surface must be considered: the Si wafer is covered by a native oxide, which is covered by a so-called contamination layer, containing a wide variety of hydrocarbons. The desorption of metal atoms may be enhanced by the formation of volatile metallo-organic compounds on this type of surface. This is supported by the observation that UV/ozone treatment, resulting in the transformation of all metal atoms on the surface into corresponding oxides, stops this loss of metal atoms.

(iv) The metal covered Si wafers were calibrated by TXRF. This technique features an information depth of about 10 monolayers, compared with an information depth of the sputter based techniques (Laser-SNMS and TOF-SIMS) of less than 3 monolayer. This means that the results of TXRF and Laser-SNMS/TOF-SIMS are only comparable if the metal atoms are present only in the uppermost monolayer and not in or under the oxide layer. This has been confirmed by appropriate depth profiling. The sputter deposited material was only detected in the uppermost monolayer. As a result of

this observation diffusion of the sputter deposited metal atoms into the native oxide could be excluded even after oxidation by UV/ozone treatment.

With the exception of W, the surface coverages by all deposited metals are adequately described by Equation (3), i.e., they are in good agreement with the TXRF results. For W, the TXRF measurements give a much lower surface concentration then calculated. From the experimental results, it could not be decided if this difference originates from a sticking probability of W atoms on the Si wafer surface being below 1, or if the TXRF results are not correct in the case of W. The latter may be possible, since for the TXRF measurements, the samples were excited by Mo-$K\alpha$ radiation, and therefore, in contrast to other elements, W must be determined by the weak $L\beta$ line.

2. Laser-SNMS features

The useful yield $Y_u(\mathrm{A},\mathrm{X}_i^q)$ is the most important quantity describing the fraction of removed particles A which are transformed into ions X_i^q available for the measurement. It is defined by following equation:

$$Y_u(\mathrm{A},\mathrm{X}_i^q)=\frac{\text{number of detected ions of type } i \text{ and charge } q \text{ formed from A}}{\text{number of removed particles A}}. \tag{5}$$

The RSFs give the relationship between the useful yields of the respective metal and the useful yield of Si (Si^+ plus SiO^+), if one neglects effects like selective sputtering. The useful yield of Si from a UV/ozone treated wafer surface is assumed to be about 2×10^{-4} in Laser-SNMS. The useful yields can be improved by one order of magnitude by reducing the primary ion pulse length considerably below 1 μs. The fraction of sputtered particles in the ionization volume at the moment of interaction with the laser pulse is then optimized. With reduction of the primary ion pulse length, however, the absolute number of particles in the ionization volume decreases accordingly. This results in a longer acquisition time for a spectrum if the same dynamic range is required. A primary ion pulse length of 1 μs used in our experiments is a compromise between a reasonable acquisition time and a high useful yield.

The differences in RSFs for different metals stem from three phenomena: preferential sputtering, differences in angular and energy distribution, and variations in ionization probabilities (see Sec. II). Because the surface coverage by the metals on the samples is below 1×10^{13} cm^{-2}, the sputter process is dominated by the silicon oxide matrix. Therefore, differences in the RSFs for different metals can be explained mainly by different densities of sputtered particles in the ionization volume and by different ionization probabilities:

(i) The density of sputtered particles is inversely proportional to the velocity of the sputtered particles. This velocity is inversely proportional to the square root of the mass, assuming the same energy distribution for the sputtered particles. Therefore, the RSFs of elements with higher atomic number are in general larger than those of lighter elements.

(ii) To achieve equal ionization probabilities for all elements it is necessary to reach ionization saturation levels. In that case, all particles in the interaction volume are available for the mass analysis, independent of the respective ionization potential E or the generalized n-photon ionization cross section $\sigma^{(n)}$. Since we are limited by the maximum laser power density of approximately 1×10^{10} W cm^{-2} (248 nm) in the focus it was not possible to reach saturation levels for the ionization of all elements, e.g., Al.

(iii) When the ionization saturation level is obtained, the ionization probability is mainly determined by the size of the ionization volume. The laser power density in the wings of the focus outside the maximum focus point can be high enough to ionize elements with a high ionization cross section $\sigma^{(n)}$. Therefore, the ionization volume for these elements is larger than that for elements with a low ionization cross section.

The RSFs are determined by calibration with the standards. Problems with quantification will arise if the investigated sample deviates too much from the calibrated standard, or the experimental conditions are changed. Because only Si wafer surfaces with metal contaminants below 1×10^{13} cm^{-2} are investigated, the first limitation is not important. Since the number of sputtered particles is proportional to the primary ion dose for all elements, variation of the primary ion dose has no influence on the RSFs. On the other hand, variations of the laser power density between different measurements may influence the RSFs. The ionization probability of elements, whose level of ionization is not saturated, varies more with the laser power density than that of elements with saturated levels. Another influence on the RSFs is given by the dependence of the size of the ionization volume on the laser power density. The ionization volume increases with greater laser power density in the wings of the focus. This effect is more pronounced for elements with a high ionization cross section $\sigma^{(n)}$, once the ionization is saturated in the focus. A third effect influencing the sensitivity of an element is the degree of dissociative ionization if this element is sputtered as a molecule, particularly as an oxide, e.g., SiO and Si. The extent of dissociation increases with higher laser power density. Under defined conditions (laser energy 95–105 mJ at 248 nm, the same focusing optic for each measurement) changes in the RSFs has been found to be below 20%.

3. TOF-SIMS features

The RSFs for TOF-SIMS describe the relationship between the useful yield of the respective metal and the useful yield of Si (Si^+ plus SiO^+), if the differences in the sputter yields for different metal atoms are neglected. The estimated useful yield for Si (UV/ozone treated wafer surface) is found to be about 5×10^{-3}. In contrast to Laser-SNMS, this yield strongly depends on the chemical environment of the sputtered atoms. By UV/ozone treatment this chemical environment was stabilized in all the described experiments. Reproducible oxidation of the Si wafer surface as well as of the metal atoms on its surface was achieved by UV/ozone oxidation, resulting in a high metal ion yield. Under these conditions, the uncertainty of RSFs was below 10%.

B. Analytical aspects

The usefulness of a microanalytical technique is mainly determined by its detection limits, the possibility of quantification, the lateral resolution, and the acquisition time. The detection limits will be discussed in more detail, and Laser-SNMS, TOF-SIMS, and the applied reference technique TXRF will be compared with regard to the features mentioned above. For a discussion of quantification see Secs. V A 2 and 3.

1. Detection limits

Absolute detection limits in Laser-SNMS and SIMS are determined by the total amount of sample material A available for the analysis. The fraction of this sample material which can be transformed into detected ions X_i^q—after sputtering, ionization, and transfer through the spectrometer—is described by the useful yield $Y_u(A,X_i^q)$. The minimal detectable surface coverage (detection limit) by an element A in a matrix M by sputtering of one monolayer is given by

$$\Theta_{\min}(A)=\frac{N_{\min}(X_i^q)}{Y_u(A,X_i^q)F}. \tag{6}$$

$N_{\min}(X_i^q)$ is the minimal number of detected ions X_i^q required to identify the signal, and F is the analyzed area. By sputtering of one monolayer from a relatively large analyzed area, $F=1\times10^{-3}$ cm^2, and an useful yield of $Y_u(A,X_i^q) = 1 \times 10^{-3}$, the detection limit is $\Theta_{\min}\approx10^7$ cm^{-2}, assuming that 10 ions will be sufficient to identify the corresponding peak. For a given amount of sample material, this detection limit can be improved by two means: by improving the overall transmission of the spectrometer, which means the probability that a generated ion is detected after mass separation, or by increasing the probability for a sample atom to be transformed into an ion (transformation probability).

Additional limitations in sensitivities are given by background or peak interferences in the spectrum, and the acquisition time required for sputtering of the sample material available. These effects can lower the detection limits for some elements considerably.

In Laser-SNMS the quantification of the main isotopes of Ti, Cr, and Fe is hampered by the following mass interferences: $^{48}Ti^+\leftrightarrow{}^{32}S^{16}O^+$ [Fig. 13(a)], $^{52}Cr^+\leftrightarrow{}^{28}Si^{12}C_2^+$, and

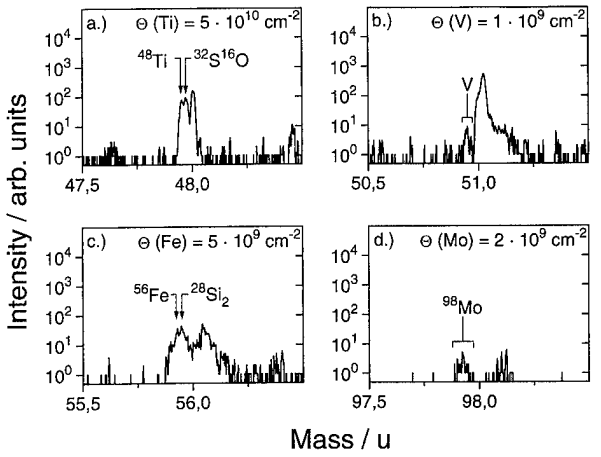


FIG. 13. Details of a Laser-SNMS spectrum (primary ion dose 4×10^{10}, analyzed area 1×10^{-3} cm^2, postionization: λ=248 nm, $I = 1 \times 10^{10}$ W cm^{-2}). The detection limits of *V* and *Mo* are limited by the useful yields of these elements and the available data rates. The detection limits of Ti and Fe are limited by mass interference.

$^{56}Fe^+\leftrightarrow{}^{28}Si_2^+$ [Fig. 13(c)]. For elements with higher masses, mass interferences become less important since the transformation probabilities of complex molecules into large molecular ions decrease. Therefore, one is not restricted by mass interferences for As^+, Mo^+ [Fig. 13(d)], and W^+. The detection limits of elements which are unhampered by background and mass interferences are only limited by the available data rates ($\approx4\times10^3$ counts/s) [Fig. 13(b)]. For the experiments reported here we were restricted by the time needed to process the data. With a faster data registration and processing, it is now possible to improve the detection limits for Laser-SNMS by more than one decade, provided one has enough sample material and is not restricted by one of the other effects mentioned above.

In TOF-SIMS, one is hampered by a significant background in the spectrum caused by metastable ions (Fig. 10). To minimize this effect, the surface concentration of organic compounds can be reduced by UV/ozone treatment. Also, cleaning by moderate presputtering (primary ion fluence $\leq2\times10^{13}$ cm^{-2}) has proven to be useful and acceptable. The detection limits in TOF-SIMS are then restricted by the maximal achievable dynamic of six decades in the spectrum.

2. Comparison of Laser-SNMS, TOF-SIMS, and TXRF

Regarding quantification of metal trace impurities on Si wafers, there are only marginal differences between TOF-SIMS and Laser-SNMS. The detection limits are comparable with some advantages for the heavier elements ($m > 70$ u) in Laser-SNMS. The accuracy of the measurements are slightly better in TOF-SIMS (error $<10\%$) than in Laser-SNMS (error $<20\%$). This is mainly caused by the more critical experimental setup of Laser-SNMS: variations in the laser power density or in the location and size of the laser focus may have effects on the measurements. TOF-SIMS

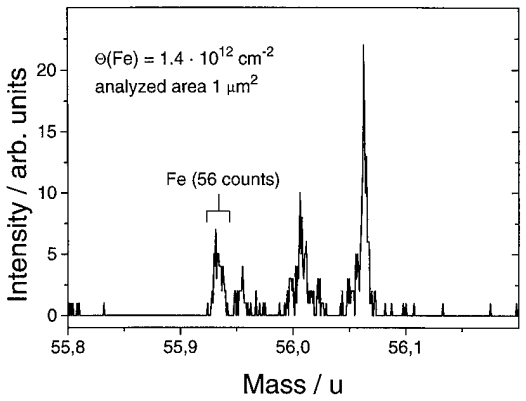


FIG. 14. High mass resolution secondary ion spectrum at 56 u (Ga^+ liquid metal ion source, primary ion fluence 2×10^{15} cm^{-2}) from a 1 μm^2 area of a Si wafer covered by 1.4×10^{12} cm^{-2} Fe atoms.

additionally supplies information on molecular surface species, but spectral interpretation may be more difficult for complex surface compositions. The main advantage of Laser-SNMS is that even for drastic changes in the chemical environment of the sputtered particles, this technique still allows a straightforward quantification in element analysis.

Comparing the surface mass spectrometries Laser-SNMS and TOF-SIMS with TXRF, the detection limits of the sputtering based techniques for metal impurities on Si wafers are lower than those of TXRF. The sensitivities are only comparable for the transition metals Fe up to Zn. For the light elements, TXRF is restricted by the signal-to-noise ratio due to the tailing of the Si peak. The sensitivity for the heavy elements is low because they must be detected by their $L\beta$-line when the sample is excited by Mo-$K\alpha$ or W-$L\beta$. The accuracy of TXRF is slightly better than that of Laser-SNMS and TOF-SIMS. Laser-SNMS and TOF-SIMS have monolayer sensitivity whereas the information depth of TXRF is about 3 nm. In addition, TOF-SIMS also gives important information on the molecular surface composition. An advantage of TXRF is the easy implementation of the technique. There is no need to place the sample under vacuum, and the analysis is almost automatic. Additionally, TXRF provides a nearly standard-free calibration. Besides their high sensitivities for almost all elements, an important advantage of the sputtering based techniques is the high lateral resolution which is only limited by the primary ion spot diameter (<100 nm for a Ga^+liquid metal ion source) in contrast to the large area (about 1 cm^2) required in TXRF. The detection limits in Laser-SNMS and TOF-SIMS are linearly scaled by the respective size of the analyzed area (Fig. 14). Both techniques allow surface imaging,[19] depth profiling and 3D-analysis of micro structures.

VI. OUTLOOK

The calibration procedure described in this paper is not restricted to metal contaminants on silicon wafer surfaces. It can easily be extended to the quantitative determination of trace metals on other semiconductors like GaAs.[20] For preliminary quantification, it has been shown to be sufficient to calibrate standards for TOF-SIMS by Laser-SNMS with the RSFs determined for metal contaminants on Si. For more accurate quantification, it is necessary to calibrate the standards by an additional reference technique, such as TXRF. For other substrates like biological samples or polymers,[21] a similar calibration procedure may be applied. Future investigations have to show the possibility of producing metal standards by sputter deposition on these kinds of surfaces. A further lowering of detection limits of trace metals on wafer surfaces may be achieved by the combination of TOF-SIMS and Laser-SNMS with the technique of vapor phase decomposition (VPD) as it is done with TXRF.[22]

ACKNOWLEDGMENTS

The authors gratefully acknowledge support by Dr. D. Gräf, Wacker Chemitronic GmbH, Burghausen, FRG, for donating the Si wafers, and Dr. Pahlke, Wacker Chemitronic GmbH, Burghausen, FRG, who carried out the TXRF measurements.

[1]H.-G. Cramer, U. Jürgens, E. Niehuis, M. Terhorst, Zhiyuan Zhang, and A. Benninghoven, in *Proceedings of the 9th International Conference on Secondary Ion Mass Spectrometry, SIMS IX, Yokohama 1993*, edited by A. Benninghoven, Y. Nihei, R. Shimizu, and H. W. Werner (Wiley, Chichester, 1994), p. 449.
[2]J. Knoth, H. Schwenke, and U. Weisbrod, Spectrochim. Acta Part B **44**, 477 (1989).
[3]E. Niehuis, T. Heller, U. Jürgens, and A. Benninghoven, J. Vac. Sci. Technol. B **7**, 512 (1989).
[4]U. Jürgens, H.-G. Cramer, T. Heller, E. Niehuis, Zhiyuan Zhang, and A. Benninghoven, in *Proceedings of the 8th International Conference on Secondary Ion Mass Spectrometry, SIMS VIII, Amsterdam 1991*, edited by A. Benninghoven, K. T. F. Janssen, J. Tümpner, and H. W. Werner (Wiley, Chichester, 1992), p. 277.
[5]M. Terhorst, R. Möllers, A. Schnieders, E. Niehuis, and A. Benninghoven, in *Proceedings of the 9th International Conference on Secondary Ion Mass Spectrometry, SIMS IX, Yokohama 1993*, edited by A. Benninghoven, Y. Nihei, R. Shimizu, and H. W. Werner (Wiley, Chichester, 1994), p. 434.
[6]A. Schnieders, F. Schröder-Oeynhausen, B. Burkhardt, F. Kötter, R. Möllers, L. Wiedmann, and A. Benninghoven, in *Proceedings of the 10th International Conference on Secondary Ion Mass Spectrometry, SIMS VIII, Münster 1995*, edited by A. Benninghoven, B. Hagenhoff, and H. W. Werner (Wiley, Chichester, in press).
[7]C. H. Becker and K. T. Gillen, Anal. Chem. **56**, 1671 (1984).
[8]G. Kampwerth, M. Terhorst, E. Niehuis, and A. Benninghoven, in *Proceedings of the 8th International Conference on Secondary Ion Mass Spectrometry, SIMS VIII, Amsterdam 1991*, edited by A. Benninghoven, K. T. F. Janssen, J. Tümpner, and H. W. Werner (Wiley, Chichester, 1992), p. 563.
[9]M. Wahl, D. Koch, W. Berthold, and A. Wucher (unpublished).
[10]S. Hayashi, in *Proceedings of the 9th International Conference on Secondary Ion Mass Spectrometry, SIMS IX, Yokohama 1993*, edited by A. Benninghoven, Y. Nihei, R. Shimizu, and H. W. Werner (Wiley, Chichester, 1994), p. 339.
[11]J. Schwieters, H.-G. Cramer, T. Heller, U. Jürgens, E. Niehuis, J. Zehnpfenning, and A. Benninghoven, J. Vac. Sci. Technol. A **9**, 2864 (1991).
[12]R. Stein, Int. J. Mass Spectrom. Ion Phys. **14**, 205 (1974).
[13]N. Matsunami, Y. Yamamura, Y. Itikawa, N. Itoh, Y. Kazumata, S. Miyagawa, K. Morita, R. Shimizu, and H. Tawara, Atom. Data Nucl. Data Tables **31**, 1 (1984).
[14]P. Sigmund, Phys. Rev. **184**, 383 (1969).
[15]J.-F. Hennequin, J. Phys. (Paris) **29**, 957 (1968).
[16]P. Lambropoulos and X. Tang, J. Opt. Soc. Am. B **4**, 821 (1987).
[17]S. Hayashi, Y. Hashiguchi, K. Suzuki, T. Ohtsubo, and B. J. McIntosh, Surf. Interface Anal. **17**, 773 (1991).

[18]H. van der Wel, P. N. T. van Velzen, U. Jürgens, and A. Benninghoven, in *Analysis of Microelectronic Materials and Devices,* edited by M. Grasserbauer and H. W. Werner (Wiley, Chichester, 1991), p. 461.
[19]M. Terhorst, R. Möllers, E. Niehuis, and A. Benninghoven, Surf. Interface Anal. **18**, 824 (1992).
[20]F. Schröder-Oeynhausen, B. Burkhardt, L. Wiedmann, A. Schnieders, R. Möllers, F. Kötter, and A. Benninghoven (unpublished).
[21]A. Karen and A. Benninghoven, in *Proceedings of the 9th International Conference on Secondary Ion Mass Spectrometry, SIMS IX, Yokohama 1993,* edited by A. Benninghoven, Y. Nihei, R. Shimizu, and H. W. Werner (Wiley, Chichester, 1994), p. 788.
[22]L. Fabry, S. Pahlke, and L. Kotz, Fres. J. Anal. Chem. **354**, 266 (1996).

Phonon scattering in novel superlattice-asymmetric double barrier resonant tunneling structure

K. Banoo and T. Daniels-Race
Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708-0291

C. R. Wallis and S. W. Teitsworth
Department of Physics, Duke University, Durham, North Carolina 27708-0305

(Received 19 December 1995; accepted 10 May 1996)

The scattering effects (specifically LO-phonon scattering) in a 45 Å AlAs/80 Å GaAs/33 Å AlAs asymmetric double barrier resonant tunneling (ADBRT) structure with a short period GaAS/$Al_{0.3}Ga_{0.7}As$ superlattice incorporated on one side of the double barrier have been studied and characterized. Enhanced levels of current conduction were produced in the ADBRT due to the superlattice miniband electron transport under forward bias. And the effect of the said superlattice on the phonon scattering phenomena exhibited by the entire device was subsequently examined. Magnetic field fan diagrams at 4.2 K under reverse bias showed a new feature at an energy of 22 meV that could be explained on the basis of previously unreported GaAs LO-phonon scattering processes from the first excited emitter level. Finally, quenching of phonon-assisted tunneling in reverse bias on decreasing the period of the superlattice was also observed. © *1996 American Vacuum Society.*

I. MOTIVATION

The concept of double barrier resonant tunneling (DBRT) devices has undergone significant development since the first theoretical work by Esaki and Tsu in 1970 and the first prototype GaAs/$Al_{0.4}Ga_{0.6}As$ resonant tunneling diode in 1974.[1] Since then, DBRT structures have found potential use in many devices such as high-frequency oscillators,[2] resonant tunneling transistors for binary as well as multivalued logic, quantum interference devices,[3] photodetectors, light-emitting diodes, and laser diodes.[4] The models used to analyze and explain the physics behind the unique property of carrier (electron) tunneling have grown more sophisticated, and attempts have been made to include all the various effects that contribute to the experimentally observed current–voltage characteristics—particularly the property of negative differential resistance (NDR).

The numerically calculated values of the peak-to-valley current ratios (PVCR) of DBRT devices are normally higher than the experimentally observed values by an order of magnitude (or more). This is due to the fact that the theoretical models of double barrier structures do not generally take into account all of the real scattering processes during tunneling. Not only do these processes produce lower peak currents but they also have been thought to be the cause of excess valley currents—a phenomenon that has not yet been completely understood. Some phenomenological[5,6] and some highly rigorous models[6,7] have recently achieved moderate success in explaining these processes. Therefore, characterization and study of these phenomena are needed in order to correctly incorporate their effects into the theory of resonant tunneling. The study of phonon scattering is important because it is the dominant scattering mechanism in electron transport.[8,9] The study of phonon scattering effects using asymmetric double barrier (ADBRT) structures is particularly suitable because these structures have been found to exhibit enhanced phonon-assisted tunneling.[10] It has been seen that the PVCRs of symmetric double barrier structures can be increased by engineering their emitter state—e.g., by growing a short-period superlattice in the emitter region.[11] Although ADBRT structures normally do not exhibit high PVCRs, they show enhanced PVCRs when grown with a short period superlattice in their emitter structure.[12] Under this new configuration, however, the question of how phonon scattering is manifested has not been examined. Therefore, this article presents a first-time study of phonon scattering in an asymmetric double barrier structure grown with variable period superlattice on one side of the double barrier. In this configuration, under forward bias, the electron current flows through the superlattice *before* the ADBRT structure—thus the superlattice is in the emitter region of the device. However, under reverse bias, the electron current flows through the superlattice *after* the ADBRT structure—the superlattice now falls in the collector region of the device. The effect of both bias conditions has been studied in this experiment.

II. THEORETICAL BACKGROUND

Phonon-assisted tunneling (PAT) can be explained as the tunneling of electrons with emitter state energies greater than that of the well state by losing the excess energy through the emission of a phonon. These are optical phonons and their dispersion curve exhibits a nonzero wave vector at zero energy. These phonons are primarily longitudinal in character, as shown in Ref. 13. In the $I-V$ characteristics, the presence of satellite peaks (shoulders) beside the main resonant peak has been attributed to PAT.[14] This phonon contribution in the valley current can be further studied under the effect of a

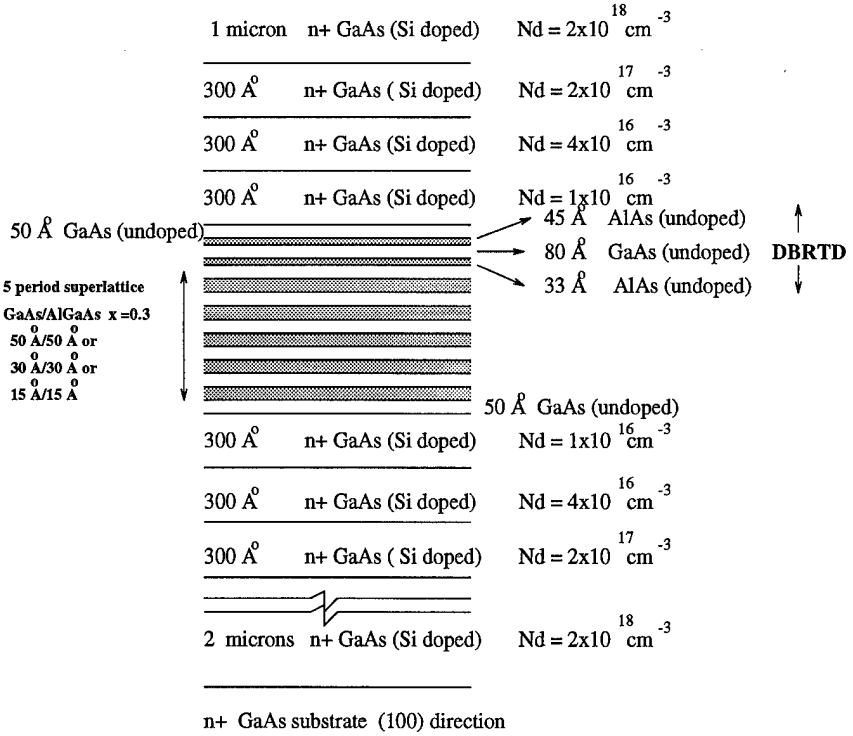


FIG. 1. Layer schematic of the asymmetric DBRTDs with superlattice emitters. Three types of five-period $GaAs/Al_{0.3}Ga_{0.7}As$ superlattices were grown with layer thicknesses of 50, 30, and 15 Å, respectively.

quantizing magnetic field.[8,15] The Schrödinger equation in the presence of a magnetic field **B** in the growth direction z is written as follows:

$$\left(\frac{[P_x-(e/c)BY]^2}{2m^*(z)}+\frac{P_y^2}{2m^*(z)}+\frac{P_z^2}{2m^*(z)}\right)\psi + [E-V(z)]\psi=0, \quad (1)$$

where **P** is the momentum operator and Y is the y component of the position operator.

The total energy is now given as follows:

$$\epsilon=\epsilon_l+(n+1/2)\hbar\omega_c, \quad n=0,1,2\ldots, \quad (2)$$

where $\omega_c=(\hbar e/m^*)B$ is the cyclotron frequency and ϵ_l is the quantum well state.

When phonon emission (scattering) takes place, it can be modeled as a change $p=n-n'$ in the eigenstate of the harmonic oscillator derived above. The difference in the energies of the phonon scattering events and the resonant tunneling events is related to the phonon energy as follows:

$$\Delta\epsilon=\Delta p\hbar\omega_c+\Delta i\hbar\omega_{\mathrm{LO}}, \quad (3)$$

where i is an integer equal to 0 for elastic scattering and 1 for LO phonon scattering and p is an integer corresponding to the change in Landau level due to scattering effects. The resonant tunneling peak in the $I-V$ characteristics corresponds to $i=0$ and $p=0$. (The case $p=1$ in resonant tunneling has also been observed and has been attributed to interface scattering.) Phonon scattering peaks, on the other hand, correspond to $i=1$ and any integral p. Multiphonon peaks are also possible ($i>1$) but no experimental evidence has been found for them as yet.

The energy of the electron is proportional to the total voltage drop across the device (taking the accumulation and depletion drops into account). Thus, the difference $\Delta\epsilon$ is proportional to the difference in the applied voltage (ΔV) between the resonant tunneling and PAT peaks, thus eliminating the accumulation and depletion drops provided the space-charge effects and charge buildup in the well are not large.[8,15] On plotting the applied voltage at the position of PAT peaks versus the magnitude of the magnetic field (B), we obtain experimental fan diagrams that yield the linear relation between $\Delta\epsilon$ and ΔV. In addition, the phonon energies calculated from the fan diagrams (discussed in a later section) represent the different types of phonon scattering processes taking place in these structures.

III. EXPERIMENT

In this study, three sample structures were grown, as shown in Fig. 1. Briefly, each structure consisted of a graded doped emitter region, a variable composition five-period superlattice, the ADBRT structure of an 80 Å GaAs well with 45 and 33 Å AlAs barriers, and a graded doped collector (listed in the order of growth). The superlattice had five periods of 50 Å GaAs/50 Å $Al_{0.3}Ga_{0.7}As$ layers or 30 Å GaAs/30 Å $Al_{0.3}Ga_{0.7}As$ layers or 15 Å GaAs/15 Å $Al_{0.3}Ga_{0.7}As$ layers in the three samples, respectively. Therefore, the structures are essentially distinct from one another

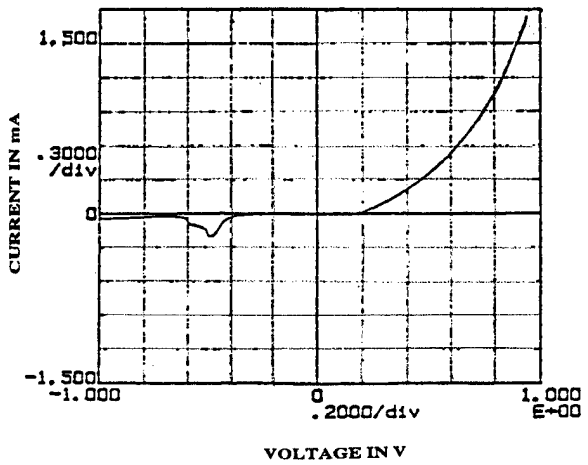


FIG. 2. Measured $I-V$ curve of the ADBRT structure with a 50 Å GaAs/50 Å $Al_{0.3}Ga_{0.7}As$ superlattice at 77 K showing the forward and reverse bias characteristics.

as each has a uniquely designed superlattice grown on one side of the double barrier. As the final step, a 1 μm heavily doped GaAs cap layer was deposited to allow ohmic contact to be made to the device.

$I-V$ characteristics were measured for all the samples described above both at 77 (without magnetic field) and 4 K (with magnetic field). The bias conditions used in these measurements have been defined at the end of Sec. I. In the case of forward bias, the superlattice forms a part of the emitter and in the case of reverse bias it falls in the collector region. All structures probed had a device area (it is necessary to know the area of the device in order to calculate current densities from the absolute values of the measured current) of 1.963×10^{-3} cm^2 defined by the deposition of Ge–Ni–Au contacts. Measurements at 77 K showed resonant tunneling and phonon-assisted tunneling in the reverse bias. None of the samples showed resonant tunneling or phonon-assisted tunneling in the forward bias (see Fig. 2).

Detailed magnetic field measurements were taken at 4.2 K under reverse bias to study the effect of the superlattice on phonon scattering. These measurements were not done in forward bias due to the enhanced forward current conduction through the samples—which has been further described in the results.

Under reverse bias the effect of the quantizing magnetic field on phonon-assisted tunneling in the 50 Å superlattice sample can clearly be seen in Fig. 3. The position of the peaks in the $I-V$ characteristics is obtained by calculating the second derivative d^2I/dV^2 to more clearly distinguish the features in current variation, as shown in Fig. 4. Thus the positions of the phonon peaks can be noted simply from the position of the maxima in the second derivative. These calculations are repeated for each $I-V$ curve obtained with different magnitudes of **B**, and the evolution of the phonon peaks with magnetic field is plotted as a fan diagram. The fan diagram of the ADBRT structure with a 50 Å GaAs/50 Å $Al_{0.3}Ga_{0.7}As$ superlattice is shown in Fig. 5. Three fans are observed starting at ΔV of 196, 343, and 489 mV, respectively, from the main resonant peak. These correspond to energies $\Delta\epsilon$ of 20.2, 36.7, and 47.3 meV using the numerical method described in Ref. 16. Here we clearly see evidence of GaAs LO-phonon scattering ($\hbar\omega_{LO}\approx36$ meV) and AlAs LO-phonon scattering ($\hbar\omega_{LO}\approx50$ meV) events.[15]

Phonon-assisted tunneling in the ADBRT structure with a 30 Å GaAs/30 Å $Al_{0.3}Ga_{0.7}As$ superlattice is illustrated in its fan diagram (Fig. 6). Two fans are observed starting at 184.8 and 351.2 mV from the main resonant peak which yields $\Delta\epsilon$ of 21.9 and 36.6 meV, respectively. Again $\hbar\omega_{LO}\approx36$ meV corresponds to GaAs LO-phonon scattering events. Likewise, phonon-assisted tunneling in the ADBRT structure with a 15 Å GaAs/15 Å $Al_{0.3}Ga_{0.7}As$ superlattice is illustrated in its fan diagram (Fig. 7). Two fans are observed starting at 137 and 254 mV yielding $\Delta\epsilon$ of 22.3 and 33.9 meV, respectively. The presence of GaAs LO-phonon scattering is shown by the second fan.

In each case, an anomalous fan of energy around 22 meV

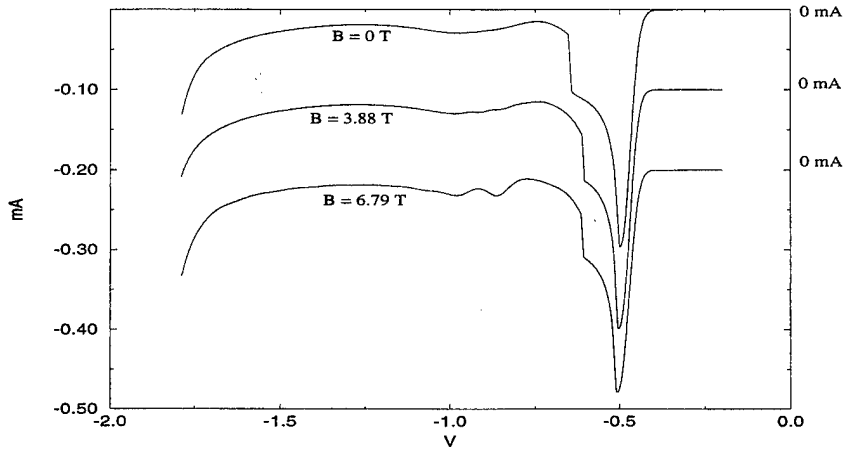


FIG. 3. Reverse bias $I-V$ characteristics of the ADBRT structure with a 50 Å GaAs/50 Å $Al_{0.3}Ga_{0.7}As$ superlattice at 4.2 K for magnetic fields varied from 0 to ≈7 T.

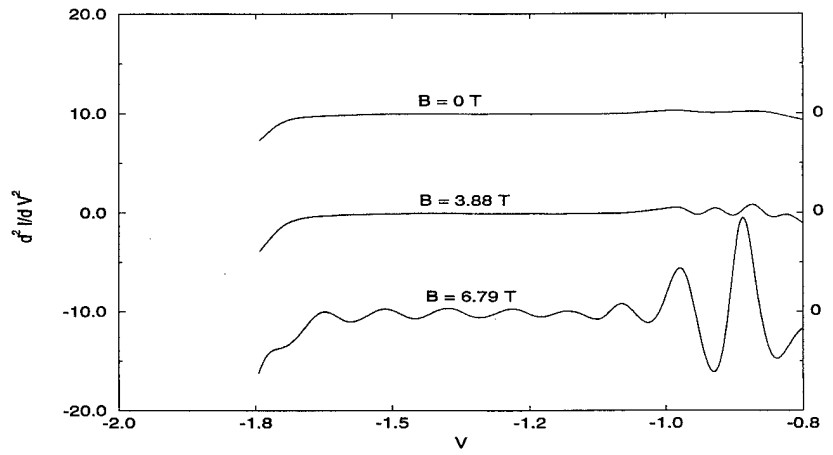


FIG. 4. Plots of d^2I/dV^2 of the ADBRT structure with a 50 Å GaAs/50 Å $Al_{0.3}Ga_{0.7}As$ superlattice in reverse bias at 4.2 K with different magnetic fields.

(20.2, 21.9, and 22.3 meV for the 50, 30, and 15 Å samples, respectively) was observed. These fan energies have not been previously reported in the literature. The presence of an anomalous line in the fan diagram may be explained as an artifact of the mathematical analysis of the data. It is possible for the second derivative of the current to show a local maximum when the phonon-assisted tunneling starts to contribute to the valley current. These peaks in the second derivative would therefore appear as a line in the fan diagram. However, the anomalous effect observed in our study was a *collection* of lines that showed the characteristic structure of a fan exhibiting an energy of 22 meV. Clearly, this indicates a phonon scattering process. The difference in energy can be explained by considering PAT not from the ground emitter energy level but from the first excited emitter level into the well state[17] when the superlattice is in the collector. It is known that the energies of the incoming electrons are those of the bound states formed due to band bending in the emitter. By accounting for the energy of the first excited state in the emitter of the ADBRT structure (≈12 meV with respect to the ground state in the emitter), the phonon energy obtained is 34 meV—which is evidently the GaAs LO-phonon energy. Therefore, this analysis strongly suggests the occurrence of phonon scattering events from the *first* excited state of the emitter into the well state. Such an effect is almost insignificant in the absence of a superlattice in the collector.

IV. DISCUSSION

Realistically, the resonant tunneling process through a complex structure such as a superlattice-ADBRT diode is extremely complicated.[18] It occurs under the alignment of the energy levels in different regions of the entire structure—a critical condition that is affected by many factors besides the applied bias. First, the energy level configuration of the entire system has a crucial dependence on the quality of the interface (its roughness) and fluctuations in layer thicknesses. Variations in these parameters tend to prevent the formation of miniband and localize the electron wave functions in the wells.[19] The electron tunneling in such

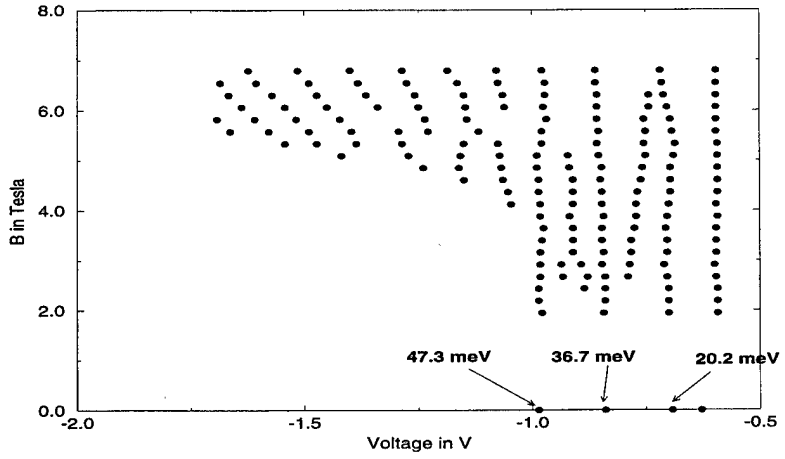


FIG. 5. Fan diagram of the phonon peaks of the ADBRT structure with a 50 Å GaAs/50 Å $Al_{0.3}Ga_{0.7}As$ superlattice calculated at 4.2 K. The arrows indicate phonon energies of the respective fans obtained after calculating them from the voltage intercept.

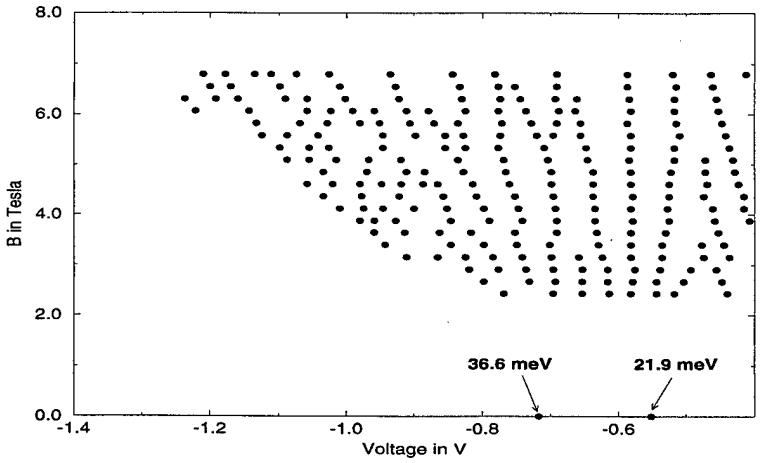


FIG. 6. Fan diagram of the phonon peaks of the ADBRT structure with a 30 Å GaAs/30 Å $Al_{0.3}Ga_{0.7}As$ superlattice calculated at 4.2 K. The arrows indicate phonon energies of the respective fans obtained after calculating them from the voltage intercept.

a case will be incoherent (or sequential) if at all. Another aspect to consider is the coupling between the superlattice and the ADBRT structure. From the measured reverse bias characteristics of the device, we can conclude that the superlattice and the ADBRT structure are essentially decoupled. (For example, this decoupling is very important in the use of the superlattice as an electron launching mechanism, independent of the operation of the ADBRT diode.)

Under forward bias, the absence of NDR may be because of increased current levels due to collective transport through the superlattice minibands. If the NDR peaks at a current much higher than the current measuring capability of our apparatus then only the *onset* of resonant tunneling in the forward direction can be observed as an exponentially increasing $I-V$ curve. This is seen to be the case for the $I-V$ characteristics of all our samples in the forward bias.

Under reverse bias, all the samples show resonant tunneling which is to be expected as the ADBRT structure has a predominant effect on the electron current under the reverse bias configuration. The presence of the superlattice in the collector region makes the "effective" collector of the double barrier well weakly transmitting. Small collector transmissivity has been found to produce enhanced scattering in the ADBRT structure[20] and, therefore, this could explain the prominent PAT observed in the reverse bias in our samples. Enhanced scattering could also explain the indications of phonon scattering from the first excited emitter level that has not been observed previously.

The fan diagrams also show decreasing levels of PAT in reverse bias with decreasing layer thickness of the $GaAs/Al_{0.3}Ga_{0.7}As$ superlattice. Therefore, it can be concluded that an optimally thin period superlattice can quench PAT in the reverse bias while producing high PVCRs in the forward bias.

V. CONCLUSIONS

We studied the $I-V$ characteristics of an ADBRT structure with three superlattice structures of different periods in

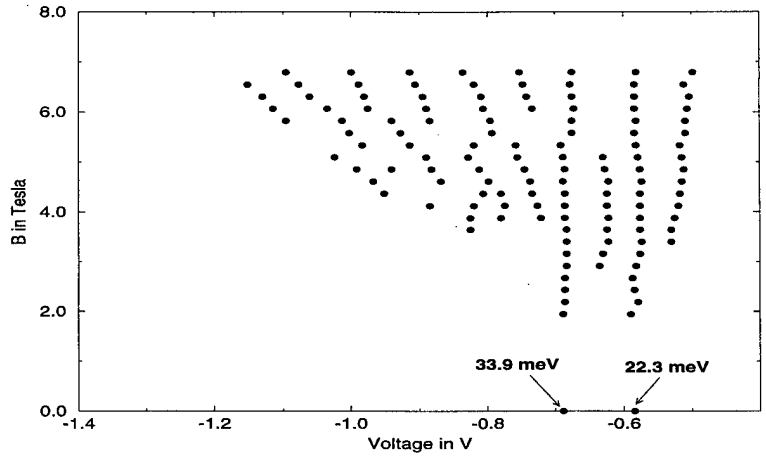


FIG. 7. Fan diagram of the phonon peaks of the ADBRT structure with a 15 Å GaAs/15 Å $Al_{0.3}Ga_{0.7}As$ superlattice calculated at 4.2 K. The arrows indicate phonon energies of the respective fans obtained after calculating them from the voltage intercept.

the emitter as well as the collector. Collective transport through the superlattice minibands in the emitter was found to produce enhanced current conduction in the forward bias and thus no NDR could be observed. Resonant tunneling and PAT were observed in the reverse bias with the quenching of PAT as the period of the superlattice in the collector was decreased. Magnetic field studies of the reverse bias PAT exhibited a new feature in the fan diagram that could be explained on the basis of phonon scattering processes from the first excited emitter level.

ACKNOWLEDGMENTS

The authors wish to thank Richard Kendall for technical assistance. This work has been supported by NSF Grant No. DMR 9208381.

[1]L. L. Chang, L. Esaki, and R. Tsu, Appl. Phys. Lett. **24**, 593 (1974).
[2]T. C. L. G. Sollner, *Physics of Quantum Electron Devices*, edited by F. Capasso (Springer, New York, 1990).
[3]F. Capasso and S. Datta, Phys. Today, Vol. 43, No. 2, February (1990), pp. 74–82.
[4]F. Capasso, Nucl. Instrum. Methods A **265**, 112 (1988).
[5]T. G. Van de Roer, *Negative Differential Resistance and Instabilities in 2D Semiconductors*, edited by N. Balkan, B. K. Ridley, and A. J. Vickers (Plenum, New York, 1993).
[6]H. Mizuta and T. Tanoue, *The Physics and Applications of Resonant Tunneling Diodes* (Cambridge University Press, Cambridge, 1995).
[7]R. Lake, G. Klimeck, M. P. Anantram, and Supriyo Datta, Phys. Rev. B **48** 15132 (1993).
[8]P. J. Turley, C. R. Wallis, S. W. Teitsworth, and P. K. Bhattacharya, Phys. Rev. B **47**, 12640 (1993).
[9]R. A. Davies, D. J. Newson, T. G. Powell, M. J. Kelly, and H. W. Myron, Semicond. Sci. Technol. **2**, 61 (1987).
[10]M. L. Leadbeater, E. S. Alves, F. W. Sheard, L. Eaves, M. Henini, O. H. Hughes, and G. A. Toombs, J. Phys. Condensed Matter **1**, 10605 (1989).
[11]R. E. Pritchard, P. C. Harness, L. Cury, J. C. Portal, B. Khamsehpour, W. S. Truscott, and K. E. Singer, Semicond. Sci. Technol. **6**, 626 (1991).
[12]K. Banoo and T. Daniels-Race, in *IEEE Proceedings of the SouthEast Conference* (IEEE, New York, 1995).
[13]P. J. Turley and S. W. Teitsworth, Phys. Rev. B **50**, 8423 (1994).
[14]V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Phys. Rev. B **36**, 7635 (1987).
[15]M. L. Leadbeater, E. S. Alves, L. Eaves, M. Henini, O. H. Hughes, A. Celeste, J. C. Portal, G. Hill, and M. A. Pate, Phys. Rev. B **39**, 3438 (1989).
[16]P. J. Turley, Ph.D. thesis, Duke University (1994); see Chap. 6.
[17]P. J. Turley (private communication).
[18]W. Pötz, J. Appl. Phys. **66**, 2458 (1989).
[19]F. Capasso, K. Mohammed, and A. Y. Cho, IEEE J. Quantum Electron. **QE-22**, 1853 (1986).
[20]R. Lake, G. Klimeck, and S. Datta, Phys. Rev. B **47**, 6427 (1993).

Facet generation during molecular beam epitaxy of GaAs/AlGaAs multilayers on GaAs (001) patterned substrates

T. Takebe,[a] M. Fujii,[b] T. Yamamoto,[c] K. Fujita, and T. Watanabe
ATR Optical and Radio Communications Research Laboratories, 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

(Received 15 March 1996; accepted 10 May 1996)

Extra facet generation during molecular beam epitaxy of GaAs/AlGaAs multilayers on (001) GaAs substrates patterned with stripes running in the [$\bar{1}$10], [110], and [100] directions and having various slopes, designated as "[$\bar{1}$10] stripe," "[110] stripe," and "[100] stripe," respectively, has systematically been investigated for the first time. It has been confirmed that extra (111)A and (114)A facets are generated on the [$\bar{1}$10] stripes, extra ($\bar{1}$11)B and ($\bar{1}$13)B facets on the [110] stripes, and extra (031), (011), (045), and (013) facets on the [100] stripes, depending on the intersection angle θ of the sidewall and the substrate plane. Among them, the (114)A, ($\bar{1}$13)B, and (013) facets are important because they persist over a wide range of θ and, therefore play a detrimental role in forming flat and uniform sidewall layers on the (001) patterned substrates. It has been made clear that no extra facets are generated on the sidewalls and flat and uniform layers maintaining the initial as-etched stripe patterns can be grown for the [$\bar{1}$10] stripes with $10° \geqslant \theta$, the [110] stripes with $24° \geqslant \theta \geqslant 20°$, and the [100] stripes with $19° \geqslant \theta$. Extra (111)A-related facets have developed on the (111)A-related intersection of the equivalent [100] and [010] stripes, while no extra facets have developed on the ($\bar{1}$11)B-related intersection. The facet generation behavior has been compared between the (001) patterned substrates and previously studied (111)A patterned substrates. © *1996 American Vacuum Society.*

I. INTRODUCTION

In our recent article, facet generation behavior during molecular beam epitaxy (MBE) of GaAs/AlGaAs multilayers on GaAs (111)A substrates patterned with ridge-type triangles was investigated for the first time, and how the behavior could be utilized for the formation of novel device structures on the (111)A patterned substrates under the sidewall control was elucidated.[1] In this respect, there are many articles published on crystal growth and microstructure formation on (001) patterned substrates.[2–26] Almost all of these articles have dealt with the growth on stripes running in the [$\bar{1}$10] direction with an exact (111)A sidewall[2–6,10,11,13,15,17,22,23] or (112)A–(113)A sidewalls.[7–9,14,26] Several articles have discussed the growth on stripes running in the [110] direction[2,3,5,11,12,16,23,24] and very recently on stripes running in the [100] direction with the (010) sidewall.[18–21,25] To the best of our knowledge, however, no articles have published systematically studying growth on stripes with a wide range of sidewall slopes or on intersections of two stripes running in the orthogonal directions from a point of view of facet generation control for device application.

In the present article, facet generation during MBE of GaAs/$Al_{0.3}Ga_{0.7}As$ multilayers on the stripes running in the [$\bar{1}$10], [110], and [100] directions with various slopes and their intersections on (001) substrates is investigated for the first time. By comparing the present results with those on (111)A patterned substrates, clearer image and deeper understanding of crystal growth on nonplanar substrates are gained on the basis of the differences in the surface properties, the position of a sidewall relative to the substrate plane, and the incident effective molecular beam fluxes.

II. EXPERIMENT

Stripes running in the [$\bar{1}$10], [110], and [100] directions with heights of 5–7 μm whose sidewalls are composed of (111)A-related, ($\bar{1}$11)B-related, and (010)-related surfaces, respectively, were formed on the (001) substrates using photolithography and selective etching techniques. They are briefly designated as [$\bar{1}$10] stripe, [110] stripe, and [100] stripe, respectively. These stripes are schematically shown with respect to the configuration of the (001) surface in Fig. 1. The [010] stripe is a mirror image of the [100] stripe with respect to the ($\bar{1}$10) plane and has sidewall orientations equivalent to those of the [100] stripe. The intersection angle θ of the sidewall and the (001) substrate plane was varied in a wide range using H_2O_2-excess $HF+H_2O_2+H_2O$ mixtures of various compositions.[27] After pregrowth chemical and thermal treatments identical to those for the (111)A patterned substrates, the (001) patterned substrates were subjected to the same MBE growth runs as the (111)A patterned substrates discussed in Ref. 1: The Ga and Al cells were located side by side and the As_4 cell was located opposite to them with respect to the sample substrate. All the molecular beams were incident on the substrate at an angle of 60°. Five pairs

[a]Present address: Semiconductor Materials R&D Department, Itami Research Laboratories, Sumitomo Electric Industries, Ltd., 1-1-1 Koya-kita, Itami, Hyogo 664, Japan; Electronic mail: takebe@ritami.sumiden.co.jp
[b]Present address: Fundamental Research Department, Central Research Laboratory, Technology Division, Mitsubishi Cable Industries, Ltd., 4-3 Ikejiri, Itami, Hyogo 664, Japan.
[c]Present address: Development Section I, Devices Development Department, Technical Administration Division, Murata Manufacturing Co., Ltd., 2288 Oshinohara, Yasu-cho, Yasu-gun, Shiga 520-23, Japan.

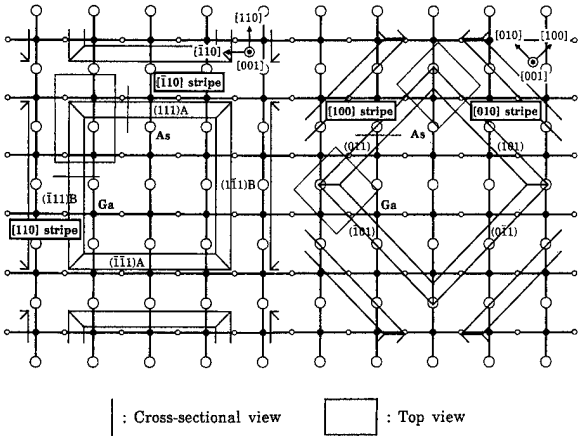


FIG. 1. Schematic presentation of [$\bar{1}$10], [110], [100], and [010] stripes formed on the GaAs (001) surface.

of undoped 0.2-μm-thick GaAs/0.2-μm-thick $Al_{0.3}Ga_{0.7}As$ layers, and an undoped 0.3-μm-thick GaAs cap were successively grown by MBE on the (001) patterned substrates at a substrate temperature of 620 °C, an As_4 pressure of 3.3×10^{-5} Torr, a V/III flux ratio of 7.4 for GaAs and 6.2 for AlGaAs, and a substrate rotation speed of 60 rpm. The growth rate was 0.77 μm/h for GaAs and 1.10 μm/h for AlGaAs, corresponding to 0.76 and 1.08 ML/s, respectively. The AlGaAs layers were used as "markers" in order to observe how the growth proceeded. The layers grown on the stripes were closely examined by scanning electron microscopy (SEM) at the positions shown in Fig. 1. No variations in the as-etched sidewall profile, growth rate, and aftergrowth surface morphology were confirmed across the whole sample (20×25 mm^2) for any of the samples investigated.

III. RESULTS

Although the epilayer–substrate interface was not clearly observed since the first layer grown was GaAs, a definite interface maintaining the initial etching profile was confirmed on Si-doped GaAs single layers grown under similar conditions and deeply stain-etched.

A. [$\bar{1}$10] stripe

Figure 2 shows ($\bar{1}$10) cross-sectional views of the aftergrowth profiles of the (111)A-related sidewalls of the [$\bar{1}$10] stripes with various values of θ. The surface index (lmn) corresponding to each value of θ is also shown (within ±2°) with l, m, $n\leqslant9$. The brighter layers correspond to GaAs and the darker layers to AlGaAs. Generation of extra facets on the sidewalls is described as follows: (a)–(d) For $\theta>21°$, two extra facets with facet angles $\theta_f=21°$ and 55° to the (001) substrate plane were identified and were assigned to (114)A and (111)A, respectively. The geometrical value of θ_f for the (114)A facet is 19° and not 21° measured in the present study. The value of 21° was obtained by averaging the values measured over all the samples with the (114)A facet. The scatter of the measured values was very small, within ±1°. The value of θ_f for the (114)A facet on the (111)A patterned

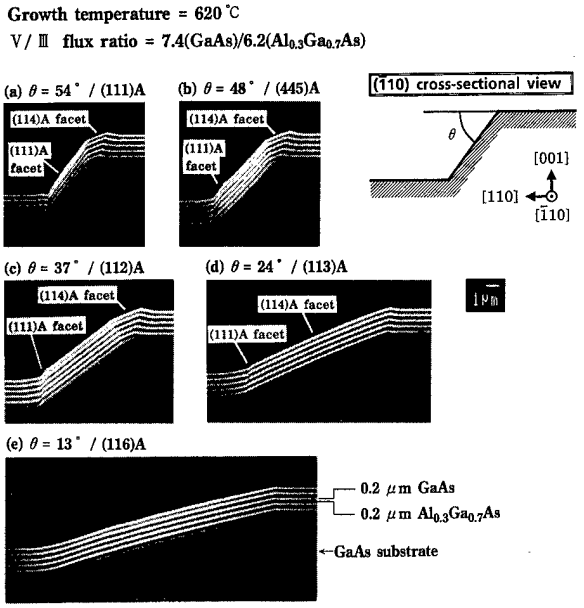


FIG. 2. ($\bar{1}$10) cross-sectional views of the aftergrowth profiles of the (111)A-related sidewalls of the [$\bar{1}$10] stripes with various slopes.

substrates grown simultaneously with the (001) substrates was 33°±1° [to the (111)A substrate plane] and not the geometrical value of 35°.[1] The sum of these two values is 54°, just corresponding to the angle of the (001) plane to the (111)A plane. The source of the deviation is not clear but it is possible that the (114)A facet misoriented by 2° towards the [111]A direction is really grown, for example, in a step-flow mode. The values of θ_f for the other facets evaluated in the same way were, except for the ($\bar{1}$11)B facet discussed below, in accordance with the geometrical values within the uncertainty in evaluation mentioned above. The measured values of θ_f are used and the present (114)A facet misoriented by 2° towards the [111]A direction is simply expressed as the (114)A facet in the present article.

As θ approached 21°, θ_f of the (114)A facet, the (114)A facet developed more towards the sidewall, while the (111)A facet receded towards the lower (001) substrate plane.

(e) For $\theta\leqslant21°$, both (114)A and (111)A facets disappeared and no extra facets grew, leading to a flat and uniform sidewall that maintained the initial as-etched profile.

Table I summarizes the facet generation behavior for the [$\bar{1}$10] stripe. In order to obtain flat and uniform layers on the sidewall that maintain the initial as-etched pattern, it is necessary to keep θ smaller than 21°. This result is close to that suggested in Refs. 7 and 8 under a higher growth temperature of 720 °C and a lower V/III flux ratio of 1, which indicates that (114)A facet generation cannot be well controlled by the growth conditions.

B. [110] stripe

Figure 3 shows (110) cross-sectional views of the aftergrowth profiles of the ($\bar{1}$11)B-related sidewalls of the [110] stripes with various values of θ. The $HF+H_2O_2+H_2O$ mixtures used produced an inverted ($\bar{1}\bar{1}$1)A-related mesa to-

TABLE I. Summary of facet generation behavior during MBE on GaAs (001) patterned substrates.

Substrate	Sidewall slope	Extra facet on sidewall		Extra facet on corner	
	[$\bar{1}$10] stripe	Slope	Orientation	Slope	Orientation
	≥21°	55°	(111)A		
		21°	(114)A[a]		
	<21°	No facets			
	[110] stripe	Slope	Orientation	Slope	Orientation
	≥53°	53°	($\bar{1}$11)B[b]		
(001)	53°~24°	24°	($\bar{1}$13)B		
	<24°	No facets			
	[100] stripe	Slope	Orientation	Slope	Orientation
	≥45°	72°	(031)	No facets	
		45°	(011)		
	45°–19°	38°	(045)	No facets	
		19°	(013)		
	<19°	No facets		≤13°	(116)A, etc.

[a]Actually, (114)A misoriented by 2° towards the [111]A direction.
[b]Actually, ($\bar{1}$11)B misoriented by 2° towards the [001] direction.

gether with the normal ($\bar{1}$11)B-related sidewall for large θ.[27] Generation of extra facets on the sidewalls is described as follows: (a),(b) For $\theta \geq 53°$, an extra ($\bar{1}$11)B facet with $\theta_f = 53°$ was observed at the edge of the inverted mesa. For this facet, the measured value of θ_f is smaller by 2° than the geometrical value of 55°, that is, the facet is really misoriented by 2° towards the [001] direction. This may be related to the spontaneous evolution of the ($\bar{1}\bar{1}\bar{1}$)B facets tilted by 2.5° towards the [$\bar{1}\bar{1}$2]A direction during MBE of GaAs on exact ($\bar{1}\bar{1}\bar{1}$)B substrates.[28] The situation is similar to the (114)A facet but θ_f of the (11$\bar{1}$)B facet on the (111)A patterned substrates was exactly the geometrical value of 71°.[1] The discrepancy is still unclear.

(c)–(e) For $53° > \theta \geq 24°$, the ($\bar{1}$11)B facet disappeared and an extra ($\bar{1}$13)B facet with $\theta_f = 24°$ developed as θ decreased.

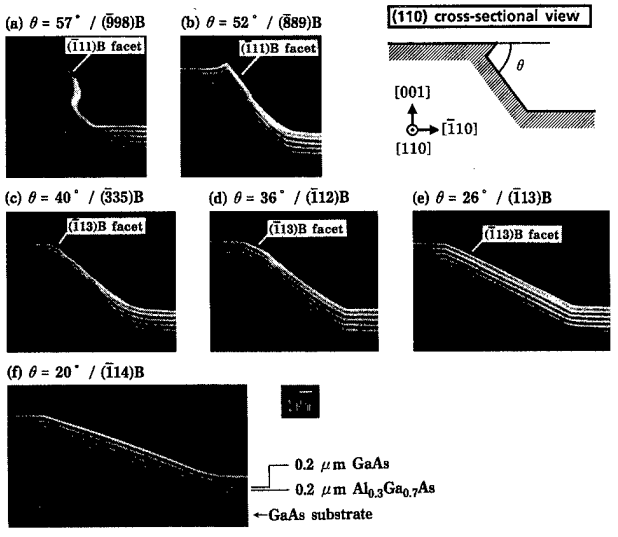


FIG. 3. (110) cross-sectional views of the aftergrowth profiles of the (111)B-related sidewalls of the [110] stripes with various slopes.

(f) For $24° > \theta$, the ($\bar{1}$13)B facet disappeared and no extra facets were observed.

The generation of the ($\bar{1}$11)B facet on the [110] stripe with an inverted mesa has been reported in the literature under different growth conditions.[2,12,16,24] The present study provided the first systematic data on the growth behavior for the [110] stripes over a wide θ range, especially for those without inverted mesas.

Table I summarizes the facet generation behavior for the [110] stripe. In order to obtain layers on the sidewall that maintain the initial as-etched pattern, it is necessary to keep θ smaller than 24°. In fact, taking the surface morphology of the sidewall layers into account, the usable θ range is further restricted as discussed below.

C. [100] stripe

Figure 4 shows (110) cross-sectional views of the aftergrowth profiles of the sidewalls of the [100] stripes with various values of θ. The $HF+H_2O_2+H_2O$ mixtures used produced round intersection profiles between the sidewall and the lower substrate plane for large θ. Since the [100] stripe does not intersect the (110) cleavage plane at right angles, as shown in Fig. 1, the actual value of θ was evaluated from the experimentally observed apparent value θ' on the SEM photographs using the equation $\theta = \tan^{-1}(\tan\theta'/\cos 45°)$. Generation of extra facets on the sidewalls is described as follows: (a)–(c) For $\theta > 45°$, extra (031) and (011) facets with $\theta_f = 72°$ and 45° were clearly identified. As θ became closer to 45°, the (011) facet developed more towards the sidewall, while the (031) facet diminished and finally disappeared.

(d)–(f) For $45° \geq \theta > 19°$, the (011) facet disappeared and an extra (013) facet with $\theta_f = 19°$ appeared. As θ approached 19°, the (013) facet also developed more towards the sidewall. The generation of an extra (045) facet with $\theta_f = 38°$, as shown in (e), is interesting in that the facet exists on the sidewalls with $\theta < 38°$, that is, θ_f of the facet, whereas the other facets observed for $\theta \geq \theta_f$ vanish for $\theta < \theta_f$, except for the (111)A facet. We confirmed from a close inspection of

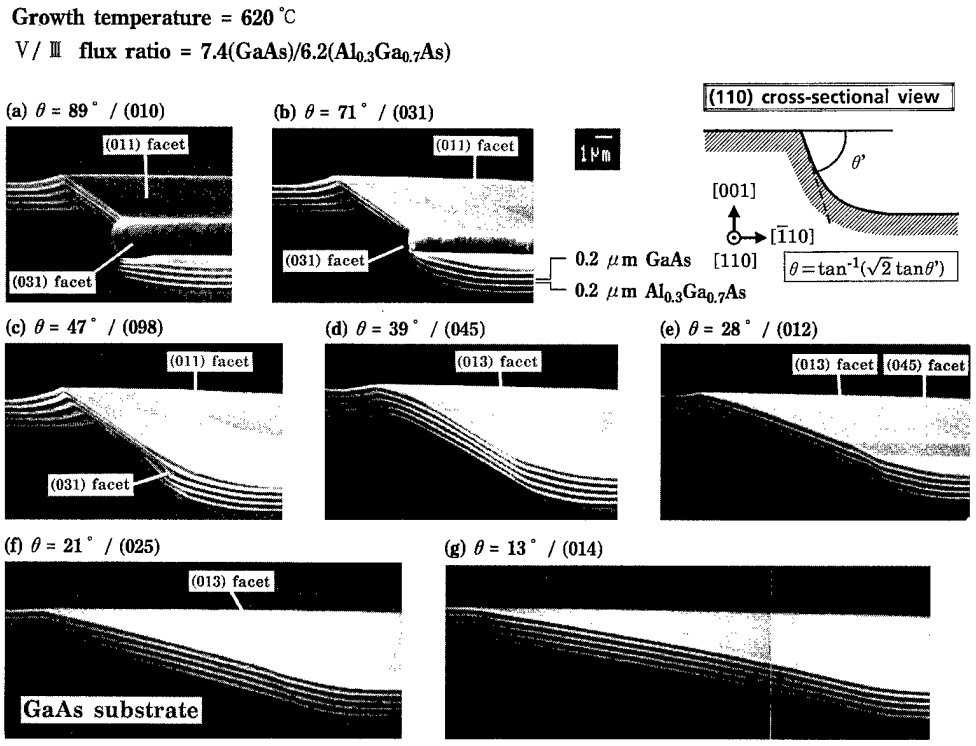


FIG. 4. (110) cross-sectional views of the aftergrowth profiles of the (010)-related sidewalls of the [100] stripes with various slopes.

the SEM photographs that the (045) facet actually exists for $\theta=43°$ (the growth profile was similar to that for $\theta=39°$ and not shown here), 39°, and 21° [(d) and (f) of Fig. 4], that is, the (045) facet accompanies the (013) facet.

(g) For $\theta\leqslant 19°$, the (013) facet [and the (045) facet] also disappeared and no other extra facets were observed.

Table I summarizes the facet generation behavior for the [100] stripe. In order to maintain the initial as-etched profile of the sidewall after growth, it is necessary to keep θ smaller than 19°. It was, however, impossible to obtain flatness and uniformity of the sidewall layers comparable to those for the [$\bar{1}$10] stripes with $\theta\leqslant 21°$, probably due to the high-index nature of the sidewalls.

D. Intersections of orthogonal stripes

In Ref. 1, the growth behavior on the corners of ridge-type triangles on the (111)A substrates, in other words, on the intersection of two crystallographically equivalent sidewalls, was investigated. Since the (001) surface has twofold rotational symmetry (Fig. 1), it will be interesting to study the growth behavior on the intersection of the nonequivalent [$\bar{1}$10] and [110] stripes and on the two nonequivalent intersections of the equivalent [100] and [010] stripes, that is, the "(111)A-related corner" and "($\bar{1}$11)B-related corner."

Figure 5 shows top views of the aftergrowth profiles of the intersections of the [$\bar{1}$10] and [110] stripes with various values of θ. Generation of extra facets on the corners was observed. Although they could not be definitely identified due to the nonflat and different as-etched profiles of the orthogonal sidewalls and to the occurrence of complicated sidewall profiles at the intersection,[27] the following assignments will be reasonable from the discussion in the previous section: the extra facet that appeared for (a),(b) seems to be (011) and the extra facet that appeared for (c),(d) seems to be (013). No extra facets are observed for (e),(f).

Figure 6 shows top views of the aftergrowth profiles of the intersections of the [100] and [010] stripes with various values of θ. There was no difference observed between the as-etched (111)A-related and ($\bar{1}$11)B-related corners. Poor morphologies of the sidewall layers compared to those for the [$\bar{1}$10] stripes, the equivalency of the growth behaviors on the [100] and [010] stripes and, to be in sharp contrast to this, the nonequivalency of the growth behaviors on the (111)A-related and ($\bar{1}$11)B-related corners can be well confirmed from Fig. 6. It should be especially noted in (d),(e) that extra facets are generated on the (111)A-related corners but not on the ($\bar{1}$11)B-related corners.

IV. DISCUSSION

A. (114)A and (111)A facets

The generation of the (114)A facet on the [$\bar{1}$10] stripe has also been reported in the literature under various growth conditions but only for exact (111)A[2,3,5,6,13,22] and (112)A–(113)A[7–9,14,26] sidewalls. In Ref. 1, the (114)A facet was shown to also develop on the (111)A patterned substrates. Therefore, it was confirmed for the first time in the present study that the generation of the (114)A facet is a common feature independent of the substrate and sidewall orientations. The development of the (114)A facet towards

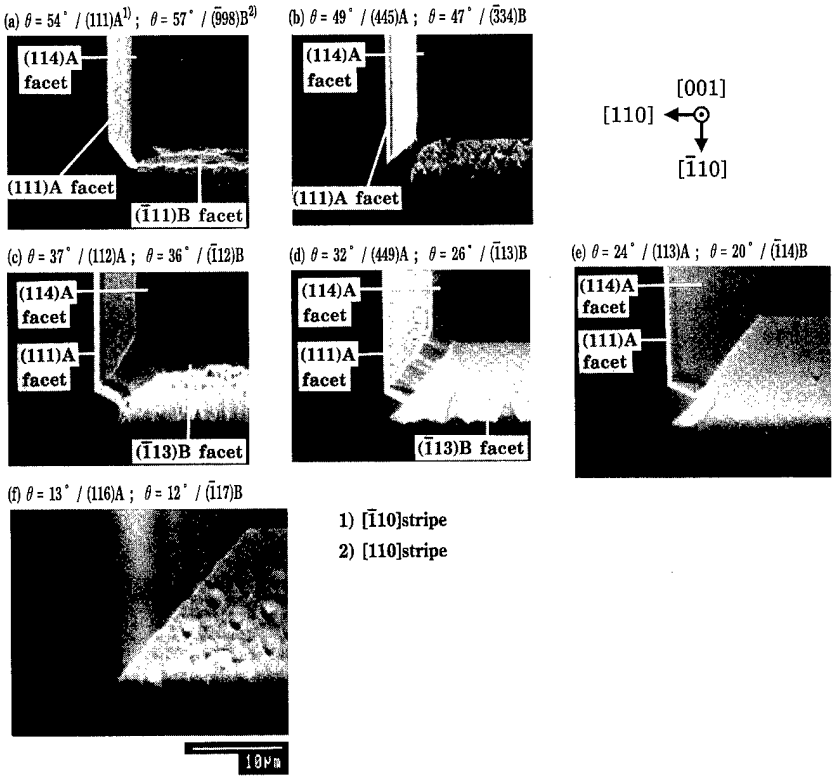


FIG. 5. Top views of the aftergrowth profiles of the intersections of the [1̄10] and [110] stripes with various sidewall slopes.

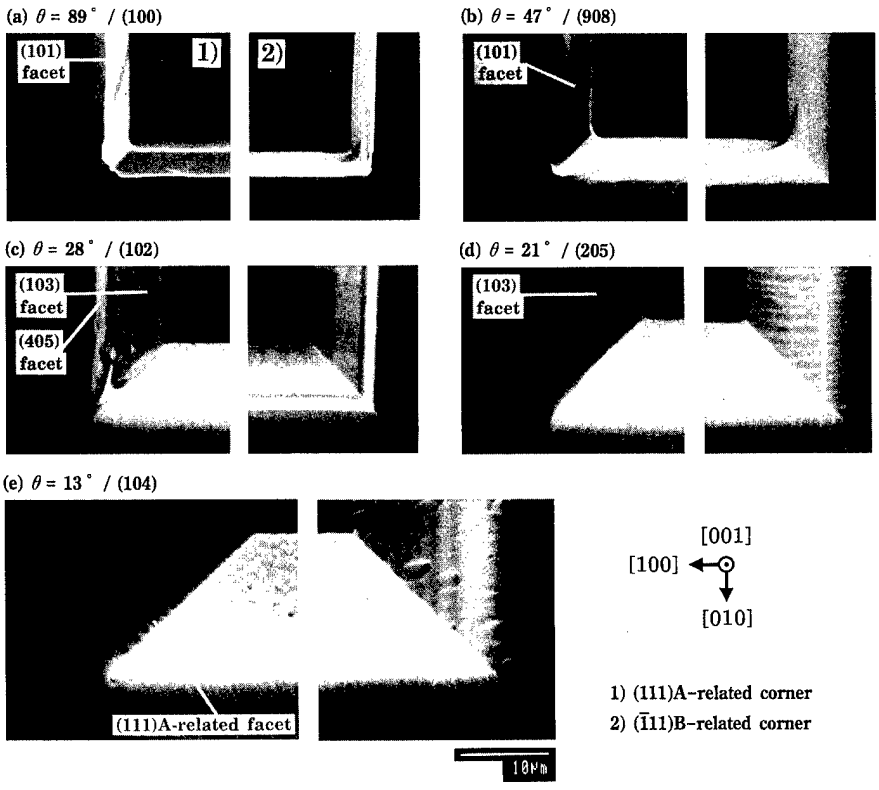


FIG. 6. Top views of the aftergrowth profiles of the intersections of the [100] and [010] stripes with various sidewall slopes.

the sidewall with decreasing θ [Figs. 2(a)–2(d)] was also observed on the (111)A patterned substrates.[1] As was mentioned in Ref. 1, it seems to be caused by enhancement of the lateral growth due to the increase in the terrace width of microsteps formed on the as-etched sidewall.

For $54° \geqslant \theta \geqslant 48°$, the layers grown on the sidewall exhibited step structures composed of the equivalent (011)-related and (101)-related facets. The average interstep distance shortened from 1.5 to 0.5 μm as θ decreased from 54° to 48°. We also observed an Al composition variation (seen as a lateral gradation in the AlGaAs layers) corresponding to the step structures. Since these sidewalls correspond to the (111)A planes misoriented by 1°–7° towards the [001] direction, it is natural that the sidewall layers exhibited the step structures identical with those grown on the planar (111)A substrates misoriented towards the [001] direction.[29] For $41° \geqslant \theta \geqslant 21°$, the surface morphology of the sidewall layers was poor except for the (114)A facet showing a smooth surface. After vanishing the (114)A facet for $\theta \leqslant 21°$, the sidewall layer exhibited a smooth surface suitable for device application.

The development of the (111)A facet at the intersection of the sidewall and the lower substrate plane was mentioned in Ref. 26 and is obvious in the SEM photographs shown in, for example, Refs. 7, 8, and 14. The reason for the appearance of the (111)A facet for θ smaller than θ_f of the facet is the low growth rate in the [111]A direction due to the long Ga diffusion length on the (111)A plane and the location at the concave intersection of the substrate and the sidewall.

In contrast to the (001) triangles on (111)A substrates,[1] the exponential thickness variation of the layers grown on the (001) substrate plane persists after flat and uniform layers on the sidewall are grown for $\theta \leqslant 21°$, as shown in Figs. 2(e) and 5(f), due to the lateral Ga diffusion from the sidewall to the (001) substrate plane. Exponential thickness variations of the sidewall layers adjacent to the (111)A facet on the basis of the same mechanism are also observed in Figs. 2(b) and 2(c). In order to obtain flat and uniform layers across the sidewall–substrate plane intersection, θ must actually be smaller than 10°. This point is critically important for device application, such as lateral p–n junctions and carrier confinement structures using the amphoteric Si doping characteristics,[4,7,8,26,30–33] because they need a steep p–n transition across the junction. For such small angles, that is, a small difference in the orientation between the substrate plane and the sidewall, steep p–n transitions will not be easily obtained without good consideration of growth conditions. This point has been ignored so far. For (111)A patterned substrates, combination of the (111)A substrate plane and the (113)A sidewall (θ=30°) gave rise to a simple p–n junction with a planar interface and without extra facets or exponential thickness variations, which showed excellent current blocking characteristics.[1,30–33]

B. ($\bar{1}11$)B and ($\bar{1}13$)B facets

It should be noted that the layers grown on the sidewalls with θ around θ_f of the ($\bar{1}11$)B plane exhibited a very rough surface morphology, whereas the ($\bar{1}11$)B facet with a relatively smooth surface grew at the edge of the (001) surface without underlying ($\bar{1}11$)B surfaces [Fig. 5(a)]. One reason for that is because the present growth conditions of high temperature and high As_4 pressure probably did not match the growth on the ($\bar{1}11$)B surface favoring high temperatures, low As_4 pressures, and low growth rates.[34–41] The round as-etched profiles of the sidewalls[27] may be another source of generating such a faceted surface. As to the ($\bar{1}11$)B facet, a lateral flow of excess Ga adatoms not incorporated on the initially formed ($\bar{1}11$)B plane to the substrate plane resulted in the low growth rate in the [$\bar{1}11$]B direction, hence the formation of the ($\bar{1}11$)B facet with a relatively smooth surface. In Ref. 1, the ($11\bar{1}$)B facet has been reported to also develop on the (111)A patterned substrates.

It is observed in Figs. 5(c) and 5(d) that the ($\bar{1}13$)B facet showed a smooth surface and it develops towards the sidewall with the rough surface as θ decreased. The whole sidewall is expected to be covered with the smooth ($\bar{1}13$)B facet for θ around 24°, θ_f of the ($\bar{1}13$)B facet (not checked experimentally). The generation of the ($\bar{1}\bar{1}3$)B facet has been reported in Ref. 42 in the lateral growth of GaAs on ($\bar{1}11$)B patterned substrates by metalorganic MBE (MOMBE) under high As_4 pressures. This and the discussions in Refs. 35–37 clearly indicate that the growth proceeds more easily on the ($\bar{1}\bar{1}3$)B surface than the ($\bar{1}11$)B surface due to freedom from the growth-hampering surface As trimer formation specific to the ($\bar{1}11$)B surface. The smooth surface was maintained after the ($\bar{1}13$)B facet vanished for θ=20°, as shown in Fig. 5(e), but the surface morphology gets worse again as θ decreased further. Consequently, the [110] stripes with flat and uniform sidewall layers maintaining the initial as-etched pattern can be obtained in the narrow θ range of 24° to (at least) 20°. The apparently irregular cross-sectional growth pattern reflecting the rough surface, as shown in Fig. 3(c), is really composed of the ($\bar{1}11$)B and ($\bar{1}13$)B facets.

C. (011) and (013) facets

Recently, growth of the (011) facet on the [100] stripe with the (010) sidewall and quantum wires formed taking advantage of the (011) facet growth were reported in Refs. 18–21 and 25. In these references, the as-etched profile of the sidewall has been limited to the type shown in Fig. 4(a). Therefore, we provided the first experimental results on the θ dependence of the growth on the [100] stripe in the present study. In the present study, the (013) facet developed after the (011) facet vanished for θ<45° under high temperature and As_4 pressure, while it has been shown in Ref. 18 that the (011) and (013) facets coexist under low temperatures and/or high As_4 pressures. It has been shown in Refs. 10, 11, 13, 23, 25 that the Ga diffusion length on the (001) surface shortens as the As_4 pressure increases and/or the temperature decreases. If this holds for the (011) and (013) surfaces and the Ga diffusion length on the (011) surface shortens more steeply than that on the (013) surface as the temperature decreases, the growth rate of the (011) facet relative to that

of the (013) facet increases, hence the contribution of the (013) facet becomes larger. This qualitatively explains the difference in the two results.

The (011) facet has been shown to also develop on the (111)A patterned substrates.[1]

D. Growth behavior contrast on [$\bar{1}$10] and [110] stripes

The growth behaviors on the [$\bar{1}$10] and [110] stripes can be more distinctly compared in Fig. 5. The surface morphology is better for the layers on the (111)A-related sidewalls than for those on the ($\bar{1}$11)B-related sidewalls. This may be attributed, first, to the present growth conditions optimized for the (111)A surface and, second, to the anisotropic structure of the (2×4)-reconstructed (001) surface[43] during MBE. The latter has demonstrated by a distinct difference in the straightness of microsteps produced by misorientations towards the [$\bar{1}$10] and [110] directions using scanning tunneling microscopy.[44,45] It should, however, be noted from (e) that flat and uniform layers can be simultaneously grown on both types of sidewalls for θ around 20° only, that is, on the (114)A and ($\bar{1}$14)B sidewalls. This suggests that the ($\bar{1}$13)B and ($\bar{1}$14)B surfaces may have wider optimum growth conditions, which are closer to the optimum growth conditions for the (001) surface, than the ($\bar{1}$11)B surface does.

E. Intersection of [100] and [010] stripes

The identification of the facets on the (111)A-related corners was made on the basis of their orientations and the values of θ_f evaluated as follows: the angle θ_c of the intersection line to the (001) substrate plane can be evaluated from the experimentally determined value of θ using the simple relation $\theta_c=\tan^{-1}(\tan\theta/\sqrt{2})$. The value of $\theta=19°$, estimated as corresponding to the onset boundary of the facets, gives the value of $\theta_c=13°$. This value corresponds to θ_f of the (116)A plane. The (111)A-related facet presented in Fig. 6(e) corresponds to (119)A. Table I also summarizes the facet generation behavior at the intersections of the [100] and [010] stripes. This result implies that it is impossible to grow GaAs and AlGaAs layers by MBE on ridge-type squares formed by the [100] and [010] stripes without modifying initial as-etched pattern at least under the present growth condition. In Ref. 21, a ridge-type square of the same type encircled with equivalent four (011) facets was proposed for a novel quantum dot. The structure shown in Fig. 1 of Ref. 21 exhibits the nonequivalent (111)A-related and ($\bar{1}$11)B-related corners although the authors did not mention it.

F. Phenomena common to three stripe patterns

Although the three stripes showed their respective facet generation behaviors as discussed above, the following point is worth mentioning as common to the three stripes: the persistent facets generated on the three stripes, that is, the (114)A facet on the [$\bar{1}$10] stripe, the ($\bar{1}$13)B facet on the [110] stripe, and the (013) facet on the [100] stripe, have θ_f in the narrow range of 19°–24°. This point seems to be related to the crystal structure of GaAs composed of the tetrahedral Ga–As bonding, as was also suggested for the (111)A patterned substrate in Ref. 1.

G. Comparison with (111)A patterned substrates

Comparing the facet generation behavior during MBE between the (001) patterned substrate described in the present article and the (111)A patterned substrate described in Ref. 1, the following common and different features are pointed out:

(1) Growth of extra (114)A, ($\bar{1}$11)B, and (110) facets and their lateral development with θ approaching their θ_f s are common to the (111)A and (001) patterned substrates. The directions of the lateral development of the (114)A, ($\bar{1}$11)B, and (110) facets are [22$\bar{1}$]B, [$\bar{1}\bar{1}$2]A, and [01$\bar{1}$], respectively, for the (001) substrate and [$\bar{2}$21]A, [$\bar{1}\bar{1}$2]B, and [00$\bar{1}$], respectively, for the (111)A substrate. Thus, the lateral development does not depend on the difference in the microstep properties of the as-etched sidewalls.

(2) Growth of extra (001), (113)A, (159), ($\bar{2}$38), and ($\bar{1}$25) facets are specific to the (111)A patterned substrates and growth of extra (111)A, ($\bar{1}$13)B, (031), (045), and (013) facets are specific to the (001) patterned substrates.

(3) There is a critical value for θ, θ_{crit}, below which flat and uniform layers on the sidewall and substrate plane grow maintaining the initial as-etched profile with no extra facets or no exponential thickness variations of layers on the substrate plane near the sidewall. The value of θ_{crit} is located in the narrow range of 33°–35° independent of the sidewall orientation for the (111)A substrate, whereas it varies between 10° and 24° depending on the sidewall orientation for the (001) substrate.

(4) As long as $\theta>\theta_{crit}$, the interfaces between the layers on the (001) substrate plane and the adjacent extra facets are curved towards the layers on the substrate plane, whereas the interface was flat and perpendicular to the substrate plane, irrespective of θ, for the (111)A substrate. This point may hinder the formation of lateral $p-n$ junctions having a steep $p-n$ transition region with structural simplicity by utilizing the amphotericity of Si doping on the (001) patterned substrate.

(5) The steps with $\theta_f=16°-17°$ appear near the sidewall–substrate plane boundary independent of the sidewall orientation for the (111)A substrates while no steps are observed for the (001) substrates.

These points reflect not only differences in the crystallographic and chemical properties between the two substrates but also differences in the interaction of adatoms between the substrate plane and adjacent sidewall. These points will be discussed in more detail in Ref. 46.

V. SUMMARY AND CONCLUSION

Extra facet generation during MBE growth of GaAs/AlGaAs multilayers on (001) GaAs substrates patterned with stripes running in the [$\bar{1}$10], [110], and [100] directions and having various slopes, designated as [$\bar{1}$10] stripe, [110] stripe, and [100] stripe, respectively, has been systematically

investigated. It has been confirmed that extra (111)A and (114)A facets are generated on the [$\bar{1}$10] stripes, extra ($\bar{1}$11)B and ($\bar{1}$13)B facets on the [110] stripes, and extra (031), (011), (045), and (013) facets on the [100] stripes, depending on the intersection angle θ of the sidewall and the substrate plane.

Among them, (114)A, ($\bar{1}$13)B, and (013) facets are important because they persist for a wide range of θ and, therefore, play a detrimental role in forming flat and uniform sidewall lavers on the (001) patterned substrates. It has been made clear that no extra facets are generated on the sidewalls and flat and uniform layers maintaining the initial as-etched stripe patterns can be grown for the [$\bar{1}$10] stripes with $10° \geqslant \theta$, the [110] stripes with $24° \geqslant \theta \geqslant 20°$, and the [100] stripes with $19° \geqslant \theta$. Extra (111)A-related facets have developed on the (111)A-related intersection of the equivalent [100] and [010] stripes, while no extra facets have developed on the ($\bar{1}$11)B-related intersection. The (114)A, ($\bar{1}\bar{1}\bar{1}$)B, and (110) facets have been confirmed as common to the (111)A and (001) patterned substrates.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Y. Furuhama and Dr. H. Inomata for their encouragement throughout this work.

[1]T. Takebe, M. Fujii, T. Yamamoto, K. Fujita, K. Kobayashi, and T. Watanabe, J. Cryst. Growth **162**, 31 (1996).
[2]W. T. Tsang and A. Y. Cho, Appl. Phys. Lett. **30**, 293 (1977).
[3]J. S. Smith, P. L. Derry, S. Margalit, and A. Yariv, Appl. Phys. Lett. **47**, 712 (1985).
[4]D. L. Miller, Appl. Phys. Lett. **47**, 1309 (1985).
[5]T. Yuasa, M. Manno, T. Yamada, S. Narituka, K. Shinozaki, and M. Ishii, J. Appl. Phys. **62**, 764 (1987).
[6]E. Kapon, M. C. Tamargo, and D. M. Hwang, Appl. Phys. Lett. **50**, 347 (1987).
[7]H. P. Meier, R. F. Broom, P. W. Epperlein, E. van Gieson, Ch. Harder, H. Jäckel, W. Walter, and D. J. Webb, J. Vac. Sci. Technol. B **6**, 692 (1987).
[8]H. P. Meier, E. van Gieson, R. F. Broom, W. Walter, D. J. Webb, C. Harder, and H. Jäckel, Inst. Phys. Conf. Ser. **91**, 609 (1988).
[9]S. Nilsson, E. van Gieson, D. J. Arent, H. P. Meier, W. Walter, and T. Forster, Appl. Phys. Lett. **55**, 972 (1989).
[10]M. Hata, T. Isu. A. Watanabe, and Y. Katayama, Appl. Phys. Lett. **56**, 2542 (1990).
[11]M. Hata, T. Isu. A. Watanabe, and Y. Katayama, J. Vac. Sci. Technol. B **8**, 692 (1990).
[12]Y. Nakamura, S. Koshiba, M. Tsuchiya, H. Kano, and H. Sakaki, Appl. Phys. Lett. **59**, 700 (1991).
[13]X. Q. Shen, M. Tanaka, and T. Nishinaga, 10th Symposium Record on Alloy Semiconductor Physics and Electronics, 1991 (unpublished), p. 65.
[14]S. Shimomura, S. Ohkubo, Y. Yuba, S. Namba, S. Hiyamizu, M. Shigeta, T. Yamamoto, and K. Kobayashi, Surf. Sci. **267**, 13 (1992).
[15]Y. Morishita, Y. Nomura, S. Goto, Y. Katayama, and T. Isu, Inst. Phys. Conf. Ser. **120**, 19 (1992).
[16]M. Walther, T. Röhr, G. Böhm, G. Tränkle, and G. Weimann, J. Cryst. Growth **127**, 1045 (1993).
[17]H. Saito, M. Sugimoto, M. Anan, and Y. Ochiai, Jpn. J. Appl. Phys. **32**, L1034 (1993).
[18]M. López, T. Ishikawa, and Y. Nomura, Jpn. J. Appl. Phys. **32**, L1051 (1993).
[19]K. Eberl, P. Grambow, A. Lehmann, A. Kurtenbach, K. v. Klitzing, D. Heitmann, M. Dilger, and M. Hohenstein, Appl. Phys. Lett. **63**, 1059 (1993).
[20]Y. Liu, S. Shimomura, N. Sano, K. Gamo, A. Adachi, and S. Hiyamizu, Semicond. Sci. Technol. **8**, 2197 (1993).
[21]M. López, T. Ishikawa, and Y. Nomura, Electron. Lett. **29**, 2225 (1993).
[22]X. Q. Shen, M. Tanaka, K. Wada, and T. Nishinaga, J. Cryst. Growth **135**, 85 (1994).
[23]X. Q. Shen, D. Kishimoto, and T. Nishinaga, Jpn. J. Appl. Phys. **33**, 11 (1994).
[24]S. Koshiba, Y. Nakamura, M. Tsuchiya, H. Noge, Y. Nagamune, T. Noda, and H. Sakaki, J. Appl. Phys. **76**, 4138 (1994).
[25]M. López and Y. Nomura, J. Cryst. Growth **150**, 68 (1995).
[26]T. Takamori and T. Kamijo, Appl. Phys. Lett. **77**, 187 (1995).
[27]T. Takebe, T. Yamamoto, M. Fujii, and K. Kobayashi, J. Electrochem. Soc. **140**, 1169 (1993).
[28]L. J. Schowalter, K. Yang, and T. Thundat, J. Vac. Sci. Technol. B **12**, 2579 (1994).
[29]T. Yamamoto, M. Inai, T. Takebe, M. Fujii, and K. Kobayashi, J. Cryst. Growth **127**, 865 (1993).
[30]M. Fujii, T. Yamamoto, M. Shigeta, T. Takebe, K. Kobayashi, S. Hiyamizu, and I. Fujimoto, Surf. Sci. **267**, 26 (1992).
[31]K. Kobayashi, T. Takebe, T. Yamamoto, M. Fujii, M. Inai, and D. Lovell, J. Electron. Mater. **22**, 161 (1993).
[32]M. Inai, T. Yamamoto, M. Fujii, T. Takebe, and K. Kobayashi, Jpn. J. Appl. Phys. **32**, 523 (1993).
[33]T. Yamamoto, M. Inai, T. Takebe, and T. Watanabe, Jpn. J. Appl. Phys. **32**, L28 (1993).
[34]K. Tsutsui, H. Mizukami, O. Ishiyama, S. Nakamura, and S. Furukawa, Jpn. J. Appl. Phys. **29**, 468 (1990).
[35]P. Chen, K. C. Rajkumar, and A. Madhukar, Appl. Phys. Lett. **58**, 1771 (1991).
[36]T. Hayakawa, M. Nagai, M. Morishima, H. Horie, and K. Matsumoto, Appl. Phys. Lett. **59**, 2287 (1991).
[37]T. Hayakawa, M. Morishima, and S. Chen, Appl. Phys. Lett. **59**, 3321 (1991).
[38]T. Hayakawa, M. Morishima, M. Nagai, H. Horie, and K. Matsumoto, Surf. Sci. **267**, 8 (1992).
[39]J. Fu, K. Zhang, and D. L. Miller, J. Vac. Sci. Technol. B **10**, 779 (1992).
[40]D. A. Woolf, Z. Sobiesierski, D. I. Westwood, and R. H. Williams, J. Appl. Phys. **71**, 4908 (1992).
[41]D. A. Woolf, J. P. Williams, D. I. Westwood, Z. Sobiesierski, J. E. Aubrey, and R. H. Williams, J. Cryst. Growth **127**, 913 (1993).
[42]T. Isu, Y. Morishita, S. Goto, Y. Nomura, and Y. Katayama, J. Cryst. Growth **127**, 942 (1993).
[43]M. D. Pashley, K. W. Habarern, W. Friday, J. M. Woodall, and P. D. Kirchner, Phys. Rev. Lett. **60**, 2176 (1988).
[44]M. D. Pashley, K. W. Habarern, and J. M. Gaines, Appl. Phys. Lett. **58**, 406 (1991).
[45]M. D. Pashley, K. W. Habarern, and J. M. Gaines, Surf. Sci. **267**, 153 (1992).
[46]T. Takebe, M. Fujii, T. Yamamoto, K. Fujita, and T. Watanabe (in preparation).

InP/InGaAs single heterojunction bipolar transistors grown by solid-source molecular beam epitaxy using a phosphorus valved cracker

W. L. Chen[a)]
Solid-State Electronics Laboratory, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan 48109-2122

T. P. Chin and J. M. Woodall
School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907-1285

G. I. Haddad
Solid-State Electronics Laboratory, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan 48109-2122

(Received 1 March 1996; accepted 1 May 1996)

In this work, InP/InGaAs single heterojunction bipolar transistors (SHBTs) with excellent performance were grown by solid-source molecular beam epitaxy using a phosphorus valved cracker. The maximum dc and differential current gain are ~27 and 32 with breakdown voltages V_{ceo}~6 V and V_{cbo}~10 V for the SHBT with a 5000 Å InGaAs collector layer doped at 4×10^{16} cm^{-3}. The ideality factors for Ib and Ic are ~1.12 and 1.07 and the transistors also show very uniform current gain down to 10^{-10} A range. For high frequency performance, the maximum f_T and f_{max} are around 80 and 125 GHz for the SHBT with an emitter area of 32 μm^2. © *1996 American Vacuum Society.*

I. INTRODUCTION

InP/InGaAs heterojunction bipolar transistors (HBTs) have demonstrated excellent high frequency performance, and they are promising candidates for integration into optoelectronic light wave communication systems.[1–4] Most of the reported InP-based HBTs were grown by gas-source MBE (GSMBE), chemical beam epitaxy (CBE) and metalorganic chemical vapor deposition (MOCVD) using highly toxic hydrides, e.g., PH_3 and AsH_3, requiring complicated and costly utilities to handle. It is desirable to use solid-source phosphorus to grow P-containing compounds for device and circuit applications to improve safety concerns. There have been some attempts to use solid phosphorus as source material in a MBE system. However, the main problems were (1) the impurities in the source material and (2) fire hazard caused by white phosphorus. To reduce fire hazard, even GaP and InP chunks were evaporated to obtain P_2 but accompanied with Ga and In impurity atoms. Recently, a valved phosphorus cracker with improved designs was used to generate P_2 successfully with reduced impurities and fire hazard. Therefore, good device results of optoelectronic and electronic devices were reported without using hydrides. There have been few results of InP/InGaAs HBTs reported using solid phosphorus in a MBE system.[5,6]

In this article, a solid-source MBE system is used to grow InP/InGaAs HBTs which show excellent dc and high frequency performance. The article starts with the growth and processing techniques. Then material characterization results are discussed, including the quality of the InP/InGaAs heterojunction and the etching characteristics of the heterojunction. After the discussions of material quality, dc and microwave device results are described with the discussions of device quality. Finally, a brief summary will be given at the end of the article.

II. GROWTH AND FABRICATION

The growth was performed in a modified Varian GEN-II CBE system, which has been described elsewhere.[7] A valved cracker is used as the phosphorus source while standard effusion cells are used for As, Be, and group-III metals. A Si filament is used as the n-type doping source. P_4 molecules are thermally cracked into P_2 at 900 °C before injected into the growth chamber. The typical phosphorus beam equivalent pressure is $1–2\times10^{-5}$ Torr. Reflection high-energy electron diffraction (RHEED) oscillations are used to calibrate growth rates, compositions, and surface V/III ratio. The typical growth temperature for InP and InGaAs are 480 °C. Double crystal x-ray rocking curve linewidth of an 1 μm InGaAs layer is 22 arcsec.

Special care is taken at the InP(top)/InGaAs(bottom) interface. A shutter sequence is designed so that the InGaAs is not exposed to phosphorus flux, which means the arsenic shutter closes first, then one monolayer of indium metal is deposited and finally the phosphorus shutter opens. As shown in Fig. 1(a), the InP/InGaAs interface of the 40 Å well is grown with such a sequence while the 50 Å well is grown with the normal shutter sequence, which means the phosphorus shutter opens before the In shutter. Low temperature (10 K) photoluminescence (PL) of the 40 and 50 Å InP/InGaAs wells are shown in Fig. 1(b). The significant improvement of the PL intensity and narrow linewidth (4.7 meV) is attributed to the smooth InP/InGaAs interface without the formation of InGaAsP. The improved InP/InGaAs

[a)]Present address: Central Research Laboratories, Texas Instruments Inc., 13588 N. Central Expressway MS 134, Dallas, TX 75265; Electronic mail: WLCHEN@CRLMSM.CSC.TI.COM

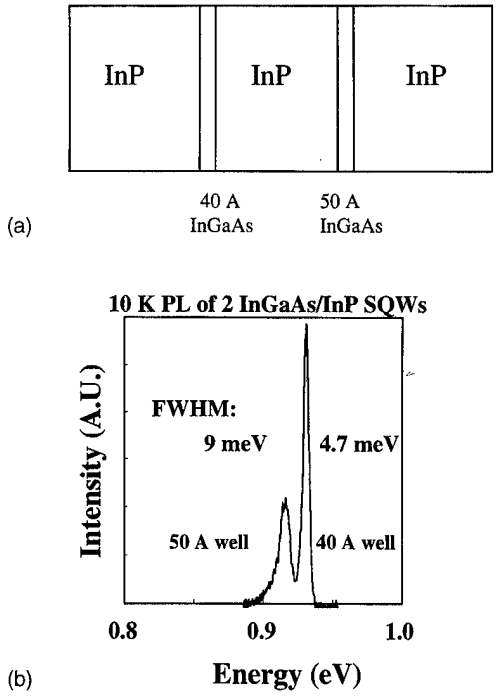


FIG. 1. (a) InP/InGaAs single quantum wells (SQWs) with 40 and 50 Å wells which are grown using different shutter sequences. (b) 10 K PL of the two InP/InGaAs SQWs shown in (a).

interface also results in very uniform surface after selectively etching InP away from InGaAs.

The typical InP/InGaAs HBT structure, as shown in Fig. 2, consists of a 1500 Å InGaAs $n+$ emitter contact layer, a 500 Å $n+$ InP layer, a 1000 Å InP emitter layer doped at 5×10^{17} cm^{-3}, a 50 Å InGaAs undoped spacer, a 650 Å InGaAs base layer doped at 5×10^{19} cm^{-3}, a 5000 Å InGaAs collector layer doped at 4×10^{16} cm^{-3}, and a 6000 Å $n+$ InGaAs collector contact layer. Si and Be were used as n-type and p-type dopants.

The InP/InGaAs HBTs were processed using a self-aligned process. In the process, $HCl:H_3PO_4$ solution was used for etching the InP layer selectively over the InGaAs layer, and $H_3PO_4:H_2O_2:H_2O$ solution was adopted for etching the InGaAs layer selectively over the InP layer. The separation between the emitter mesa and the base metal is less than 0.2 μm. Then, Ti/Pt/Au was used as the emitter and collector ohmic contacts and Pt/Ti/Pt/Au was used as the base ohmic contact, resulting in contact resistivities of $\sim7\times10^{-7}$ Ω cm^2 (emitter), 6×10^{-7} Ω cm^2 (collector), and 3×10^{-6} Ω cm^2 (base). After the deposition of the ohmic contacts, PECVD SiO_2 were deposited all over the wafer followed by via hole openings using RIE. Finally, Ti/Al/Ti/Au was used as interconnect metals for dc and microwave probing.

n+	InGaAs	2E19	1500 A
n+	InP	2E19	500 A
n-	InP	5E17	1000 A
	InGaAs	undoped	50 A
p+	InGaAs	5E19	650 A
n-	InGaAs	4E16	5000 A
n+	InGaAs	2E19	6000 A
InP Semi-insulating Substrate			

FIG. 2. The InP/InGaAs HBT structure grown by solid-source MBE.

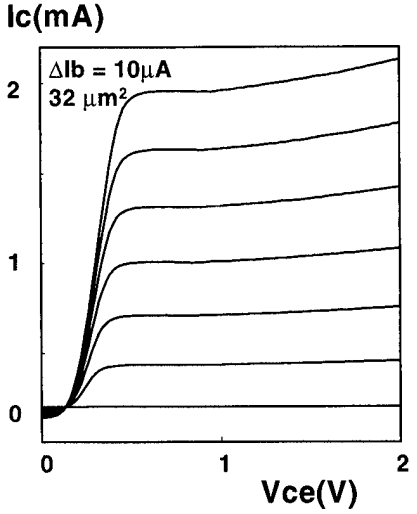


FIG. 3. The common emitter $I-V$ characteristics.

III. TRANSISTOR CHARACTERISTICS

The typical common emitter $I-V$ characteristics are shown in Fig. 3 for the HBTs with emitter sizes of 32 μm^2. The maximum dc current gain Ic/Ib and differential current gain (dIc/dIb) are around 27 and 32, which are commonly observed for HBTs grown by other growth techniques with similar base doping and base thickness. The current gain is also very uniform down to 10^{-10} A as shown in Fig. 4. The ideality factors for Ib and Ic are $\sim$1.12 and 1.07. Based on the common emitter $I-V$ characteristics and the Gummel plots, the surface recombination current is not a dominant factor even in the range of low current densities and there is no significant leakage path in the emitter layer, the base layer or the emitter/base junction, indicating good quality of InP/InGaAs HBTs grown by solid-source MBE. If there is any significant leakage path caused by either poor quality of the emitter/base junction or high defect densities in the emitter or base layers, the current gain would be very nonuniform in the common emitter $I-V$ characteristics and high intrinsic base leakage current should be seen in the Gummel plots.

The other important device parameters are the offset and breakdown voltages. The offset voltage is less than 0.2 V in the common emitter $I-V$ characteristics, indicating a normal emitter/base junction. The breakdown voltages BV_{ceo} and BV_{cbo} are $\sim$6 and 10 V and both voltages are defined at

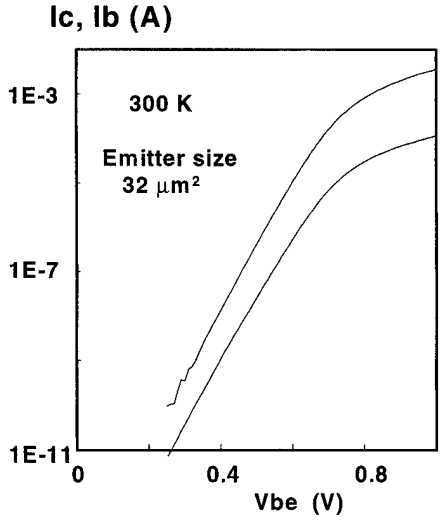


FIG. 4. The Gummel plot.

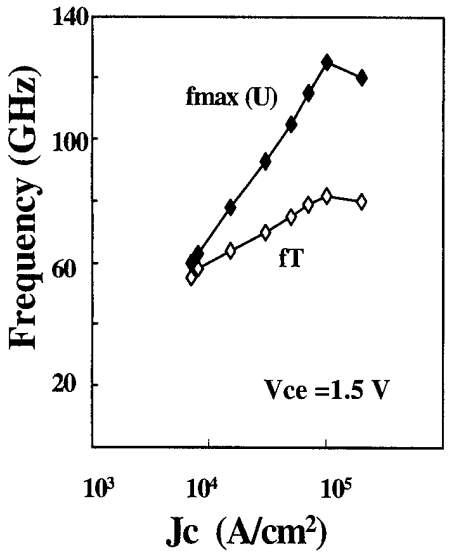


FIG. 6. f_T and f_{max} at different bias conditions.

$Ic = 100$ μA for the HBTs with 32 μm^2 emitter sizes. Moreover, temperature dependent measurements are also performed to verified the breakdown mechanism as a tunneling procedure which is normally observed for InP/InGaAs HBTs grown by other techniques.

For high frequency performance, the HBTs with 32 μm^2 emitter sizes were biased at different output current densities. As shown in Figs. 5 and 6, the highest f_T (80 GHz) and f_{max} (125 GHz) are obtained at the output current density of around 1×10^5 A/cm^2. Such a combination of f_T and f_{max} corresponds to a R_bC_{bc} time constant of 0.21 ps. Biasing the HBT beyond this current density, f_T and f_{max} start to fall-off due to the Kirk effect. The HBTs seem to have stable characteristics even under high output current densities and there is no rapid degradation of transistor characteristics observed during the dc and microwave measurements.

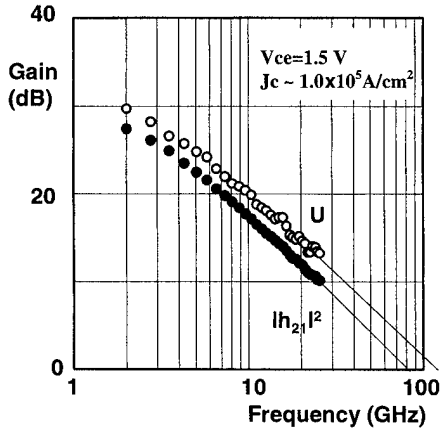


FIG. 5. Gain vs frequency plots.

IV. CONCLUSIONS

High performance InP/InGaAs HBTs have been grown successfully by solid-source MBE using a phosphorus valved cracker. The HBT performance is as competitive as those obtained by other growth techniques. To our knowledge, the high frequency performance (f_{max}) of the HBTs is the highest value reported for devices grown using this growth technique. In conclusion, the P-valved cracker could be used as an alternative source material to improve the safety concerns significantly about using conventional P source materials such as PH_3.

ACKNOWLEDGMENTS

This work was supported by ARO under URI program Grant No. DAAL03-92-G-0190 and ARPA under Contract No. DAAH-04-93-G-0242 at the University of Michigan. The work at Purdue University was supported by the Material Research Science and Engineering Center from the National Science Foundation Grant No. DMR-9400415.

[1]Y. K. Chen, R. N. Nottenburg, M. B. Panish, R. A. Hamm, and D. A. Humphrey, IEEE Electron Device Lett. **EDL-10**, 267 (1989).
[2]H. F. Chau and E. A. Beam, IEEE Electron Device Lett. **14**, 388 (1993).
[3]J. C. Cowles, W. L. Chen, G. O. Munns, and G. I. Haddad, Proceedings of IEEE/Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits, Ithaca, NY, 2–4 August 1993 (unpublished), p. 62.
[4]M. Ida, S. Yamahata, K. Kurishima, H. Ito, T. Kobayashi, and Y. Matsuoka, *High-fmax InP/InGaAs HBTs with Extrinsic Base Layers Selectively Grown by MOCVD*, the proceedings of 1995 IEEE 53rd annual device research conference, June 19–21, 1995, Charlottesville, Virginia, p. 78.
[5]Willen, H. Asonen, and M. Toivonen, Electron. Lett. **31**, 1514 (1995).
[6]Application Note of EPI MBE Products Group, *State-of-the-Art Device Results Using EPI Valved Crackers for Phosphorus and Arsenic*, Note No. 5/95, September 1995.
[7]T. P. Chin, W. L. Chen, G. I. Haddad, and J. M. Woodall, J. Vac. Sci. Technol. B **13**, 750 (1995).

Chemistry of arsenic incorporation during GaAs/GaAs(100) molecular beam epitaxy probed by simultaneous laser flux monitoring and reflection high-energy electron diffraction

Adina K. Ott, Sean M. Casey,[a] April L. Alstrin,[b] and Stephen R. Leone[c]
JILA, National Institute of Standards and Technology and University of Colorado and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0440

(Received 15 December 1995; accepted 19 April 1996)

Arsenic incorporation during GaAs/GaAs(100) molecular beam epitaxy is studied *in situ* with laser single-photon ionization time-of-flight mass spectrometry and reflection high-energy electron diffraction (RHEED). Incident and scattered fluxes of Ga and As_n species in front of the growing GaAs wafer are ionized repetitively by a pulsed laser beam of 118 nm (10.5 eV) photons. The methods to obtain and interpret time-of-flight mass spectra and the simultaneous RHEED measurements are described. The real time behaviors of incident Ga and desorbing As_2 and As_4, obtained without mass spectral cracking, are studied during growth of GaAs layers with As_4 and when growth is arrested as a function of substrate temperature and Ga/As_4 flux ratio. During growth only with As_4, both As_2 and As_4 are desorbed or scattered in varying amounts depending on flux and substrate temperature conditions. Without an incident gallium flux, desorbing As_4 decreases while desorbing As_2 increases with increasing surface temperature. During gallium deposition and GaAs growth, the amounts of desorbing arsenic fluxes decrease linearly with increasing Ga/As_4 flux ratio, but the arsenic incorporation rate saturates at a Ga/As_4 flux ratio ≥ 2, i.e. (Ga/As$\geq\frac{1}{2}$). The total integrated incorporation of arsenic increases linearly with increasing Ga/As_4 flux ratio when the surface is allowed to recover with an incident arsenic flux after the gallium flux is terminated. In the range of substrate temperatures optimum for layer-by-layer GaAs growth with As_4, As_4 incorporation dominates at low temperatures, while As_4 and As_2 incorporations contribute equally at high temperatures. Surface reaction sequences and mechanisms of arsenic incorporation are discussed and compared with measured RHEED results and previous experimental and theoretical results. © *1996 American Vacuum Society.*

I. INTRODUCTION

The growth of GaAs by molecular beam epitaxy (MBE) has been studied extensively by *in situ* tools over the past thirty years. After the discovery that oscillations in the reflection high-energy electron diffraction (RHEED) specular beam intensity correspond to layer-by-layer epitaxy of GaAs,[1] the field exploded with new investigations of Ga surface diffusion,[2] Ga-stabilized growth,[3–5] surface coverages,[6–10] and kinetics of surface relaxation after growth.[1,11–13] These experiments measured microscopic properties based on surface reconstructions, which give information about the arrangement of surface species averaged over a wide temperature range and surface area. Specular beam intensity, which is a measure of the surface roughness, however, is not necessarily a linear function of the surface coverage.[14] While much information about the surface growth has been deduced from RHEED and other methods, basic mechanisms and rates for GaAs growth are still inaccessible: more work is needed to understand the reaction kinetics and incorporation mechanisms of gallium and arsenic surface species on a molecular level.

In situ tools that are species specific are important to gain a better understanding of the microscopic picture of GaAs growth mechanisms. Early studies using modulated molecular beam techniques provide the base of kinetic information on GaAs growth, including proposed orders of the reactions involved for growth with As_2 (Ref. 15) and As_4 (Ref. 16). Considerable insight on growth mechanisms of GaAs since these early studies has been provided by Monte Carlo simulations[17–20] and other modeling work based on RHEED reconstructions.[21–23] Results from these studies linked Ga migration to reactivity, indicating that formation of a surface Ga cluster is key to GaAs growth, but the availability of basic rate data of Ga-As_4/Ga-As_2 growth is still limited.

Recent mass spectrometric data show promise that nonintrusive real time monitors of specific surface species can give a better microscopic picture of the epitaxy, providing surface incorporation mechanisms and rates. Reflection mass spectrometry[24,25] (REMS) has emerged as one such method to obtain information by monitoring gaseous fluxes scattering or desorbed from a surface. REMS has been used to study gallium desorption, growth of multiple quantum well structures,[26] InAs growth,[27,28] cation incorporation in other III–V structures,[29] and GaAs growth.[30,31] A comprehensive study on GaAs growth by Brennan *et al.*[31] has provided insight into As_4 and As_2 incorporation during GaAs MBE growth. Drawbacks of this method include the difficulty to interpret mass spectra due to fragmentation of molecules by the electron impact ionizer and the inability to measure incident growth species.

[a]National Research Council Postdoctoral Fellow, National Institute of Standards and Technology.
[b]Present address, Quantum, 2270 S. 88th St., Louisville, CO 80028.
[c]Staff member, Quantum Physics Division, National Institute of Standards and Technology; Electronic mail: srl@jila.colorado.edu

This article describes a similar mass spectrometric study of the GaAs growth interface using single-photon ionization time-of-flight mass spectrometry (SPI-TOFMS) concurrently with RHEED. The single-photon ionization method provides a non-intrusive monitor that is species-specific and complements the macroscopic information provided by RHEED. The method detects both incident and scattered or desorbed fluxes. Mass spectra are easily interpreted since the photons used for ionization provide enough energy to ionize, but not fragment, the molecules involved in growth. Quantification of mass signals is provided by relative ionization probabilities and detection efficiencies of incident Ga and As_4 and scattered As_2 and As_4 species. The real time changes in mass signals during growth are related to incorporation of species into the growing wafer and to changes in reactive sticking probabilities on the surface.

In this article, we examine these parameters as a function of both substrate temperature and Ga/As_4 flux ratio. During gallium deposition in the range of substrate temperatures studied, the reactive sticking of As_4 dominates arsenic incorporation at lower surface temperatures, while the incorporation of both As_2 and As_4 during the growth with As_4 contributes to arsenic incorporation at higher surface temperatures. The reactive sticking of both As_2 and As_4 increase with increasing incident Ga flux until a gallium flux is reached in which the gallium rich surface saturates the arsenic sticking kinetics on the surface. An additional arsenic flux is required after Ga deposition is stopped to consume the excess surface gallium and produce a stoichiometric film of GaAs. Mechanisms are proposed to describe the observed changes in sticking probabilities and desorption and incorporation rates. The SPI-TOFMS measurements are compared with both the RHEED data as well as with prior theoretical and experimental results.

II. EXPERIMENTAL CONFIGURATION

A schematic of the experiment is shown in Fig. 1. The ultra-high vacuum chamber is equipped with a 400 ℓ/s ion pump and a Ti sublimator to obtain base pressures of 1×10^{-8} Pa (8×10^{-11} Torr). The chamber is equipped with an Auger electron spectrometer, RHEED spectrometer, laser ionization source and time-of-flight mass spectrometer, an Ar ion sputter gun, a residual gas analyzer, a nude ionization gauge, and a load lock for sample transfer.

The cryoshielded epitaxial source region contains three shuttered effusive ovens that are differentially pumped with a 30 ℓ/s ion pump. The arsenic source contains a second stage of heating for cracking of the As_4 molecular beam to an As_2 molecular beam. This oven is positioned 22 degrees below the surface normal. Although only As_4 is used for the GaAs growth studies, this cracking feature is used to calibrate the laser ionization efficiencies of As_4 vs. As_2. The two ovens containing gallium and silicon (not used) are positioned at an angle of 42 degrees below the surface normal.

The RHEED gun is positioned so that high energy (20 keV) electrons graze the sample surface and strike a custom fluorescent screen. The fluorescent screen is positioned opposite the gun for monitoring the diffraction pattern and specular beam intensity during growth, but the screen has a cutaway portion to transmit the laser beam in the opposite direction to the RHEED electrons. A charge coupled device (CCD) camera and television monitor allow for the real time measurement of changes in the specular RHEED beam intensity with a photodiode positioned to sample the monitor screen.

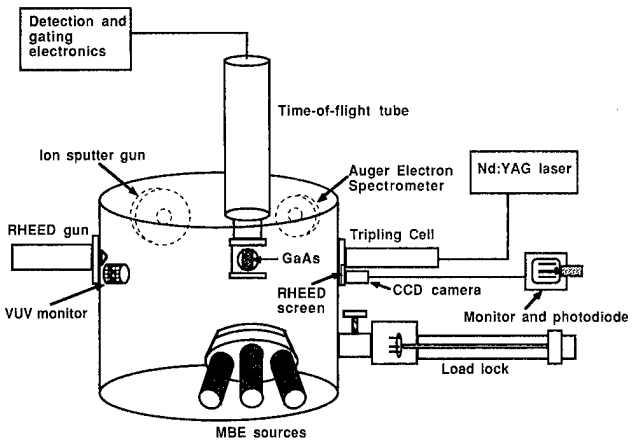


FIG. 1. Schematic illustration of the experimental apparatus.

Single-photon ionization is accomplished by passing a pulsed beam of 118 nm photons (which have 10.5 eV of energy) approximately 1 cm in front of the GaAs substrate at a repetition rate of 10 Hz. Here, the photons ionize a small fraction of the gaseous incoming, scattering, and desorbing growth species. The substrate is positioned between the extraction plates of a Wiley–MacLaren TOFMS.[32] The ions created are extracted perpendicular to the substrate and up a 1 meter time-of-flight tube, which is held at a −2000 V bias. The ions then come into contact with a focused mesh electron multiplier detector, which is differentially pumped by a 30 ℓ/s ion pump and is biased at −2600 V. This contact causes a cascading of electrons whose amplification is directly proportional to the number of ions. The signal is amplified, converted to a voltage by a charge-sensitive pre-amp, and the signal is sent to the gated integrating boxcars for real time monitoring of mass peaks or to a digital oscilloscope and laboratory computer for recording of a mass spectral trace resulting from a single laser pulse.

The photons used for ionization are created by first tripling the 1064 nm output of a Nd:YAG laser in KDP crystals. The characteristics of the resulting 355 nm light are 5 ns pulse duration, 20–60 mJ/pulse, and 10 Hz repetition rate. The 355 nm light is focused into a static cell of Xe:Ar (1400:10300 Pa), where the photon frequency is again tripled by a four wave mixing process.[33–35] The conversion efficiency of this process is estimated from the Pt photodiode signal to be approximately 10^{-6}, resulting in hundreds of nJ of 118 nm light. The 118 nm photons pass through a LiF lens and window into the chamber and are either focused or collimated into the extraction region of the TOFMS by the LiF lens. The remaining 355 nm photons are dispersed in the

UHV chamber. The 118 nm photons enter the chamber next to the RHEED screen and opposite the RHEED gun. This coaxial arrangement allows for simultaneous SPI-TOFMS and RHEED monitoring of epitaxial growth. Other geometries are also possible in the future, such as a remote location of the TOFMS to avoid surrounding the GaAs wafer region with grids and extraction plates.

Fluctuations in the power of the 118 nm beam are monitored by a platinum photodiode placed near the RHEED gun. The photons impinge on a piece of platinum, creating electrons via the photoelectric effect. The electrons are collected by a surrounding stainless steel mesh. The fluctuations in the signal output of this photodiode are used to normalize the amplitude fluctuations in the intensity of the mass signals.

Integrating boxcars are gated on the arrival time of Ga^+, As_4^+, and As_2^+ ions and the output of the platinum photodiode. The area under each peak is calculated by the boxcar integrator on each laser pulse and is recorded by the laboratory computer. Normalization of the mass signals is discussed in detail below. The output of the RHEED photodiode is simultaneously recorded by the computer to produce a real time record of the surface smoothness oscillations during growth from the corresponding behavior of the RHEED specular beam intensity.

Samples of GaAs (100) are approximately 8 mm x 21 mm in dimension and are 0.4–0.5 mm thick. The wafers are cleaned by Ar ion sputtering of the surface to remove carbon and oxygen contaminants and then heating the wafers radiatively under a flux of As_4 to remove the remaining oxide layer. Auger electron spectra confirm the removal of carbon and oxygen from the GaAs surface. The RHEED pattern is monitored until a 2×4 arsenic-stable reconstruction appears. A buffer layer of GaAs is grown before each experiment to create a smooth substrate. The temperature is measured by a type K thermocouple clamped to the GaAs surface.

III. RESULTS

A. Mass spectra and flux calibrations

Figure 2(a) shows a time-of-flight mass spectrum resulting from a summation average of 200 laser pulses monitoring the gaseous species in front of a GaAs substrate bathed in As_4 and held at 757±5 K. The As_4 peak at mass 300 results from four contributions: incident, specularly scattering, desorbing, and background tetramer species in the chamber. The As_2 peak results from desorption from the heated substrate and a small amount of cracking of the As_4 by scattered 355 nm photons (<0.4% of As_4).[36] In this trace, the As_4 incoming flux comprises (97±3%) of the As_4 signal, while the scattered and desorbed portion represents (3±3)% of the signal. The signal from background As_4 has been subtracted. Because the laser beam is collimated and positioned low in front of the substrate in this experiment, the photons intersect a large portion of the incoming As_4 flux. The small peaks to the right of the arsenic signals are due to ringing in the electronics during the mass sweep.

Figure 2(b) shows a time-of-flight mass spectrum taken with 200 laser pulses during gallium deposition. The two

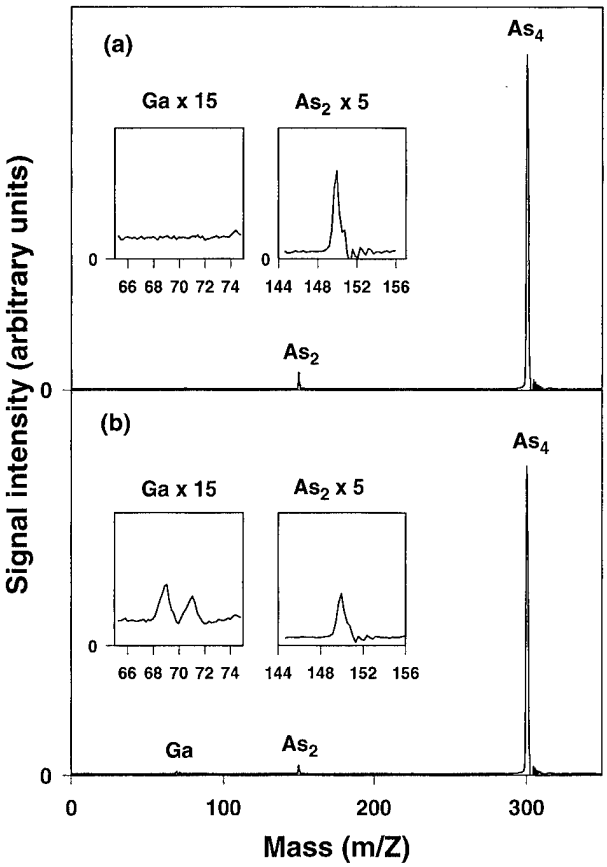


FIG. 2. (a) Time-of-flight mass spectrum of a GaAs wafer held at 757 K and bathed in an As_4 flux of 0.1 ML/s, and (b) time-of-flight mass spectrum of a GaAs wafer held at 757 K under As_4 and Ga fluxes (Ga flux=0.2 ML/s, As_4 flux=0.1 ML/s).

isotopes of Ga appear at masses 69 and 71 u when the shutter is opened. Arsenic dimer signal is still present, but this signal has decreased considerably due to incorporation of arsenic into the growing wafer. The As_4 signal has also decreased due to the increased sticking and incorporation of As_4 on available surface gallium.

In the limit of weak photoionization, the mass signals can be related to the density of molecules according to the equation:

$$I = n \times \sigma \times p \times l \times E,$$

where I=number of ions detected, n=density of molecules, σ=118 nm ionization cross section, p=the number of photons per laser pulse, l=the path length of the photons, and E=collection efficiency of the time-of-flight detector. Typical values of these parameters for an incoming flux of As_4 are as follows: $n(As_4)=3\times10^9$ molecules/cm^3, $(As_4)=5\times10^{-17}$ cm^2, $p=7.5\times10^{10}$, $l=1$ cm, $E=0.005$, and $I=50$ ions detected. The number of ions detected was calculated from the current produced by the electron multiplier using its specified gain and the efficiency was derived from this result.

Note that the TOF signal is a measure of the density of molecules. To determine the corresponding flux, one must apply a velocity correction:

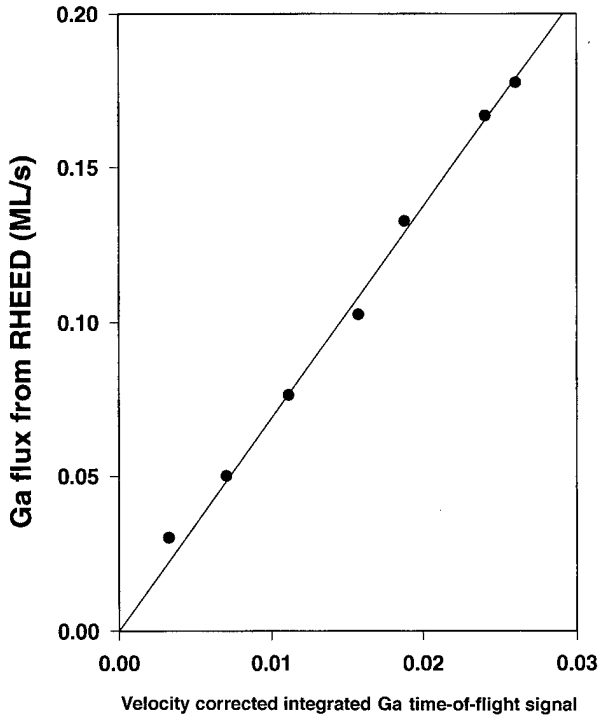


FIG. 3. Plot of Ga flux determined by the period of RHEED oscillations vs velocity corrected integrated time-of-flight gallium signal.

$$j = n \times v,$$

where j=flux and v=velocity. The As_4 molecules emanating from the oven source travel at a velocity of approximately 200 m/s. In typical experiments, the As_4 flux is equivalent to a deposition rate at the GaAs wafer of 0.1 ML s^{-1}.

An advantage of the laser TOF method is that the TOF signal is sensitive to slight changes in the flux that are not visible using the thermocouple readings on the oven sources or the chamber pressure ion gauge. One can therefore closely track sudden pressure bursts from a source or depletion of an oven species.

Because the collection efficiency of the experimental setup can vary from day to day, calibration of the fluxes is necessary. The flux of Ga atoms impinging on the surface was determined using the RHEED oscillation periods for GaAs growth measured over a range of Ga fluxes. Figure 3 is a plot of the RHEED oscillation period as a function of integrated gallium time-of-flight signal, corrected for velocity as determined from the gallium source temperature. The linear behavior of this plot indicates that the time-of-flight signal accurately measures the incoming flux of gallium.

The incident As_4 flux and contribution of scattered As_4 signal were calibrated using the saturation of the As_4 sticking probability on a gallium-rich surface. Second order kinetics dominates the process, which dictates a maximum sticking probability for the As_4 of 0.5 on a gallium rich surface.[16] The change to a gallium-rich surface, in this case, occurs when the As_4 incident flux is only half the incident Ga flux. The decrease in scattered As_4 on a gallium-rich surface should be one-half of the total amount of scattered As_4. The rest of the signal is assumed to result from incoming and background tetramer signal.

Based on the incident fluxes of Ga and As_4 determined by RHEED and their respective integrated TOF signals, a relative flux sensitivity ratio measured with the laser is obtained for As_4:^{69}Ga; this value is (375±25) for one specific geometry (see, for example, Fig. 2). The result is reasonable since the laser is positioned to intersect a greater portion of the incident As_4 beam than the incident Ga beam and the cross section for ionization of Ga is smaller than for As_4. Note, however, that this ratio is not the difference in ionization cross sections.

The quantification of scattered fluxes of As_2 and As_4 is accomplished by using a measured relative density sensitivity ratio that relates the TOF signals to the product of their 118 nm ionization cross sections (σ), densities (n), and collection efficiencies (E):

$$\mathrm{TOF}(As_4)/\mathrm{TOF}(As_2) = [\sigma(As_4) \times E(As_4) \times n(As_4)]/[\sigma(As_2) \times E(As_2) \times n(As_2)].$$

The As_4:As_2 sensitivity ratio has been determined using two methods in our laboratory. In the first method, described previously,[36] the temperature of the cracker stage of the arsenic oven is varied to produce different relative amounts of As_2 and As_4 coming from the oven. Two As_2 molecules are formed from the dissociation of each As_4 molecule. When the temperature of the cracker is raised, any decrease in the As_4 signal is assumed to correspond to a factor of two increase in the As_2 signal. Using this method under a variety of oven settings, a sensitivity ratio of 2.1±0.4 is measured. This value differs substantially from previously reported numbers from our laboratory due to modifications to the time-of-flight tube and extraction region, the removal of an orifice in the arsenic source, and the current use of lower arsenic fluxes and higher laser powers. When prior experimental conditions were matched, a sensitivity ratio of 9±1 was measured, much closer to the previously reported density sensitivity ratio of 7.[36] The previously measured ratio is discarded due to the many chamber modifications, and we believe this newly measured ratio reflects the approximate ratio of ionization cross sections.

The second method employed to determine the sensitivity ratio involves varying the surface temperature of a molybdenum foil sample and comparing the relative amounts of As_4 and As_2 scattering or desorbing from the substrate. Again, a decrease in scattering or desorbing As_4 is assumed to correspond to an increase in desorbing As_2, due to increased cracking of As_4 to As_2 on the heated surface. Using this method, the sensitivity ratio As_4:As_2 was found to be 2.2±0.3. The measured density sensitivity ratio for As_4:As_2 is also consistent with arguments based on the electron configurations of the molecules.[37]

To convert this sensitivity ratio to a flux ratio, the relative velocities of scattering As_2 and As_4 are needed. Corrections for differences in mass, translational temperature, and angular distributions of desorbing As_2 and As_4 are needed to cal-

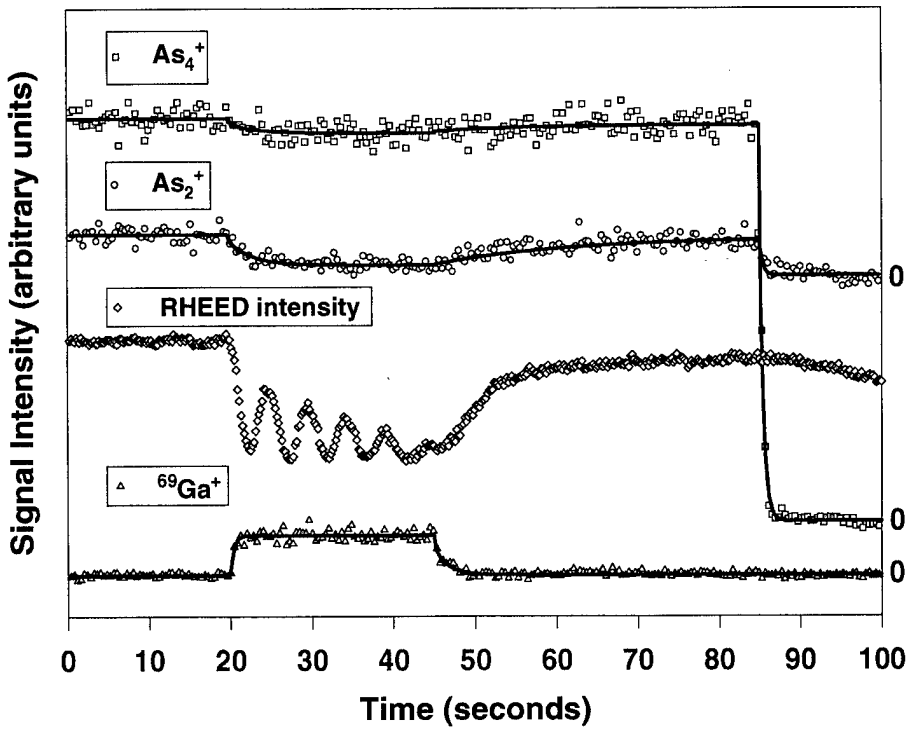


FIG. 4. Real time monitoring of As_4^+, As_2^+, and $^{69}Ga^+$ time-of-flight and RHEED photodiode signals. The gallium oven shutter is opened at 20 s and closed at 45 s. The wafer is bathed in As_4 until 85 seconds when the arsenic oven shutter is closed.

culate the relative velocities accurately. The angular distributions of desorbing As_2 and As_4 probably do not differ enough to affect the results within experimental uncertainty. The temperature of the molecules, however, can affect the relative velocities more strongly. Foxon and Joyce reported that desorbing As_4 molecules in the temperature range of 300–600 K are completely thermally accommodated to the GaAs (100) surface.[16] Arthur and Brown investigated the velocities of As_2 and As_4 desorbing from a GaAs (111) surface and found that the velocity distribution of As_2 is at a lower temperature than the surface, while As_4 is even lower in temperature.[38] Despite this result, over the temperature range we studied, the difference in As_2 and As_4 velocities relative to the surface temperature would not affect the final sensitivity ratio outside the experimental uncertainty, provided we assume the translational temperatures of the molecules are within 100 K of one another. The mass correction is applied and a relative flux sensitivity ratio of 1.5 ± 0.3 is used to convert the desorbed As_2 data to provide the correct relative fluxes of arsenic. This allows the direct comparison of trends in As_2 and As_4 behavior as well as accurate measurement of the behavior of the total arsenic signal.

In summary, quantification of the relative fluxes has been accomplished by calibration of the incoming Ga and As_4 fluxes using RHEED and known surface kinetics. Relative fluxes of As_2 and As_4 were determined using the measured density sensitivity ratio and appropriate velocity corrections.

B. Real time monitoring of growth species

The gated mass signals provide information on the real time behavior of the surface growth process. Figure 4 shows a typical average of four data runs recorded in real time under closely reproduced conditions of constant oven fluxes ($Ga/As_4 = 2.3$) and constant substrate temperature (757 ± 5 K). Boxcar integrators are used to measure the ^{69}Ga, As_2, As_4 TOFMS signals, the Pt photodiode signal, and the RHEED photodiode signal. The mass signals shown have been normalized to the 118 nm laser power fluctuations. In this run, the GaAs substrate is initially bathed in a flux of As_4. A desorbed arsenic dimer signal is present due to the elevated temperature of the GaAs wafer. Both As_4 and As_2 signals are constant in time, indicating the steady-state concentration of the surface desorbed and scattered species. At time=20 s, the gallium shutter is opened and growth commences. The gallium time-of-flight signal increases to a steady-state level, corresponding directly to the flux from the oven. The As_2 and As_4 signals both decrease and, after passing through the transient change, reach new steady state levels. The RHEED intensity oscillations shown indicate that Ga-stable growth is occurring, which is further evidenced by a change from the arsenic-stable 2×4 reconstruction to the gallium-stable 4×2 reconstruction. When the gallium oven is closed at time=45 s, the gallium signal drops to zero and the arsenic signals recover to their initial steady-state values. This recovery is slow due to the excess surface gallium present when growth is terminated, and indicates the continuing incorporation of arsenic into the growing layers of GaAs. At 85 s, the arsenic shutter is closed to obtain a measure of background As_4 and As_2 in the chamber. After each run, the surface is annealed to 810 ± 5 K in a flux of arsenic and is cooled to the temperature of interest at a rate of ap-

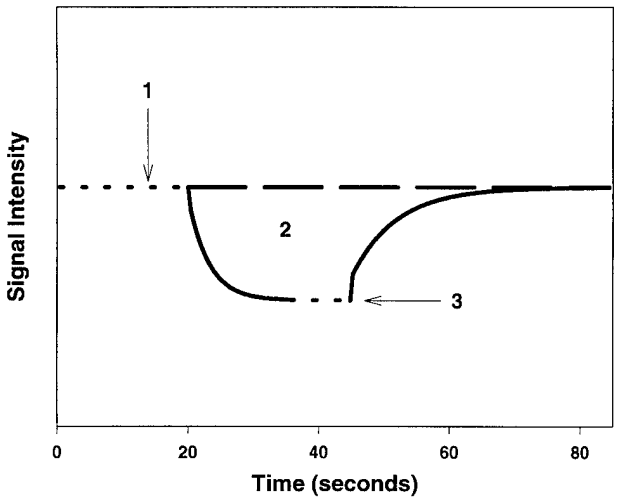


FIG. 5. Real-time behavior of a hypothetical arsenic time-of-flight signal. The gallium shutter is opened at 20 seconds and closed at 45 s. Level 1−Level 3=the steady-state decrease in arsenic signal and the area 2=the total incorporation of arsenic into the growing wafer.

proximately 1 K/s. Subsequent runs of data under nearly identical conditions are performed after the initial RHEED intensities match. Note that the scatter in the As_4 signal is clearly larger than that in the As_2 signal. We believe that the large flux of the tetramer molecule is affecting the noise created at the multiplier detector, possibly due to breakup of the larger tetramer molecules into multiple fragments. The gain on the detector, however, must be maintained at a level at which the As_2 and Ga mass signals can also be observed. Therefore, all of the data presented in this article will show a relatively large uncertainty for the tetramer molecules.

Figure 5 shows an example of a hypothetical arsenic TOFMS signal during deposition to explain the terminology used for the remainder of the results. The signal shown is analogous to either of the As_2 or As_4 signals displayed in Fig. 4. The "initial steady-state signal" is shown by the line marked 1 and describes the amount of arsenic present in the gas phase while the wafer is being bathed in an As_4 flux prior to gallium deposition. The level of the signal marked 3 is termed "steady-state signal during growth" and refers to the amount of arsenic present in the gas phase while GaAs growth is occurring. The term "steady-state decrease in signal" refers to the quantity obtained when the level of the dashed line marked 3 is subtracted from the level of the dashed line marked 1. This level can be likened to a change in the reactive sticking on the surface or a change in desorption rate, both of which indicate incorporation into the growing wafer. Note the transient region of arsenic incorporation, but this will not be considered here. Level 3 only includes the portion of signal that has reached a new steady-state level with time. The term "total integrated incorporation" refers to area 2 in the figure. This value is a measure of all the arsenic incorporated into the GaAs layers grown during and after gallium deposition. The empirical units of measured steady-state levels of arsenic are integrated TOFMS signal. Arsenic signals are normalized using the relative sensitivity ratios described earlier. The empirical units of incorporation are integrated TOFMS signal × time. The relative flux and mass corrections are made so that the steady-state levels of arsenic are directly proportional to the flux of arsenic atoms and the measured total integrated incorporation of arsenic is proportional to coverage of arsenic atoms.

The ^{69}Ga, As_2, and As_4 TOF and RHEED signals were recorded for each growth run in real time at a constant substrate temperature of 757 K and over a gallium flux range of 0.025–0.3 ML/s. As shown in Fig. 4, the shuttering scheme begins with bathing the surface in a 0.1 ML/s flux of As_4, which is sufficient to maintain an arsenic-stable RHEED reconstruction. The gallium oven shutter is opened at time=10 s and closed at time=45 s. The surface is then allowed to "recover," if necessary, under a flux of As_4 until time=85 s, then the As_4 shutter is closed to measure background arsenic species. An average of four runs of data under identical flux conditions followed by an average of every ten data points is made and these are used to obtain the steady-state decreases in arsenic species and integrated arsenic incorporation.

Figures 6 (a)–(c) demonstrate the behavior of the steady-state decrease in As_4 (a), As_2 (b), and total arsenic (c) as a function of velocity-corrected and integrated gallium TOF signal (which is directly proportional to gallium flux). The results indicate that the reactive sticking of As_4, As_2, and total arsenic increase with increasing gallium flux, but that the sticking is eventually saturated when enough oversupply of gallium is added to produce a gallium stable surface. Panels (d)–(f) of Fig. (6) show the total integrated incorporation of As_4 (a), As_2 (b), and total arsenic (c) at each gallium flux. These panels display an increasing linear trend in arsenic incorporation with increasing gallium flux. This linear behavior is a reflection of the additional incorporation of arsenic that occurs after the gallium oven shutter is closed and growth is terminated. The incoming arsenic tetramer flux eventually consumes the excess surface gallium at the higher gallium fluxes to produce stoichiometric GaAs.

The ^{69}Ga, As_2, and As_4 TOF and RHEED signals were also recorded in real time over the substrate temperature range of 620–830 K. The surfaces were initially bathed in a 0.2 ML/s flux of As_4, which is sufficient to maintain an arsenic-stable RHEED reconstruction over the temperature range studied. The Ga oven shutter was opened at time=20 s producing a Ga flux of 0.28 ML/s and was closed at time=65 s. The As_4 shutter was closed at time=85 s to measure the background arsenic. At each surface temperature, two runs of data were collected and averaged together.

Figure 7 is a plot of the initial As_2 steady-state signal level as a function of substrate temperature. This figure shows that the rate of desorption of As_2 increases with increasing surface temperature. Figures 8 (a)–(c) include plots of the steady-state decreases in As_4 (a), As_2 (b), and total arsenic (c) as a function of substrate temperature during growth. Panels (d)–(f) display the total integrated incorporation of As_4 (d), As_2 (e), and total arsenic (f) into the growing wafer as a function of substrate temperature. Panels (a) and (d) of Fig. 8 show that reactive sticking and incorporation

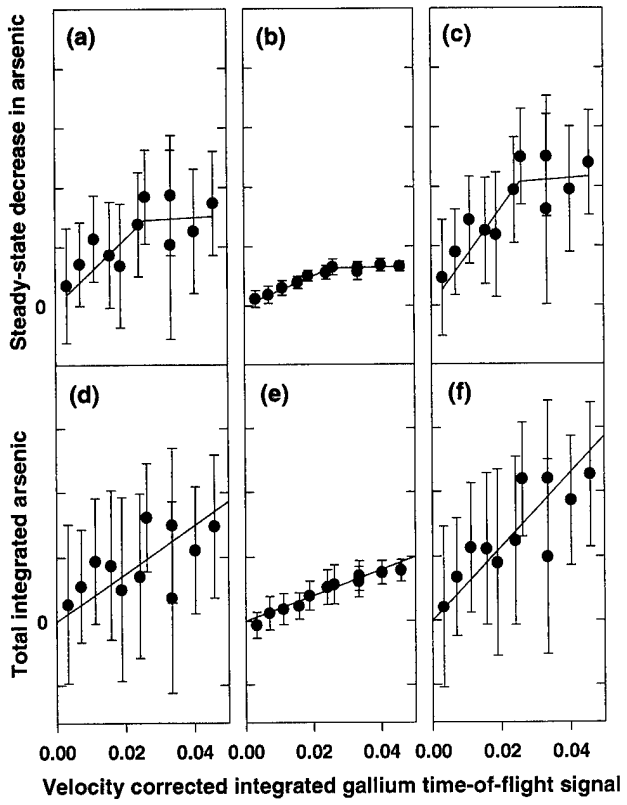


FIG. 6. Steady-state decreases in (a) As_4, (b) As_2 and (c) total arsenic, and total incorporation of (d) As_4, (e) As_2 and (f) total arsenic as a function of velocity corrected integrated Ga TOF signal when growing GaAs with As_4.

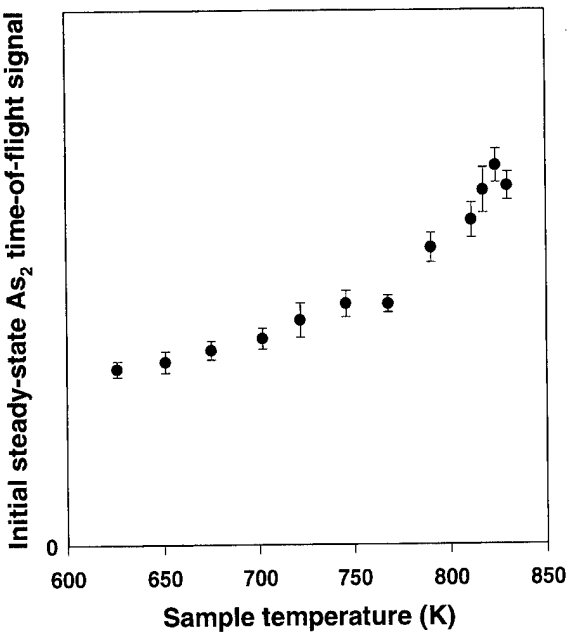


FIG. 7. The effects of varying the substrate temperature on the initial steady-state As_2 time-of-flight signal under a constant As_4 flux.

rate of As_4 dominate the incorporation of arsenic into the growing wafer at lower surface temperatures, while panels (b) and (e) of Fig. 8 show that reactive incorporation of As_2 increases to be almost equal to the As_4 at higher surface temperatures. The observed trends indicate a change in the mechanisms for growth of GaAs over this temperature range, and the data reflect a correlation in the incorporation probabilities of As_4 vs. As_2. Panels (c) and (f) show that the total arsenic reactive sticking and incorporation rate has a slight downward slope with increasing surface temperature, which suggests that the surface Ga population decreases at higher surface temperatures. Figure 9 shows the integrated Ga TOF signal over the temperature range studied at a constant Ga flux, which is a measure of the incident gallium atoms. The upward slope of this plot indicates that there may be an increase in the gas phase gallium atoms at elevated surface temperatures, also implying that the sticking probability of gallium decreases slightly at higher surface temperatures.

IV. DISCUSSION

We now turn to a more detailed discussion of Figs. 6–9. In Section A, a discussion of the GaAs surface prior to gallium deposition includes a proposed reaction scheme and suggests mechanisms to describe the observed trends in the initial steady-state arsenic signals. The surface during gallium deposition and growth is covered in Section B. Another reaction scheme is proposed with a discussion of mechanisms for each reaction. The effect of Ga/As_4 flux ratio is described in Section B (1) along with a more detailed discussion of Fig. 6. Section B (2) covers the effect of substrate temperature on the growth of GaAs. Figures 8 and 9 are described in more detail in terms of the proposed reactions and mechanisms.

A. The surface prior to gallium deposition

In the temperature range of our measurements and in the absence of the gallium flux, As_2 can desorb from the surface and/or the bulk of the GaAs wafer. The incoming As_4 molecules can then replace that arsenic lost from the wafer. The following reaction scheme is developed to describe these events:

$$\mathrm{As_4(in)} \rightarrow \begin{matrix} \nearrow\ \mathrm{As_4(g)}, \\ \mathrm{As_4(ph)} \rightarrow \mathrm{As_4(g)}, \end{matrix} \tag{1}$$

$$\mathrm{2As_4(ph)} \rightarrow \begin{matrix} \nearrow\ \mathrm{2As_2 + As_4(g)}, \\ \searrow\ \mathrm{2As_4(g)}, \end{matrix} \tag{2}$$

$$\mathrm{GaAs(b)} \rightarrow \mathrm{Ga(ch) + As_2(g)}, \tag{3}$$

$$\mathrm{As_2(ch)} \rightarrow \mathrm{As_2(g)}, \tag{4}$$

where in=incoming flux, ch=a chemisorbed state of the

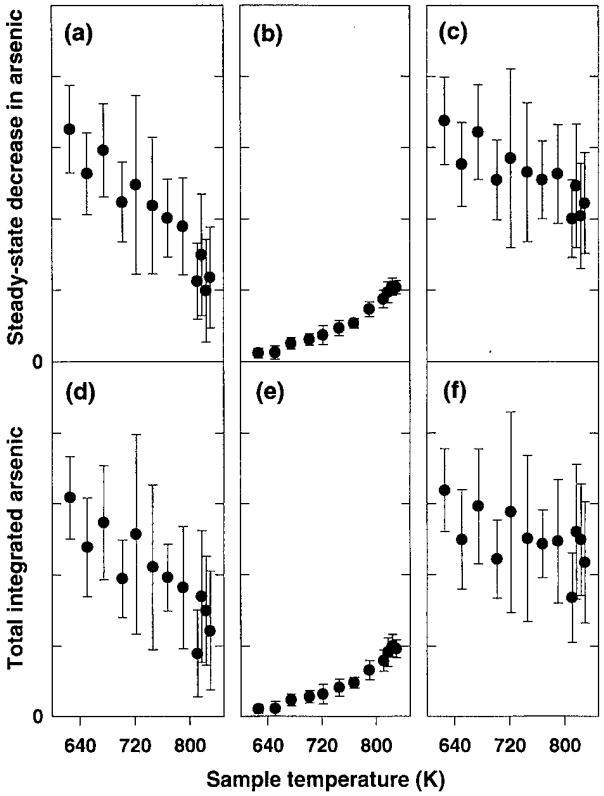


FIG. 8. Steady-state decreases in (a) As_4, (b) As_2 and (c) total arsenic, and total incorporation of (d) As_4, (e) As_2 and (f) total arsenic as a function of substrate temperature.

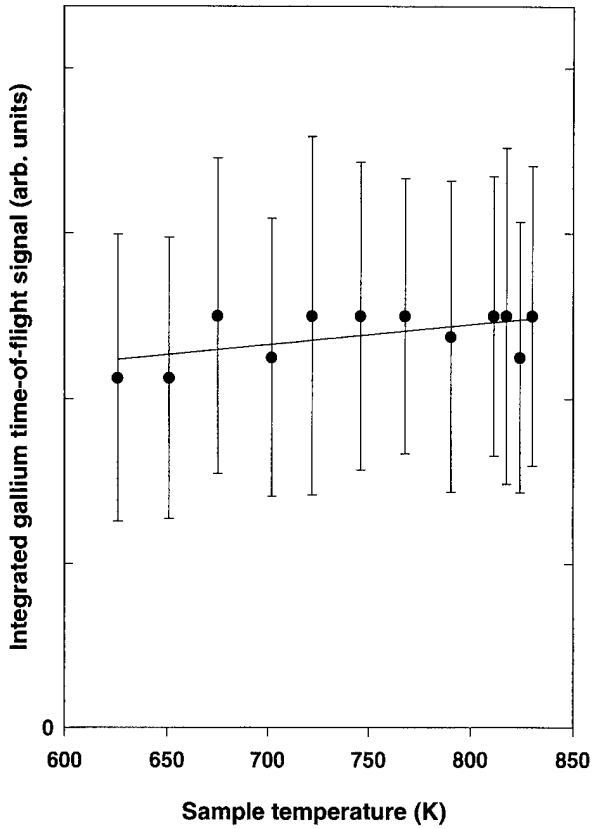


FIG. 9. Integrated gallium time-of-flight signal as a function of substrate temperature.

molecule, ph=a physisorbed state of the molecule, b=bulk, and g=a molecule in the gas phase (either specularly scattered or desorbed from the surface).

First, it is important to discuss which species are measured with the TOFMS. Reaction (1) describes what the incoming tetramer species can do once in contact with the surface. In these experiments we cannot distinguish between As_4 species that are specularly reflected [in this equation, denoted by $As_4(g)$] and those which are desorbed from the surface [here, denoted by $As_4(ph) \rightarrow As4(g)$]. The differences between these two gaseous As_4 molecules involve the geometry with which they scatter from the surface, their surface residence time, and the temperature with which they leave the surface. Specularly reflected or directly "scattered" arsenic tetramers would leave the surface at or near the specular angle from the incoming beam and move with a velocity similar to that of the incoming tetramer molecules, assuming there is little loss of energy to the wafer during contact. According to a standard model of surface desorption, desorbing tetramer molecules that are completely thermally accommodated would leave the surface in a cos (θ) distribution (where θ is the angle measured from the normal to the surface) and travel at a velocity determined by the temperature of the surface. The barrier to desorption of physisorbed As_4 has been measured to be 0.38 eV.[16] At the surface temperatures involved in this study, the surface lifetime of As_4 would be $< 10^{-5}$ s,[16] which is not measurable experimentally using our technique. Further, the spatial difference between the two types of molecules cannot be resolved because the laser beam probes a broad angular region in our system. A decrease in the detection of gas phase tetramer species is assumed to correspond to an increase in chemisorbed As_4. The wafer is therefore bathed in excess As_4 so that physisorbed As_4 is in constant supply as a precursor to chemisorption. The As_2 molecules detected must all result from desorption, since As_2 is not provided externally.

Reaction (2) describes how physisorbed As_4 can chemisorb to the surface. In the absence of gallium, As_4 has a sticking coefficient of zero on an arsenic stable surface at these surface temperatures. However, when gallium is present, As_4 binds to the surface via precursor [$As_4(ph)$] mediated chemisorption, which is dissociative and is known to obey second order kinetics, as demonstrated by Foxon and Joyce.[16] Reaction (3) describes the behavior of a GaAs wafer that is heated in the absence of an incoming arsenic flux. When heated, arsenic desorbs from the bulk as dimers, leaving gallium to form islands on the wafer surface, in turn driving the chemisorption of As_4. Therefore, as the surface temperature increases, one would expect a decrease in the amount of $As_4(g)$ signal, corresponding to increased chemisorption of the tetramer. Initial signal levels of As_4 do indeed decrease with increasing surface temperature.

Reaction (3) describes one source of the $As_2(g)$ signal. At an elevated surface temperature, arsenic from bulk GaAs has

enough energy to overcome the barrier to desorption. Second, because the initial surface is kept under sufficient As_4 flux to maintain the arsenic-stable 2×4 reconstruction, arsenic is assumed to exist as dimers on the surface.[39,40] These dimers will also desorb at an elevated surface temperature [Reaction (4)]. One would therefore expect an increase in As_2 signal as the surface temperature is raised. The data indeed show that the initial gas phase As_2 increases (Fig. 7) with increasing surface temperature.

A fit of the data displayed in Fig. 7 was attempted to elucidate an activation energy for desorption. Measurement of the amount of As_2 desorbing from the surface was assumed to be a measure of the overall rate of the Reaction (4):

$$\text{rate} \propto k \times (\theta^n),$$

where n=the order of the reaction with respect to coverage of arsenic, θ=coverage and k=rate constant for As_2 desorption. However, an Arrhenius plot using either a reaction order of one or two showed non-linearity, indicating that the activation energy and, therefore, the rate constant most likely depend on coverage. Further experiments that measure the absolute value of the rate constant or coverage are necessary for a complete kinetic analysis. In summary, the initial TOFMS levels of As_2 and As_4 increase and decrease, respectively, with increasing surface temperature.

B. The surface during gallium deposition and growth

The growing surface of GaAs is considerably more complicated than the initial wafer surface. The following reactions describe the interaction of incoming gallium with the arsenic species present on the surface before growth. Reactions (1)–(4) will still be referred to in this section and additionally:

$$\text{Ga(in)} \rightarrow \text{Ga}_n\text{(ch)} + \text{Ga(g)}, \tag{5}$$

$$(1/2)\text{As}_2\text{(ch)} + (1/n)\text{Ga}_n \rightarrow \text{GaAs(b)}, \tag{6}$$

$$2\text{As}_4\text{(ph)} + \text{Ga}_n \rightarrow 2\text{As}_2\text{(ch)} + \text{As}_4\text{(g)}. \tag{7}$$

The term Ga_n(ch) is used here to represent the fact that gallium can form different sized islands on the surface after deposition. Reaction (5) describes the behavior of the incoming flux of gallium atoms [Ga(in)]. At the surface temperatures used in this study, we assume that the incoming gallium chemisorbs on the surface with a sticking coefficient of unity at the lowest substrate temperatures.[41] The entire Ga TOFMS signal below this temperature is assumed to result from the incident flux of gallium atoms, while at higher surface temperatures the scattered flux from the heated surface additionally contributes to the signal. The surface temperature can also affect the size of these Ga clusters and the mobility of Ga atoms. The interaction of the surface gallium with As_2 and As_4 is described by Reactions (6) and (7).

Reaction (6) describes the incorporation of As_2 present on the surface with available Ga to form GaAs. A possible mechanism for this reaction includes a slow step in which the gallium must come into contact with As_2 either by landing on top of, or diffusing to, As_2 sites. The fast step of GaAs formation would occur provided enough energy is available to break As-As bonds and form Ga-As bonds. This proposed model assumes, however, that As_2 is relatively static on the arsenic-stable surface and not appreciably mobile. In other words, the barrier to diffusion for As_2 is assumed to be approximately equal to or greater than the barrier to desorption of As_2. A bonding argument can support this assumption. Four bonds to underlying Ga atoms must be broken or lengthened for each arsenic dimer to induce diffusion. It may be electronically more feasible for a second and third bond to form within the arsenic dimer species once these surface bonds are broken than for four new surface bonds to form elsewhere on the arsenic-stabilized GaAs. In terms of Reaction (6), an increase in chemisorbed gallium would increase the GaAs formation rate and decrease the surface desorption rate of As_2(g) via Reaction (4).

Reaction (7) describes a reaction that is governed by the same factors as Reaction (3). Available surface gallium still drives the chemisorption of As_4, but the arsenic is now supplied by incoming tetramers rather than by available dimers on the surface. The same precursor mediated dissociative chemisorption mechanism [As_4(ph)] applies in this case. Once arsenic dimers are chemisorbed, formation of GaAs with underlying Ga_n via Reaction (6) is rapid. Therefore, an increase in chemisorbed gallium is expected to drive the chemisorption of As_4 and result in a decreased rate of scattered or desorbed As_4(g). Both Reactions (6) and (7) predict that decreases in the As_4 or As_2 TOFMS signals can be directly related to incorporation of arsenic into the growing GaAs wafer.

1. Effect of Ga/As_4 flux ratio

During deposition, the steady-state decreases in arsenic TOFMS signals are affected by the incoming gallium flux. Figures 6 (a)–(c) show the change in these values with increasing gallium flux. As the Ga flux is raised, the steady-state decreases of As_4 and As_2 also rise. Available gallium sites increase the sticking probability of As_4 on the surface and decrease the desorption rate of As_2 from the surface. Eventually, the steady-state decreases are saturated by excess available Ga. The second order reaction of As_4 on a Ga-covered surface predicts that the As_4 sticking coefficient will become 0.5 on a gallium-rich surface.[16] The Ga-rich condition occurs when the flux (Ga)$\geq$2 × flux (As_4) and establishes the break in the linear increase of the steady-state signals shown in Figs. 6 (a)–(c). The As_2 desorption also is saturated at this Ga flux value, indicating that the usual desorption sites have been blocked by available gallium, and this also coincides with the As_4 turnover point when excess gallium is being supplied to the surface.

The change in the RHEED diffraction pattern at this point shows an alteration in the surface reconstruction from the arsenic-stable 2×4 to the gallium-stable 4×2 during growth, which further supports that the surface is now gallium rich. The period of the RHEED oscillations also indicates that growth of GaAs on a Ga-stable surface is occurring. Once the surface is gallium rich, formation of GaAs layers is limited by the available incident arsenic flux. The period of the

RHEED oscillation begins to reflect the rate of arsenic atom incorporation rather than the rate of gallium deposition, as it does at lower Ga/As_4 flux ratios. Growth of GaAs can no longer proceed stoichiometrically during deposition of gallium at these flux ratios.

Once the gallium shutter is closed, any excess surface Ga continues to react with incoming arsenic since chemisorption of As_4 is driven by available Ga sites. After the gallium shutter is closed, arsenic is still incorporated into the wafer. "Tails" appear in the As_2 and As_4 mass signals vs. time, indicating that reactions of arsenic species with excess surface gallium are still occurring. Figures 6 (d)–(f) describe the changes in the total incorporation of arsenic species as a function of Ga flux. Panel (f) illustrates that the incorporation of total arsenic is indeed linearly increasing through the entire Ga flux and indicates that growth will proceed stoichiometrically if the surface remains under sufficient As_4 flux. Because the Ga and As_4 ovens were shuttered at identical times for all data runs, the surfaces under the highest gallium fluxes were not given enough time to recover before the As_4 was closed for background arsenic readings. Therefore, the two points at the highest gallium fluxes in panels (d)–(f) lie below the line used to fit the arsenic incorporation as a function of gallium flux. The behavior of desorbing arsenic mass signals vs Ga flux matches the arsenic incorporation results obtained by Brennan, Tsao, and Hammons[31] in their REMS studies as well as early studies by Foxon and Joyce.[16] In both studies, saturation of reactive sticking of arsenic species and additional sticking on gallium-rich surfaces after termination of growth were observed.

2. Effect of surface temperature

The effect of surface temperature on the incorporation of arsenic is less straightforward than the effect of gallium flux. Figures 8 (a) and (b) show the steady-state decreases in As_4 and As_2 signals during gallium deposition as a function of the surface temperature. The figure shows a decrease in As_4 and an increase in the steady-state decrease of As_2 with increasing surface temperature, indicating that the rate of reactive sticking and incorporation of As_4 dominates at lower surface temperatures, while incorporation of As_2 becomes more significant at the higher surface temperatures.

This observed trend can be explained in terms of relative rates of diffusion of surface species and surface Ga cluster size. At low surface temperatures, Ga on the surface is unable to diffuse [E_{act} for Ga diffusion ~1.3 eV (Ref. 2)] appreciably, but the mobility of the gallium atoms increases with increasing surface temperature. Arsenic dimers are still assumed to be static on the surface and will desorb rather than diffuse at elevated surface temperatures. The arsenic tetramer molecules, however, have a very low activation barrier to diffusion, 0.25 eV,[16] and can diffuse appreciably at all surface temperatures studied.

The surface temperature can also affect the size of the gallium clusters on the surface. At low surface temperatures, gallium may be able to diffuse across the surface and find another gallium atom or cluster, bind there, and form a two-dimensional cluster. The desorption energy of Ga from a Ga cluster [~2.7 eV (Ref. 42)] can be assumed to be an upper limit on the energy of activation for Ga-Ga cluster bond breakage and subsequent Ga diffusion. Because this energy is still substantially greater than the barrier to diffusion, low surface temperatures may not allow gallium atoms to break away from the cluster and diffuse again as atoms on the surface. At higher surface temperatures, the gallium atoms can overcome this barrier and diffuse across the surface forming Ga-As bonds.

Monte Carlo simulations[17–20] and other theoretical studies have postulated that the formation of two-dimensional gallium deposits on the surface is key to arsenic incorporation. Although the details of incorporation mechanisms differ, most studies involve As_2 as the reactive incident arsenic species and require formation of a two-dimensional gallium cluster composed of up to four gallium atoms[20] for reaction with incoming arsenic. Steric factors may lead one to suspect that gallium cluster size is also an important factor for formation of GaAs using an incident flux of As_4.

The relative diffusion rates of Ga, As_2, and As_4 could affect the reaction to form GaAs in the following way. At low temperatures, one might expect that Ga would not only form large clusters, but would be more accessible to reaction with incoming As_4, since Ga and As_2 are relatively static compared to diffusing As_4. The next step would be the diffusion of physisorbed As_4 to Ga reaction centers [Reaction (7)], followed by GaAs formation via Reaction (6), in which As_2 is provided by dissociative chemisorption of As_4 on gallium. One would then expect to see a greater contribution to incorporation of arsenic in GaAs from As_4 at low temperatures and a subsequently larger relative decrease in As_4(g) as seen in Fig. 8 (a).

At higher temperatures, gallium is able to diffuse to the static As_2 sites on the surface. The pathway of Reaction (6) contributes more to the GaAs growth. Gallium atoms diffuse directly to As_2 sites and insert into dimer bonds, forming GaAs and decreasing the desorption of As_2(g) considerably. The lack of availability of Ga reaction centers for As_4 chemisorption, in turn, would result in a smaller sticking coefficient of As_4 and small relative steady-state decrease in As_4(g).

The total steady-state decrease in arsenic described in panel 8 (c) has a slightly negative slope ($-3.7 \pm 0.7 \times 10^{-5}$ [integrated TOF total arsenic signal/temperature (K)]), indicating that at the lowest temperatures the reactive sticking coefficient of arsenic on the surface during growth is increased. The chemisorption of both arsenic dimer and tetramer is driven by the amount of gallium on the surface [Reactions (6) and (7)]. Figure 9 shows a slight increasing trend in the amount of gas-phase gallium as the surface temperature is increased, indicating that there may be some scattered gallium from the heated GaAs surface. This correlates with a decrease in the sticking probability of incident gallium atoms on GaAs at higher surface temperatures. Arthur also observed a decrease in the sticking coefficient of gallium on

GaAs at surface temperatures higher than 750 K.[41] This decrease in available surface gallium may, in turn, decrease the reactive sticking of both arsenic species.

The total incorporation of arsenic is affected by these proposed changes in reactive sticking. Figure 8 (f) has a slope of $-1.3 \pm 0.6 \times 10^{-3}$ [integrated TOF total arsenic signal×time (s)/temperature (K)], indicating that the total arsenic incorporation decreases with increasing surface temperature. The total incorporation is limited by the amount of surface gallium available during the entire growth run. If the sticking coefficient of gallium decreases, then the total amount available for film growth also decreases.

The changes in the signals observed in Figs. 8 (c), and (f), and 9 reflect very small changes in the actual fluxes of gallium and arsenic species. The measured change in period of the RHEED oscillations from 0.26 to 0.28 (bilayers GaAs)/s over this temperature range is an indication that only small changes in the scattered flux of gallium and the total incorporated fluxes of arsenic species occur.

The TOFMS results are shown to be quite sensitive to changes in the surface temperature during growth, allowing us to explore possible mechanisms for incorporation. The proposed mechanisms are supported by arguments based on previously measured activation energies of diffusion and formerly obtained theoretical results.

V. CONCLUSIONS

We have measured changes in the gas phase steady-state levels of As_2, As_4, and Ga species during homoepitaxy of GaAs with simultaneous SPI-TOFMS and RHEED. Results obtained by varying the Ga/As_4 flux ratio are consistent with observed RHEED reconstructions and intensity oscillations as well as previously obtained results measured with modulated molecular beam experiments and REMS. Results obtained by varying the substrate temperature are also consistent with RHEED observations. A proposed model based on theoretical results and formerly measured activation energies for diffusion describes the incorporation of arsenic into GaAs in terms of gallium surface mobility and sticking probability.

SPI-TOFMS is a non-intrusive tool that complements the macroscopic information obtained by RHEED measurements by providing real time species-specific data on surface molecules. This method also permits measurement of growth rates and incorporation over a larger range than that of RHEED alone. Growth mechanisms and kinetic data on other MBE systems should also be easily elucidated by using this technique.

ACKNOWLEDGMENTS

Support of this research by the National Science Foundation, National Institute of Standards and Technology, and the donors of the Petroleum Research Fund, administered by the American Chemical Society, is gratefully acknowledged. One author (A. K. O.) gratefully acknowledges the support of an AT&T Ph.D. Fellowship, another author (S. M. C.) is supported by a National Research Council Postdoctoral Fellowship, and a third author (A. L. A.) thanks the Boettcher Foundation for a Graduate Fellowship.

[1] J. H. Neave, B. A. Joyce, P. J. Dobson, and N. Norton, Appl. Phys. A **31**, 1 (1983).
[2] J. H. Neave, P. J. Dobson, B. A. Joyce, and J. Zhang, Appl. Phys. Lett. **47**, 100 (1985).
[3] B. F. Lewis, R. Fernandez, A. Madhukar, and F. J. Grunthaner, J. Vac. Sci. Technol. B **4**, 560 (1986).
[4] R. Chow and R. Fernandez, Mater. Res. Soc. Symp. Proc. **145**, 13 (1989).
[5] A. Bosacchi, S. Franchi, Yu. O. Kanter, and S. I. Chikichev, J. Cryst. Growth **96**, 899 (1989).
[6] C. W. Farley and B. G. Streetman, J. Vac. Sci. Technol. B **6**, 749 (1988).
[7] C. W. Farley, G. J. Sullivan, M. J. Mondry, and D. L. Miller, J. Cryst. Growth **96**, 19 (1989).
[8] C. Deparis and J. Massies, J. Cryst. Growth **108**, 157 (1991).
[9] H. Nörenberg and N. Koguchi, Surf. Sci. **296**, 199 (1993).
[10] S. Yu. Karpov, Yu. V. Kovalchuk, V. E. Myachin, Yu. V. Pogorelsky, Surf. Sci. **314**, 79 (1994).
[11] J. C. Garcia, C. Neri, and J. Massies, J. Cryst. Growth **98**, 511 (1989).
[12] B. A. Joyce, T. Shitara, A. Yoshinaga, D. D. Vvedensky, J. H. Neave, and J. Zhang, Appl. Surf. Sci. **60/61**, 200 (1992).
[13] A. Yoshinaga, M. Fahy, S. Dosanjh, J. Zhang, J. H. Neave, and B. A. Joyce, Surf. Sci. Lett. **264**, L157 (1992).
[14] H. Yamaguchi and Y. Horikoshi, Appl. Surf. Sci. **60/61**, 224 (1992).
[15] C. T. Foxon and B. A. Joyce, Surf. Sci. **64**, 293 (1977).
[16] C. T. Foxon and B. A. Joyce, Surf. Sci. **50**, 434 (1975).
[17] S. V. Ghaisas and A. Madhukar, J. Vac. Sci. Technol. B **3**, 540 (1985).
[18] S. V. Ghaisas and A. Madhukar, Phys. Rev. Lett. **56**, 1066 (1986).
[19] S. V. Ghaisas and A. Madhukar, J. Appl. Phys. **65**, 3872 (1989).
[20] B. K. Medvedev, V. G. Mokerov, and N. V. Peskov, Phys. Dokl. **38**, 124 (1994).
[21] H. H. Farrell, J. P. Harbison, and L. D. Peterson, J. Vac. Sci. Technol. B **5**, 1482 (1987).
[22] J. P. Harbison and H. H. Farrell, J. Vac. Sci. Technol. B **6**, 733 (1988).
[23] H. H. Farrell and C. J. Palmstrom, J. Vac. Sci. Technol. B **8**, 903 (1990).
[24] J. Y. Tsao, T. M. Brennan, and B. E. Hammons, Appl. Phys. Lett. **53**, 288 (1988).
[25] A. J. SpringThorpe and P. Mandeville, J. Vac. Sci. Technol. B **6**, 754 (1988).
[26] F. G. Celii, Y. C. Kao, E. A. Beam III, W. M. Duncan, and T. S. Moise, J. Vac. Sci. Technol. B **11**, 1018 (1993).
[27] Y. C. Kao, F. G. Celii, and H. Y. Liu, J. Vac. Sci. Technol. B **11**, 1023 (1993).
[28] K. R. Evans, C. E. Stutz, D. K. Lorance, and R. L. Jones, J. Cryst. Growth **95**, 197 (1989).
[29] K. R. Evans, C. E. Stutz, D. K. Lorance, and R. L. Jones, J. Vac. Sci. Technol. B **7**, 259 (1989).
[30] J. Y. Tsao, T. M. Brennan, J. F. Klem, and B. E. Hammons, Appl. Phys. Lett. **55**, 777 (1989).
[31] T. M. Brennan, J. Y. Tsao, and B. E. Hammons, J. Vac. Sci. Technol. A **10**, 33 (1992).
[32] W. C. Wiley and I. H. McLaren, Rev. Sci. Instrum. **26**, 1150 (1955).
[33] A. H. Kung, J. F. Young, and S. E. Harris, Appl. Phys. Lett. **22**, 301 (1973).
[34] R. Mahon, T. F. McIlrath, V. P. Myerscough, and D. W. Koopman, IEEE J. Quantum Electron. **15**, 444 (1979).
[35] R. Hilbig and R. Wallenstein, IEEE J. Quantum Electron. **17**, 1566 (1981).
[36] P. G. Strupp, A. L. Alstrin, R. V. Smilgys, and S. R. Leone, Appl. Opt. **32**, 842 (1993).
[37] J. Berkowitz (private communication).
[38] J. R. Arthur and T. R. Brown, J. Vac. Sci. Technol. **12**, 200 (1975).
[39] M. D. Pashley, K. W. Haberern and W. Friday, Phys. Rev. Lett. **60**, 2176 (1988).
[40] A. R. Avery, D. M. Holmes, J. Sudijono, T. S. Jones, B. A. Joyce, Surf. Sci. **323**, 91 (1995).
[41] J. R. Arthur, J. Appl. Phys. **39**, 4032 (1968).
[42] E. M. Gibson, C. T. Foxon, J. Zhang, and B. A. Joyce, Appl. Phys. Lett. **57**, 1203 (1990).

Solid source molecular beam epitaxy of low threshold 1.55 μm wavelength GaInAs/GaInAsP/InP semiconductor lasers

F. G. Johnson,[a] O. King, F. Seiferth, and D. R. Stone
Laboratory for Physical Sciences, University of Maryland, College Park, Maryland 20740

R. D. Whaley and M. Dagenais
Department of Electrical Engineering, University of Maryland, College Park, Maryland 20742

Y. J. Chen
Department of Electrical Engineering, University of Maryland, Baltimore, Maryland 21228

(Received 6 March 1996; accepted 17 May 1996)

We report the growth and characterization of 1.55 μm wavelength GaInAsP based semiconductor lasers grown by solid source molecular beam epitaxy. Quaternary compositions were reproducible over time. Photoluminescence and x-ray diffraction spectra indicate abrupt quantum well interfaces. Separate confinement heterostructure laser diodes with four quantum wells had threshold current densities as low as 580 A/cm^2 and 275 A/cm^2 for unstrained $Ga_{0.47}In_{0.53}As$ and strained $Ga_{0.27}In_{0.73}As_{0.8}P_{0.2}$ wells, respectively. These results are as good as the best results reported for similar lasers grown by any growth technique. © *1996 American Vacuum Society.*

I. INTRODUCTION

Thin films of $Ga_xIn_{(1-x)}As_yP_{(1-y)}$ can be grown with lattice constants equal to InP and direct band gap energies between 0.75 and 1.35 eV. This material system has been studied extensively for its applications to 1.3 and 1.55 μm wavelength semiconductor lasers. A number of crystal growth techniques, gas source molecular beam epitaxy[1] (GSMBE), chemical beam epitaxy[2] (CBE), and metalorganic vapor phase epitaxy[3] (MOVPE), are successful in producing high quality unstrained 1.55 μm separate confinement heterostructure (SCH) lasers. Graded index separate confinement heterostructures (GRINSCH)[4] and compressive strain[5] lead to further improvements in performance. By optimizing these parameters, threshold current densities as low as 100 A/cm^2 have been achieved for single quantum well GRINSCH Lasers.[6]

Valved effusion cells containing solid arsenic[7] and phosphorus[8] allow for precise control of group V fluxes in MBE. These cells replace the toxic gases arsine and phosphine commonly used for $Ga_xIn_{(1-x)}As_yP_{(1-y)}$ epitaxy, by their less toxic, solid counterparts. By combining two valved effusion cells, arsenide/phosphide heterostructures with abrupt interfaces[9] and $Ga_xIn_{(1-x)}As_yP_{(1-y)}$ films lattice matched to InP[10] have been grown by solid source MBE. These advances led to the first report of a $Ga_xIn_{(1-x)}As_yP_{(1-y)}$ based SCH laser diode, operating at 1.35 μm, grown by solid source MBE.[11] In this article, we report $Ga_xIn_{(1-x)}As_yP_{(1-y)}$ based 1.55 μm SCH laser diodes grown by solid source MBE and compare their performance to laser diodes grown by GSMBE, CBE, and MOVPE.

The number of quantum wells, amount and sign of strain, and GRINSCH versus SCH core structure all affect the performance of laser diodes. To compare laser diodes grown by solid source MBE to those grown using other techniques, we chose structures similar to those reported by other researchers. We first grew a simple double heterostructure (DH) laser diode with an emission wavelength of 1.35 μm. Next, we grew two SCH laser diodes designed for emission at 1.55 μm. The first contained four unstrained $Ga_{0.47}In_{0.53}As$ quantum wells; whereas, the second contained four compressively strained $Ga_{0.27}In_{0.73}As_{0.8}P_{0.2}$ quantum wells. The threshold current densities J_{th} for unstrained SCH broad area laser diodes are comparable to reports in the literature. Broad area lasers made from the SCH with four compressively strained quantum wells had J_{th} values as good as or better than reports of similar laser diodes grown by GSMBE,[1] CBE,[2] or MOVPE.[3,5,12]

II. EXPERIMENT

The material was grown in a VG V80 MBE machine pumped by a 3000 ℓ/s cryopump and a 400 ℓ/s ion pump. Standard effusion cells containing elemental sources supplied the fluxes of gallium and indium as well as the *n*- and *p*-type dopants, silicon and beryllium. The InP growth rate was 1.3 monolayers/s, and the quaternary growth rates were near 1.8 monolayers/s. Two independent valved effusion cells (EPI Components, Inc.) were used to supply and regulate group V fluxes of As_4 and P_2 with cracker temperatures of 500 and 1000 °C, respectively. One cell contained elemental arsenic. The other cell was loaded with elemental red phosphorus that was later converted to white phosphorus before growth.[13]

The laser diode structures were grown on 2 in. sulfur-doped (100) InP substrates misoriented 4° towards [111]A. Wafers were etched for 1 min in a solution of 4 H_2SO_4: 1 H_2O: 1 H_2O_2, rinsed, and exposed to an ultraviolet-ozone cleaning step[14] for 3 min. In the growth chamber, the substrate was exposed to a phosphorus flux (P_2) with a beam equivalent pressure (BEP) of 1.5×10^{-5} mbar and heated to 500 °C. At this temperature, the surface reconstruction

[a]Electronic mail: FRED@LPS.UMD.EDU

changed from a diffuse 1×1 to a 2×4 pattern as seen by reflection high energy electron diffraction (RHEED). The temperature was increased an additional 25 °C and held for several minutes before dropping it to the growth temperature of 475 °C, which was constant during growth.

The layer structure for the DH laser diode consisted of 0.5 μm Si-doped (2×10^{18}) InP, 0.2 μm unintentionally doped $Ga_{0.29}In_{0.71}As_{0.65}P_{0.35}$, 2.0 μm Be-doped (2×10^{18}) InP, and a cap layer of 0.2 μm Be-doped (5×10^{18}) $Ga_{0.47}In_{0.53}As$. The unstrained SCH laser diode structure was 1 μm of Si-doped (5×10^{17}) InP, 0.24 μm unintentionally doped core, 1.5 μm of Be-doped (5×10^{17}) InP, and a 0.2 μm Be-doped (1×10^{19}) $Ga_{0.47}In_{0.53}As$ cap. The core consisted of $Ga_{0.27}In_{0.73}As_{0.57}P_{0.43}$ (band gap wavelength $\lambda_g=1.27$ μm) and four lattice matched 65 Å $Ga_{0.47}In_{0.53}As$ quantum wells separated by 125 Å. The strained SCH laser diode had a similar structure except for the core, which contained four compressively strained ($\epsilon\approx+0.7\%$) 105 Å $Ga_{0.27}In_{0.73}As_{0.8}P_{0.2}$ quantum wells separated by 100 Å. Interruptions of 10 s were inserted at each interface during the growth to change the group V fluxes.

III. RESULTS AND DISCUSSION

The compositions of the layers in the laser diodes were confirmed from the characterization of $Ga_{0.27}In_{0.73}As_{0.57}P_{0.43}$ and $Ga_{0.47}In_{0.53}As$ reference layers grown before and after the lasers. The precise compositions were determined from the lattice mismatch and the band gap. Symmetric (004) rocking curves acquired by a double-crystal x-ray diffractometer were used to examine the lattice mismatch of the quaternary reference films. All films were nearly lattice matched with InP. The full width at half-maximum (FWHM) linewidths of the quaternary peaks (30 arcsec) were broader than the linewidths of the InP substrate peaks (20 arcsec). Band gaps were measured using room temperature photoluminescence (PL). The quaternary PL peaks had FWHM linewidths between 42 and 48 meV, which are similar to the best results reported for MBE[10,15] and comparable to other growth techniques.

The reproducibility of the group V composition over time was remarkable. Lattice matched quaternary films were grown by MBE by simply adjusting the group V fluxes to a calibrated value as measured by an ion gauge. Three lattice matched quaternary films with compositions near $Ga_{0.27}In_{0.73}As_{0.57}P_{0.43}$ were grown, each separated from the others by a period over a week. Before each growth, the As_4 and P_2 fluxes were adjusted to give BEPs, measured by an ion gauge, of 1.3×10^{-5} mbar for As_4 and 1.7×10^{-5} mbar for P_2. The lattice mismatch ($\Delta a/a$) varied within ±0.0005, corresponding to a change in As mole fraction of less than ±0.015. By converting the BEPs into atomic fluxes,[20] the relative atomic incorporation coefficient of As_4 to P_2 for these growth conditions was found to be 9:1 and is similar to the values reported by others.[15]

The laser diode samples were also characterized using photoluminescence and x-ray diffraction. The cap and the upper cladding layer were removed from the SCH lasers before collecting the PL spectra. The PL spectra from the unstrained SCH laser is shown in Fig. 1(a). The peak at 1.57 μm (FWHM 27 meV) corresponds to the electron to heavy-hole transition. The electron to light hole transition appears as a higher energy shoulder. Figure 1(b) is the corresponding spectrum at low temperature (15 K). The peak position and linewidth are 1.45 μm and 6 meV, respectively. This narrow linewidth is evidence of abrupt interfaces and uniformity of the four quantum wells.[16] A smaller peak appears at 22 meV lower energy and is likely associated with an acceptor bound exciton transition.[17] The PL results from the strained SCH are similar. The room and low temperature PL peaks appear at 1.55 and 1.43 μm, respectively, with FWHM linewidths of 25 and 7 meV.

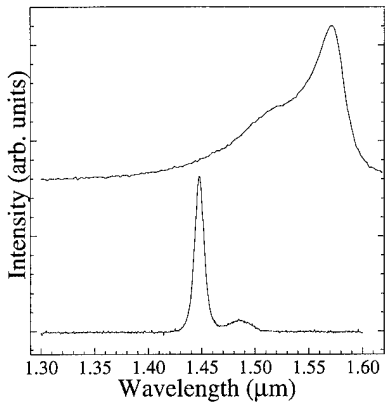


FIG. 1. Photoluminescence spectra at (a) 300 K and (b) 15 K from a SCH with unstrained quantum wells.

Symmetric (004) x-ray rocking curve spectra were analyzed for each laser diode sample. The spectrum from the strained SCH laser structure is shown in Fig. 2(a). The central peak results from the InP substrate and cladding layers. The $Ga_{0.47}In_{0.53}As$ cap and $Ga_{0.27}In_{0.73}As_{0.57}P_{0.43}$ core layers are nearly lattice matched to InP and appear as a weak shoulder on the central peak. Three superlattice satellite peaks can be resolved to the left of the InP central peak. These result

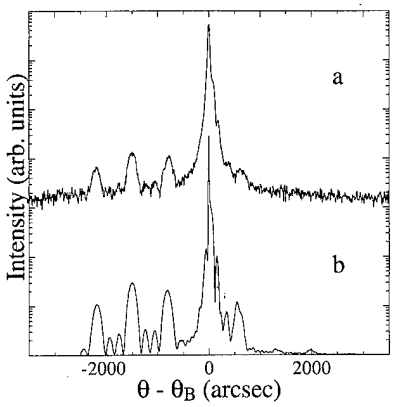


FIG. 2. Symmetric (004) x-ray diffraction spectra from a SCH with compressively strained wells: (a) experiment; (b) simulation.

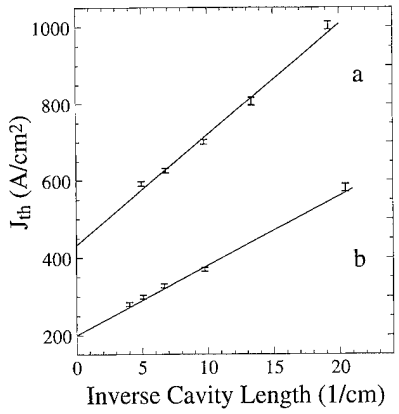


FIG. 3. Plot of Jth vs inverse cavity length for SCH laser diodes with (a) unstrained and (b) compressively strained quantum wells.

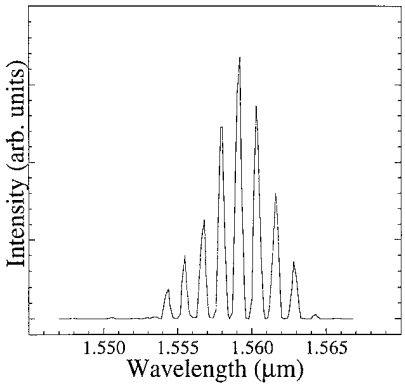


FIG. 4. Emission spectrum of a 500-μm-long SCH laser diode.

from the long range periodicity introduced within the core by the four strained quantum wells. A simulated spectrum was obtained using a commercially available dynamical simulation[18] and appears in Fig. 2(b). The best fit was found for a model structure with abrupt interfaces and quaternary compositions given above. A superlattice periodicity of 220 Å, which is close to the 205 Å value expected, gave the best fit of simulated to experimental data.

Broad area laser diodes were processed from the three laser samples. The n contact was formed by blanket metalization with AuGe/Ni/Au, and the p contact was defined by 100-μm-wide Ti/Pt/Au stripes. A selective etch was used to remove the 0.2 μm $Ga_{0.47}In_{0.53}As$ contact layer to obtain electrical isolation. The facets were formed by cleaving and were uncoated. Threshold current densities were determined using the dimensions of the p contacts, and no correction was made for current spreading.

Laser diodes produced from the DH laser structure had performance characteristics typical of high quality material.[19] Diodes with cavity lengths of 1 mm had an emission wavelength of 1.35 μm and a threshold current density J_{th} of 900 A/cm^2. The characteristic temperature T_0 was 62 K. Unstrained SCH laser diodes also performed well. The emission wavelength for a cavity length of 1 mm was 1.57 μm, and J_{th} versus inverse cavity length is plotted in Fig. 3(a). J_{th} was 580 A/cm^2 for a cavity length of 2 mm. This value is comparable to values from the literature of 500 to 700 A/cm^2 for similar devices.[1-3,5] The internal quantum efficiency η_i was 0.75 and the internal loss α was 20 cm^{-1}. T_0 was measured for temperatures between 15 and 55 °C and found to be 55 K.

The J_{th} values for laser diodes processed from the SCH containing four compressively strained ($\epsilon \approx +0.7\%$) quantum wells are plotted as a function of inverse cavity length in Fig. 3(b). Diodes had J_{th} values as low as 275 A/cm^2 for a cavity length of 2.5 mm. The values of T_0 (47 K), η_i(0.67), and α (12 cm^{-1}) are also similar to published results. The spectrum of a 500-μm-long device is shown in Fig. 4.

The J_{th} values reported above are better than results reported for GSMBE[1] and CBE[2] and as good as the best results reported for four quantum well SCH laser diodes with compressive strain grown by MOVPE.[5,12] The best J_{th} values reported for these devices decrease with increasing strain: 324 A/cm^2 ($\varepsilon \approx +0.9\%$),[5] 286 A/cm^2 ($\varepsilon \approx +1.2\%$),[12] and 268 A/cm^2($\varepsilon \approx +1.8\%$).[5] We expect further reductions in J_{th} below the 275 A/cm^2 ($\varepsilon \approx +0.7\%$) reported in this letter, by increasing the strain.

IV. CONCLUSION

We have shown that solid source MBE can produce high quality GaInAsP based laser diodes on InP. Quaternary compositions can be reproduced by simply adjusting the group V fluxes, as measured by an ion gauge, to a calibrated value before each growth. Quantum well interfaces are abrupt as determined by photoluminescence and x-ray diffraction measurements. The threshold current densities of 1.55 μm SCH laser diodes grown by solid source MBE are comparable with the best results reported to date for any growth technique.

ACKNOWLEDGMENTS

The authors are grateful to W. T. Beard, T. L. Filemyr, K. Mobarhan, R. A. Wilson, and C. E. C. Wood for useful discussions and technical assistance.

[1]C. Starck, J. Y. Emery, R. J. Simes, M. Matabon, L. Goldstein, and J. Barrau, J. Cryst. Growth **120**, 180 (1992).
[2]W. T. Tsang, F. S. Choa, M. C. Wu, Y. K. Chen, A. M. Sergent, and P. F. Sciortino, Jr., Appl. Phys. Lett. **58**, 2610 (1991).
[3]A. Ougazzaden, J. D. Ganiere, Y. Gao, E. V. K. Rao, B. Sermage, C. Kazmierski, and A. Mircea, J. Cryst. Growth **107**, 761 (1991).
[4]H. Temkin, T. Tanbun-Ek, R. A. Logan, J. A. Lewis, and N. K. Dutta, Appl. Phys. Lett. **56**, 1222 (1990).
[5]J. S. Osinski, P. Grodzinski, Y. Zou, and P. D. Dapkus, IEEE J. Quantum Electron. **29**, 1576 (1993).
[6]N. Yamamoto, K. Yokoyama, T. Yamanaka, and M. Yamamoto, Electron. Lett. **30**, 243 (1994).
[7]D. L. Miller, S. S. Bose, and G. J. Sullivan, J. Vac. Sci. Technol. B **8**, 311 (1987).
[8]G. W. Wicks, M. W. Koch, J. A. Varriano, F. G. Johnson, C. R. Wie, H. M. Kim, and P. Colombo, Appl. Phys. Lett. **59**, 342 (1991).

[9]F. G. Johnson and G. W. Wicks, Appl. Phys. Lett. **63**, 2402 (1993).
[10]J. N. Baillargeon, A. Y. Cho, F. A. Thiel, R. J. Fischer, P. J. Pearah, and K. Y. Cheng, Appl. Phys. Lett. **65**, 207 (1994).
[11]M. Toivonen, A. Salokatve, M. Jalonen, J. Nappi, H. Asonen, M. Pessa, and R. Murison, Electron. Lett. **31**, 797 (1995).
[12]P. J. A. Thijs, L. F. Tiemeijer, P. I. Kuindersma, J. J. M. Binsma, and T. Van Dongen, IEEE J. Quantum Electron. **27**, 1426 (1991).
[13]W. E. Hoke, D. G. Weir, P. J. Lemonias, H. T. Hendriks, G. S. Jackson, and P. Colombo, J. Vac. Sci. Technol. B **13**, 733 (1995).
[14]J. A. McClintock, R. A. Wilson, and N. E. Byer, J. Vac. Sci. Technol. **20**, 241 (1982).
[15]C. C. Wamsley, M. W. Koch, and G. W. Wicks, J. Vac. Sci. Technol B (to be published).
[16]M. S. Skolnick, K. J. Nash, M. K. Saker, S. J. Bass, P. A. Claxton, and J. S. Roberts, Appl. Phys. Lett. **50**, 1885 (1987).
[17]K. H. Goetz, D. Bimberg, H. Jurgensen, J. Selders, A. V. Solomonov, G. F. Glinskii, and M. Razeghi, J. Appl. Phys. **54**, 4543 (1983).
[18]Bede Scientific, computer code RADS, Bede Scientific, Durham, UK.
[19]T. Sudersena Rao, C. Lacelle, S. J. Rolfe, M. Dion, J. Thompson, P. Marshall, P. Chow-Chong, D. Ross, M. Davies, and A. P. Roth, Appl. Phys. Lett. **65**, 1015 (1994).
[20]C. E. C. Wood, D. Desimone, K. Singer, and G. W. Wicks, J. Appl. Phys. **53**, 4230 (1982).

Molecular beam epitaxy grown ZnSe studied by reflectance anisotropy spectroscopy and reflection high-energy electron diffraction*

J. -T. Zettler,[a] K. Stahrenberg, and W. Richter
Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstr. 36, D-10623 Berlin, Germany

H. Wenisch, B. Jobst, and D. Hommel[b]
Physikalisches Institut der Universität Würzburg, Germany

(Received 1 April 1996; accepted 10 May 1996)

In this work we report on the *in situ* investigation of MBE grown ZnSe surfaces performed simultaneously by reflectance anisotropy spectroscopy (RAS) and reflection high-energy electron diffraction (RHEED). Reconstructions and anisotropic reflectance of the ZnSe(001) surface are studied in the temperature range from 50 °C to 380 °C. With increasing temperature the ZnSe surface evolves from a Se-rich $c(2\times2)_{Se}$ via a Se-rich (2×1) to a Zn-rich $c(2\times2)_{Zn}$ surface reconstruction with the transition temperatures depending on whether the surface is Se stabilized or not. Each surface reconstruction as verified by RHEED is accompanied by a characteristic RAS spectrum. Time resolved measurements of the RAS signal at fixed photon energies allowed to determine the activation energy (0.7 eV) for the Se desorption from the (2×1) reconstructed ZnSe(001) surface. © *1996 American Vacuum Society.*

I. INTRODUCTION

Since reflectance anisotropy spectroscopy (RAS/RDS — reflectance difference spectroscopy) has been introduced as a new surface sensitive optical technique[1] many studies on III–V semiconductor surfaces have been performed by a number of groups establishing correlations between the surface reconstructions measured by reflection high-energy electron diffraction (RHEED) and the reflectance anisotropy spectra.[2–4] In this work we report on a comparative RHEED and RAS investigation on the II–VI semiconductor surface of ZnSe(001).

Wide gap ZnSe-based II–VI compounds are of increasing interest due to their application in blue-green laser diodes operating at room temperature.[5–8] One of the reasons for the rather short life time of these laser diodes are dislocations formed at the ZnSe/GaAs interface.[9] It is therefore of importance to analyze the growth of wide gap II–VI compounds on GaAs substrates in more detail. One of the mostly used *in situ* characterization methods during ZnSe molecular beam epitaxy (MBE) is reflection high-energy electron diffraction (RHEED). Two dominating ZnSe(001) surface reconstructions have been found, a (2×1) and a $c(2\times2)$, assigned to Se-rich and Zn-rich surface reconstructions, respectively.[10] Recently, applying the electron counting rule[11] and performing first principles total-energy calculations it was stated in Refs. 12 and 13, that the Zn-rich $c(2\times2)$ surface reconstruction is characterized by a $\Theta_{Zn}=0.5$ monolayer (ML) Zn-coverage that is not forming any dimer bonds. If the ZnSe surface is exposed to a sufficiently high Se beam pressure a Se-rich (2×1) reconstruction is typically formed[10] characterized by a $\Theta_{Se}=1.0$ ML Se-coverage with one Se-Se dimer per surface unit mesh.[12,13] This is supported by recently published STM results[14]: Se dimers oriented along $[\bar{1}10]$ have been identified on (2×1) reconstructed ZnSe overlayers (2 ML thick) on GaAs(001). However, high resolution STM images of relaxed ZnSe(001) surfaces to our knowledge are still lacking.

Recently, the existence of an additional Se-rich $c(2\times2)$ reconstruction at low temperatures and under extremely Se-rich conditions was predicted from total energy calculations by Garcia and Northrup.[12,13] [For distinction and clarity we use in the following the notation: Zn-rich $c(2\times2)$ $\equiv c(2\times2)_{Zn}$ and Se-rich $c(2\times2)\equiv c(2\times2)_{Se}$.]

Reflectance anisotropy spectroscopy (RAS), which we have applied to the ZnSe (001) surfaces additionally to RHEED, is an optical *in situ* technique highly sensitive to surface effects on cubic III–V and II–VI semiconductors while suppressing their isotropic bulk contribution. RAS spectra yield information on type, direction and concentration of surface dimers with only limited and indirect response to long range order (see, e.g., Ref. 3 and references therein). Therefore RAS is complementary to RHEED measurements but for II–V semiconductors it has been applied up to now only *ex situ* for the contactless determination of carrier concentration in doped and oxidized ZnSe(001) layers.[15]

In this work the Se-rich $c(2\times2)_{Se}$ reconstruction of ZnSe (001) predicted by Garcia and Northrup[12,13] is experimentally verified by RHEED. A RAS data base is gained for the complete set of reconstructions characteristic for the (001) surface of relaxed ZnSe in the 50 °C to 380 °C substrate temperature range. The RHEED patterns which are well proven as an effective MBE growth monitor[16] are used here to assign surface reconstructions to the related characteristic RAS features. Because RAS can monitor local changes at the surface before long range order is established it is utilized in this work also to study the dynamics of Se desorption from the ZnSe (001) surface in UHV.

*Published without author corrections.

[a]Electronic mail: zettler@marie.physik.tu

[b]Present address: Institut für Festkörperphysik, Universität Bremen, Germany.

II. EXPERIMENT

The experiments were performed in a II/VI growth chamber for the ZnSe material system that is connected via an UHV module track to a separate GaAs growth chamber (within a 6-chamber RIBER 2300 system). Elemental Zn and Se source materials of 6N purity have been used. The RHEED system is equipped with a 35 keV electron gun and the reflection patterns can be monitored with a CCD camera. Substrates were exactly oriented ($\pm 0.5°$ off) GaAs (001) wafers, being chemically etched before loading them into the system. GaAs-buffers (200 nm, undoped) were grown in the III–V chamber at 600 °C to obtain an atomically flat and clean surface. After transferring them into the II–VI MBE chamber a well ordered As-rich $c(4\times4)$ reconstruction was verified by RHEED and RAS at a substrate temperature of $T=50$ °C. Subsequently, at $T=300$ °C a ZnSe buffer was grown thick enough (about 1.5 μm) to suppress Fabry–Perot interferences at photon energies above the E_0 gap (about 2.7 eV) of ZnSe. Below 2.7 eV ZnSe is optically transparent and therefore the RAS spectra are dominated by strong interference patterns due to the multiple reflections between the ZnSe/GaAs interface and the sample surface. Because of the GaAs-ZnSe lattice mismatch the ZnSe relaxes when a buffer thickness of about 200 nm is reached.[17] Therefore, in the about 300 nm thick substrate near region of the ZnSe layer the density of dislocations is very high yielding a significant optical bulk anisotropy. Otherwise, once the ZnSe buffer is grown thick enough (>1 μm) according to cross-sectional transmission electron microscopy images[17] only a few dislocations are reaching the surface. This is in line with the observation of RHEED patterns typical for two-dimensional growth.

For photon energies less than 2.7 eV the superposition of dislocation induced bulk anisotropy, surface anisotropy, and interference effects complicates the analysis of the RAS spectra. Therefore we discuss in the following only the 2.7 eV to 5.0 eV RAS features. These are predominantly surface related, i.e., they probe only the uppermost monolayers of the sample. The RAS system which follows the standard design of Aspnes[1] was mounted at a strain-reduced pyrometer window of the MBE chamber.

The real part of the reflectance anisotropy

$$\frac{\Delta r}{r}=2\frac{r_{[\bar{1}10]}-r_{[110]}}{r_{[\bar{1}10]}+r_{[100]}} \tag{1}$$

is presented in the following. Below the E_1 bulk critical point, i.e., below about 4.5 eV for ZnSe, the real part of the reflectance anisotropy is directly related to the imaginary part of the surface dielectric anisotropy ($\epsilon_{[\bar{1}10]}-\epsilon_{[110]}$).[2] Therefore it mirrors the anisotropic absorption properties of the reconstructed ZnSe surface. The imaginary part of $\Delta r/r$ was also measured and used to check the Kramers-Kronig consistency between the real and imaginary part of the measured data thus minimizing systematic errors.[18] The influence of residual window strain on the RAS signal was compensated by taking spectra in two perpendicular sample orientations. The window anisotropy cancels out when the two spectra are subtracted from one another.[19]

Light induced thermal heating effects and UV-photon activated reactions can cause growth modifications as it is known from ZnSe photo-MBE.[20] The 75 W Xe-arc lamp used in our RAS setup yields a total energy flux of about 0.3 W/cm^2 to the sample with only about 0.1 W/cm^2 in the spectral region above the ZnSe E_0 gap. Therefore, and because no change in the spectra was observed when part of the light was blocked with filters, photo-induced modifications of the ZnSe surface can be excluded.

III. RESULTS AND DISCUSSION

The meanwhile well studied GaAs(001) surface is known to form a variety of As-rich reconstructions [$c(4\times4)$, $\gamma(2\times4)$, $\beta(2\times4)$, $\alpha(2\times4)$] with As coverages decreasing with increasing substrate temperature (1.75, 1.0, 0.75, and 0.5 monolayers, respectively).[21,22] In order to check if similarly on ZnSe several Se-rich surface reconstructions can be found, we studied the ZnSe(001) surface in the 50 °C to 380 °C substrate temperature range. For starting from a very Se rich surface the sample was initially *in situ* capped at $T=50$ °C with an amorphous (a-) Se layer. The results of these experiments are summarized in Fig. 1. The spectra were taken with increasing substrate temperature under two conditions: without Se stabilization [Fig. 1(a)] and under a Se beam pressure of $p_{Se}=2.8\times10^{-6}$ Torr [Fig. 1(b)]. At about 100 °C the Se cap desorbs and a $c(2\times2)$ RHEED pattern appears with clear reconstructions in the [100] and [010] directions. To our knowledge this is the first observation of a $c(2\times2)_{Se}$ reconstruction at low temperatures and under extremely Se-rich conditions which was predicted recently from total energy calculations by Garcia and Northrup.[12,13] The RAS signal is positive in the complete spectral range and features a characteristic minimum (termed B in Fig. 1) at about 3.7 eV. This RAS feature found in our experiments to be typical for this reconstruction should be related to the Se trimers which according to[12,13] form an energetically preferred $c(2\times2)_{Se}$ reconstruction with 1.5 ML Se surface coverage. The authors in Ref. 13 stated that this $c(2\times2)_{Se}$ is unlikely to be prepared from Se capped surfaces. This is apparently incorrect, at least for the *in situ* capped and decapped samples used in our experiments. However, the RAS feature B is more pronounced under Se flux indicating that a sufficiently high Se supply is essential for the stabilization of the $c(2\times2)_{Se}$ reconstruction.

Further increase in temperature transforms the surface into a (2×1) reconstruction with the transition temperatures depending on whether a stabilizing Se-flux is utilized or not (about 220 °C and 150 °C, respectively). The RAS spectrum of this well known Se-rich (2×1) reconstruction is characterized by a minimum at about 3.0 eV (termed A in Fig. 1) and a positive RAS signal increasing with photon energy to a maximum at about 4.8 eV. We attribute this characteristic RAS minimum A to Se dimers found to be typical for the (2×1) surface by first principles total-energy calculations[12,13]

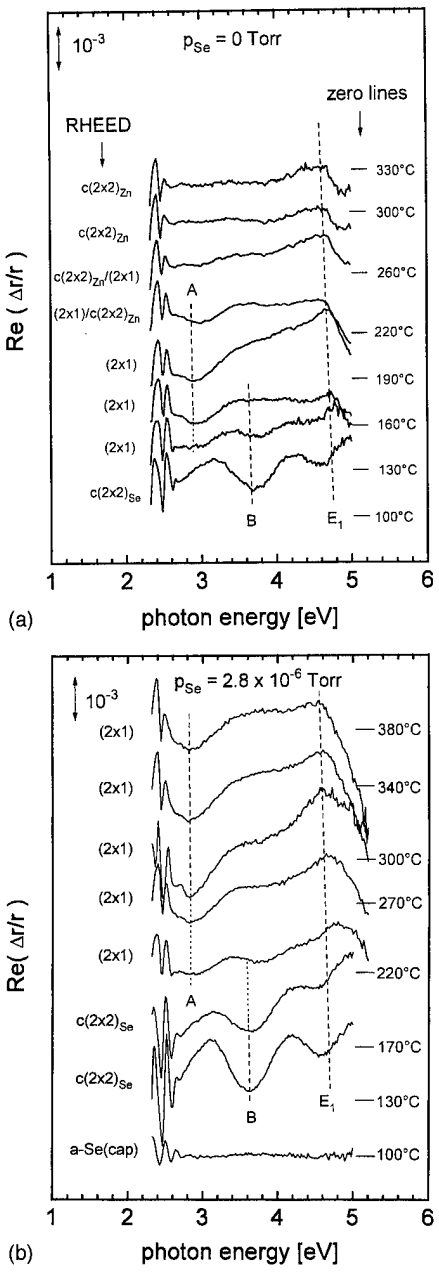


FIG. 1. RA spectra taken during heating up a just grown and subsequently Se capped thick ZnSe layer on GaAs(001) (a) without Se stabilization and (b) under a constant Se pressure of $p_{Se}=2.8\times10^{-6}$ Torr. The characteristic RAS features are labeled by A, B, and E_1. We assume A and B to be dimer and trimer related resonances, respectively. E_1 is close to a bulk critical point. The RAS zero level and the substrate temperature belonging to each spectrum are given in the right part of the figure. The surface reconstructions measured by RHEED synchronously to the RAS spectra are given in the left part of the figure.

and by relative peak intensity analysis in grazing incidence x-ray scattering experiments.[23] In the 4.8 eV to 5.5 eV photon energy range a sharp decrease to negative RAS values is found. This feature was labelled E_1 in Fig. 1 because its energy position is close to the ZnSe E_1 bulk critical point.[24] which might be modified by surface effects or remaining strain of the incompletely relaxed ZnSe layer.

Under Se-stabilization the (2×1) reconstruction is stable in the complete temperature range from 220 °C to 380 °C [Fig. 1(b)] with the RAS features A and E_1 shifting to lower energies with increasing temperature. Without Se stabilization the surface transforms at about 260 °C to the Zn-rich $c(2\times2)_{Zn}$ reconstruction. The related RAS spectrum is quite flat in the complete spectral range except for the 4.8 eV E_1 feature. This is again consistent to the total-energy calculations in Refs. 12 and 13 where it was concluded that the $c(2\times2)_{Zn}$ surface reconstruction is characterized by a Zn-coverage of $\Theta=0.5$ ML not forming any dimer bonds.

After monitoring the transformations of the ZnSe surface occurring with increasing substrate temperature the inverse process was studied by slowly cooling the sample down to 50 °C under constant Se flux. Both the $c(2\times2)_{Se}$ and the (2×1) reconstruction were found again.

The Fabry–Perot interferences at photon energies below the E_0 bulk critical point of ZnSe at about 2.6 eV could not be quantitatively interpreted even for known ZnSe thicknesses due to the superposition of surface anisotropy and bulk anisotropy effects. The contribution of bulk anisotropy is obvious in the a-Se RAS spectrum in Fig. 1(b). The surface of amorphous cap layers is known to be isotropic[25] and in fact gives a zero line signal in the high photon energy part of the spectrum. Below 2.7 eV, where the Se-cap is sufficiently transparent, predominantly ZnSe bulk anisotropies originating from the partially relaxed and dislocation rich region close to the ZnSe/GaAs interface should contribute to the interference modified RAS features. Forthcoming experiments with ZnSe layers grown on ZnSe (001) substrates should allow for the investigation of surface anisotropy also in this low-energy spectroscopic range.

In a further set of experiments we studied dynamic changes of the ZnSe(001) surface stoichiometry. For this purpose the Se desorption with increasing substrate temperature was monitored with 100 ms time resolution. The 3.1 eV photon energy was chosen for this experiment because it is sensitive to both the presumably dimer related RAS minimum A and the trimer related feature B [Fig. 1(a)]. The result of the desorption experiments is given in Fig. 2. Obviously the surface phase transition from $c(2\times2)_{Se}$ to (2×1) reconstruction is shifted towards higher temperatures by about 50 K in the case of Se stabilization.

At substrate temperatures above 250 °C this phase transition from $c(2\times2)_{Se}$ to (2×1) reconstruction can be driven just by switching off the Se flux and thus the Se desorption can be studied. The slope of the logarithmically plotted transients from the (2×1) to the $c(2\times2)_{Zn}$ RAS level at 3.1 eV yields the Se desorption rate. From the slope of the Arrhenius plot of the desorption rate (insert in Fig. 2) an activation energy of 0.7 eV for the Se desorption from a (2×1) reconstructed ZnSe surface is determined. This result is in ±0.1 eV agreement with our recently published RHEED intensity transient experiments in the same temperature range.[17] In general, however, RHEED experiments yielded activation energies in the relatively wide 0.5 eV to 1.2 eV range[26–30]

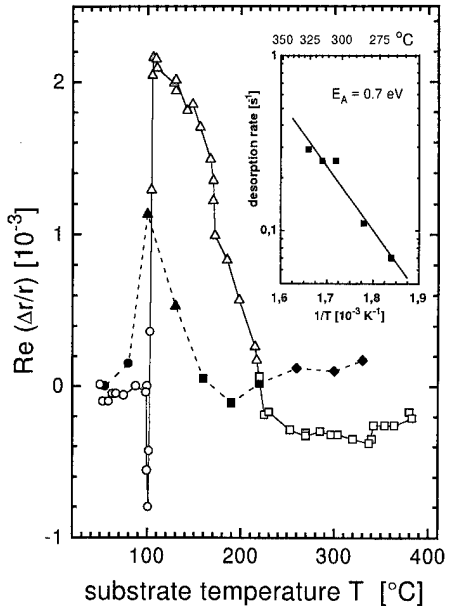


FIG. 2. Se desorption as monitored by RAS at 3.1 eV while gradually heating up the ZnSe sample which is *a*-Se capped (circles) at the beginning of the experiment. The reconstruction as measured by RHEED is indicated by symbols: $c(2\times2)_{Se}$ (triangles), (2×1) (squares) and $c(2\times2)_{Zn}$ (diamonds). The filled symbols indicate $p_{Se}=0$ Torr and the open symbols are used for $p_{Se}=2.8\times10^{-6}$ Torr. Above 250 °C the ZnSe surface can be switched from the Se-rich (2×1) reconstruction to the Zn-rich $c(2\times2)_{Zn}$ reconstruction simply by closing the Se shutter. The insert gives the substrate temperature dependence of the related Se desorption rate that was determined from RAS transients taken at 3.1 eV during this Se desorption process.

depending on substrate temperature, electron energy and obviously other experimental details.

Finally, in preliminary experiments RAS and RHEED have been used to monitor the ZnSe on ZnSe growth. At our standard growth conditions (300 °C substrate temperature, growth rate 500 nm/h) RHEED gave a (2×1) pattern, i.e., the same reconstruction as before growth under Se flux only. Otherwise RAS sensitively monitors the changing surface status due to growth initiation: the RAS signature *A* disappears and a fairly flat RAS spectrum is measured during growth similar to that of the $c(2\times2)_{Zn}$ reconstructed ZnSe(001) surface in Fig. 1(b). We assume that the Se dimers causing the RAS signature *A* of the (2×1) pregrowth reconstruction are mostly broken when growth is started.

IV. SUMMARY

In this work RAS reference spectra have been determined for ZnSe (001) surfaces characterized by RHEED under MBE conditions. The Se-rich $c(2\times2)_{Se}$ surface predicted recently by Garcia and Northrup in Refs. 12 and 13 was verified. A reflectance anisotropy feature at 3.6 eV characteristic for this surface reconstruction was related to the Se trimers which are energetically preferred according to the total-energy calculations in Refs. 12 and 13. Under Se flux and at substrate temperatures above 200 °C we found the well known Se-rich (2×1) reconstruction featuring a probably Se-dimer related RAS minimum at 3.0 eV and a strong RAS maximum close to the ZnSe E_1 bulk critical point. Without Se stabilization and at substrate temperatures above 260 °C the ZnSe surface is Zn-rich $c(2\times2)_{Zn}$ reconstructed. The related RAS spectrum is flat with a minimum close to the E_1 bulk critical point.

ACKNOWLEDGMENTS

J. Rumberg is thanked for optimizing the RAS system. U. Frotscher and J. Weise performed the ellipsometry experiments for reference. The authors thank D. Kisker, T. Yasuda, and Itaru Kamiya for helpful discussions. Part of this work was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 296).

[1]D. E. Aspnes, J. P. Harbison, A. A. Studna, L. T. Florez, and M. K. Kelly, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[2]I. Kamiya, D. E. Aspnes, L. T. Florez, and J. P. Harbison, Phys. Rev. B **46**, 15 894 (1992).
[3]D. E. Aspnes, Mater. Sci. Eng. B **30**, 109 (1995).
[4]W. Richter and J.-T. Zettler, Appl. Surf. Sci. (to be published).
[5]G. Landwehr and D. Hommel, Phys. Status Solidi B **187**, 269 (1995).
[6]M. Haase, I. Qin, I. DePuydt, and H. Cheng, Appl. Phys. Lett. **59**, 1272 (1991).
[7]H. Jeon, J. Ding, W. Patterson, A. Nurmikko, W. Xie, and D. Grillo, Appl. Phys. Lett. **59**, 3619 (1991).
[8]N. Nakayama, S. Itoh, K. Nakano, H. Okuyama, M. Ozawa, A. Ishibashi, M. Ikeda, and Y. Mor, Electron. Lett. **29**, 1488 (1993).
[9]J. Petruzello, S. Herko, J. Gaines, and T. Marshall, Phys. Status Solidi B **187**, 297 (1995).
[10]T. Yao, Z. Zhu, K. Uesugi, S. Kamiyama, and M. Fujimoto, J. Vac. Sci. Technol. A **8**, 997 (1990).
[11]M. Pashley, Phys. Rev. B **40**, 10481 (1989).
[12]A. Garcia and J. Northrup, J. Vac. Sci. Technol. B **12**, 2678 (1994).
[13]A. Garcia and J. Northrup, Appl. Phys. Lett. **65**, 708 (1994).
[14]D. Li and M. Pashley, J. Vac. Sci. Technol. B **12**, 2547 (1994).
[15]H. Farrel, M. Tamargo, T. Gmitter, A. Weaver, and D. E. Aspnes, J. Appl. Phys. **70**, 1033 (1991).
[16]L. Däweritz and K. Ploog, Semicond. Sci. Technol. **9**, 123 (1994).
[17]P. Ruppert, D. Hommel, T. Behr, H. Heinke, A. Waag, and G. Landwehr, J. Cryst. Growth **138**, 48 (1994).
[18]S. M. Scholz, K. Jacobi, J.-T. Zettler, and W. Richter, Surf. Sci. (to be published).
[19]D. E. Aspnes, Y. C. Chang, A. A. Studna, L. T. Florez, H. H. Farrell, and J. P. Harbison, Phys. Rev. Lett. **64**, 192 (1990).
[20]Y. Martaing, Semicond. Sci. Technol. **6**, A60 (1991).
[21]D. K. Biegelsen, R. D. Brigans, J. E. Northrup, and L.-E. Swartz, Phys. Rev. B **41**, 5701 (1990).
[22]T. Hashizume, Q. K. Xue, A. Ichimiya, and T. Sakurai, Phys. Rev. B **51**, 4200 (1995).
[23]D. Kisker, P. Fuoss, S. Brennan, G. Renaud, K. Tokuda, and J. Kahn, J. Cryst. Growth **101**, 43 (1990).
[24]S. Adachi and T. Taguchi, Phys. Rev. B **43**, 9569 (1991).
[25]U. Resch, S. M. Scholz, U. Rossow, A. B. Müller, and W. Richter, Appl. Surf. Sci. **63**, 106 (1993).
[26]K. Menda, I. Takayasu, T. Minato, and M. Kawashima, Jpn. J. Appl. Phys. **26**, L1326 (1987).
[27]T. Yao and T. Takeda, J. Cryst. Growth **81**, 43 (1987).
[28]Z. Zhu, M. Hagino, K. Uesugi, S. Kamiyama, M. Fujimoto, and T. Yao, J. Cryst. Growth **99**, 441 (1990).
[29]N. Matsumura, K. Senga, J. Kakuta, and J. Saraie, J. Cryst. Growth **117**, 129 (1992).
[30]M. Ohishi, H. Saito, H. Torihara, Y. Fujisaki, and K. Ohimori, Jpn. J. Appl. Phys. **30**, 1646 (1991).

Solvent effect on the properties of sulfur passivated GaAs

Vasily N. Bessolov, Elena V. Konenkova, and Mikhail V. Lebedev[a)]
A.F. Ioffe Physico-Technical Institute, Russian Academy of Sciences, Politekhnicheskaya 26, St. Petersburg 194021, Russia

(Received 22 March 1996; accepted 19 April 1996)

Photoluminescence and x-ray photoelectron spectroscopy have been used to study the solvent effect on surface properties of GaAs passivated in different sulfide solutions. It has been found that sulfur passivation could be made more efficient by decreasing the solution dielectric constant through the use of various alcohols as solvents. Specifically, the band edge photoluminescence intensity is much higher, the total amount of oxides is lower, and sulfur coverage is higher compared with a GaAs surface treated in an aqueous sulfide solution. The role of the solvent in sulfur passivation is discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

It is known that the surface treatment of III–V semiconductors in sulfide-containing solutions or gases results in a dramatic decrease of the surface state density in the middle of the band gap. Attendant reduction of the surface recombination velocity allows one to improve the performance and reliability of many devices.[1–5] However, this improved surface electronic quality is not long lived, so the problems of an increase of passivation efficiency and stability remain unsolved.

In order to improve properties of passivating coats, different approaches to sulfur passivation were developed. Among these are passivation from a gas phase,[6–8] electrochemical passivation,[9–11] and photochemical passivation.[12–14] Besides, different sulfide-containing substances such as sodium sulfide,[1] ammonium sulfide,[2] sulfur chloride,[15] thiols,[16,17] and other complex organic compounds[18] were used for passivation.

In recent years, investigations focused on studies of the interaction of sulfur atoms or ions with the surface of III–V semiconductors and, in particular, on chemical processes which lead to changes of surface atomic and electronic structures in the course of sulfidizing. It was shown[19] that to perform a photoelectrochemical reaction of sulfidizing in solutions, the presence of protons is necessary. Therefore, any liquid which contains protons after self-dissociation can be used as a solvent in passivating sulfide solution. Being a protic liquid, water is often used to prepare solutions of inorganic sulfides. Other protic liquids including different alcohols also could be considered as a solvent in passivating solutions. For instance, in electrochemical passivation, the sodium sulfide solution in ethylenglycol was used.[11]

The solvent should have a great effect on both the kinetics and the pathway of any chemical reaction involving ions.[20] We have shown that the sulfide treatment in isopropanol-based solutions results in more effective passivation of GaAs compared with aqueous solutions.[21] Besides, the photoluminescence properties of GaAs treated in isopropanol-based sulfide solutions remain constant even after intensive laser irradiation. Additionally, the sulfur passivation of InGaAs/AlGaAs lasers in isopropanol-based solutions markedly increases their catastrophic optical damage level.[22]

In this article we investigated the role of the solvent in the course of sulfur passivation of III–V semiconductors. Using photoluminescence and x-ray photoelectron spectroscopy we studied the properties of GaAs surfaces treated in solutions of sodium sulfide in different protic solvents.

II. EXPERIMENT

The investigations were performed on n-GaAs(100) ($n=4\times10^{15}$ cm^{-3}) 10 μm thick epilayers grown on semi-insulating substrates by metal-organic chemical-vapor deposition (MOCVD). Immediately before sulfidizing, some specimens were etched in a 1:8:500 solution of H_2SO_4:H_2O_2:H_2O at room temperature for 30 s.

The sulfide treatment was carried out by immersing the samples in different sulfide solutions under illumination with a 200 W incandescent lamp. The solutions used were saturated solutions of sodium sulfide (Na_2S:$9H_2O$) in water or in different alcohols such as ethylenglycol, ethanol, isopropanol, butanol, and tert-butanol. The immersion time was about 1 min as a rule. In aqueous solutions, the samples were treated at different temperatures (from 0 to 100 °C). In alcohol-based solutions, the treatment was carried out at room temperature.

The intensity of band-edge photoluminescence of GaAs at room temperature ($h\nu=1.412$ eV) was measured. Photoluminescence was excited by a Xe laser ($I=3.0$ kW/cm^2).

To investigate the GaAs surface composition x-ray photoelectron spectroscopy was used. Photoelectron spectra were recorded with a Perkin–Elmer PHI 5400 spectrometer using the Mg K_α ($h\nu=1253.6$ eV) x-ray source operated at 300 W. Angles of incidence of x-ray radiation on the crystal surface were 45° and 5°. The samples were put into the vacuum chamber not later than 10 min after completing the sulfidizing process. Data analysis were carried out using Perkin–Elmer software. The identification of the chemical bonds corresponding to different peaks in spectra was accomplished according to Ref. 23.

[a)]Electronic mail: nsab@les.ioffe.rssi.ru

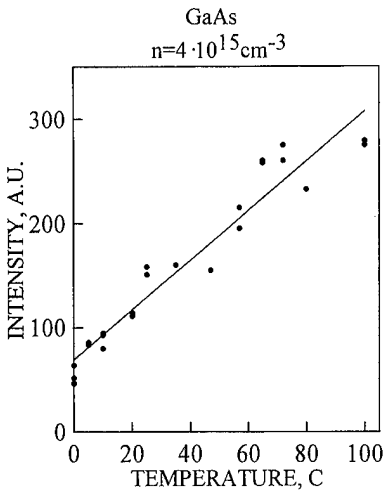


FIG. 1. Dependence on the treatment temperature of the photoluminescence intensity of GaAs(100) ($n=4\times10^{15}$ cm^{-3}) treated in saturated aqueous sodium sulfide solution for 1 min. The intensity of the untreated sample was used as a reference.

III. RESULTS

A. Photoluminescence measurements

Following the sulfide treatment in any of the solutions, the photoluminescence properties of GaAs improved considerably. The band-edge photoluminescence intensity of GaAs treated in aqueous sulfide solutions became higher when the treatment temperature was increased for the same immersion time (Fig. 1).

The photoluminescence intensity of GaAs sulfidized in any alcohol-based solution is always considerably higher then that of GaAs sulfidized in an aqueous solution (Fig. 2).

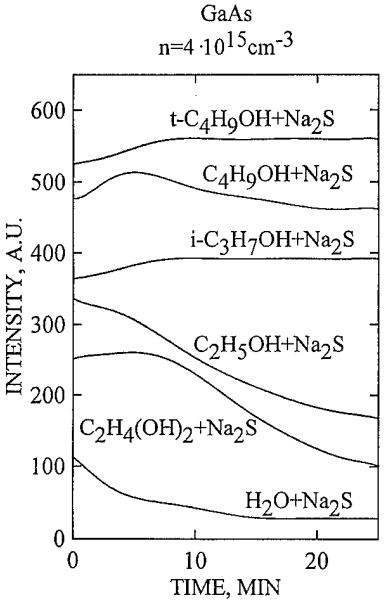


FIG. 2. Variation of the photoluminescence intensity of GaAs(100) ($n=4\times10^{15}$ cm^{-3}) treated in different saturated sulfide solutions. The intensity of the untreated sample was used as a reference.

The greatest increase of the intensity of sulfide treated GaAs relative to that of the untreated one has been achieved after treatment in tert-butanol-based solution. It should be noted that the pretreatment of the semiconductor in H_2SO_4:$8H_2O_2$:$500H_2O$ etch solution before sulfidizing did not lead to any increase of the photoluminescence intensity of solution-treated GaAs.

The degradation of the properties of the sulfide-treated GaAs surface also depended on the solvent used (Fig. 2). The photoluminescence intensity of GaAs, treated in aqueous solutions at any temperature, rapidly (in 15–20 min) dropped down to typical intensity values of untreated GaAs. The photoluminescence intensity of the GaAs surface treated in ethylenglycol and ethanol-based solutions also decreased with time of the laser irradiation but to a lesser extent. On the other hand, the Xe-laser irradiation did not affect the photoluminescence properties of GaAs treated in isopropanol, butanol, or tert-butanol-based solutions.

B. X-ray photoelectron spectroscopy

The survey photoelectron spectra contained Ga, As, C, and O peaks regardless of the angle of incidence. After sulfidizing, the S peaks appeared in spectra. In structures treated in alcohol-based solutions, the Na peak was also seen in the spectrum. Figures 3 and 4 show high resolution spectra of Ga $3d$ and As $3d$ peaks for untreated structures and structures treated in different sulfide solutions (namely, aqueous, at two temperatures, ethylenglycol-, ethanol-, and isopropanol-based solutions).

Deconvolution of the Ga $3d$ spectra showed the following (Fig. 3): In the spectrum of untreated GaAs two components are clearly seen, namely, a Ga–As bond (binding energy 19.9 eV) and a Ga–O bond (binding energy 21.2 eV) [Fig. 3(a)]. In the spectra of sulfide treated GaAs, the Ga–O bond component has a smaller chemical shift (binding energy 21.0 eV); moreover, its intensity in spectra of GaAs treated in aqueous and ethylenglycol-based solutions was practically the same as in untreated GaAs [Figs. 3(a)–3(d)]. In the spectra of GaAs treated in ethanol- and isopropanol-based solutions, the intensity of the Ga–O bond component was considerably reduced but simultaneously one more component was clearly seen (binding energy 20.55 eV), presumably associated with the formation of the Ga–S bonds on the surface [Figs. 3(e) and 3(f)].

Deconvolution of the As $3d$ spectra showed the following (Fig. 4): The spectrum of untreated GaAs can be resolved into three components, i.e., an As–Ga bond (binding energy 40.8 eV), elemental As^0 (binding energy 41.55 eV), and an As–O bond (binding energy 43.7 eV) [Fig. 4(a)]. After treatment in any solution an additional peak with binding energy 42.5 eV appeared. This peak could be associated with the formation on the surface of As–S bonds [Figs. 4(b)–4(f)].

According to our analysis of the spectra, the content of As–O decreases after the sulfide treatment and depends on the solution used. Among sulfide treated structures, it was highest in GaAs treated in the aqueous solution at room temperature and lowest in GaAs treated in the isopropanol-

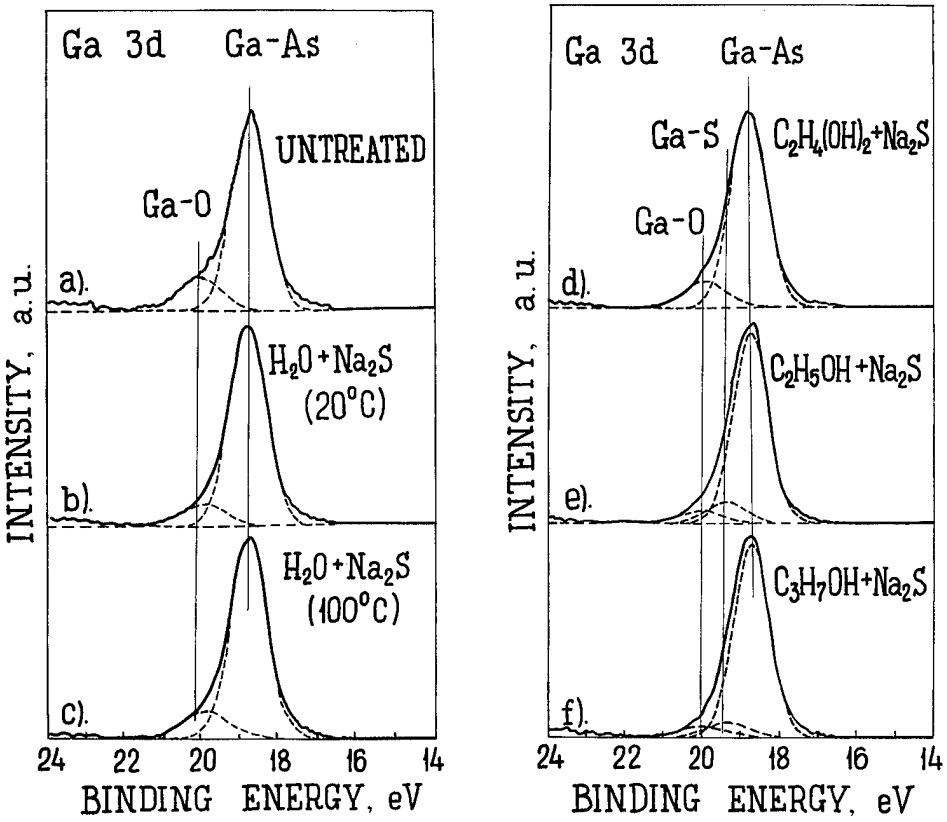


FIG. 3. Ga 3*d* x-ray photoelectron spectra of GaAs(100) preliminary etched in a 1:8:500 solution of H_2SO_4:H_2O_2:H_2O at room temperature for 30 s after various surface treatments taken at an electron take-off angle of 45 deg. (a) Untreated surface, (b) treated in saturated aqueous sodium sulfide solution at room temperature, (c) treated in saturated aqueous sodium sulfide solution at 100 °C, (d) treated in saturated sodium sulfide solution in ethylenglycol, (e) treated in saturated sodium sulfide solution in ethanol, and (f) treated in saturated sodium sulfide solution in isopropanol.

based solution (Table I). The intensity of the peak relating to the As–S bonds also depended on the solution used: it was highest in GaAs treated in the isopropanol-based solution and lowest in the GaAs treated in the aqueous solution at room temperature (Table I).

The elemental As^0 content on the surface slightly decreased after the treatment in the aqueous solutions. On the contrary, treatment in the alcohol-based solutions did not lead to any reduction of As^0 content (Table I).

It should be stressed that the total amount of Ga and As oxides decreased after the treatment in any solution. This reduction was greatest after the treatment in the isopropanol-based solution.

Spectra obtained at angle of x-ray incidence of 5° showed that elemental As^0 and sulfides were concentrated in near-surface 10 Å thick region (their content increased by a factor of 2 compared with the spectra at a take-off angle of 45°).

The estimates carried out according to Ref. 2 showed that the sulfur coverage of the GaAs surface depended on the solvent used (Table I).

IV. DISCUSSION

The experimental data testify that the treatment of the GaAs surface in solutions of sodium sulfide in different alcohols allows one to obtain more effective passivation than the usual treatment in aqueous sulfide solutions. Indeed, the passivation in alcohol-based solutions leads to the greatest enhancement of the photoluminescence intensity. On the other hand, the treatment in alcohol-based solutions leads to more effective removal of oxides from the semiconductor surface and to the formation of sulfides there (Figs. 3 and 4 and Table I).

In order to explain these facts, it is necessary to consider the role of the solvent in electronic and chemical processes at the interface between the semiconductor and the solution in the course of sulfidizing. First, let us consider the role of the solvent in the process of modification of the surface states spectrum.

The enhancement of the GaAs photoluminescence in the course of sulfur passivation can be attributed to the decrease of the surface recombination velocity, i.e., to the lowering of the activity of the surface states. From the chemical point of view such states represent so-called soft acid Lewis centers,[16] since they can easily capture electrons from the conduction band of a semiconductor. To passivate these surface states, it is necessary at least to weaken their ability to capture electrons of the semiconductor. It could be achieved, for instance, by covalent bonding of the acid Lewis centers of the surface with appropriate bases. According to the principle of hard and soft acids and bases,[24] the soft acid can

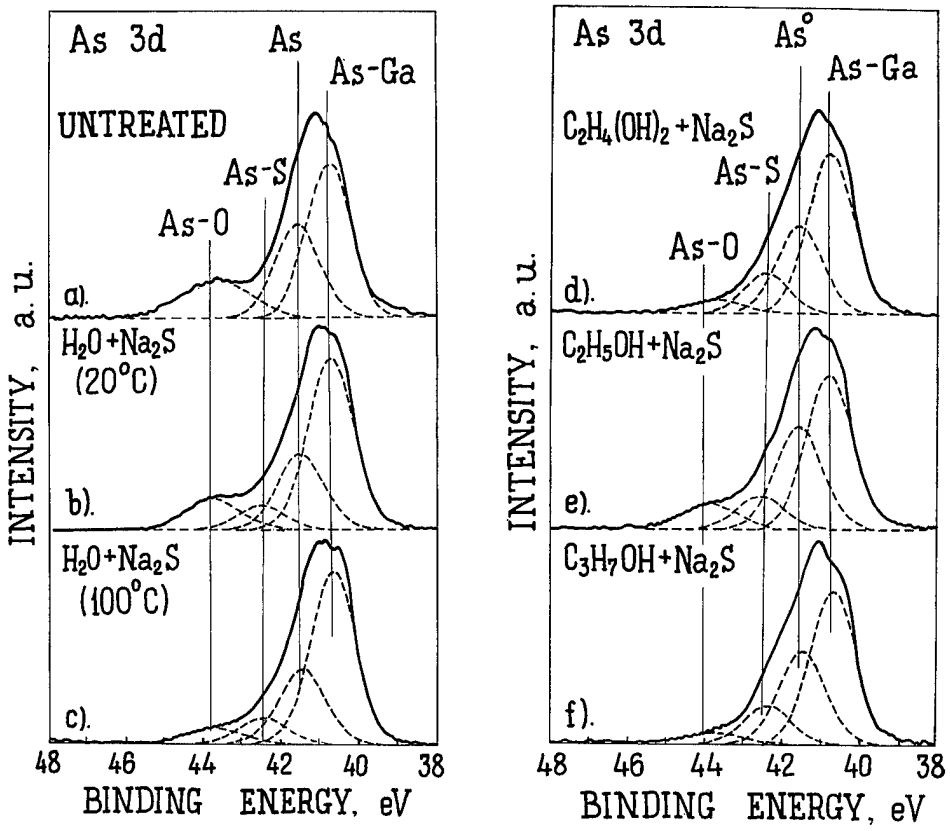


FIG. 4. As 3*d* x-ray photoelectron spectra of GaAs(100) preliminarily etched in a 1:8:500 solution of H_2SO_4:H_2O_2:H_2O at room temperature for 30 s after various surface treatments taken at an electron take-off angle of 45 deg. (a) Untreated surface, (b) treated in a saturated aqueous sodium sulfide solution at room temperature, (c) treated in saturated aqueous sodium sulfide solution at 100 °C, (d) treated in saturated sodium sulfide solution in ethylenglycol, (e) treated in saturated sodium sulfide solution in ethanol, and (f) treated in saturated sodium sulfide solution in isopropanol.

form the covalent bond with a soft base. The sulfide ions are exactly such soft bases[16] and therefore quite suitable for III–V surface passivation. When a covalent bond is formed, the hardness of the cloud of electrons shared by the sulfide-ion and the surface state, according to Ref. 25, will be higher than the hardness of the separate entities, resulting in a less active recombination center.

If the covalent bond formation occurs in a liquid polar medium, a solvate shell consisting of oriented dipole molecules of the solvent would be formed around the ion and near semiconductor surface. The solvate shell should greatly affect both the properties of the electron cloud of the sulfur ion and the electronic distribution in near-surface region of the semiconductor and the parameters of the resulting covalent bonds. Indeed, it has been shown by computations[26] that the hardness of the covalent bond formed in a liquid medium depends on the solvation strength in such a way that the bond hardness increases with the decrease of the solvent dielectric constant. That is, more hard surface states are formed after passivation in solutions with a lower dielectric constant value (i.e., in alcohols) than is the case with the usual passivation from aqueous sulfide solutions. Therefore, the efficiency of nonradiative surface recombination of GaAs treated in alcohol-based solutions with low dielectric constant value should be lower.

Indeed, with the decrease of the dielectric constant value of the solvent being used, the photoluminescence intensity of the sulfidized GaAs, which can be considered as a measure

TABLE I. Relative intensities of different peaks and sulfur coverage for GaAs treated in different solutions.

Treatment	S 2*s* / Ga 3*d*+As 3*d*	As–O / Ga 3*d*+As 3*d*	As–S / Ga 3*d*+As 3*d*	As^0 / Ga 3*d*+As 3*d*	Ga–O / Ga 3*d*+As 3*d*	Ga–S / Ga 3*d*+As 3*d*	Sulfur coverage (MLs)
Untreated	0	0.12	0	0.19	0.07	0	0
Na_2S+H_2O 20 °C	0.04	0.07	0.04	0.14	0.05	0	0.4
Na_2S+H_2O 100 °C	0.05	0.03	0.05	0.13	0.07	0	0.7
$Na_2S+C_2H_4(OH)_2$ 20 °C	0.07	0.03	0.07	0.16	0.08	0	1.1
$Na_2S+C_2H_5OH$ 20 °C	0.10	0.05	0.06	0.19	0.02	0.04	0.8
$Na_2S+C_3H_7OH$ 20 °C	0.12	0.02	0.09	0.19	0.02	0.03	1.3

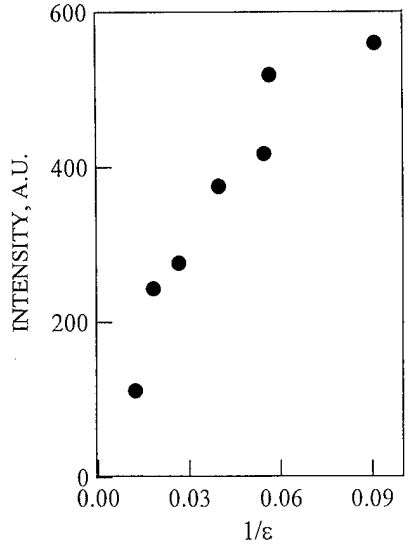


FIG. 5. Dependence of the photoluminescence intensity of sulfide treated GaAs(100) ($n=4\times10^{15}$ cm^{-3}) on the reciprocal dielectric constant of the solvent used. The intensity of the untreated sample was used as a reference.

of the passivation efficiency, increases (Fig. 5).

Consider now the role of the solvent in chemical reactions in the course of sulfidizing of the semiconductor surface. The photoelectrochemical reaction of the sulfur coat formation on a III–V surface during sulfidizing in solution[19] proceeds with participation of different ions (protons, sulfide ions, etc.). In the course of this reaction, the electrostatic interaction of ion reagents with the semiconductor surface should play an important role. Indeed, on a semiconductor surface there is a constant dipole moment due to band bending and adsorbed foreign atoms. Moreover, because of the partly ionic character of semiconductor chemical bonds, Ga and As atoms in a crystal lattice are in fact ions and give a contribution to this electrostatic interaction.[27]

In particular, the choice of the solvent affects the dielectric constant of the medium in which photoelectrochemical reaction of sulfidizing takes place. The replacement of the aqueous solution with an alcohol-based one leads to a significant decrease of the dielectric constant and hence to the increase of the electrostatic interaction. On the one hand, this should lead to more effective association of the sulfide ions with the surface atoms of the semiconductor[20] and, on the other, to the increase of the charge transfer rate between the semiconductor and the passivating solution.[28] Hence the decrease of the solvent dielectric constant should lead to the increase of the equilibrium constant of the photoelectrochemical reaction of sulfidizing. It should be noted that the increase of the temperature of the aqueous solution also leads to the decrease of its dielectric constant.

Based on the x-ray photoelectron spectroscopy data, the equilibrium constant of this reaction was estimated as the ratio of the relative intensity of the S 2s peak (Table I, column 1) to the concentration of sulfur ions in the solution. It is seen that with the decrease of the solvent dielectric constant, the equilibrium constant of the reaction increases (Fig. 6). Moreover, the equilibrium constant is an exponential function of the reciprocal dielectric constant of the solvent, and exactly such dependence should be observed in any reaction of ion association in a solution.[20]

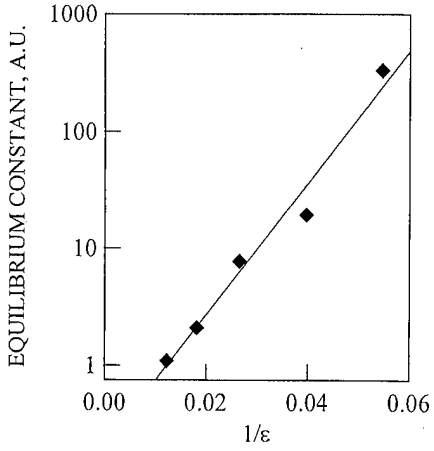


FIG. 6. The estimated value of the equilibrium constant of the reaction of sulfur coat formation on GaAs surface treated in different sulfide solutions vs the reciprocal dielectric constant of the solvent used.

On the other hand, the decrease of the solvent dielectric constant leads to the reduction of the total amount of oxides on the sulfur-treated surface of GaAs (Fig. 7). This fact suggests that the removal of oxides from the surface in the solution occurs as a result of dissociation. During the dissociation, complex oxide ions from the surface get into the solution and on the surface some charge is formed. This charge enforces the electrostatic interaction of sulfide ions with the surface atoms of the semiconductor.

Consequently, the decrease of the dielectric constant value of sulfide solution by appropriate choice of the solvent leads to a more effective passivation of the surface as evidenced by

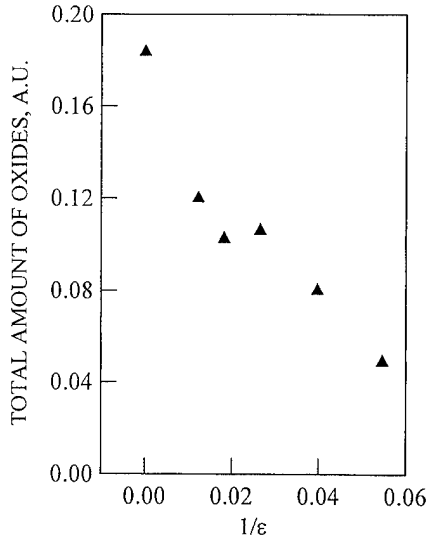


FIG. 7. The dependence of the total amount of oxides remaining after GaAs surface passivation in solutions on the reciprocal dielectric constant of the solvent used. The $1/\epsilon=0$ correspond to the untreated sample.

the photoluminescence enhancement of the semiconductor after sulfur treatment and by the increase of efficiency of both oxide removal and sulfide formation.

Thus the sulfur passivation of the GaAs surface in alcohol-based solutions with low dielectric constants is more effective than traditionally used passivation from aqueous sulfide solutions.

V. CONCLUSION

Properties of GaAs treated in solutions of sodium sulfide in different protic solvents were studied by photoluminescence and x-ray photoelectron spectroscopy. It has been observed that the decrease of the dielectric constant value of the solution through the use of alcohols as solvents allows one to obtain a more effective surface passivation. It has been found that with the decrease of the solution dielectric constant, the band-edge photoluminescence intensity of the semiconductor is enhanced, the total amount of oxides on the surface is reduced, and the amount of sulfides is increased as well. These results testify that the sulfur coverage formation during solution passivation occurs by association of sulfide ions with surface atoms of the semiconductor. The formation of covalent bonds between sulfur and surface atoms in solutions having a low dielectric constant leads to the formation of harder surface states, which are less active in nonradiative recombination.

ACKNOWLEDGMENTS

The authors would like to thank Professor B. V. Tsarenkov for helpful discussions. They are deeply indebted to A. F. Ivankov and V. S. Strykanov for their help with the measurements. The research described in this article was made possible in part by Grant No. R1V300 from the International Science Foundation and by Grant No. 95-03-09330 from the Russian Foundation for Basic Research.

[1]C. J. Sandroff, R. N. Nottenburg, J.-C. Bischoff, and R. Bhat, Appl. Phys. Lett. **51**, 33 (1987).
[2]M. S. Carpenter, M. R. Melloch, B. A. Cowans, Z. Dardas, and W. N. Delgass, J. Vac. Sci. Technol. B **7**, 845 (1989).
[3]S. Shikata, H. Okada, and H. Hayashi, J. Appl. Phys. **69**, 2717 (1991).
[4]S. Kamiyama, Y. Mori, Y. Takahashi, and K. Ohnaka, Appl. Phys. Lett. **58**, 2595 (1991).
[5]W. S. Hobson, U. Mohideen, S. J. Pearton, R. E. Slusher, and F. Ren, Electron. Lett. **29**, 2199 (1993).
[6]T. Tiedje, K. M. Kolbow, D. Rogers, Z. Fu, and W. Eberhardt, J. Vac. Sci. Technol. B **7**, 837 (1989).
[7]G. J. Hughes, L. Roberts, M. O. Henry, K. McGuigan, G. M. O'Connor, F. G. Anderson, G. P. Morgan, and T. Glynn, Mater. Sci. Eng. B **9**, 37 (1991).
[8]Y. Fukuda, M. Shimomura, N. Sanada, and M. Nagoshi, J. Appl. Phys. **76**, 3632 (1994).
[9]S. Weiguo, Appl. Phys. A **52**, 75 (1991).
[10]X. Y. Hou, W. Z. Cai, Z. Q. He, P. H. Hao, Z. S. Li, X. M. Ding, and X. Wang, Appl. Phys. Lett. **60**, 2252 (1992).
[11]J. Yota and V. A. Burrows, J. Vac. Sci. Technol. A **11**, 1083 (1993).
[12]A. S. Weling, K. K. Kamath, and P. R. Vaya, Thin Solid Films **215**, 179 (1992).
[13]Y. Tao, A. Yelon, E. Sacher, Z. H. Lu, and M. J. Graham, Appl. Phys. Lett. **60**, 2669 (1992).
[14]J. L. Leclercq, E. Bergignat, and G. Hollinger, Semicond. Sci. Technol. **10**, 95 (1995).
[15]Z. S. Li, W. Z. Cai, R. Z. Su, G. S. Dong, D. M. Huang, X. M. Ding, X. Y. Hou, and X. Wang, Appl. Phys. Lett. **64**, 3425 (1994).
[16]S. R. Lunt, P. G. Santangelo, and N. S. Lewis, J. Vac. Sci. Technol. B **9**, 2333 (1991).
[17]J. F. Dorsten, J. E. Maslar, and P. W. Bohn, Appl. Phys. Lett. **66**, 1755 (1995).
[18]K. Asai, T. Miyashita, K. Ishigure, and S. Fukatsu, Surf. Sci. **306**, 37 (1994).
[19]V. N. Bessolov, M. V. Lebedev, E. B. Novikov, and B. V. Tsarenkov, J. Vac. Sci. Technol. B **11**, 10 (1993).
[20]E. A. Moelwyn-Hughes, in *The Chemical Statics and Kinetics of Solutions* (Academic, London, 1971).
[21]V. N. Bessolov, A. F. Ivankov, E. V. Konenkova, and M. V. Lebedev, Tech. Phys. Lett. **21**, 20 (1995); Mater. Res. Soc. Symp. Proc. **378**, 1019 (1995).
[22]V. N. Bessolov, M. V. Lebedev, Y. M. Shernyakov, B. V. Tsarenkov, J. S. Yu, J. H. Lee, G. T. Pak, and T. I. Kim, Tech. Phys. Lett. **21**, 561 (1995).
[23]J. F. Moulder, W. F. Stickle, P. E. Sobol, and K. D. Bomben, *Handbook of X-ray Photoelectron Spectroskopy*, edited by J. Chastain (Perkin-Elmer, Eden Prairie, 1992).
[24]R. G. Pearson, J. Am. Chem. Soc. **85**, 3533 (1963).
[25]R. G. Pearson, J. Chem. Educ. **64**, 561 (1987); R. G. Parr and P. K. Chattaraj, J. Am. Chem. Soc. **113**, 1854 (1991).
[26]J. Cioslowski and M. Martinov, J. Chem. Phys. **103**, 4967 (1995).
[27]V. N. Bessolov, A. F. Ivankov, and M. V. Lebedev, J. Vac. Sci. Technol. B **13**, 1018 (1995).
[28]H. Gerischer, Surf. Sci. **18**, 97 (1969).

RAPID COMMUNICATIONS

This section is reserved for short submissions which contain important new results and are intended for accelerated publication.

Silicon dioxide chemical vapor deposition using silane and hydrogen peroxide

Frédéric Gaillard
France Telecom CNET Grenoble, BP 98, 38243 Meylan, Cedex, France

Pascal Brault[a)]
Groupe de Recherches sur l'Energétique des Milieux Ionisés, URA 831 CNRS, Université d'Orléans BP 6759, rue de Chartres, F-45067 Orleans, Cedex, France

Pierre Brouquet
France Telecom CNET Grenoble, BP 98, 38243 Meylan, Cedex, France

(Received 10 April 1996; accepted 24 May 1996)

SiH_4 and H_2O_2 have been successfully used for the deposition of silicon dioxide for shallow trench isolation. With this chemistry, it is possible to fill up trenches without voids up to an aspect ratio 2.3:1. The very good gap filling is due to the presence of SiOH groups in the oxide film. To obtain a density close to that of thermal oxide, the film must be annealed at high temperature. The electrical characteristics are equivalent to those obtained using thermal silicon dioxide. © *1996 American Vacuum Society.*

For shallow trench isolation (STI) by silicon dioxide (SiO_2), it is important to obtain a homogeneous gap filling (no voids) with a density and stoichiometry close to that of thermal oxide for both physical (etching...) and electrical properties (interface trap and charge densities).[1]

The conventional SiO_2 deposition process uses low-pressure chemical vapor deposition (LPCVD),[2] or plasma-enhanced CVD (PECVD).[2,3] The main problem arising in such methods is the appearance of voids due to nonconformal step coverage.[2,4] New ways are currently being investigated to minimize this spurious effect resulting in homogeneous oxide layers. For example, subatmospheric CVD (SACVD) operating with a tetraethylorthosilicate (TEOS) and O_3 mixture[5,6] provide a very good gap filling, but this technique depends critically on the state of the surface.[5] Indeed, for vertical trenches, it is not possible to use plasma treatment to promote the growth. A silicon rich buffer layer is necessary to avoid this problem.

In the present article, we propose a new solution for depositing SiO_2 to fill up vertical shallow trenches. To obtain homogeneous SiO_2 films, the film should be liquid at some step in the deposition processes for a complete filling of the trench. Brinker and Scherer[7] reported that $Si(OH)_4$ is a liquid obtained as a byproduct in the reaction of TEOS with H_2O. In this case, the main reaction step is controlled by the OH production. Our intention was to obtain $Si(OH)_4$ from the gas phase in a CVD process. We demonstrate below, that a mixture of $SiH_4+H_2O_2$ is the best way to obtain a growth of SiO_2 and fill up the trenches. This layer must be annealed, however, in order to recover a quality equivalent to that of thermal oxide.

Patterned 8 in. *p*-type silicon wafers with a resistivity of 8–15 Ω cm were used to study the gap filling. For electrical and physical characterization, the wafers were unpatterned. The experimental setup is described in detail elsewhere.[8,9] A cross section of the CVD deposition chamber (Electrotech Planar 200) is presented in Fig. 1. The SiH_4 and H_2O_2 gases are introduced through a duplex shower head arrangement which avoids any gas cross contamination until they enter the reaction chamber. The H_2O_2 liquid is controlled by weight measurement and is flash-evaporated before entering the CVD chamber. Hydrogen peroxide is usually available in industrial 30% concentrated water solution.

All process parameters were optimized in order to obtain a thickness uniformity better than 2% (1σ). The substrate temperature was varied from 0 to 40 °C in different experiments. The total pressure was controlled via a throttle at 1 Torr. The SiH_4 flow was 40 sccm and the hydrogen peroxide solution flow was 1 g/min. Annealing after deposition was carried out in one step: a conventional furnace annealing at 1050 °C for 30 min under N_2 atmosphere. The deposited films were characterized by Fourier transform infrared (FTIR) spectroscopy on bare wafers. The gap filling properties were checked by scanning electron microscopy (SEM) observation on STI test structures. For the $C(V)$ electrical characteristics, metal oxide semiconductors (MOS) structures were made according to the following sequence: silicon substrate, thermal oxide (250 Å), CVD oxide (1000 Å), and aluminum contact.

A FTIR spectrum of the oxide layer (4000 Å) deposited at

[a)]Corresponding author; Electronic mail: brault@univ-orleans.fr

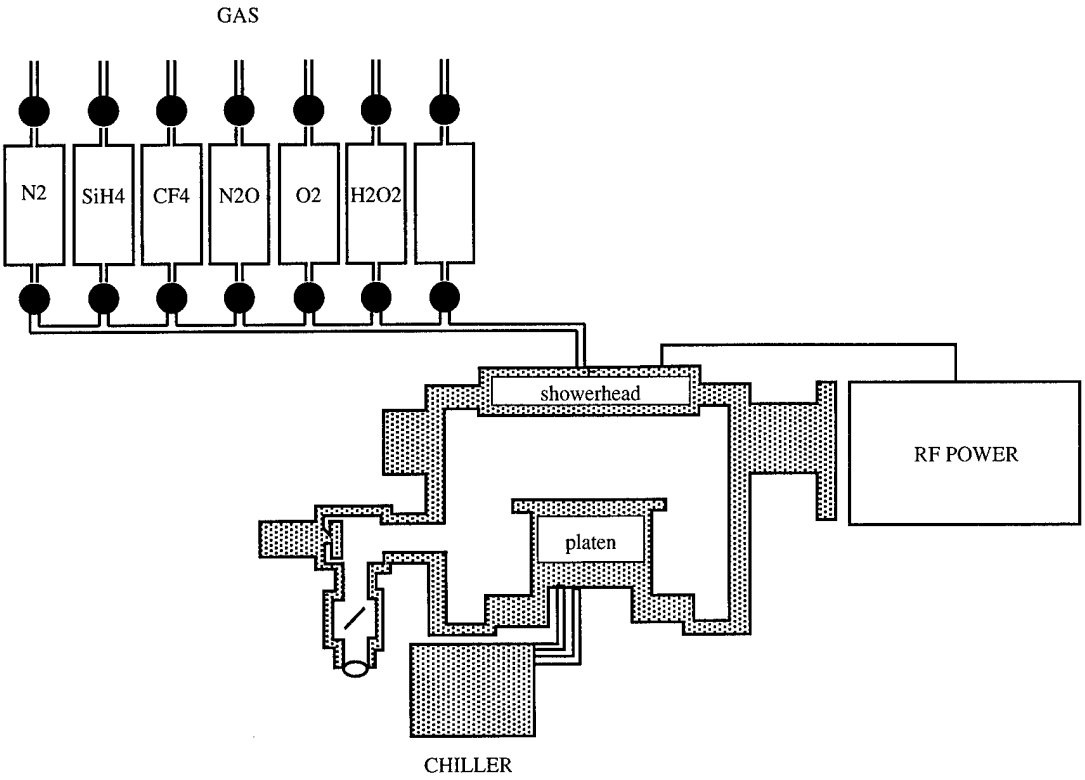


FIG. 1. Cross section of CVD chamber.

0 °C substrate temperature is shown in Fig. 2(a). Three main peaks are present: the SiO stretching is located at 1075 cm^{-1} which compares well with the thermal oxide (1077 cm^{-1}). In addition, we can observe a peak at 3380 cm^{-1}. This 3380 cm^{-1} peak often results from OH vibration in water.[2,4,10,11]

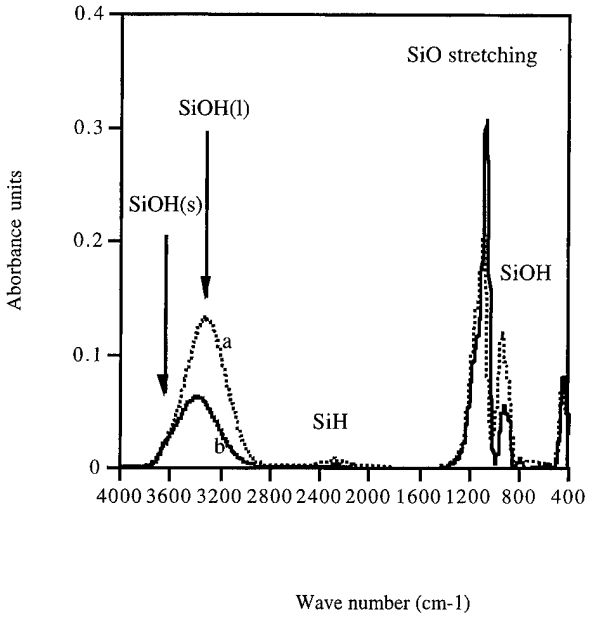


FIG. 2. CVD oxide IR spectra: (a) substrate temperature equal to 0 °C; (b) substrate temperature equal to 40 °C.

However, if it originates from OH, we should be able to observe a companion peak at 1640 cm^{-1},[2,10] which is not the case here. We therefore propose the following explanation for this peak:[10] in solid film, the Si–OH vibration [from silanol ($Si(OH)_4$)] occurs at 3680 cm^{-1}; but for a liquid phase, it is shifted to about 3380 cm^{-1}. Moreover, the peak at 930 cm^{-1} in the PECVD or LPCVD oxides corresponds to the vibration between a hydrogen atom and an oxygen atom itself linked to a silicon atom.[10] This explanation confirms the presence of SiOH groups in the SiO_2 film.

If the substrate temperature is increased to 40 °C during deposition, the 3380 and 930 cm^{-1} peaks [Fig. 2(b)] shift to lower values which is consistent with the evolution of a "liquid" vibration versus temperature: the film becomes "dry" when the temperature increases. We can conclude that the film deposited in the temperature range of the experiments is in a liquidlike state just after deposition because of the presence of $Si(OH)_4$ groups.

To correlate this FTIR analysis with the gap filling properties, we performed SEM measurements on STI structures. Figure 3 presents SEM micrographs showing the influence of a liquid SiOH content on the gap filling. Clearly with substrate temperature during deposition at 0 °C, the gap filling is very good for aspect ratio up to 2.3:1 [the trenches are 0.3 μm wide and 0.7 μm deep (Fig. 4)], while at 40 °C the gap filling becomes poorer.

After annealing, the film has the same Si/O concentration ratio as stoichiometric silicon dioxide (for as deposited films, Si/O was already 0.5). This is confirmed by the fact that the

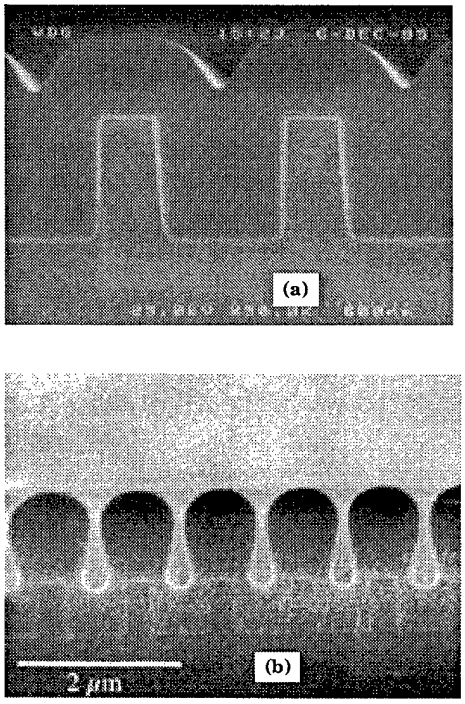


FIG. 3. SEM view for substrate temperature (a) at 0 °C; (b) at 40 °C.

film becomes more stable because there are no SiOH groups in the liquid state in the SiO_2 film. Secondly, the density of the film measured by mass measurement varies from 1.65 g/cm^3 as deposited to 2.1 g/cm^3 after annealing, which compares favorably with 2.20 g/cm^3 for thermal oxide. Finally, the electrical properties after annealing at 1050 °C for 30 min under N_2 are closer to those of a thermal oxide. Conventional high-frequency capacitor tests and a quasi-static test have been used for fixed charge and interface trap density measurements.[12,13] The fixed electrical charge density is about 5×10^9 cm^{-2} and the interface trap density is about 3×10^9 cm^{-2} eV^{-1} (for as deposited films, these values were, respectively, 2×10^{11} cm^{-2} and 10^{11} cm^{-2} eV^{-1}). These values are very low. This SiO_2 filling technique can be used for the STI process.

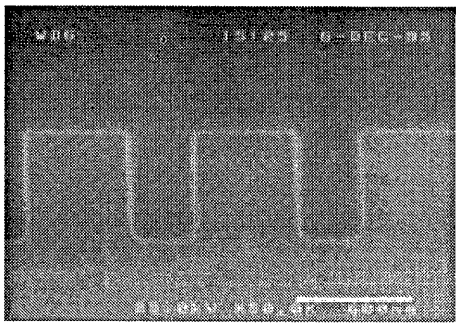


FIG. 4. SEM view with trenches 0.3 μm wide, 0.7 μm deep.

In conclusion, using this low-energy CVD process after moderate temperature annealing, we can thus obtain all the parameters close to those of a thermal oxide: The density and the electrical characteristics are similar. Contrary to the standard CVD process, the oxide film is homogeneous in the trenches (no voids) and no cracks. The silanol plays the key role for gap filling. To obtain a higher aspect ratio, we expect that the SiOH group number in the oxide layer must be increased. Moreover, annealing leads to a thermal oxidelike film.

Acknowledgments: This work has been carried with the framework of the GRESSI consortium between France Telecom-CNET and the Commissariat à l'Energie Atomique-LETI. The authors would like to thank G. Bomchil for helpful discussions and M. Timmins for reading the manuscript.

[1]P. Sallagoïty, M. Ada-Hanifi, M. Paoli, and M. Haond, ESSDERC Conference, 1995 (unpublished), p. 371.
[2]S. Rojas, L. Zanotti, A. Borghesi, A. Saselle, and G. U. Pignatel, J. Vac. Sci. Technol. B **11**, 2081 (1993).
[3]H. M. Dauplaise, K. Vaccaro, B. R. Bennett, and J. P. Lorenzo, J. Electrochem. Soc. **139**, 1684 (1992).
[4]A. C. Adams, Solid State Technol. 135 (1983).
[5]K. Kwok, E. Yieh, S. Robles, and B. C. Nguyen, J. Electrochem. Soc. **141**, 2172 (1994).
[6]A. Perera, J. Lin, Y. Ku, M. Azrak, B. Taylor, J. Hayden, M. Thompson, and M. Blackwell, IEDM Conference, 1995 (unpublished).
[7]C. J. Brinker and G. W. Scherer, *The Physics and Chemistry of Sol-Gel Processing, Sol-Gel Science* (Academic, New York, 1990).
[8]C. D. Dobson, A. Kiermasz, K. Beekman, and R. J. Wilby, Semicond. Int. 86 (1994).
[9]M. Matsuura, Y. Hayashide, H. Kotani, T. Nishumura, H. Iuchi, C. D. Dobson, A. Kiermasz, K. Beeckmann, and R. Wilby, IEDM Conference, 1994 (unpublished), p. 117.
[10]L. J. Bellamy, *The Infrared Spectra of Complex Molecules*, 3rd ed. (Chapman and Hall, London, 1986).
[11]H. M. Dauplaise, K. Vaccaro, B. R. Bennett, and J. P. Lorenze, J. Electrochem. Soc. **139**, 1684 (1992).
[12]E. H. Nicollian and J. R. Brews, *MOS Physics and Technology* (Wiley, New York, 1975).
[13]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1968).

SHOP NOTE

These are "how to do it" papers. They should be written and illustrated so that the reader may easily follow whatever instruction or advice is being given.

Analytical method of gigabit trench doping uniformity by secondary ion mass spectrometry*

Naoto Matsuo
Department of Electrical and Electronic Engineering, Yamaguchi University, Tokiwadai, Ube 755, Japan

Kazuyoshi Tsukamoto
Matsushita Technoresearch, Inc. Co., Ltd., Yakumo-nakamachi, Moriguchi, Osaka 570, Japan

Tadaki Miyoshi
Department of Electrical and Electronic Engineering, Yamaguchi University, Tokiwadai, Ube 755, Japan

(Received 10 October 1995; accepted 17 May 1996)

To increase the integration of dynamic random access memory (DRAM), the area of the memory cell must be decreased. For 1 or 4 Gbit DRAM, it becomes 0.21 μm^2 or 0.05 μm^2, respectively. Despite the shrinkage of the cell area, the capacitance of the memory cell must be kept 20–25 fF. One method of achieving these requirements is to form a trench capacitor that has a high aspect ratio. Some types of trench capacitors with high aspect ratios have been developed.[1] However, there are two difficulties to fabricate the trench capacitors: one is the doping method to the trench side wall, and the other is the quantitative measuring method of the uniformity for the trench doping. For the latter problem, we proposed the new measuring method of the uniformity for the trench side-wall doping corresponding to 64 Mbit integration by using the secondary ion mass spectrometry (SIMS).[2] The purpose of this article is to show that the measuring method has high resolution because of the polycrystalline Si (poly-Si) that fills up the trench, and therefore it is effective to the analysis of the uniformity of the As concentration for the trench side wall corresponding to 1–4 Gbit and beyond integrations. First, the uniformity of the trench sputtering by filling up the trench with a poly-Si is examined. Second, whether the method detects precisely the As concentrations at the random two points of the trench side wall which ratio for the concentration is approximately two is discussed. Figure 1 shows the schematic cross section of the specimen prepared for analysis by SIMS. A trench array is formed in a p type Si wafer by the reactive ion etching, and the size of one trench is 0.5–0.6 μm in width and 2–4 μm in depth. The As planar-type solid-diffusion method[2,3] is used for the trench side-wall doping of the first experiment. The tilt-angle ion implantation method[4] is used for the trench side-wall doping of the second experiment: Double implantation of As ions is carried out under both conditions of 5×10^{15} cm^{-2}, 150 keV, 7° off the normal incidence to the surface of the sample and of 2×10^{15} cm^{-2}, 150 keV, 20° off angle as shown by the inset of Fig. 3. For the first and the second experiments, the poly-Si is deposited on the wafer by the low-pressure chemical vapor deposition (LPCVD) method without forming SiO_2 on the surface of the trench. The reason of this is described in Ref. 2. The wafer surface is leveled by sputtering the poly-Si with oxygen radicals. To examine the effect of the poly-Si which fills up the trench, the specimen without filling up the poly-Si in it is also prepared for the experiment. By sputtering the specimen with or without the poly-Si in the trench from the wafer surface to the level of the trench bottom by Cs^+, the uniformity of the As concentration along the trench side wall is measured. The detected ions are either As^- or $AsSi^-$. The relationship between the resolution of the measuring method and the sampling time of these ions by the detector of the SIMS is also examined. Figure 2 shows the profiles of As concentration along the trench depth. Sputtering the As ions by Cs^+ occurs parallel to the substrate surface for the specimen with poly-Si in the trench, and it occurs nonuniformly for the specimen without poly-Si. It is found that the precise measurement is not carried out if the trench is not filled up with the poly-Si. For A, the peak of As concentration is observed at the depth of 3.5 μm. This peak is due to the bottom As diffusion layer with the thickness of 0.1 μm. The sampling time of the As

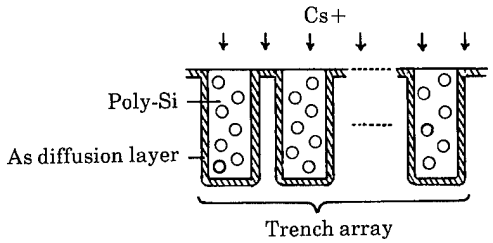


FIG. 1. Schematic cross section of the specimen prepared for analysis by SIMS.

*Published without author corrections.

 0734-211X/96/14(4)/2770/2/$10.00 ©1996 American Vacuum Society

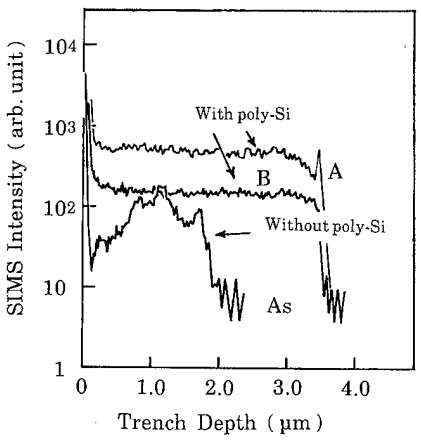


FIG. 2. Profiles of As concentration along the trench depth. As planar-type solid-diffusion method is used for the trench side-wall doping.

ions for A is larger than that for B. It is found that the measuring method detects the thin As diffusion layer by fixing the sampling time to the optimum value.

Next, whether the present method detects precisely the As concentration at the random two points which ratio for the concentrations is approximately two and whether the method is feasible for the analysis of the trench doping with 1–4 Gbit and beyond integrations are discussed. Figure 3 shows the profile of the As concentration along the trench depth by the present SIMS measurement. The reason why the SIMS intensity is weak from the surface to the depth of approximately 0.5 μm is due to the residual poly-Si on the wafer surface that is not removed by the leveling with the oxygen radicals. And the opening of the trench corresponds to the peak of the first spike that is approximately 0.65 μm from the surface. The two spikes of the SIMS profile correspond to the concentrations of the diffusion layers at the wafer surface and at the trench bottom, respectively. The relationship between the ion currents and the As concentration of the trench side wall is discussed. From the relationship between the ion implantation conditions and the geometric shape of the trench, the ratio of the As concentration of the side wall (C_1) from the opening of the trench to the depth of 1.6 μm and that (C_2) from the depth of 1.6 μm to the bottom of the trench, C_1/C_2 becomes 2.1. The ion current (I_1) at the depth of 0.8 μm from the opening of the trench corresponds to C_1 and that (I_2) at the depth of 2.2 μm corresponds to C_2. It is found that the ratio I_1/I_2 of the measured values of the ion currents becomes 2.0, and it is almost the same as the calculated value of C_1/C_2. At the depth of approximately 1.5 μm from the opening, a broad peak is observed. This peak is considered to be due to the reflected As ions from the trench side wall to the opposite one. The ratio of the ion current at the peak and I_1 becomes 1.1. It is found that the measuring method detects precisely the As concentrations at the random two points of the trench side wall which ratio for the concentrations is 1.1 to 2.0. If the total quantity of As ions flying to the detector of the SIMS from the trench side wall in the detection area is more than the detection limit of it, this method would be applied to the analysis of the uniformity for the As concentration of the trench side wall whose dimension is smaller than the present one. Because the trench depth becomes larger and the trench width becomes smaller for the higher integration, there will not be much difference of the total As ions at the trench side wall for the present and higher integrations. Therefore, even if the integration of the device increases up to 1–4 Gbit and beyond, it would be possible for the method to detect As ions as accurately as the measured value shown in Fig. 3.

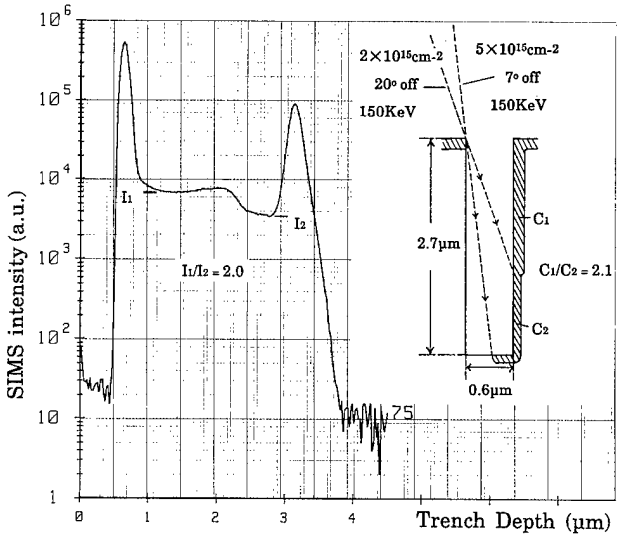


FIG. 3. Profiles of As concentration along the trench depth. Double implantation of As ions to the trench side wall is carried out by the tilt-angle ion implantation method.

In conclusion, it is found that the present measuring method of As ions at the trench side wall using SIMS detects precisely the concentrations at the random two points whose ratio for the concentrations is 1.1 to 2.0. This method is thought to be effective for the analysis of the uniformity of the trench doping for the 1–4 Gbit and beyond.

Acknowledgment: This work was partly supported by The Murata Science Foundation.

[1]See, for example, T. V. Rajeevakumar, T. Lii, Z. A. Weinberg, G. B. Bronner, P. Mcfarland, P. Coane, K. Kwietniak, A. Megdanis, K. J. Stein, and S. Cohen, IEDM Tech. Dig. **835** (1991); T. V. Rajeevakumar and G. B. Bronner, Tech. Dig. VLSI Symp. **7** (1991).

[2]N. Matsuo, S. Okuda, Y. Hirofuji, K. Tsukamoto, H. Iwasaki, T. Takemoto, and Y. Yoshioka, Jpn. J. Appl. Phys. **28**, 1866 (1989).

[3]N. Matsuo, T. Miyoshi, and K. Matsubara, Microelectron. J. **27**, 73 (1996).

[4]G. Fuse, M. Fukumoto, A. Shinohara, S. Odanaka, M. Sasago, and T. Ohzone, IEEE Trans. Electron. Devices **34**, 356 (1987).

ERRATA

Erratum: Measurement of interfacial shear (friction) with an ultrahigh vacuum atomic force microscope [J. Vac. Sci. Technol. B 14, 1289 (1996)]

R. W. Carpick
Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720 and Department of Physics, University of California at Berkeley, Berkeley, California 94720

N. Agraït
Instituto Universitario de Ciencia de Materiales "Nicolás Cabrera," Laboratorio de Bajas Temperaturas, C-III, Universidad Autónoma de Madrid, 28049 Madrid, Spain

D. F. Ogletree and M. Salmeron
Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

The exponent in Eq. (13) in the Appendix is incorrect. The correct equation is

$$F_c=\tau\cdot A_c=\tau\cdot\pi\cdot\left[\left(\frac{6\pi\gamma}{K}\right)\cdot\frac{1}{(n(n+1)c\Psi)^2}\right]^{2/(2n-1)}. \quad (13)$$

It is then clearly seen that Eq. (13) correctly reduces to Eq. (6) for the case $n=2$.

 0734-211X/96/14(4)/2772/1/$10.00 ©1996 American Vacuum Society

Papers from the 23rd Annual Conference on the Physics and Chemistry of Semiconductor Interfaces

21–25 January 1996
Scripps Institute of Oceanography
La Jolla, California

Sponsored by
Office of Naval Research
American Vacuum Society

Editor for the Conference:
D. E. Aspnes

Published for the American Vacuum Society by
the American Institute of Physics, New York, 1996

PREFACE

This volume contains papers presented at the 23rd Annual Conference on the Physics and Chemistry of Semiconductor Interfaces (PCSI-23). The Conference was held at the Sumner Auditorium of the Scripps Institute of Oceanography and the Sea Lodge on La Jolla Shores Beach, from 21 to 25 January 1996. La Jolla featured outstanding winter weather and views of the Pacific, with opportunities to carry out discussions while walking the beach.

The Conference continued to emphasize fundamental understanding of interface problems relevant to emerging areas of technology. The size of the meeting, as measured by the number of attendees and abstracts submitted, increased about 10% relative to 1995. A total of 81 papers were presented in 8 sessions of invited talks intermixed with oral and poster presentations. Silicon-related topics, including SiGe and Si/SiO_2 interfaces, accounted for about 40% of the Conference Program, III–V semiconductor topics represented about 30%, and the rest was distributed among a variety of topics including wide-bandgap semiconductors, metal–semiconductor interfaces, new techniques, etc. Silicon-related topics continued to gain attention, as did wide-bandgap materials. A particularly stimulating session was held on the atomic structure of the Si/SiO_2 interface, where Reed McFeely and Alfredo Pasquarello provided experimental and theoretical, respectively, perspectives that fueled a spirited debate concerning the interpretation of experiments that probe the microscopic structure of the interface. A Tuesday evening rump session focused on atomic structure at buried interfaces, one of the themes to be highlighted in 1997. Also continuing a trend was the increase in theoretical contributions. Decreasing activity was evident in metal–semiconductor interfaces and in III–V epitaxy, although heteroepitaxy continues to receive substantial attention.

The Conference succeeded through the efforts of many individuals. We would like to recognize in particular the many anonymous reviewers who performed their work carefully, thoroughly, and efficiently; ensuring that the high standards of the Journal of Vacuum Science and Technology be maintained; Jack Dow for conference management, Carole Dow for registration management, and Becky York, who provided a smooth interface between the Editor and the Journal of Vacuum Science and Technology. The commitment of Gary McGuire for making this collection a part of JVST is also appreciated, along with that of Galen Fisher, who provided an effective link with the American Vacuum Society. Financial support from Larry Cooper of the Office of Naval Research is gratefully acknowledged. The Conference was held under the sponsorship of the American Vacuum Society through the Electronic Materials and Processing Division, and the Office of Naval Research (this work relates to Department of Navy Grant No. N00014-96-1-0134 issued by the Office of Naval Research. The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein).

D. E. Aspnes
Conference Editor

ORGANIZING COMMITTEE

PROGRAM COMMITTEE

R. M. Feenstra
IBM

M. S. Hybertsen
AT&T Bell Laboratories Research

A. Kahn
Princeton University

G. Lucovsky
North Carolina State University

C. G. Van de Walle
Xerox Palo Alto Research Center

CONSULTANTS

D. K. Ferry
Arizona State University

T. C. McGill
California Institute of Technology

CHAIR

C. G. Van de Walle
Xerox Palo Alto Research Center

EX OFFICIO

D. E. Aspnes (Conference Editor)
North Carolina State University

L. R. Cooper
Office of Naval Research

J. D. Dow
Arizona State University

D. J. Wolford
Iowa State University

ARRANGEMENTS

K. L. Kavanagh
UC San Diego

CONFERENCE SECRETARY

Carole Dow

Effect of dislocations in strained Si/SiGe on electron mobility

K. Ismail[a)]
Department of Communication and Electronics, Faculty of Engineering, Cairo University, Giza, Egypt

(Received 22 January 1996; accepted 24 March 1996)

In this article the presence of threading and misfit dislocations in strained modulation-doped Si/SiGe heterostructures and their effect on electron mobility are investigated. It is found that the low-temperature electron mobility is sensitive to threading dislocations when their density exceeds 3×10^8 cm^{-2}, and decreases by two orders of magnitude when the threading dislocation density is 1×10^{11} cm^{-2}. The room-temperature mobility is reduced under the same conditions by 10% and 50%, respectively. Misfit segments in the graded Ge-content buffer limit the mobility when the Si channel is 0.4 μm or less away from that buffer. Misfit dislocations at the Si-channel bottom interface very strongly scatter the electrons in the channel once a continuous network of misfit segments is created. © *1996 American Vacuum Society.*

I. INTRODUCTION

In the past decade there has been increasing interest in the growth of strained Si/SiGe layers for studying the material properties as well as for potential high-speed device applications.[1,2] Due to the large lattice mismatch between Si and Ge, a relaxed SiGe buffer layer can only be achieved at the expense of creating misfit dislocations (md). A certain fraction of these md results in dislocations threading to the surface of the sample. In the case of growing a fully relaxed $Si_{0.7}Ge_{0.3}$ buffer directly on Si the resulting density of threading dislocations (td) is typically in the range of 1×10^{10} and 1×10^{11} cm^{-2}. The use of a graded Ge-content buffer layer[3] grown on a Si substrate has allowed the growth of strained Si imbedded between relaxed SiGe layers with a low density of threading dislocations (1×10^5–1×10^8 cm^{-2}).

Extremely high electron mobilities in modulation-doped Si/SiGe both at room temperature[4] and at low temperature[5–7] have been reported by various groups. Figure 1 shows the evolution of low-temperature electron mobility since 1985. It is important to stress that there are two major advancements, one in 1990 and the other in 1994, both due to the reduction in dislocation scattering, as will be described in this article.

II. LAYER DESIGN

The layers under study were grown by ultrahigh vacuum chemical vapor deposition (UHV-CVD) in the temperature range of 500–560 °C. A schematic of the layer structure is shown in Fig. 2. First, a step-graded $Si_{1-x}Ge_x$ layer is grown, reaching a 30%–40% Ge composition. This is followed by the growth of a relaxed $Si_{1-x}Ge_x$ buffer layer, where x is in the range of 0.25–0.35. A strained Si channel is then grown to a thickness in the range of 8–15 nm, followed by an undoped 15 nm $Si_{1-x}Ge_x$ spacer layer, an n-type doped $Si_{1-x}Ge_x$ supply layer, and a 4-nm-thick Si cap layer. The Ge fraction and the degree of relaxation of the buffer layer were determined by Auger analysis and high-resolution x-ray diffraction, respectively. The measurement of layer thickness and the identification of defects were achieved by cross-sectional and planar-view transmission electron microscopy (TEM), and the mobility measurements were performed in the range of 0.4–300 K using standard Van der Pauw and Hall bar geometries.

In order to investigate the effect of scattering by the dislocations we have systematically varied the density of threading and misfit dislocations. The former can be controlled by changing the grading rate of the Ge content in the buffer. In the extreme limit of an abrupt SiGe buffer (i.e., infinite grading rate) the density of threading dislocations is about 1×10^{11} cm^{-2}. A slow grading rate of 30% Ge per micron, on the other hand, results in a density of threading dislocations in the Si channel of approximately 1×10^5 cm^{-2}.

III. EFFECT OF THREADING DISLOCATIONS

In the original work of Abstreiter *et al.*[8] the measured electron mobility at 4.2 K was about 2000 cm^2/V s. In that case, a relaxed $Si_{0.75}Ge_{0.25}$ was grown on a Si substrate. From our own experiments, layers of similar design have a density of td of about 1×10^{11} cm^{-2}. At low temperature, where the thermal effects and scattering by phonons is dramatically reduced, the electron mobility becomes very sensitive to residual scattering mechanisms due to background charge impurities, roughness, and dislocations.

A threading dislocation forms a strong barrier to electron transport and, hence, depletes the electrons around it. When the distance between dislocations is significantly longer than the electron mean free path, the electrons can move freely in the regions between dislocations. Once the density increases to the extent that there is an overlap between the depletion regions of adjacent dislocations, the electrons are forced to scatter off the dislocations and the mobility decreases dramatically. If the distance between dislocations approaches the electron Fermi wavelength (about 30–50 nm in our case) localization effects[9] become clear and modify the electron transport properties.

More quantitatively, the depletion region, which is on the order of the Debye length (and, hence, depends on the electron density) is estimated to be in the range of 0.2–0.4 μm.

a)Present address: IBM T.J. Watson Research Center, Yorktown Heights, NY 10598.

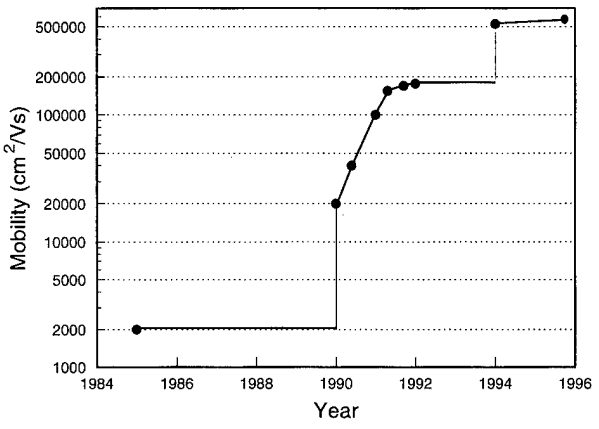


FIG. 1. Evolution of low-temperature electron mobility in modulation-doped Si/SiGe heterostructures.

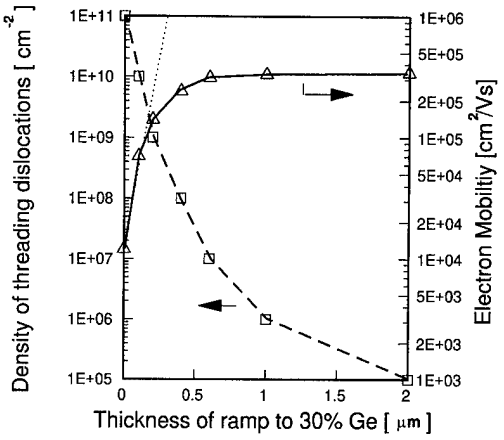


FIG. 3. Effect of grading rate of the Ge content in the buffer on the density of threading dislocations and the electron mobility measured at 0.4 K.

Thus the overlap is expected to take place when the density of threading dislocations is about $0.6–2.3\times10^9$ cm^{-2}.

In order to change the density of dislocations, the grading rate of the buffer was changed (keeping the relaxed SiGe buffer layer thickness at 1 μm). Figure 3 shows that the dislocation density can be reduced in a controlled way from 1×10^{11} cm^{-2} down to 1×10^5 cm^{-2}. The corresponding increase in low-temperature electron mobility is also shown. Extrapolating the mobility at thin graded Ge-content buffers (dotted line) shows that the threading dislocations start limiting the mobility when their density exceeds 1×10^8 cm^{-2}, in agreement with our simple calculation above. When the density of dislocation is below this limit, the mobility is no more sensitive to any further reduction and is limited by other scattering mechanisms.

At room temperature, on the other hand, thermal diffusion and phonon scattering obscure the effect of scattering by td. For a td density in the range of $1\times10^5–1\times10^{11}$ cm^{-2} the channel electron mobility varied from 2700 to 1300 cm^2/V s.

IV. EFFECT OF MISFIT DISLOCATIONS

During the relaxation of the graded Ge-content buffer, misfit dislocations are generated and are responsible for the strain relaxation. Each misfit segment is associated with a strain field, which locally perturbs the band structure around it and, thus, forms a strong scattering potential. Once the graded Ge-content buffer has reached its final Ge content and relaxes to the lattice constant thereof, no additional misfit dislocations are created. However, the strain field decays over a length scale comparable to the thickness of the layer it is trying to relax. Hence, if the electron channel is grown within that range, the electron mobility senses the presence of the dislocations.[9]

In order to systematically study this effect, we have fixed the graded buffer layer thickness and changed the thickness of the following relaxed SiGe buffer, separating the Si channel from the misfit dislocations. Figure 4 shows the dependence of the mobility on this separation. There is a very fast increase in mobility for relaxed buffer layer thickness up to about 0.5 μm, which agrees with the expected range of the

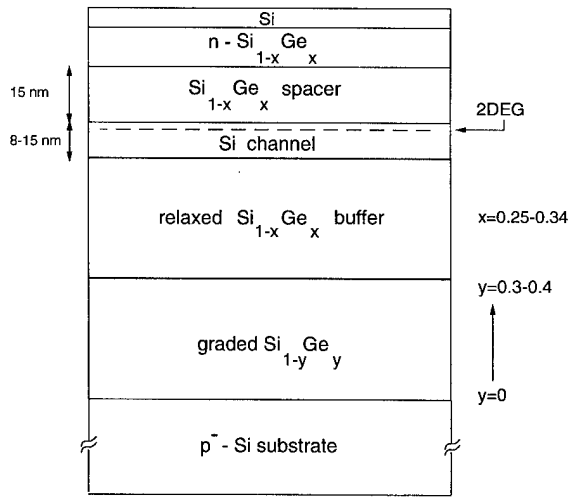


FIG. 2. Schematic of the layer structure grown by UHV-CVD.

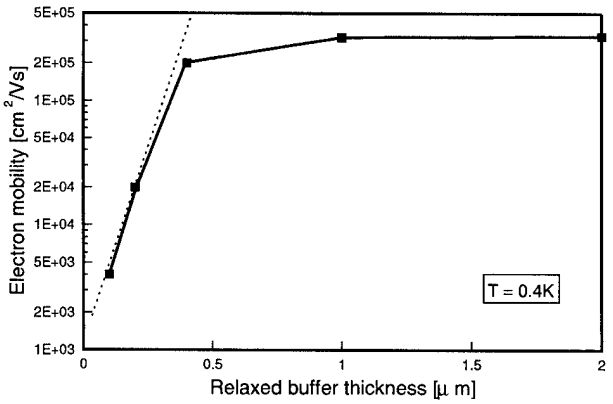


FIG. 4. Electron mobility as a function of the relaxed $Si_{0.7}Ge_{0.3}$ buffer thickness grown on top of a 0.7-μm-thick graded Ge-content buffer.

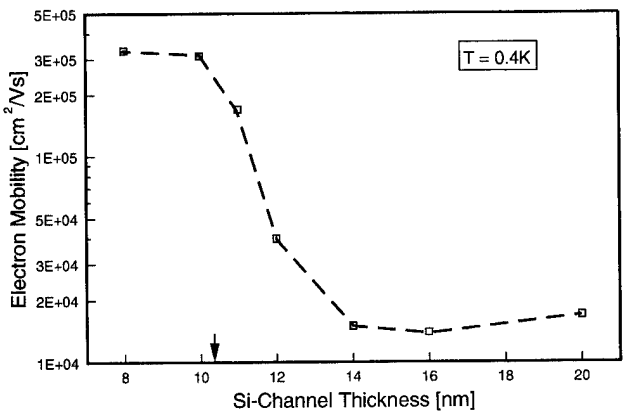


FIG. 5. Electron mobility as a function of Si-channel thickness grown on a relaxed $Si_{0.7}Ge_{0.3}$ buffer. The arrow indicates the Matthews and Blakeslee critical thickness.

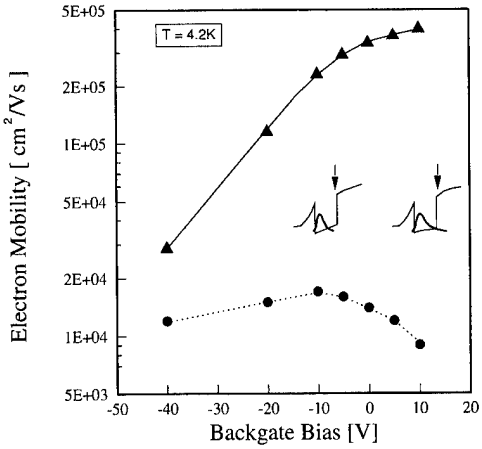


FIG. 6. Effect of backgating on electron mobility in a high mobility sample with no misfit dislocations at the Si-channel interface (▲) and a low mobility sample with 1×10^8 cm^{-2} misfits at the bottom of the Si channel (●). The insert shows the electron wave function at −10 and +10 V on the backgate.

strain field perturbation (the graded buffer is 0.7 μm in this case). Any further increase in buffer layer thickness beyond 1 μm does not seem to change the mobility noticeably. Some reduction in mobility in the case of even thicker buffer may be due to the increase in surface roughness, which is not the subject of this study. At room temperature, the mobility does not change (within the experimental error) for buffer layers thicker than 0.1 μm, indicating that for device applications the thickness of the buffer (and the corresponding growth time) can be safely reduced.

Another source of misfit dislocations, which has been ignored so far, is due to the strain in the active Si channel. Since the buffer layers have threading dislocation densities in the order of 10^5 to 10^6 cm^{-2}, each of these will glide as soon as the Si-channel thickness exceeds the Matthews–Blakeslee critical thickness t_c,[10] resulting in a misfit dislocation at the bottom interface of the Si channel.

Due to the proximity of these dislocations and their attendant strain field to the centroid of the electron gas, it induces a very strong scattering potential. Thus the electron mobility becomes a very sensitive tool to detect the presence of such dislocations. Figure 5 shows the mobility as a function of Si-channel layer thickness at 0.4 K. A dramatic reduction in mobility starts at a channel layer thickness which is in good agreement with t_c (shown by the arrow). It is interesting to note that there are two competing effects which influence the mobility. The first is the increase in dislocation density with the Si-channel thickness. At the same time, the dislocations are separated more from the centroid of the electron gas. This competition may explain the slight increase in mobility at a channel thickness of 20 nm compared to 15 nm.

In order to confirm the above results, two samples corresponding to a Si-channel thickness of 10 and 15 nm were backgated, which allowed varying the charge density in the channel as well as modifying the position of the charge centroid within the quantum well.[11] When a positive backgate voltage is applied, the electron density is increased and the centroid of the electron wave function is moved away from the top towards the bottom interface (Fig. 6). In the sample with no misfit segments at the bottom interface, the electron mobility increases as a function of positive gate bias. In the sample with misfit dislocations, on the other hand, there is a clear drop in mobility at positive gate voltage associated with the reduction in the spacing between the charge centroid and the dislocations. Although we expect a shift of only 2 nm in the position of the centroid, there is a reduction by more than a factor of 2 in mobility.

In summary, we have systematically investigated the effect of dislocations on the electron mobility in modulation-doped Si/SiGe layers. It is found that a density of threading dislocations higher than 1×10^8 cm^{-2} starts affecting the low-temperature mobility and that a density of 1×10^{11} cm^{-2} can reduce it by more than two orders of magnitude. For the same range of threading dislocation density, the room-temperature mobility is reduced from around 2700 to 1300 cm^2/V s. Misfit dislocations in the graded Ge-content buffer have a negligible effect at room (low) temperature when the electron channel is more than 0.1 (0.5) μm away from the graded buffer. As it gets closer than this limit a dramatic reduction in mobility is observed. Misfit segments at the bottom of the Si-channel interface can be avoided when the Si channel is grown thinner than the critical thickness. Once that thickness is exceeded the mobility degrades very quickly due to the proximity of the induced strain field.

ACKNOWLEDGMENTS

The author is grateful to M. Arafa, J. O. Chu, R. Feenstra, F. K. LeGoues, B. S. Meyerson, P. Mooney, and F. Stern for their help and support.

[1]D. Harame, L. Larson, M. Case, S. Kovacic, S. Voinigescu, T. Tewksbury, D. Nguyen, K. Stein, J. Cressler, S. J. Jeng, J. Malinowski, R. Groves, E. Eld, D. Sunderland, D. Rensch, M. Gilbert, K. Schonenberg, D. Ahlgren, S. Rosenbaum, J. Glenn, and B. Meyerson, IEDM Technical Digest 731 (1995).
[2]K. Ismail, IEDM Technical Digest 509 (1995).
[3]B. S. Meyerson, K. J. Uram, and F. K. LeGoues, Appl. Phys. Lett. **53**, 2555 (1988).

[4]K. Ismail, S. F. Nelson, J. O. Chu, and B. S. Meyerson, Appl. Phys. Lett. **63**, 660 (1993).
[5]F. Schaffler, D. Tobben, H. Herzog, G. Abstreiter, and B. Hollander, Semicond. Sci. Technol. **7**, 260 (1992).
[6]Y. J. Mii, Y. H. Xie, E. A. Fitzgerald, D. Monroe, F. A. Thiel, and B. E. Weir, Appl. Phys. Lett. **59**, 1611 (1991).
[7]K. Ismail, M. Arafa, K. Saenger, J. Chu, and B. Meyerson, Appl. Phys. Lett. **66**, 842 (1995).
[8]G. Abstreiter, H. Brugger, T. Wolf, H. Jorke, and H. J. Herzog, Phys. Rev. Lett. **30**, 226 (1985).
[9]D. Monroe, Y. H. Xie, E. A. Fitzgerald, P. J. Silverman, and J. P. Watson, J. Vac. Sci. Technol. B **11**, 1731 (1993).
[10]J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).
[11]K. Ismail, M. Arafa, F. Stern, J. O. Chu, and B. S. Meyerson, Appl. Phys. Lett. **66**, 842 (1995).

Quantum transport: Silicon inversion layers and InAlAs–InGaAs heterostructures

D. Vasileska, T. Eldridge,[a] and D. K. Ferry
Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6206

(Received 21 January 1996; accepted 26 April 1996)

Using a real-time Green's functions formalism, we investigate the influence of depletion charge scattering on the room temperature mobility in scaled silicon metal-oxide-semiconductor field-effect transistors and low-temperature transport in $In_{0.4}Al_{0.6}As$-$In_{0.4}Ga_{0.6}As$ modulation doped heterostructures. Our simulation results suggest that depletion charge scattering, which is usually ignored, has considerable impact on the electron transport in silicon inversion layers near the threshold gate voltage, even at room temperature. We also find that the weighting coefficients a and b (for the inversion and depletion charge densities) strongly depend on the substrate doping and deviate from that reported in the literature. In the case of modulation doped heterostructures, the low-temperature mobility is limited by alloy and Coulomb scattering. Intersubband scattering considerably affects the broadening of the states which, in turn, leads to mobility reduction.
© *1996 American Vacuum Society.*

I. INTRODUCTION

Electron transport properties of quasi-two-dimensional (Q2D) systems have been studied quite extensively, experimentally as well as theoretically. The inversion layer at a silicon interface[1] represents the most widely employed physical realization of the Q2D concepts and the electron mobility of the inversion layer electrons is one of the most important parameters for device characterization. It has been confirmed experimentally[2,3] that for substrates with low doping densities, the electron mobility follows a universal curve when plotted as a function of the effective electric field. However, pushing the devices towards scaling limits[4] necessarily leads to thinner oxide films (<10 nm) and higher substrate dopings ($>10^{17}$ cm^{-3}). For these devices, it was recently demonstrated[5] that experimental mobilities significantly deviate from the universal curves near the threshold voltage as a consequence of significant depletion charge scattering.

Since the invention of the modulation doping technique, transport properties of two-dimensional electron gases in high-mobility modulation doped AlGaAs–GaAs heterojunctions (MDHs) have also been a subject of great interest from both the technological and fundamental point of view.[6–10] The enhanced mobilities in these structures are attributed to the spatial separation of the electrons from their parent donors, which significantly reduces the scattering rate from the remote impurities.[11–13] This has stimulated much research toward their application in low-power, high-speed devices. Recently, there has been an increased interest devoted to the growth of InAlAs–InGaAs lattice matched heterojunctions[14–16] which have better carrier confinement due to a larger conduction-band discontinuity and a lighter conduction-band effective mass (compared to AlGaAs-GaAs). However, since the electron gas in the InAlAs–InGaAs MDH is mostly localized within a ternary random alloy (InGaAs), alloy scattering[17] is much more effective so that it is hard to imagine that the low-temperature mobility of InGaAs based modulation doped heterostructures may exceed the very high mobilities already reported in AlGaAs–GaAs heterostructures.

The theoretical model for the calculation of the subband structure and the electron mobilities of the Q2D electron gas, taking into consideration all major scattering mechanisms and both intrasubband and intersubband scattering, is briefly described in Sec. II. In Sec. III, devoted to transport in silicon inversion layers, we describe the influence of depletion charge scattering on the mobility results near the threshold gate voltage. We also show how the weighting coefficients for the inversion and depletion charge densities, which appear in the definition of the average electric field,[18] vary with doping density. Low-temperature transport properties of the modulation doped $In_{0.4}Al_{0.6}As$-$In_{0.4}Ga_{0.6}As$ heterostructures are discussed in Sec. IV. Some concluding comments related to the present work are given in Sec. V.

II. THEORETICAL MODEL

The subband structure is obtained from a self-consistent solution of the Schrödinger and Poisson equations, coupled with the Dyson equation for the retarded Green's function. Exchange-correlation corrections to the chemical potential are calculated within the local density approximation. The self-energy contributions from the various scattering mechanisms are evaluated by employing the self-consistent Born approximation. Screening is treated within the random phase approximation. In the calculation of the conductivity, we have taken account of a correction term due to the particle-hole ladder diagrams. More details on the calculational model and the analytical expressions for the broadening of the states and conductivity can be found elsewhere.[19]

[a]Permanent address: McDonnel Douglas Helicopter Systems, 5000 E. McDowell Rd., Mesa, AZ 85215-9797.

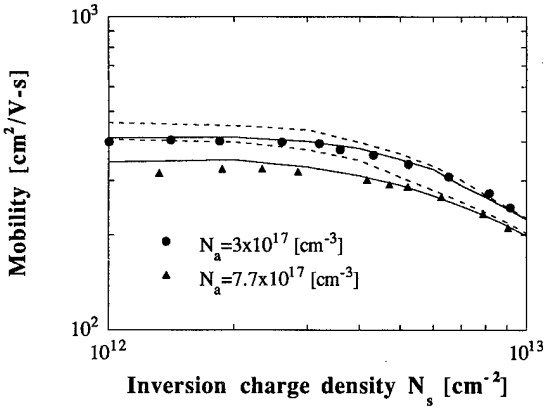


FIG. 1. Room temperature mobility vs inversion charge density for two different devices with high doping densities. For both devices, interface-trap density is $N_{it}=1.75\times10^{11}$ cm^{-2}. We use an exponential model for the surface-roughness autocorrelation function. We find that the rms height of the bumps is 0.26 nm (0.25 nm) for the device with lower (higher) doping. For both devices, roughness correlation length is estimated to be 1.4 nm.

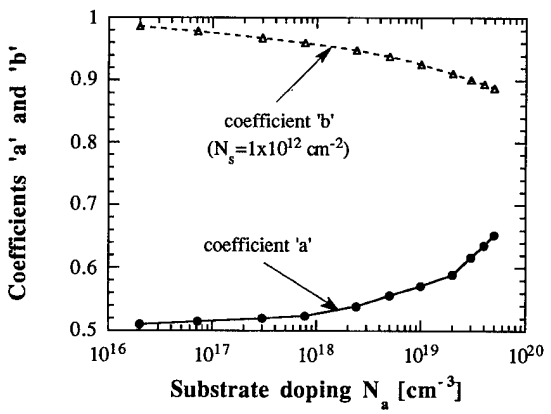


FIG. 2. Variation of the weighting coefficients a and $b=1-\langle z\rangle/W$ with substrate doping. We find that the penetration of the wave functions into the oxide is about 1 nm for very high values of N_a. Slightly higher values for the coefficient b were obtained for $N_s>1\times10^{12}$ cm^{-2} (not shown in this figure) due to the closer confinement of the carriers to the interface and increased population of the lower-lying subbands.

III. SILICON INVERSION LAYERS WITH HIGH DOPING PROFILES

The physical model employed in these calculations includes scattering with all bulk phonons (acoustic and nonpolar optical), surface-roughness scattering, and Coulomb scattering from both ionized impurities in the depletion region and interface/oxide charges. For intervalley scattering (zero-order and first-order interactions), we use the same phonon energies and coupling constants as those relevant for bulk silicon.[20] The anisotropy of the deformation potential interaction is also accounted for.[21] More details about the model can be found in the Appendix.

In Fig. 1, we compare our simulation results for the mobility with the experimental effective mobility data for n-channel metal-oxide-semiconductor field-effect transistors (MOSFETs) fabricated on (100) Si wafers. Filled circles and triangles represent the experimental effective mobility data derived from drain conductance measurements.[5] The inversion charge density has been determined from the gate-channel capacitance measurements. Solid (dashed) lines are the corresponding simulation results when depletion charge scattering is included (omitted) in the calculations. A total of four subbands (three from the unprimed and one from the primed ladder of subbands) are taken into account. We use the real density of states function (DOS) when evaluating the self-consistent potential-energy profile. In contrast to some previous studies,[18,19,22] we do not ignore the real self-energy corrections (renormalizations of the spectrum) when evaluating the broadening of the states and conductivity. We find that depletion charge scattering leads to mobility reduction of about 9% near the gate voltage thresholds (small values of N_s) for the device with $N_a=3\times10^{17}$ cm^{-3}. For the device with higher substrate doping ($N_a=7.7\times10^{17}$ cm^{-3}), the mobility reduction is about 12.5%. For both devices, the importance of surface-roughness scattering is shifted towards lower gate voltages. In addition, scattering due to acoustic and nonpolar optical phonons is found to be of lesser importance in these structures.

It has already been reported[22,23] that room temperature electron mobilities show universal behavior independent of the substrate doping when plotted as a function of the average electric field $E_{av}=(e/\epsilon_{sc})(aN_s+bN_{depl})$, where a and b are the weighting coefficients for the inversion N_s and depletion N_{depl} charge densities. In Fig. 2 we show the variation of these two coefficients with substrate doping. We find that a significantly deviates from 0.5 for substrates with very high doping densities. We also observe that the weighting coefficient b decreases with increasing substrate doping due to a decrease of the depletion region width. Results shown in Fig. 2 differ from those obtained by Matsumoto and Uemura[24] ($a=0.5$ and $b=1$), for example, because in their analytical result for the effective electric field, the terms of the order of $\langle z\rangle/W$ (suggested by Krutsick and White[23]), where $\langle z\rangle$ is the centroid of the electron density distribution and W is the depletion region width, were neglected.

IV. $In_{0.4}Al_{0.6}As$-$In_{0.4}Ga_{0.6}As$ MODULATION DOPED HETEROSTRUCTURES

At low temperatures, the dominant scattering mechanisms in the $In_{0.4}Al_{0.6}As$-$In_{0.4}Ga_{0.6}As$ MDH shown in Fig. 3 are: remote impurity (from the δ-depletion layer), interface-trap, alloy disorder, interface-roughness and deformation potential, and piezoelectric-coupled acoustic mode scattering.[12] Polar optical phonon scattering is found not to play any significant role at 1.6 K. Screening of the alloy disorder[17] and deformation potential scattering potentials by the free carriers is also neglected because of their short-range nature. This effect should certainly have been included for the long-range electron-piezoelectric acoustic phonon interaction, but since this scattering process is extremely weak at 1.6 K, its matrix element is also left unscreened. Assuming that deformation potential scattering is an elastic process, one has $\hbar\omega_Q/k_BT$

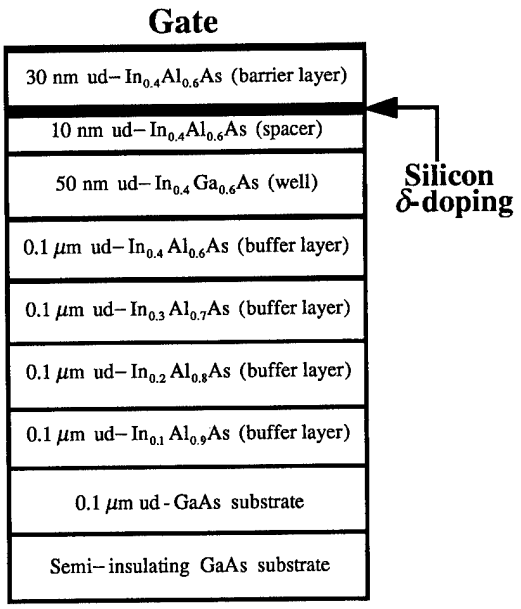


FIG. 3. Cross-sectional view of the δ-doped $In_{0.4}Al_{0.6}As$-$In_{0.4}Ga_{0.6}As$ heterostructure on (100)-oriented semi-insulating GaAs substrate with intermediate, undoped, and compositionally step-graded buffer layers.

$\approx \sqrt{3c_L m^*/\rho k_B T} = \sqrt{T_e/T}$, where ω_Q is the phonon frequency, k_B is the Boltzmann constant, c_L is the longitudinal elastic constant, m^* is the effective mass, and ρ is the crystal density. Using the material parameters summarized in Table I, we find that the equivalent temperature is $T_e \approx 0.22$ K. Thus, the equipartition approximation can be assumed to hold for temperatures above 1 K for the acoustic mode scattering of thermal electrons. The matrix elements for surface-roughness, acoustic phonon, and alloy disorder scattering can be found elsewhere.[26] For the wavevector dependent matrix element for scattering between subband n and m due to Coulomb charges located at a distance $z_i<0$ from the interface (in the InAlAs layer), we use the following expression:

TABLE I. Values of the $In_xGa_{1-x}As$ material parameters used in the calculation of the broadening of the states, density of states function, and mobility. For the electron effective mass and static dielectric constants of the $In_{0.4}Al_{0.6}As$, we use $m^*=0.99\ m_0$ and $\epsilon=11.5\ \epsilon_0$, respectively. The energy gaps of strained and unstrained $In_xGa_{1-x}As$ layers can be found in Ref. 25.

Parameter	Notation and units	Value
Relative electron effective mass	m^*	0.058 for $x=0.4$
Relative static dielectric constant	ϵ	$13.1+1.5x$
Deformation potential constant	D (eV)	$8.6-2.8x$
Piezoelectric constant	P	$0.052-0.007\,9x$
Longitudinal elastic constant	c_L ($\times 10^6$ N/cm^2)	$13.97-3.99x$
Lattice constant	a (Å)	$5.653\,25+0.405\,15x$
Alloy disorder parameter	$\langle V\rangle$ (eV)	0.5–0.55
Crystal density	ρ (g/cm^3)	5.504

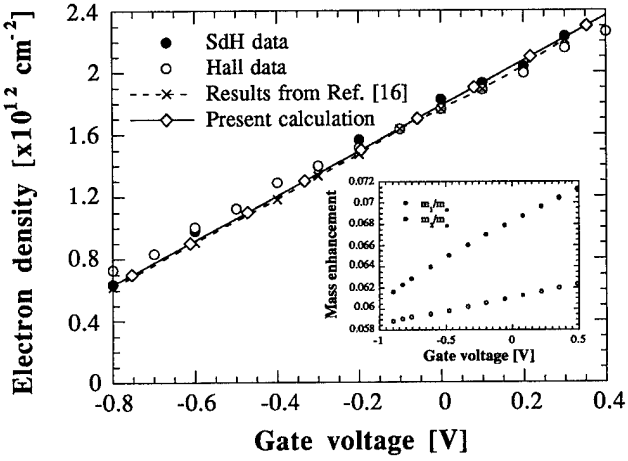


FIG. 4. Gate-voltage dependence of the total electron density for the MDH shown in Fig. 3. Also shown in this figure are the experimental data derived from the low-field Hall and resistivity measurements, Shubnikov–de Haas magnetoresistance measurements as well as the theoretical simulations reported in Ref. 16. The effective mass enhancement due to conduction band nonparabolicity for the carriers residing in the ground and first excited subband is shown in the inset.

$$\langle n|U(q,z_i)|m\rangle = \frac{e^2}{2q\epsilon_b}\int_{-\infty}^{0} dz\ \psi_n(z)\left[e^{-q|z-z_i|} + \frac{\epsilon_b-\epsilon_w}{\epsilon_b+\epsilon_w}\,e^{-q|z+z_i|}\right]\psi_m(z) + \frac{e^2}{q(\epsilon_b+\epsilon_w)} \times \int_{-\infty}^{0} dz\ \psi_n(z)e^{-q|z-z_i|}\psi_m(z), \tag{1}$$

in which the image term due to the difference of the static dielectric constants in the well (ϵ_w) and in the barrier layer (ϵ_b) is properly taken into account. The effective mass enhancement with increasing gate voltage due to conduction band nonparabolicity is also taken into account.[27]

When solving for the subband structure, the number of the self-consistent field iterations was considerably decreased when using the analytical solutions of the Poisson equation. We calculate that in equilibrium, the electron transfer from the δ-doped region to the well (see Fig. 3) is governed by

$$E_F+\Phi_0-eV_G=\Delta E_c+\frac{e^2N_DL}{\epsilon_b}-\frac{e^2}{\epsilon_b}\sum_i N_i^b\int_{-\infty}^{0} z\psi_i^2(z)dz - \frac{e^2}{\epsilon_b}(W+L)\sum_i\left\{N_i^w\left[1-\int_{-\infty}^{0}\psi_i^2(z)dz\right] + N_i^b\int_{-\infty}^{0}\psi_i^2(z)dz\right\}, \tag{2}$$

where $L=30$ nm and $W=10$ nm are the widths of the barrier and spacer layers, E_F is the Fermi level, $\Phi_0=0.8$ eV is the Shottky barrier height, V_G is the gate voltage, $\Delta E_c=0.52$ eV is the conduction band offset, $N_D=3\times10^{12}$ cm^{-2} is the Si δ-doping sheet, and $N_i^b(N_i^w)$ is the sheet electron density of the ith subband in the barrier (quantum well).

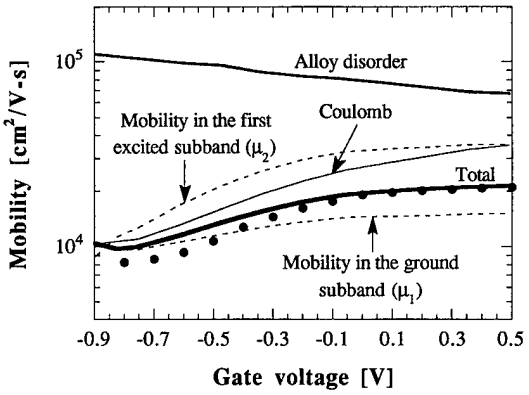


FIG. 5. Cumulative and component theoretical mobilities vs gate voltage. Filled circles are the experimental data determined from Hall effect and resistivity measurements (Ref. 16). The material parameters employed in the simulations are summarized in Table I. Also shown in the figure are the mobilities of the carriers in the ground (μ_1) and first excited subband (μ_2).

Self-consistent results for the gate-voltage dependent electron densities are shown in Fig. 4. We observe good agreement between the experimental data derived from the Hall effect and SdH measurements[16] and the theoretical calculations for all gate voltages. Our simulation results are also in agreement with the theoretical calculations reported by Wieder *et al.*[16] We find that the significant effective mass enhancement (see the inset of Fig. 4) due to conduction band nonparabolicity prevents the population of the third subband until $V_G>0.4$ V. We also find that, in contrast to silicon inversion layers where exchange-correlation effects are significant even at room temperature, many-body corrections do not considerably affect the subband structure in the MDH because of the small effective mass.

Gate-voltage dependence of the low-temperature theoretical mobilities is shown in Fig. 5. Although the average mobility is $\bar{\mu}=\Sigma_i N_i\mu_i/N_s$, measured mobilities shown in Fig. 5 are Hall mobilities[16] so that for a multisubband case, the expression $\bar{\mu}=\Sigma_i N_i\mu_i^2/\Sigma_i N_i\mu_i$ is more suitable for comparing the experimental results with the theoretical calculations. To take into account the diffusion effects, the δ-doping sheet is modeled as a 4 nm wide layer centered 8 nm away from the interface with Si doping density of 7.52×10^{18} cm^{-3}. Interface-roughness scattering is assumed to exist at both heterointerfaces of the quantum well. We assume that the rms height of the roughness is one monolayer (ML) ($\Delta=2.83$ Å). The lateral spread of the interface roughness is taken to be $\zeta=100$ Å. For this particular heterostructure, we find that surface-roughness and acoustic phonon limited component mobilities (not shown on the figure) are of the order of $(0.5–1)\times10^6$ and 10^7 cm^2/V s, respectively. By fitting the calculated mobilities at high gate voltages to the experimental data, we were able to determine the alloy-disorder scattering parameter in the system to be in the range 0.5–0.55. Component mobility results presented in Fig. 5 suggest that the inclusion of alloy scattering, in addition to Coulomb scattering (from the ionized donors in the δ-doped region and interface traps), leads to good agreement between the theoretical calculations and the experimental data. The dip in the theoretical mobility at $V_G=-0.85$ V signals that a new channel (second subband) for scattering opens in the MDH which leads to a sudden increase in phase space for scattering between the two subbands to occur. Since the carriers residing in the first excited subband are at about four times larger average distance from the heterointerface (compared to those in the ground subband), they do not feel as much the Coulomb scattering potentials and exhibit about two times higher mobility.

V. CONCLUSIONS

Our results suggest that depletion charge scattering, which has been ignored in some previous theoretical studies, is responsible for the deviation of the mobility from the universal curves near the threshold voltage. At high gate voltages (high values of N_s), due to the closer confinement of the carriers to the interface, the magnitude of the mobility in these highly doped samples is considerably reduced by surface-roughness scattering, in addition to the still significant intervalley scattering. We also find that the weighting coefficients a and b—that appear in the definition of the average electric field—show a pronounced doping dependence. For the $In_{0.4}Al_{0.6}As$-$In_{0.4}Ga_{0.6}As$ MDH, we find that at $T=1.6$ K the magnitude of the mobility is limited by Coulomb and alloy scattering. From the subband structure (gate-voltage dependence of the electron density) and mobility simulations, we calculate that the relevant effective mass at the conduction band minima of the $In_{0.4}Ga_{0.6}As$ ternary alloy is $m^*=0.058$ m_0. An observed effective mass enhancement suggests that nonparabolicity effects are considerable in this MDH and must be taken into account if accurate mobility results are desired. The effect of the many-body interactions, calculated in the local density-functional approximation, is found not to be important for the MDH.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Office of Naval Research for sponsoring this project. They also would like to thank H. H. Wieder for providing us with valuable information about the material parameters of the $In_{0.4}Al_{0.6}As$-$In_{0.4}Ga_{0.6}As$ heterostructure.

APPENDIX

The expression for the matrix elements for scattering between subbands n and m due to a Coulomb charge in the depletion region, in the oxide, or right at the interface, in the presence of a different dielectric medium, can be found in the review article by Ando, Fowler, and Stern.[1] For the surface-roughness matrix element, we use the result obtained by Ando[28] who corrected the result obtained by Matsumoto and Uemura[24] by taking into account the change in the electron density distribution and the effective dipole moment of

the deformed Si–SiO_2 surface in addition to the usual local potential energy term that equals the linear term of the Taylor expansion of the surface potential.

For anisotropic elastic continuum such as silicon, for which the deformation potential constant Ξ is a tensor,[29] matrix element squared for scattering between subbands n and m due to acoustic phonons, in the usual equipartition approximation, equals to[21]

$$|\langle n|U_\lambda^{ac}(\mathbf{q})|m\rangle|^2=\frac{k_BT}{\rho V v_{s\lambda}^2}[\Delta_{\lambda,nm}^{\text{eff}}(\mathbf{q})]^2F_{nm}, \tag{A1}$$

where

$$F_{nm}=\int_0^\infty dz\ \psi_n^2(z)\psi_m^2(z), \tag{A2}$$

$v_{s\lambda}$ is the sound velocity, and ρ is the Si mass density. The effective deformation potential constant that appears in expression (A1) is calculated from

$$[\Delta_{\lambda,nm}^{\text{eff}}(\mathbf{q})]^2=\frac{1}{F_{nm}}\int_0^\infty dq_z\ \Delta_\lambda^2(\theta_Q)|\mathscr{F}_{nm}(q_z)|^2, \tag{A3}$$

where

$$\mathscr{F}_{nm}(q_z)=\int_0^\infty dz\ \psi_n(z)e^{iq_zz}\psi_m(z) \tag{A4}$$

and $\theta_\mathbf{Q}$ is the angle between the wave vector $\mathbf{Q}$ of the emitted (absorbed) phonon and the longitudinal axis of the valley. According to Herring and Vogt,[29] the effective deformation potentials $\Delta_\lambda(\theta_\mathbf{Q})$ (λ=LA or TA) that appear in Eq. (A3) are given by

$$\Delta_{LA}(\theta_\mathbf{Q})\approx\Xi_d+\Xi_u\cos^2(\theta_\mathbf{Q}) \tag{A5a}$$

and

$$\Delta_{TA}(\theta_\mathbf{Q})\approx\Xi_u\cos(\theta_\mathbf{Q})\sin(\theta_\mathbf{Q}). \tag{A5b}$$

In Eqs. (A5a) and (A5b), Ξ_u is the uniaxial shear potential, and Ξ_d is the dilatation potential that are believed to have values of approximately 9.0 and −11.7 eV, respectively.[21,30] It is important to point out that Eq. (A5b) accounts for the contribution of both TA branches.

The matrix element for nonpolar optical phonon scattering is generally found from a deformable ion model. The zero-order process, which describes a perturbation that can be associated with the local dilatation or compression of the lattice which produces local fluctuation in the electron energy, dominates at moderately high electron energies. The matrix elements for scattering between subbands n and m (that belong to α and β valleys) due to this zero-order interaction is given by

$$|\langle n|U_\lambda^{\text{op}(0)}|m\rangle|^2=\frac{\hbar D_\lambda^2}{2\rho V\omega_{o\lambda}}F_{nm}, \tag{A6}$$

where D_λ is the deformation field and $\omega_{0\lambda}$ is the frequency of the phonon mode.

When the zero-order matrix element for the optical or intervalley interaction vanishes, then D_λ is identically zero. In this case, one has to consider the first-order term of the interaction. The squared matrix element for scattering between subbands n (α valley) and m (β valley) due to this first-order process is given by[31]

$$|\langle n|U_\lambda^{\text{op}(1)}|m\rangle|^2=\frac{\hbar D_{1\lambda}^2}{2\rho V\omega_{0\lambda}}(q^2F_{nm}+c_{nm}), \tag{A7}$$

where

$$c_{nm}=\int_0^\infty dz\left\{\frac{d}{dz}[\psi_n(z)\psi_m(z)]\right\}^2. \tag{A8}$$

In the scattering among the equivalent valleys, there are two types of phonons that might be involved in the process. The first type, the so-called g phonon couples the two valleys along opposite ends of the same axis, i.e., ⟨100⟩ to ⟨$\bar{1}$00⟩. This is an umklapp process and has a net phonon wavevector $0.3\pi/a$. The so-called f phonons couple the ⟨100⟩ valley with ⟨010⟩, ⟨001⟩, etc. Hence, the reciprocal lattice vector involved in the g-process is $\mathbf{G}_{100}$ and that for an f process is $\mathbf{G}_{111}$.

Within a four subband approximation, scattering between the two valleys in the ϵ_0, ϵ_1, and ϵ_2 subbands involves only g-type phonons. The scattering between these three minima is treated by using a high-energy phonon of 750 K activation temperature (treated as zero-order interaction) and 134 K phonon treated via first order interaction. Scattering between the ϵ_0, ϵ_1, and ϵ_2 subbands and the four ϵ_0' subbands involves f phonons with activation temperatures of 630 and 230 K treated via zero-order and first-order interaction, respectively. Scattering between the ϵ_0' subbands involves both g and f phonons with activation temperatures of 630 K (zero-order interaction) and 190 K (first-order interaction). All of the high-energy phonons are assumed to be coupled with a value of $D_\lambda=9\times10^8$ eV/cm, and all of the first-order coupled phonons are assumed to be coupled with $D_{1\lambda}=5.6$ eV. These values are consistent with the results given in Refs. 20 and 32.

[1]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[2]A. G. Sabnis and J. T. Clemens, IEDM Tech. Dig. **18** (1979).
[3]S. C. Sun and J. D. Plummer, IEEE Trans. Electron Devices **ED-27**, 1497 (1980).
[4]M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro, and H. Iwai, IEEE Trans. Electron Devices **ED-42**, 1822 (1995).
[5]S. Takagi, A. Toriumi, M. Iwase, and H. Tango, IEEE Trans. Electron Devices **ED-41**, 2357 (1994).
[6]R. Dingle, H. L. Störmer, A. C. Gossard, and W. Wiegmann, Appl. Phys. Lett. **33**, 665 (1978).
[7]H. L. Störmer, A. Pinczuk, A. C. Gossard, and W. Wiegmann, Appl. Phys. Lett. **38**, 691 (1981).
[8]A. Kastalsky, R. Dingle, K. Y. Cheng, and A. Y. Cho, Appl. Phys. Lett. **41**, 274 (1982).
[9]K. Lee, M. S. Shur, T. J. Drummond, and H. Morkoç, J. Appl. Phys. **54**, 6432 (1983).
[10]L. Pfeiffer, K. W. West, H. L. Störmer, and K. W. Baldwin, Appl. Phys. Lett. **55**, 1888 (1989).
[11]E. E. Mendez, P. J. Price, and M. Heiblum, Appl. Phys. Lett. **45**, 294 (1984).
[12]H. L. Störmer, Surf. Sci. **132**, 519 (1983).
[13]P. J. Price, Surf. Sci. **143**, 145 (1984).
[14]N. C. Tien, J. Chen, J. M. Fernàndez, and H. H. Wieder, IEEE Trans. Electron Device Lett. **EDL-13**, 621 (1992).

[15]J. Chen, H. H. Wieder, and A. P. Young, J. Appl. Phys. **76**, 4743 (1994).
[16]H. H. Wieder, R. S. Goldman, J. Chen, and A. P. Young, J. Vac. Sci. Technol. B **13**, 1853 (1995).
[17]G. Bastard, Appl. Phys. Lett. **43**, 591 (1983).
[18]D. Vasileska, P. Bordone, T. Eldridge, and D. K. Ferry, J. Vac. Sci. Technol. B **13**, 1841 (1995).
[19]D. Vasileska, T. Eldridge, P. Bordone, and D. K. Ferry, VLSI Design (to be published).
[20]D. K. Ferry, Phys. Rev. B **14**, 1605 (1976).
[21]M. V. Fischetti and S. E. Laux, Phys. Rev. B **48**, 2244 (1993).
[22]D. Vasileska, P. Bordone, T. Eldridge, and D. K. Ferry, Physica B (to be published).
[23]T. J. Krutsick and M. H. White, IEEE Trans. Electron Devices **ED-35**, 1153 (1988).
[24]Y. Matsumoto and Y. Uemura, Jpn. J. Appl. Phys. Suppl. **2**, 367 (1974).
[25]S. Adachi, J. Appl. Phys. **53**, 8775 (1982).
[26]S. M. Goodnick and D. K. Ferry, in *Physics and Chemistry of III–V Compound Semiconductor Interfaces*, edited by C. W. Wilmsen (Plenum, New York, 1985).
[27]E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957); in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), p. 75; in *Handbook on Semiconductors*, edited by W. Paul (North-Holland, Amsterdam, 1982), Vol. 1, p. 193.
[28]T. Ando, J. Phys. Soc. Jpn. **43**, 1616 (1977).
[29]C. Herring and E. Vogt, Phys. Rev. **101**, 944 (1956).
[30]B. K. Ridley, *Quantum Processes in Semiconductors* (Oxford University Press, New York, 1993).
[31]D. K. Ferry, *Semiconductors* (Macmillan, New York, 1991).
[32]D. K. Ferry, Phys. Rev. B **14**, 5364 (1976).

Ballistic electron emission microscopy of Au–InAs–GaAs system

Mao-long Ke, D. I. Westwood, C. C. Matthai,[a] and B. E. Richardson
Department of Physics and Astronomy, University of Wales Cardiff, Cardiff CF2 3YB, United Kingdom

R. H. Williams
University of Wales Swansea, Swansea SA2 8PP, United Kingdom

(Received 21 January 1996; accepted 13 April 1996)

We report an investigation of barrier formation and electron transport across InAs and GaAs interfaces by ballistic electron emission microscopy. The barrier height was found to decrease with the InAs thickness and the detailed variation correlates with the relaxation of the InAs layer. When the thickness is below one monolayer, three thresholds are observed and they are attributed to electron transmission into the Γ, L, and X valleys of the GaAs, respectively. As the thickness increases to three monolayers or beyond, only two thresholds are observed and the reasons are explained in the context of the band structure of the system. © *1996 American Vacuum Society.*

I. INTRODUCTION

The growth of InAs on GaAs (001) has recently attracted a large degree of attention since it is possible to realize zero dimensional quantum dots through strain relaxation.[1,2] It is well known that the InAs layer undergoes a number of transitions as the thickness of the InAs layer increases.[3] Initially, the InAs layer grows pseudomorphically strained onto the GaAs substrate; above ~1.6 monolayers, due to the large mismatch (~7.2%), it starts to island coherently; finally misfit dislocations are generated and the islands coalesce as the thickness increases further. The aim of this article is to study how the various stages of InAs film development affect the barrier formation between Au and GaAs, with particular attention being paid to the barrier height, electron transport, and the homogeneity of the interface.

Ballistic electron emission microscopy (BEEM) has been used for this study. The technique is a three-terminal extension of scanning tunnelling microscope (STM). It uses the tip of the STM to inject hot electrons into the metal surface and collects the fraction of them transmitted over the barrier into the semiconductor substrate. It can measure the barrier height with very high lateral resolution.[4–6] Furthermore, the semiconductor band structure near the interface can also be probed since transmission into the different conduction bands (valleys) can be measured.[7]

This article is organized as follows: the sample growth is presented in Sec. II, and is followed by the results of BEEM measurements with discussions (Sec. III). Finally our conclusions are presented in Sec. IV.

II. GROWTH DETAILS

All the samples studied were prepared in a Vacuum Generators ultrahigh vacuum (UHV) system consisting of a conventional solid source VG V80H molecular beam epitaxy (MBE) machine attached, via an UHV gate valve, to a home designed metallization chamber. Both the MBE and the metallization chamber have been described previously.[8]

The substrate used were highly doped ($n=1\times10^{18}$ cm^{-3}) on-axis GaAs (100) wafers which, following a standard etch, were In mounted onto Mo Blocks. Oxide desorption was performed in the growth chamber of the MBE machine under an As_4 flux of $F(As_4)\sim5\times10^{14}$ molecules cm^{-2} s^{-1} and followed by growth, at a sample temperature of 600 °C and a growth rate of 1.0 μm/h, of a thick (0.5–0.7 μm) low doped GaAs layer ($n=5\times10^{15}$ cm^{-3}). Following this the sample temperature was lowered to 500 °C before growth of the final doped ($n=10^{19}$ cm^{-3}) InAs layer at a rate of 0.5 μm/h. Islands formed at low coverages (equivalent to a few monolayers of InAs), undergo a complex development with time, temperature, and coverage once growth has been terminated. At a temperature of 500 °C this development is reported to take a few seconds,[9,10] therefore, in this study, all the InAs surfaces were kept at 500 °C for 1 min following cessation of growth for equilibrium to be reached. The sample was then allowed to cool to ~300 °C before removal from the growth chamber and transfer to the metallization chamber where a very thin gold film (50 Å) was deposited. The predeposition pressure was $\sim3\times10^{-10}$ mb and the pressure during the metallization was controlled to around 5×10^{-9} mb. The deposition rate was ~0.33 Å per second. The BEEM measurements were performed using a WA STM system in air at room temperature.

III. MEASUREMENTS AND DISCUSSIONS

A set of ten samples were investigated, with InAs thickness ranging from 0–40 monolayers. Figure 1 shows three typical BEEM spectra for the cases of 1, 5, and 27 monolayers of InAs. Fitting the data with a 5/2 power law,[11] it was found there were three thresholds for the case of 0–1 monolayer with the energy separation between the two lower thresholds (~300 meV) in agreement with that between the Γ and the L valleys of GaAs.[12] The separation between the first and third thresholds (~470 meV) is consistent with the $\Gamma-X$ valley separation.[12] For the cases of 3–27 monolayers of InAs, however, only two thresholds were observed with the energy separations varying in the range from 340–410 meV. These results will be discussed in detail later.

[a]Electronic mail: matthai@cardiff.ac.uk

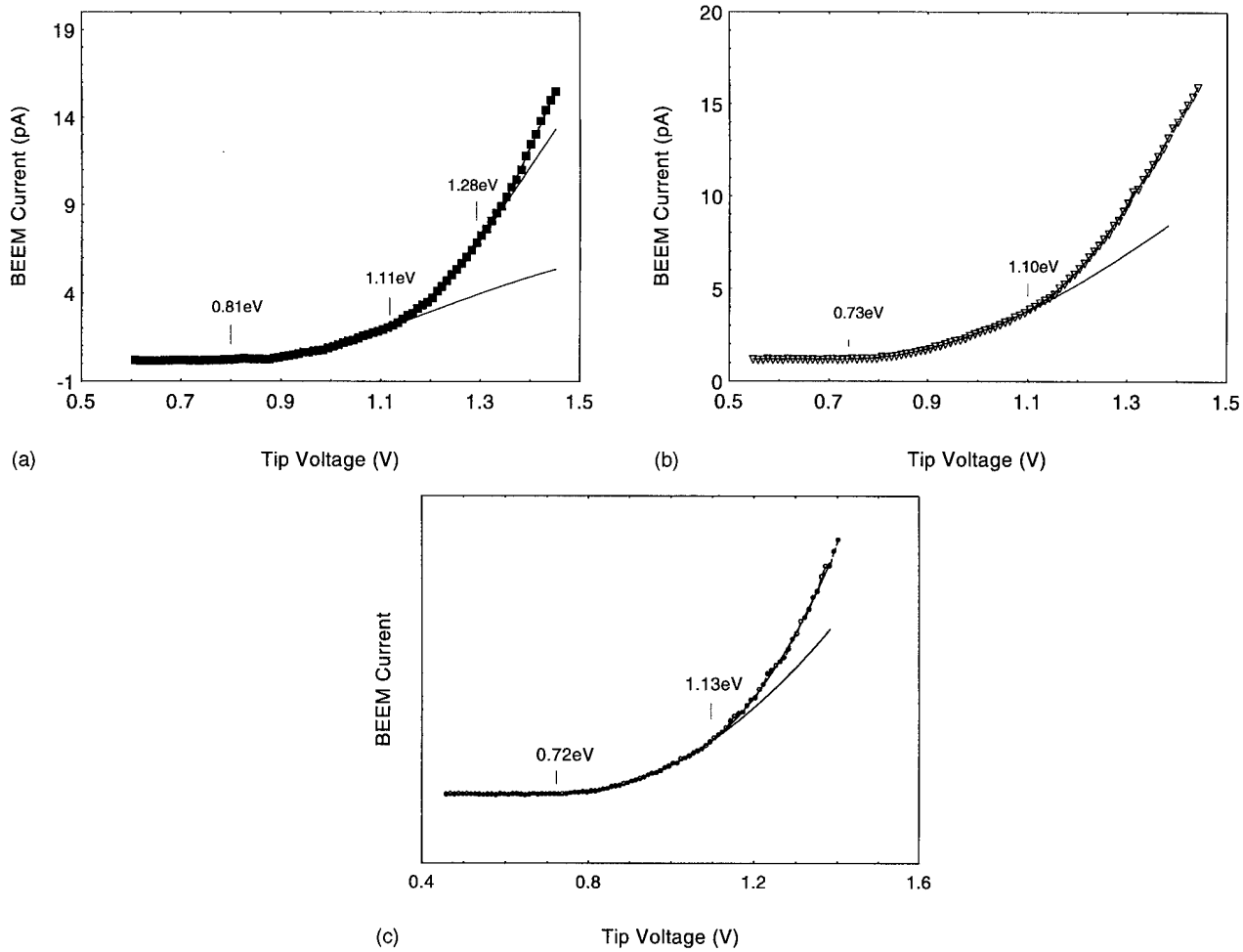


FIG. 1. Three representative BEEM current curves for the cases of 1, 5, and 27 monolayers of InAs, respectively. Three thresholds were obtained from fitting for the case of one monolayer of InAs (a) while only two were seen in the other cases (b) and (c).

Figure 2 shows the variation of the barrier heights (lowest thresholds) as a function of the InAs layer thickness. The "zero monolayer" sample is simply a Au/GaAs reference system. The data show that the barrier decreases with the thickness of InAs interlayer. One monolayer of InAs lowers the barrier from 0.90–0.80 eV, a value which is uniform across the interface. No barriers as high as that of Au/GaAs were measured in this structure, which indicates that a uniform InAs film has been produced. As the thickness of the InAs layer is increased to three monolayers the barrier was found to decrease to ~0.74 eV. The barrier height remains approximately constant for up to 27 monolayers, after which it reduces to ~0.63 eV for the case of 33 monolayers (100 Å) and beyond.

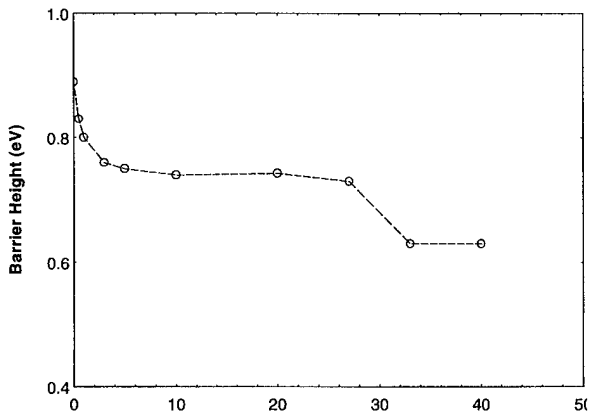


FIG. 2. The variation of barrier height as a function of the InAs layer thickness.

It is well known that the contact between Au and n^+ InAs material is ohmic,[13,14] thus the measured barrier for the thick InAs layer can be seen as the effective band offset between the InAs and GaAs. It is interesting to note that the measured barriers fall into three groups (Fig. 2), which coincide with the three transition stages of the InAs film. A thin uniformly strained InAs film (one monolayer) has a barrier height of 0.80 eV; relaxation of strain through islanding (3–27 monolayers) resulted in a smaller height of ~0.74 eV; heavily dislocated films (≥33 monolayers) have the lowest barrier of ~0.63 eV. The barrier height at each stage does not vary considerably, with the largest fluctuation of about ±50 meV.

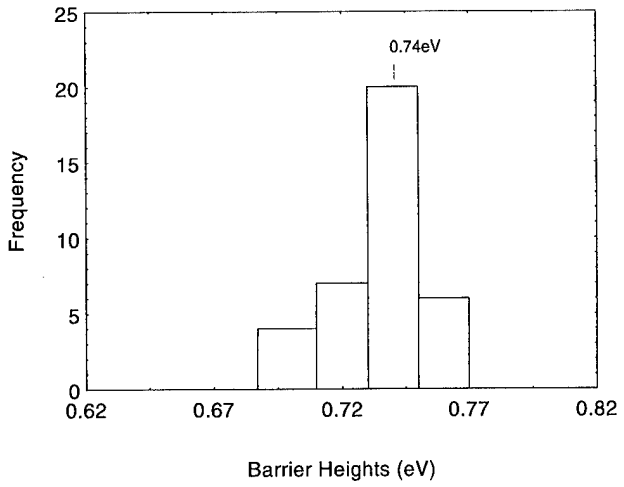


FIG. 3. The histogram of the barrier heights obtained from the sample with 27 monolayers of InAs.

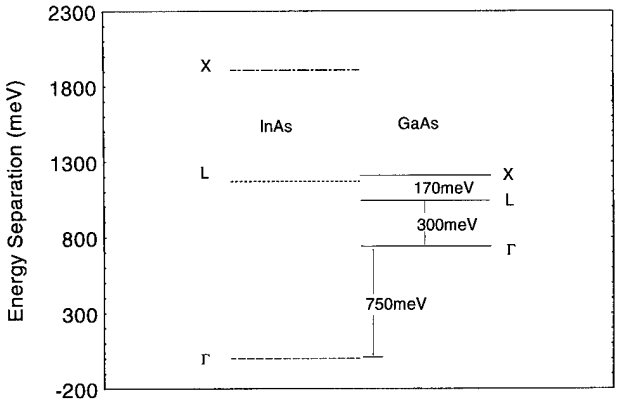


FIG. 4. The band alignment at the InAs/GaAs interface. The relative positions of the Γ, *L*, and *X* valleys for Fig. 5 are obtained from Ref. 17. Other work (such as Ref. 15) suggests a similar configuration.

Most of the observed barriers are very close to the average value. Figure 3 shows a histogram for the case of 27 monolayers. The measured barriers over different positions are largely concentrated around 0.74 eV, which indicates the interface is predominately a uniform one.

Oloumi and Matthai[15] calculated the band alignment at the InAs/GaAs interface and found that the band offsets were crucially dependent on the interlayer separation at the interface. For InAs pseudomorphically strained on GaAs substrate, they found that if the interlayer separation is fully energy minimized, which was akin to the As layer being pulled towards the In atom and away from the Ga atom, a conduction band offset of 1.0 eV (higher than that the Schottky barrier between Au and GaAs) can be achieved. However, when the interlayer separation was found to be the average of the bond lengths, the conduction band offset was found to reduce to half of the previous value. This latter configuration could be thought to represent a fully relaxed InAs film whereas in the former case, the first layer InAs atoms are strained onto the substrate. The conclusion would therefore be that when the strain is relieved at the interface (e.g., through misfit dislocation) there is a corresponding decrease in the conduction band offset. Our result of 0.63 eV ($\geqslant$33 monolayers) and 0.74 ($\leqslant$27 monolayers), although not in agreement with the calculated value, is consistent with the calculated trend, assuming that $\geqslant$33 monolayers represents the structure when the strain in the InAs film is relieved by the onset of dislocations.

The idea of the strain relaxation through islanding is consistent with above argument as well since the islands are said to have an average size of around 300 Å,[9,10] which is much larger than the expected BEEM resolution.[4–7] The barrier heights measured on top of the islands are, therefore, representing the band offset between the InAs island and the GaAs substrate. The different average value at each transition stage is probably a result of a different degree of strain relaxation. For thick InAs layers ($\geqslant$33 monolayers), the dislocations may play a dominate role in the barrier formation; while in the region from 3–27 monolayers, the measured barrier height is probably the consequence of the band offset between the strained InAs islands and the GaAs substrate. For even thinner InAs layers (below one monolayer), it is inappropriate to talk about the band structure of the InAs, the measured barrier height is essentially the Schottky barrier between Au and GaAs, but modified in the interface region by the dipoles from the InAs interlayer.

Another important feature in the BEEM spectra is that there are three thresholds for the case of the Au/GaAs contact (zero monolayer of InAs) and for the samples with very thin layer of InAs (one monolayer or less), but only two [Figs. 1(b) and 1(c)] can be found when the InAs interlayer is increased to three monolayers or beyond. For systems with three or five monolayers of InAs, the two observed thresholds have energy separations of 340 and 370 meV, respectively. When the InAs layer thickness increased to 10 monolayers, the energy separation of the two thresholds becomes 410 meV and remains at that value for up to 27 monolayers. For samples with 33 monolayers of InAs or thicker, it was difficult to ascertain exactly how many upper thresholds exist because the BEEM current is noisy, which is not surprising since the barrier height of 0.63 eV is about the low limit of room temperature operation.

As mentioned earlier, the observed three thresholds for the case of a very thin layer of InAs (below one monolayer) are identified as electron transmission into the Γ, *L*, and *X* valleys of the GaAs, respectively.[12,16] To understand why only two thresholds are observed for 3–27 monolayers of InAs, it is necessary to examine the band structure, or band alignment, of the system in more detail. The first point to note is that when considering transport along (001) direction, the *X* point can be projected onto the Γ point whereas the *L* point cannot. Thus $\Gamma-L$ valley transmission can only occur when the condition for lateral momentum conservation is relaxed. This is more likely to happen at the Au–InAs interface than at the semiconductor heterojunction (InAs/GaAs). Second, the *X*,*L* valleys of the InAs are higher in energy than those in GaAs[15,17] (see Fig. 4). Finally, the relative po-

sition of the L and X valleys in InAs do change with strain. Thus of the four possible transport routes into the higher valleys (Γ,X) of GaAs, Au–Γ(InAs)–L(GaAs), Au–Γ(InAs)–X(GaAs), Au–L(InAs)–L(GaAs), and Au–X(InAs)–X(GaAs), only the first two could show up as a threshold at the corresponding valley position of GaAs and of these only the second satisfies the conservation rule. The last two routes, if in operation, would manifest as a threshold at the corresponding InAs valley position (rather than GaAs). Since some theoretical calculations predict that the L valley of InAs is about 1.2 eV higher than the Γ point,[15,17] we therefore believe that the second threshold [at 1.13 eV in Fig. 1(c)] for 10–27 monolayers of InAs is a measure of the X valley energy in GaAs. The separation between the two thresholds (410 meV) is in a broad agreement with Γ-X spacing.[12] The second threshold observed for the thinner (three and five monolayers) InAs interlayers (340 and 370 meV) could be a result of direct electron scattering from Au into the L "valley" of GaAs. As mentioned earlier, when the InAs interlayer is only one monolayer or even thinner the direct electron scattering from Au into the L valley minimum (300 meV from the Γ point) of GaAs is strong and clearly visible [Fig. 1(a)]. As the interlayer thickness increases, the scattering strength from Au into the L valley minimum of GaAs is expected to decrease and become invisible. However, since the scattering strength is believed to be energy dependent,[16] an increase in energy would therefore help to enhance the electron scattering rate.

Finally, we would like to point out that we have not seen the current contribution from the X valley of the GaAs for these two samples (three and five monolayers). The precise reason is unknown. One possible explanation is that the X valley contribution is probably close in energy with the observed second threshold, which would be very difficult to resolve from the fitting procedure. It is therefore possible that the second threshold is in fact due to a combination of two processes. Obviously, further attention is needed for samples with a such thin interlayer (~5 monolayers) of InAs.

IV. CONCLUSIONS

Ballistic electron emission microscopy has been used to probe the barrier variation of Au/InAs/GaAs system as the thickness of the InAs was varied between 0 and 40 monolayers. The barrier height generally decreases with the InAs thickness and its variation has been found to correlate to the strain relaxation of the InAs layer. The transport properties were investigated as well, and the absence of the electron transmission into the L valley of the GaAs substrate for samples with a thicker InAs layer (10–27 monolayers) was explained by the requirement of lateral momentum conservation across the InAs and GaAs interface, a requirement that is relaxed for the less perfect Au/InAs and Au/GaAs interfaces. For systems with intermediate thickness of the InAs layer (three or five monolayers), the second threshold in BEEM spectra is believed to be a result of direct electron scattering from Au into the GaAs.

ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support from U.K. Engineering and Physical Science Research Council. The EC HCM "GOODS" program is also gratefully acknowledged.

[1]S. Guha, A. Madhukar, and K. C. Rajkumar, Appl. Phys. Lett. **57**, 2110 (1990).
[2]P. Chen, Q. Xie, A. Madhukar, C. Chen, and A. Konkar, J. Vac. Sci. Technol. B **12**, 2568 (1994).
[3]D. Leonard, K. Pond, and P. M. Petroff, Phys. Rev. B **49**, 11 687 (1994).
[4]W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988).
[5]L. J. Schowalter and E. Y. Lee, Phys. Rev. B **43**, 9308 (1991).
[6]Mao-long Ke, D. I. Westwood, S. Wilks, S. Heghoyan, A. Kestle, C. C. Matthai, B. E. Richardson, and R. H. Williams, J. Vac. Sci. Technol. B **13**, 1684 (1995).
[7]W. J. Kaiser, M. H. Hecht, L. D. Bell, F. J. Grunthermer, J. K. Liu, and L. C. Davies, Phys. Rev. B **48**, 18 324 (1993).
[8]D. I. Westwood, D. A. Woolf, and R. H. Williams, J. Cryst. Growth **98**, 782 (1989); S. Wilks, Ph.D. thesis, University of Wales (1992).
[9]S. M. Scholz, A. B. Müller, W. Richter, D. R. T. Zahn, D. I. Westwood, D. A. Woolf, and R. H. Williams, J. Vac. Sci. Technol. B **10**, 1710 (1992).
[10]J. Y. Marzin, J. M. Gérard, A. Izraël, D. Barrier, and G. Bastard, Phys. Rev. Lett. **73**, 716 (1994).
[11]R. Ludeke, M. Prietsch, and A. Samsavar, J. Vac. Sci. Technol. B **9**, 2342 (1991).
[12]J. S. Blackmore, J. Appl. Phys. **53**, R123 (1982).
[13]N. A. Washburn, J. R. Sites, and H. H. Wieder, J. Appl. Phys. **50**, 4872 (1979).
[14]D. Tsui, Solid State Commun. **9**, 1789 (1971).
[15]M. Oloumi and C. C. Matthai, J. Phys. Condensed Matter **2**, 5153 (1990).
[16]Mao-long Ke, D. I. Westwood, C. C. Matthai, B. E. Richardson, and R. H. Williams, Phys. Rev. B **53**, 4845 (1996).
[17]J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976).

Interface roughness effects on transport in tunnel structures

D. Z.-Y. Ting[a)]
Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125 and Department of Physics, National Tsing Hua University, Hsinchu, Taiwan 300, Republic of China

T. C. McGill[b)]
Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125

(Received 22 January 1996; accepted 16 April 1996)

Direct simulations of interface roughness effects on transport properties of tunnel structures are performed using the planar supercell stack method. The method allows for the inclusion of realistic three-dimensional rough interfacial geometries in transport calculations. For double barrier resonant tunneling structures, we used our method to analyze the effect of roughness at each of the four interfaces, and to test for sensitivity of transport properties to island size and height. Our simulations yields the following conclusions: (1) We find that scattering of off-resonance states into on-resonance states provides the dominant contribution to interface roughness assisted tunneling. Analyses of scattering strength sensitivity to interface layer configurations reveals preferential scattering into $\Delta k_{\parallel} \approx 2\pi/\lambda$ states, where λ is the island size. (2) We find that roughness at interfaces adjacent to the quantum well can cause lateral localization of wave functions, which increases with island size and depth. Lateral localization can result in the broadening and shifting of transmission resonances, and the introduction of preferential transmission paths. In structures with wide and tall islands, it is possible to find localization over "islands" as well as localization over "oceans." (3) The leading rough interface is the strongest off-resonance scatterer, while rough interfaces adjacent to quantum well are the strongest on-resonance scatterers. The trailing interface is the weakest scatterer. © *1996 American Vacuum Society.*

I. INTRODUCTION

While simple one-dimensional models[1,2] can explain the qualitative behavior of double barrier resonant tunneling diodes, they are less successful at reaching quantitative agreement with experimental results, particularly in predicting peak-to-valley current ratios. The discrepancy has been attributed to electron-phonon scattering, electron-electron interactions, band structure effects, impurity scattering, and interface roughness. In this work, we focus on the effect of interface roughness which, to some extent, can be manipulated by controlling the growth process. Several theoretical works have analyzed interface roughness effects previously. Liu and Coon[3] were among the first to approach the subject, modeling rough interfaces with regularly corrugated patterns. Chevoir and Vinter[4] and Rudberg[5] studied the topic with Fermi's golden rule. Leo and MacDonald[6] also treated interface roughness to leading order, but with a technique which preserved unitarity. These perturbative approaches[4–6] have been successful in dealing with small island sizes. More recently, Johansson[7] applied the coherent potential approximation which included multiple scattering events. One assumption frequently invoked by many of the previous works is that roughness scattering from the four interfaces in a double barrier structure can be treated independently. Several authors treated interface roughness only at a single interface, and there is some controversy as to which one of the four interfaces plays the more important role.[5,6] In this work we formulate the problem in real space with a tight-binding supercell model which allows for arbitrary spatial variations within laterally repeating supercells. Exact solutions are obtained in our calculations. Our model is well-suited for studying effects such as lateral spatial localization, providing a good complement to previous studies.

II. METHOD

In the planar supercell stack method (PSSM), we define the active region of a structure as a stack of N_z layers perpendicular to the z direction, with each layer containing a periodic array of rectangular planar supercells of $N_x \times N_y$ sites. A one-band nearest-neighbor tight-binding Hamiltonian is used to describe the potential over this volume of interest. Our model is formally equivalent to the one-band effective mass equation[8]

$$-\frac{\hbar^2}{2}\nabla\cdot\frac{1}{m(\mathbf{x})}\nabla\psi+V(\mathbf{x})\psi=E\psi, \tag{1}$$

discretized over a Cartesian grid, and subject to periodic boundary conditions (with supercell periodicity) in the x and y directions, and open boundary conditions in the z direction. Since the planar supercell model may be considered as a multiband model with $N_x \times N_y$ bands, transmission coefficients for structures described by the planar supercell stack can be determined exactly by the direct application of the multiband quantum transmitting boundary method.[9] Our

[a)]Electronic mail: dzt@phys.nthu.edu.tw
[b)]Electronic mail: tcm@ssdp.caltech.edu

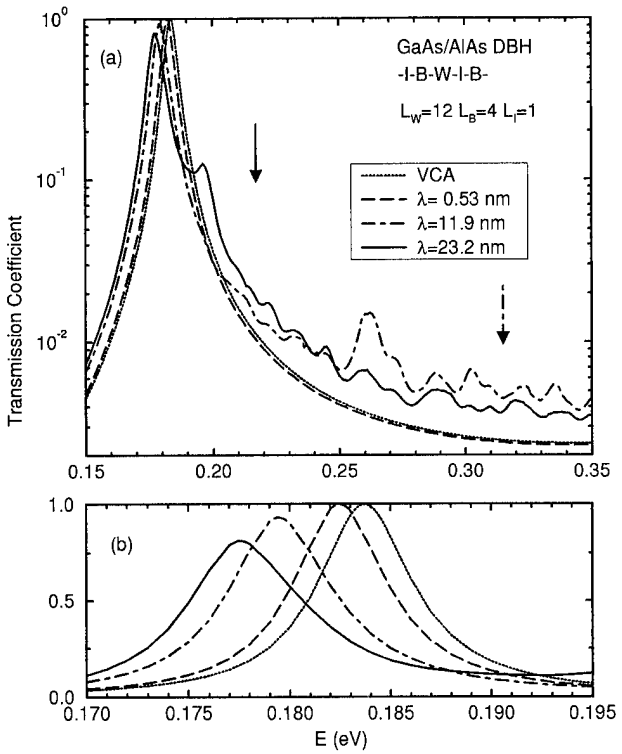


FIG. 1. Transmission coefficients for a set of GaAs/AlAs double barrier structures each containing two rough interfaces (between the leading electrode and the first barrier, and between the quantum well and the second barrier). The different structures have different island sizes, as indicated in the legend. Also included for comparison is a structure where each of the two interface layers is replaced by a VCA $Al_{0.5}Ga_{0.5}As$ alloy layer. 32×32 planar supercells are used for this calculation.

method requires accurate and efficient solutions of large sparse linear systems, which is achieved using the quasi-minimal residual method.[10]

III. RESULTS AND DISCUSSION

We examine how the transmission properties of GaAs/AlAs double barrier heterostructures are modified by the presence of interface roughness. We consider a set of structures with $L_W=12$ (monolayers) GaAs wells and $L_B=4$ AlAs barriers. For each GaAs-AlAs interface on the left (incident) side of a AlAs barriers, a 50% random coverage interface layer is inserted between the pure GaAs and AlAs layers. This configuration corresponds to having rough inverted interfaces and smooth normal interfaces, i.e., if we number the four interfaces 1 through 4, starting from the incident side, then interfaces 1 and 3 are rough, and interfaces 2 and 4 are smooth. A simulated annealing algorithm[11] can be used to generate random configurations of interfacial layers with different island sizes. Note that the actual island size, denoted by λ, depends on the supercell configuration as well as the lateral discretization size. In this article, 32×32 planar supercells are used in all calculations.

Figure 1 shows transmission spectra near the $n=1$ resonance for the structures discussed above. These structures have average island sizes (λ) of 5.3 Å, 119 Å, and 232 Å. For comparison we also show the result for a reference structure in which each of the two rough interfacial layers is replaced by a smooth virtual crystal approximation (VCA) $Al_{0.5}Ga_{0.5}As$ alloy layer. The incident plane waves are chosen to have no in-plane momentum ($\mathbf{k}_\parallel=\mathbf{0}$). Our results are similar to those we obtained in a calculation using a smaller (20×20) planar supercell.[12] We describe the main points below and refer readers to a more detailed discussion in Reference 12.

Figure 1(a) demonstrates the dependence of interface roughness scattering on island size. Evidently, if the length scale of roughness is small compared to de Broglie wavelength, as in the case of the λ=5.3 Å structure, the interface roughness potential is felt by the electron only in an averaged sense, and thus replacing the rough interfacial layer with a smooth VCA layer is a good approximation. For the structures with larger island sizes of λ=119 Å and λ=232 Å, the transmission spectra show a series of satellite peaks not found in the reference (VCA) spectrum. The presence of satellite peaks are due to interface roughness assisted resonant tunneling,[12] which can be pictured as follows: A nominally off-resonance state is elastically scattered by interface roughness into a state with a different $\mathbf{k}_\parallel$. If it so happens that the electron energy with respect to the scattered $\mathbf{k}_\parallel$ is on-resonance, then resonant tunneling occurs, leading to enhance transmission. Since coupling between different $\mathbf{k}_\parallel$ states is determined by the interface roughness configuration, we would expect to find preferential scattering into $\Delta k_\parallel \approx 2\pi/\lambda$ states.[12] Arrows in Figure 1(a) indicates positions where preferential scattering is expected to occur for the λ=119 Å and λ=232 Å structures based on the simple estimate using island sizes; the results seem to confirm our expectations.

In principle, interface roughness scattering can result in continuous range of $\Delta\mathbf{k}_\parallel$. In our model, an incident wave can only be scattered into states with $\Delta\mathbf{k}_\parallel$ equal to a reciprocal lattice vector $\mathbf{g}$ of the (finite) planar supercell. Thus, the appearance of discrete satellite peaks in our calculate transmission spectra is an artifact of the finite supercell size effect. This artifact can be mitigated by the use of larger supercells. The results of the present 32×32 supercell calculation shows closer spacing between satellite peaks than the previous 20×20 supercell calculation.[12] In the limit of very large supercell sizes, the spacings between the satellite peaks would become smaller than the resonance peak width, and the satellite peaks would then coalesce, giving a much smoother appearance.

Figure 1(b) shows that the main resonance peak broadens and shifts to lower energy as island size increases. We attribute this to lateral wave function localization. The rough interfacial layer between the quantum well and the second barrier introduces well width fluctuation. If the island sizes are sufficiently large, we could consider the quantum well as consisting of wide-well ($L_W=13$, in this example) and narrow-well ($L_W=12$) regions. If a quasibound level is localized (laterally) in the wide-well region, its resonance energy would be lowered. Evidently, the degree of lateral lo-

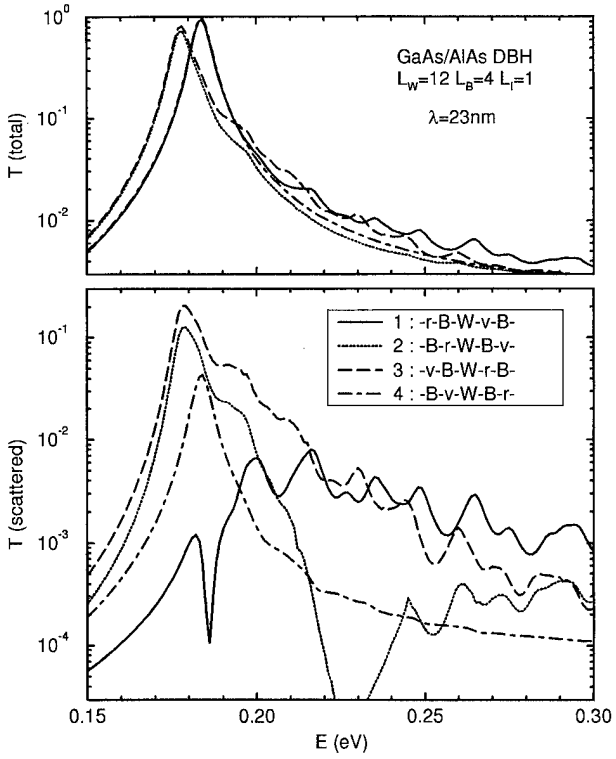


FIG. 2. Transmission coefficients for a set of GaAs/AlAs double barrier structures with interface roughness only in one of the four interfacial layers. The total transmission coefficient is shown in the top panel, while the transmission coefficient for only the scattered components (i.e., $\mathbf{k}_{\parallel}^{\text{incident}} \neq \mathbf{k}_{\parallel}^{\text{transmitted}}$) is shown in the bottom panel.

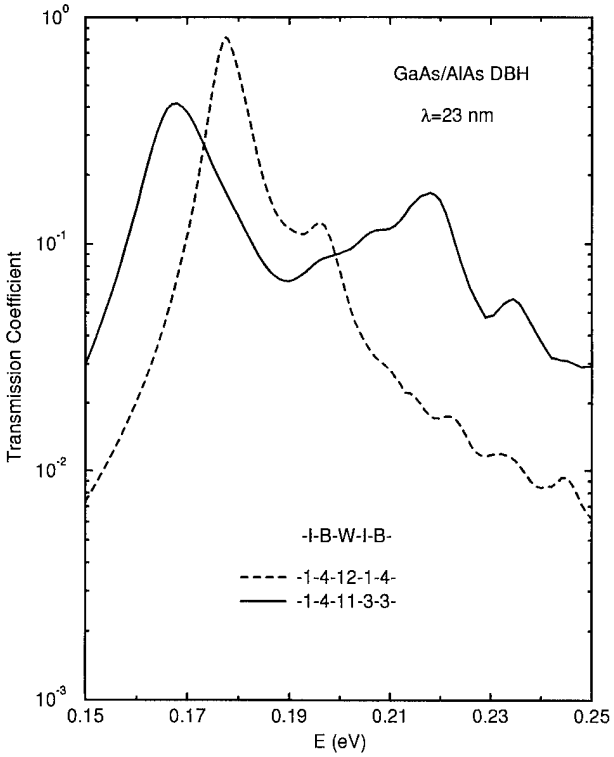


FIG. 3. Transmission coefficients for two GaAs/AlAs double barrier structures. The rough interfacial layer between the quantum well and the barrier for one structure (dashed line) is one monolayer thick, and for the other (solid line) is three monolayers.

calization increases with island size. Reference 12 examines wave function localization directly; it also demonstrates preferential transmission paths resulting from wave function localization. Note also that since the second barrier is thinner where the quantum well is wider, an electron localized in the wide-well/thin-barrier region can escape with more ease through the thinner barriers, leading to shorter quasibound state lifetime, or broader resonance, as seen in Figure 1(b).

In Figure 2 we explore the role of the individual interfaces. Strictly speaking, with wave function coherence across the entire device structure, encompassing all four interfaces, we cannot separate out the effects due to the different rough interfacial layers. Nevertheless, it is instructive to study the following. We examine four different configurations. Configuration 1 is similar to the configurations discussed earlier, except that the third interface layer (between the quantum well and the second barrier) is now replaced by a smooth VCA layer, leaving only the leading interface (1) rough. We use "-r-B-W-v-B-" to denote the layer sequence of this structure. Configurations 2, 3, and 4 are similarly constructed, each with one rough interface and one VCA interface; their layer sequences are shown in the inset of Figure 2. We note that the main resonance peak is down-shifted only in configurations 2 and 3, since lateral localization of wave function only occurs if one of the interfaces adjacent to the quantum well is rough. We also note that interface roughness scattering is the least important in configuration 4, where the rough interfacial layer is located away from the incident side. This is because unless the incident electron is on-resonance, and the transmission coefficient is reasonably large, the wave function is largely located on the incident side, with extremely small probability of penetrating across the device structure to the transmitted side to sense the roughness at the far interface. The same argument explains the fact that off-resonance scattering is the largest in configuration 1, where the leading interface is rough.

Figure 3 shows the transmission spectra of two structures with similar island sizes but different island heights. The first structure is identical to the $\lambda = 232$ Å structure from Figure 1. The second structure is similar to the first one, except that the random-alloy interfacial region between the quantum well and the second barrier is now expanded into three monolayers (at the expense of the surrounding well and barrier layers, so as to keep the effective well width the same). We note that the main resonance peak of the second structure (with taller islands) is lower than the first structure; its wave function also shows a higher degree of lateral confinement over the wide-well region. It is also interesting to note that the wave function for the transmission peak at 218 meV shows the opposite type of lateral confinement, i.e., the wave functions is localized over the islands themselves (the narrow-well regions) rather than over the "ocean." This indicates a splitting of the $n = 1$ quasibound state into orthogo-

nal states laterally localized in different regions of the quantum well.

IV. SUMMARY

We studied the effect of interface roughness on resonant tunneling in double barrier structures using a tight-binding supercell model. We find that interface roughness assisted scattering of off-resonance states into on-resonance state provides the dominant scattering channel. We showed that scattering strength is sensitive to the configuration of the roughness layer, and in particular, we find preferential scattering into $k_{\|} \approx 2\pi/\lambda$ states (λ = island size). We showed that lateral wave function localization increases with island size and depth, and is responsible for broadening and shifting of transmission resonances. Both types of lateral localization (over "islands" or over "ocean") are possible. We also explored the roles of the individual interfaces, and showed that the leading rough interface is the strongest off-resonance scatterer, while the rough interfaces adjacent to quantum well are the strongest on-resonance scatterers.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (ONR) under Grant No. N00014-92-J-1845. One of the authors (D.Z.T.) acknowledges partial support from Department of Physics, National Tsing Hua University (Hsinchu).

[1]R. Tsu and L. Esaki, Appl. Phys. Lett. **22**, 562 (1973).
[2]B. Ricco and M. Ya. Azbel, Phys. Rev. B **29**, 1970 (1983).
[3]H. C. Liu and D. D. Coon, J. Appl. Phys. **64**, 6785 (1988).
[4]F. Chevoir and B. Vinter, Appl. Phys. Lett. **55**, 1859 (1989); Phys. Rev. B **47**, 7260 (1993).
[5]B. G. R. Rudberg, Semicond. Sci. Technol. **5**, 600 (1990).
[6]J. Leo and A. H. MacDonald, Phys. Rev. Lett. **64**, 817 (1990); Phys. Rev. B **43**, 9763 (1991).
[7]P. Johansson, Phys. Rev. B **46**, 12 865 (1992).
[8]D. J. BenDaniel and C. B. Duke, Phys. Rev. **152**, 683 (1966).
[9]D. Z.-Y. Ting, E. T. Yu, and T. C. McGill, Phys. Rev. B **45**, 3583 (1992).
[10]R. W. Freund and N. M. Nachtigal, Numer. Math. **60**, 315 (1991); RIACS, NASA Ames Research Center Technical Report 90.51, Dec. 1989.
[11]N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[12]D. Z.-Y. Ting, S. K. Kirby, and T. C. McGill, Appl. Phys. Lett. **64**, 2004 (1994).

Electron transport at Au/InP interface with nanoscopic exclusions

S. Anand,[a] S-B. Carlsson, K. Deppert, L. Montelius, and L. Samuelson
Department of Solid State Physics, Lund University, Box 118, S-22100 Lund, Sweden

(Received 21 January 1996; accepted 26 April 1996)

We present an investigation of electron transport at the Au/InP metal semiconductor (MS) interface in the presence of nanoscopic barrier inhomogeneities. In particular, we focus on the transport regime wherein the low barrier inhomogeneous regions are pinched off by the depletion potential of the surrounding higher barrier region. To realize this, we have fabricated a composite MS structure comprising of a known density of nanometer sized Ag aerosol particles on InP overlayed by a uniform Au film. Temperature dependent current–voltage ($I-V$) measurements were performed to characterize the electron transport at the composite MS interface. The experimental observations are explained using a simple model for the MS junction current in which the Ag/InP regions are assumed to be identical and noninteracting. Our results clearly demonstrate that the electron transport at the MS interface is significantly affected by low barrier regions even in the pinch-off regime. In addition, the influence of the particle density and the Ag/InP barrier heights on the diode characteristics is also investigated. It is suggested that Schottky barrier inhomogeneities could be the main source of the usually observed larger-than-unity ideality factors in diodes. Furthermore, our results indicate possibilities of using such composites to explore the physics and applications of nanoinjecting contacts. © *1996 American Vacuum Society.*

I. INTRODUCTION

The formation of Schottky barrier at metal semiconductor (MS) interface has been a subject of extensive investigations for several years.[1] Pinning of the Fermi level by electronic states at the MS interface has been frequently invoked to explain the formation of Schottky barriers.[2,3] More recently, experimental investigations of epitaxial MS interfaces suggest that the barrier heights depend on the structure of the MS interface.[4] It is therefore plausible, though not certain, that nonepitaxial MS interfaces could be inhomogeneous resulting in a distribution of barrier heights at the interface. Some investigations have recognized the presence of inhomogeneities and their effect on diode characteristics.[5–7] Theoretical modeling of the effect of such inhomogeneities on the electron transport appears to be rather successful in accounting for a wide variety of properties including temperature dependence of electron transport properties at the interface as well as the variation of the ideality factor of the Schottky barriers.[8] However, it is very difficult to verify the above model in nonepitaxial MS interfaces since the exact nature and extent of the inhomogeneities is usually unknown. In cases where the inhomogeneous region is not pinched off by the surrounding barrier, a parallel conduction model wherein the current is assumed to be a sum of currents flowing in all the individual patches can be used. This model has been employed with reasonable success to explain some of the observed anomalous characteristics of diodes.[9,10] However, as pointed out in Ref. 8, the parallel conduction model is in serious error if the barrier height varies spatially on a scale less than or comparable to the depletion width. MS interfaces characterized by the above conditions present an interesting situation and electron transport at such MS junctions is relatively unexplored.[8]

In this work we address the above questions by using Au/Ag composite barriers formed on weakly n-doped epitaxial InP. The main purpose of using a composite structure is to deliberately introduce inhomogeneities such that the spatial variation of the barrier height is on a scale comparable to the depletion width. The composites were formed by first depositing a known density of Ag aerosol particles (radius≈17 nm) on InP surface, and subsequently Au is evaporated onto the sample. This results in nanoscopic and well controlled barrier inhomogeneities at the MS interface produced by the embedded Ag particles. Temperature dependent current–voltage ($I-V$) measurements under forward bias were performed to investigate the effect of the inhomogeneities introduced by the Ag particles on the electron transport at the composite MS interface. The experimental results are explained using a simple model for the junction current wherein the Ag/InP regions are assumed to be identical and noninteracting. In this work, we demonstrate that the electron transport at the MS interface is significantly affected by the low barrier regions even in the pinch-off regime.

II. EXPERIMENT

The inhomogeneous Schottky barriers were produced by forming a composite MS structure on 5-μm-thick nominally undoped InP. The InP samples were grown by metal organic chemical vapor deposition (MOCVD) on n^+ InP substrates. The composite MS interface was realized by first depositing a known density of nanometer sized Ag aerosol particles on InP. Subsequently, 340-μm-diam, 100-nm-thick Au dots were made by evaporation of Au through a metal contact mask. A schematic sketch of the composite diode structure is

[a]Electronic mail: anand.srinivasan@ftf.lth.se

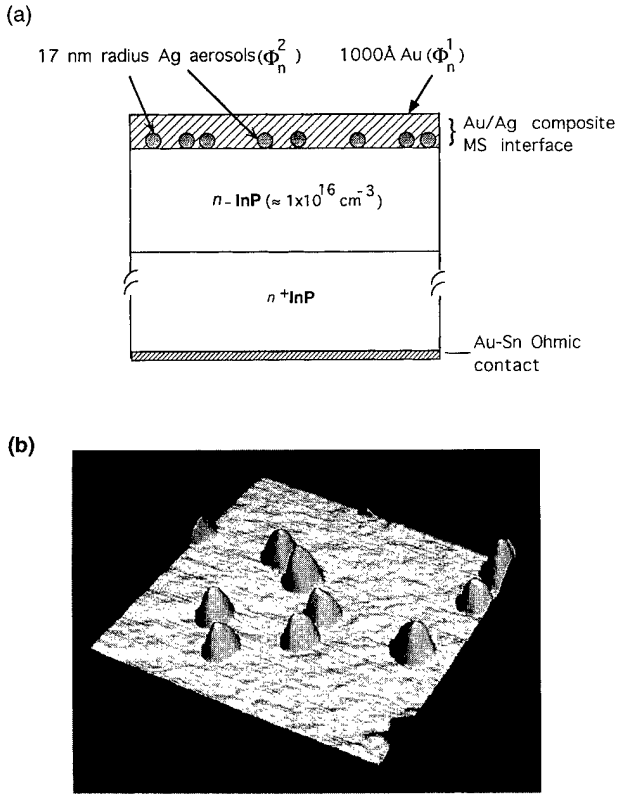


FIG. 1. (a) A schematic diagram of the composite MS diode structure showing the embedded Ag aerosol particles. (b) An (976 nm)^2 AFM image of the Ag aerosol particles for a representative sample with a nominal particle density of about 10^9 cm^{-2}. The image was obtained using a Topometrix Explorer 1000 operated in ambient air. The Ag particles are uniform and typically 17 nm in radius.

shown in Fig. 1(a). Contact mode atomic force microscopy (AFM) image of the Ag particles is shown in Fig. 1(b). Prior to deposition of the Ag aerosol particles, the InP samples were cleaned in standard organic solvents, etched in HF:H_2O (1:10) for 1 min, and rinsed thoroughly in deionized water. After deposition of the aerosols, without additional processing, the samples were loaded into the evaporation chamber keeping the time of exposure of the sample surface to ambient air as short as possible. The physical dimensions of the deposited Ag aerosols (around 17 nm in radius) and the doping of InP ($\approx 1\times10^{16}$ cm^{-3}) are such that the nanoscopic inhomogeneities produced by Ag/InP regions are pinched-off by the depletion potential of the surrounding Au/InP regions. In the above context, a region at the MS interface is said to be pinched-off if the majority carriers (originating from outside the space charge region) reach the MS interface by surmounting a potential barrier that is higher than the band edge position at the region.[8] Due to this peculiar physical situation, we refer to the nanoscopic inhomogeneities as *nanoscopic exclusions*.

Two sets of composite samples were fabricated, one with a Ag aerosol particle density of 5×10^9 cm^{-2}, and the other with a much lower density of 10^8 cm^{-2}. Further, prior to Au evaporation, some of the samples were also subjected to a moderate heat treatment of 350 °C for 5 min in nitrogen atmosphere to reduce the barrier height of the Ag/InP regions. Control samples consisting only of Au dots were also made for comparison. For a meaningful comparison, the InP surface preparation conditions for the control samples were kept as identical as possible to that of the composites. The back ohmic contact for all the samples was made by alloying Au–Sn to the n^+ InP substrates. The diodes are characterized by $I-V$ measurements in the temperature range 290–77 K. The $I-V$ measurements were performed using a Keithley 236 source–voltage current meter configured by LabVEIW software.

The aerosol generator set up, which is described in detail elsewhere,[11] consists of a furnace tube with an Ag containing ceramic boat. This first furnace for the aerosol generation was kept at a temperature of 1240 °C, resulting in a sufficiently high number of particles with diameters below 100 nm. After a size selection procedure, Ag particles were sent through a second tube furnace. This furnace was operated at 500 °C, leading to a reshaping of the aerosol particles. After this sintering step, the particles exhibit an almost spherical appearance. Then, to keep the size distribution of the particles narrow after reshaping in the second furnace, a second size selection was carried out. The resulting Ag particles were deposited onto InP surface in a deposition chamber by a perpendicular electric field. The whole process was run under nitrogen carrier flow to avoid oxidation of the particles. The physical density of the deposited particles can be varied by choosing appropriate deposition times. Contact mode AFM investigations of the Ag aerosol particles show that the particles are quite uniform,[12] and are about 17 nm in radius [Fig. 1(b)].

III. RESULTS AND DISCUSSION

Variation of the Schottky barrier height (SBH) at the MS interface implies the variation of the potential from region to region. The composite MS structure described before can be thought of as an inhomogeneous Schottky barrier with a certain density of low barrier regions (due to the Ag particles) in an otherwise uniform high SBH (Au/InP). The Ag particle–InP interface region can be approximated to a circular patch of some effective radius and due to the almost spherical shape of the particles the effective patch radius will be less than or at best equal to the geometrical radius of the particle. For a doping of 1×10^{16} cm^{-3}, the depletion width due to Au/InP (barrier height of 0.5 V) is of the order of 250 nm which is significantly larger than the radius of the Ag particles. Under these physical conditions, the potential $V(0,z)$ along the z axis and at the center of the low barrier Ag/InP circular patch of effective radius R_{Ag} can be expressed as[8]

$$V(0,z)=V_{\mathrm{bb}}\left[1-\frac{z}{W}\right]^2+V_n+V_a-\Delta\left[1-\frac{z}{(z^2+R_{\mathrm{Ag}}^2)^{1/2}}\right], \tag{1}$$

where $V_{\mathrm{bb}}=\Phi_n^1-V_n-V_a$ is the band bending due to the MS junction with a uniform SBH of Φ_n^1, W is the depletion width defined as $(2\epsilon_s V_{\mathrm{bb}}/qN_D)^{1/2}$ where ϵ_s is the dielectric constant and q the electronic charge, V_n is the difference between the

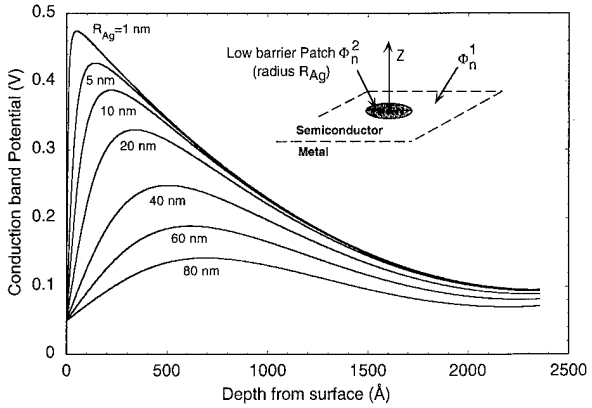


FIG. 2. Conduction band potentials along the z axis for different patch radii, calculated using Eq. (1) for zero bias. The insert shows a schematic diagram of an isolated low barrier circular patch indicating the relevant coordinates. The Ag/InP regions are approximated to a circular patch of effective radius R_{Ag}.

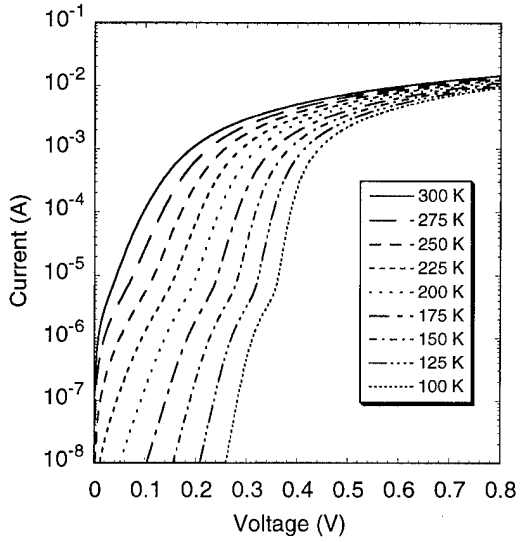


FIG. 3. Simulated current–voltage characteristics at the inhomogeneous MS (composite) interface, calculated using Eq. (2). The values of the parameters used are $A=9\times10^{-4}$ cm^2, $\Phi_n^1=0.5$ V, patch radius=7 nm, $\Delta=0.45$ V, $N_D=1\times10^{16}$ cm^{-3}, and patch density $d=5\times10^9$ cm^{-2}. Effect of the series resistances are also included in the calculations.

Fermi level and the conduction band minimum, V_a is the applied bias, and $\Delta=\Phi_n^1-\Phi_n^2$ is the difference in the barrier potential between the uniform enveloping barrier and the patch (Φ_n^2). Figure 2 shows the simulated conduction band potential using Eq. (1). The radius of the patch is varied to illustrate the pinch-off situation. Clearly, the smaller the radius of the low barrier patch, the easier it is to obtain pinch-off conditions. In addition, the results of Fig. 2 also point to interesting possibilities for obtaining injecting nano-ohmic contacts by suitable choice of Φ_n^2, R_{Ag}, and sample structure in terms of doping, etc. Detailed simulations and discussions presented in Ref. 8 demonstrate that the dependence of the potential on the applied bias [Eq. (1)] significantly affects the transport characteristics of inhomogeneous MS interfaces. Turning to the evaluation of the total current flow in such MS interfaces, the exact expression for the current is rather complicated and depends on the detailed nature of the inhomogenieties and their distributions.[8] However, under certain reasonable approximations, analytical expressions can be obtained for the total current at the MS junction. In particular, for a MS interface containing a low density of nearly identical low barrier circular patches, the total forward bias current (I_{tot}) can be expressed as[8]

$$I_{tot}=AA^*T^2\exp(-\beta\Phi_n^1)[\exp(\beta V_a)-1]\times\left[1+\frac{4\pi d\eta^{2/3}\gamma}{9\beta V_{bb}^{2/3}}\exp\left\{\frac{\beta\gamma V_{bb}^{1/3}}{\eta^{1/3}}\right\}\right], \quad (2)$$

where A is the area of the diode, A^* is the Richardson constant, T is the temperature, d is the areal density of patches, $\beta=q/kT$ is the Boltzman factor, $\eta=\epsilon_s/(qN_D)$, and $\gamma=3(\Delta R_{Ag}^2/4)^{1/3}$. The parameters γ and d are characteristic of the low barrier patches.

Within some reasonable limits, our composite MS structure can be assumed to correspond to the above situation, and the model for the electron transport [Eq. (2)] can be used to qualitatively illustrate the $I-V$ characteristics of the composite MS diodes. Figure 3 shows the simulated forward bias $I-V$ characteristics using Eq. (2) for our composite diode for some representative parameters. At higher temperatures, the electron transport is dominated by thermionic emission over the Au/InP barrier. However, at low temperatures, at low biases, the junction current is dominated by the Ag/InP regions. In fact, one also sees a development of anomalous plateaulike sections in the $I-V$ curves which are due to an ohmic effect. Since the current densities near the patches may be much higher than the average current density, the current through the patches tends to be limited by series resistance effects, resulting in plateau-like features in the $I-V$ curves. If the pinch-off is very effective (a very small patch radius), we find that variation of the local SBH at the Ag/InP interface does not have a very significant effect on the appearance of the $I-V$ curves. However, in simulations for larger patch radii, variation of Δ results in strong changes in the $I-V$ curves resulting in pronounced plateau-like features. To investigate the influence of the patch (Ag particle) density on the transport properties, we varied the patch density from 5×10^9 cm^{-2} down to 10^8 cm^{-2} while keeping the other parameters same as was used for the $I-V$ curves of Fig. 3. The $I-V$ curves (not shown here) exhibit a systematic dependence with particle density, especially at low temperatures. The higher their density, the larger their influence on the transport properties. We also note here that in simulations using larger effective patch radii, we observe that the influence of the patch density on the $I-V$ curves is even more pronounced. Below, we discuss our experimental results and show that the observed characteristics of the composite diodes are qualitatively consistent with the simulations presented before.

Figure 4 shows the measured $I-V$ characteristics at selected temperatures for a representative Au/InP diode (con-

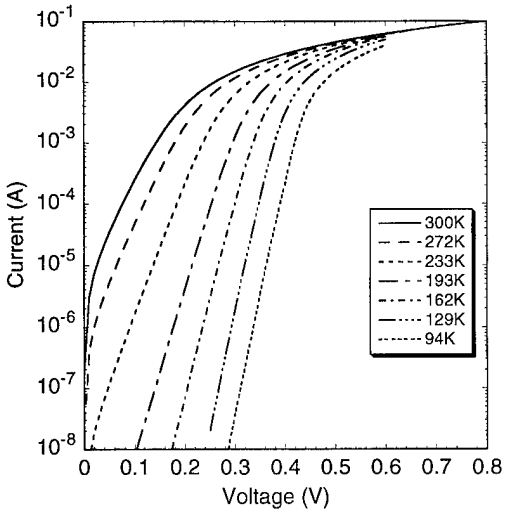


FIG. 4. Measured $I-V$ curves for a representative Au/InP control sample at selected temperatures.

trol sample). Several diodes were measured and their characteristics are similar to those shown in Fig. 4. The measured $I-V$ curves show the normal semilogarithmic behavior, and there is no perceptible departure in linearity with temperature. The barrier height obtained from an activation plot ($I-V-T$) is 0.5 V and is in good agreement with reported values in literature.[13,14] As discussed in the preceding paragraph, in the composite samples, depending on parameters such as particle density and Δ, one can expect to see a variety of behaviors in the temperature dependent $I-V$ curves. The measured $I-V$ characteristics for a representative composite diode with Ag particle density of 5×10^9 cm^{-2} are shown in Fig. 5. At high temperatures the $I-V$ curves are similar to the control samples. However, at about 120 K, there is a clear plateaulike structure at a bias of 0.3 V, which

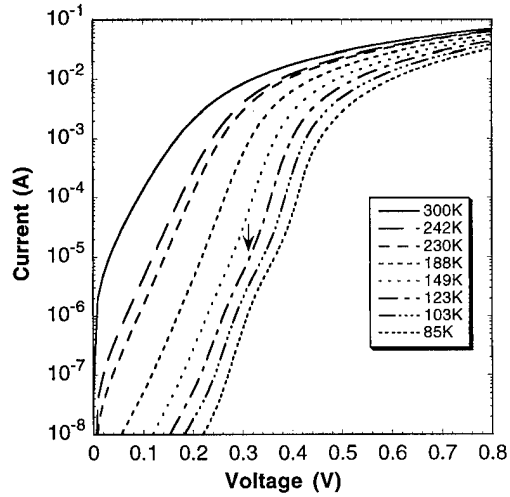


FIG. 5. Typical set of measured $I-V$ curves for the Au/Ag/InP composite samples at different temperatures. The Ag particles are as-deposited and their density is 5×10^9 cm^{-2}. The departure in the semilogarithmic behavior of the $I-V$ characteristics is indicated by the arrow.

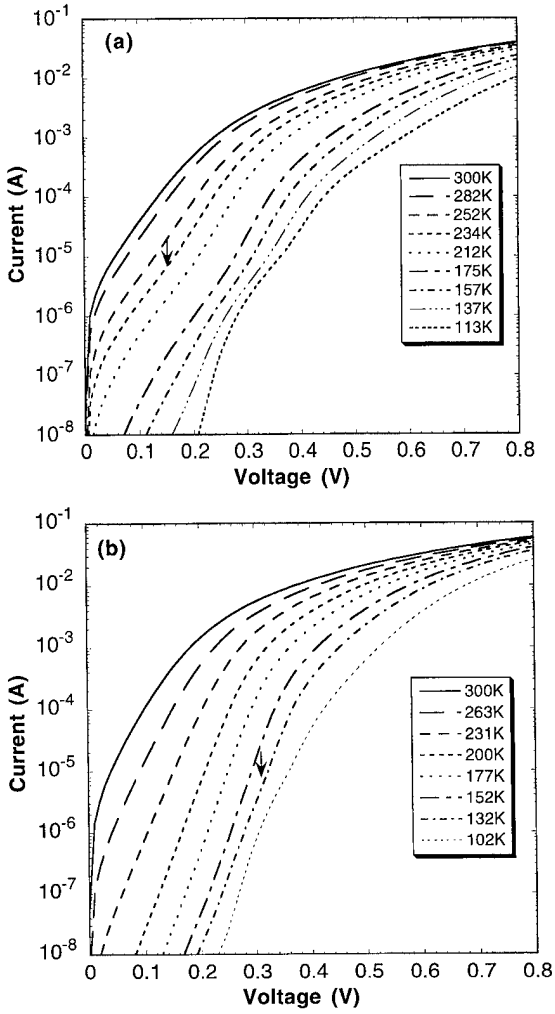


FIG. 6. (a) Typical set of measured $I-V$ curves for the Au/Ag/InP composite samples at different selected temperatures. The density of the Ag particles is 5×10^9 cm^{-2} and are alloyed to the InP surface. (b) $I-V$ scans at selected temperatures for a low particle density (10^8 cm^{-2}) composite sample. Here, also, the Ag particles were alloyed to the InP surface. The arrows indicate the departure in the semilogarithmic behavior of the $I-V$ characteristics.

persists at lower temperatures. Since this feature is clearly absent in the control samples, we can attribute this signature to electron transport at the low barrier Ag/InP regions. Further proof that this is so is obtained by looking at the characteristics of composites wherein the deposited Ag particles were alloyed on the InP surface to further reduce their barriers. The $I-V$ characteristics of these diodes are shown in Fig. 6. From Eq. (2), we can see that the currents are sensitive to the value of Δ. For a given particle density, if Δ is made large, the $I-V$ curves can be expected to show the influence of the low barrier patches at a higher temperature. And, for a given Δ, if the density is significantly reduced, the junction current will be dominated by the low barrier patches only at relatively lower temperatures [Eq. (2)]. That this is indeed the case is seen in Fig. 6. For the high density sample [Fig. 6(a)], the ohmic effect starts to show up at as high a temperature as 230 K while that in the low density sample appears at a lower temperature of about 100 K [Fig. 6(b)].

Preliminary simulations using Eq. (2) of the experimental $I-V$ curves (Figs. 5 and 6) indicate that the effective patch radius is around 7 nm, and the barrier heights are about 0.3 and 0.05 V for the as-deposited and alloyed particles, respectively.

All of the aforementioned results are qualitatively consistent with Eq. (2) and demonstrate the influence of nanoscopic SBH inhomogeneities (nanoscopic exclusions) on the electron transport. In this work, by deliberately introducing the inhomogeneities with good size control, we have also been able to demonstrate the influence of parameters such as the particle density and local barrier heights on transport properties at the MS junction. Further experiments exploring other possible composite structures using different particle sizes and semiconductor structures are underway.

Finally, we comment briefly on the ideality factors of the diodes. The ideality factor (n) for a Schottky diode is an emphirical quantity and is defined as $n=\beta\partial V_a/\partial[\ln(I_{\text{tot}})]$ for $V_a>3/\beta$. An ideal Schottky diode is characterized by an unity ideality factor. However, most practical diodes usually have ideality factors in excess of unity. Physical mechanisms such as image force lowering,[15] generation-recombination, interface states,[16,17] and SBH inhomogeneities[8] have been discussed as possible sources for the observed larger-than-unity ideality factor in diodes. For a meaningful comparison of the ideality factors for the different diodes used in this work, the ideality factors were determined in the temperature range from 300 to 250 K. In this temperature range the $I-V$ curves for all the diodes are semilogarithmic (see Figs. 4–6). The control samples (Au/InP) show nearly ideal behavior and the ideality factors are in the range 1.02–1.05. However, for all the composite samples, we find that it is invariably larger than that of the control samples. For the low density sample with alloyed Ag particles, the ideality factors were typically 1.08–1.11. Somewhat larger values of 1.12–1.15 was found for the high density samples with as-deposited Ag particles. The worst case is for the composites formed from high density alloyed Ag particles. Here, both Δ and the density are large, thereby significantly increasing the role of the Ag/InP regions in electron transport across the MS interface. In these samples, the ideality factors were typically 1.4–1.55 which is significantly larger that what is determined for the control samples. These observations suggest that the most significant mechanism for larger-than-unity ideality factors usually observed in diodes is largely due to the fact that most of MS interfaces are in general inhomogeneous. Similar observations have been made in epitaxial MS interfaces which were deliberately made inhomogeneous.[4]

IV. CONCLUSIONS

In summary, we have investigated the problem of electron transport at the Au/InP MS interface in the presence of nanoscopic barrier inhomogeneities. The investigated transport regime pertains to an interesting situation wherein the inhomogeneous regions are pinched off by the depletion potential of the surrounding higher barrier region. To study this problem, we have used composite MS structure comprising of a known density of nanometer sized uniform Ag aerosol particles on InP overlayed by a uniform Au film. Temperature dependent $I-V$ measurements have been performed to characterize the electron transport at the composite MS interface. The experimental results are explained using a simple model for the current across the inhomogeneous MS junction. All the Ag/InP regions have been assumed to be nearly identical and noninteracting. Our results clearly demonstrate that the electron transport at the MS interface is significantly affected by low barrier regions even in the pinch-off regime. It is suggested that the commonly observed larger-than-unity ideality factors in MS diodes is largely due to an inhomogeneous MS interface. Furthermore, our results indicate possibilities of using such composite MS structures to explore the physics and applications of nanoinjecting MS contacts.

ACKNOWLEDGMENTS

The work was performed within the Nanometer Structure Consortium in Lund and was supported by grants from the Swedish National Board of Technical and Industrial Development, the Swedish Natural Science Research Council, and the Swedish Technical Science Research Council. The authors are thankful to I. Lindau for fruitful discussions and to W. Seifert and N. Carlsson for providing the InP samples used in this work.

[1]E. H. Rhoderick and R. H. Williams, *Metal–Semiconductor Contacts* (Clarendon, Oxford, 1988).
[2]J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[3]W. E. Spicer, I. Lindau, P. Skeath, C. Y. Su, and P. W. Chye, Phys. Rev. Lett. **44**, 420 (1980).
[4]R. T. Tung, A. F. J. Levi, J. P. Sullivan, and F. Schrey, Phys. Rev. Lett. **66**, 72 (1991).
[5]O. Engström, H. Petterson, and B. Sernelius, Phys. Status. Solidi. A **95**, 691 (1986).
[6]W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988).
[7]A. Tanabe, K. Konuma, N. Teranishi, S. Tohyama, and K. Masubuchi, J. Appl. Phys. **69**, 850 (1991).
[8]R. T. Tung, Phys. Rev. B **45**, 13509 (1992).
[9]I. Ohdomari and K. N. Tu, J. Appl. Phys. **51**, 3735 (1980).
[10]M. V. Schneider, A. Y. Cho, E. Kollberg, and H. Zirath, Appl. Phys. Lett. **43**, 558 (1983).
[11]K. Deppert, I. Maximov, L. Samuelson, H.-C. Hansson, and A. Wiedensohler, Appl. Phys. Lett. **64**, 3293 (1994).
[12]T. Junno, S. Anand, K. Deppert, L. Montelius, and L. Samuelson, Appl. Phys. Lett. **66**, 3295 (1995).
[13]L. J. Brillson, C. F. Brucker, A. D. Katnani, N. G. Stoffel, and G. Margaritondo, Appl. Phys. Lett. **38**, 784 (1981).
[14]E. Hökelek and G. Y. Robinson, J. Appl. Phys. **54**, 5199 (1983).
[15]S. M. Sze, C. R. Crowell, and D. Kahng, J. Appl. Phys. **35**, 2534 (1964).
[16]H. C. Card and E. H. Rhoderick, J. Phys. D **4**, 1589 (1971).
[17]E. F. Broom, Solid-State Electron. **14**, 1087 (1971).

Relaxation of the step profile for different microscopic mechanisms

Da-Jiang Liu, Elain S. Fu, and M. D. Johnson
Department of Physics, University of Maryland, College Park, Maryland 20742

John D. Weeks
Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

Ellen D. Williams[a)]
Department of Physics and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

(Received 13 February 1996; accepted 24 April 1996)

Theoretical and experimental studies of the rate of decay of metastable structures are compared quantitatively. The effect of decay mechanism, size, and periodicity of the structure on the rate of decay is evaluated within both a coarse-grained step-based model and a continuum model. For high-amplitude structures, the decay scales with size (N) and time as $(t/N^{\alpha})^{-\beta}$. The exponents α and β depend on the mass transport mechanism. The size scaling is $\alpha=4$ for locally conserved diffusive flux and $\alpha=2$ for locally nonconserved flux. The time scaling exponent is $\beta=1/5$ for diffusive limited mass transport and $\beta=1/4$ for step attachment limited mass transport. Experiments were performed on metastable structures of controlled sizes 3–5 nm in height, prepared by direct current heating on Si(111). Quantitative agreement with theoretical predictions of both scaling ($\alpha=4.3\pm0.5$, $\beta=0.2-0.3$) and absolute rate of decay were obtained. © *1996 American Vacuum Society.*

I. INTRODUCTION

As the size scale of artificially structured devices shrinks into the submicrometer regime, it becomes increasingly important to understand their structural properties on an atomic level. In this work, we present the results of a combined theoretical and experimental study designed to test our understanding of the stability of nanoscale structures. This work was based on an understanding of surface mass transport in terms of the addition and removal of material from steps on surfaces. In the following sections we present first a description of the theoretical formalism used to describe the decay of nanostructures. We then present results on the fabrication and decay of structures of a similar geometry on a Si(111) surface.

II. THERMODYNAMICS OF THE STEP MODEL

Perturbations of a thermodynamically stable surface will eventually relax back toward equilibrium. It is natural to assume that at least for small disturbances the relaxation rate is proportional to the difference in free energy between the perturbed and equilibrium surface. The simplest example is a system above the roughening temperature with isotropic surface tension where the surface free energy is proportional to the total area. In several classic papers, Mullins[1–3] derived the equations of motion for one-dimensional (1D) surface profiles under different mass transport mechanisms. Because of its simplicity, Mullins' isotropic continuum model has been frequently used to compare with experiment, and the relevant physical concepts such as surface tension and surface diffusion rates are often determined from experimental data, even when the underlying physical situation is rather different from that considered by Mullins. One such example is that of surfaces below the roughening transition temperature. In this case, there is a singular orientation at which the orientational dependence of surface free energy per unit area has a cusp. This causes problems in a continuum treatment. Moreover, the system in this case is certainly not isotropic. In this article, we further develop some models proposed previously by several authors[4–6] for the relaxation of vicinal surfaces below the roughening transition. The common features of these models are the use of steps as the fundamental entities and the change from partial differential equations for a continuum model to a set of coupled ordinary differential equations describing step positions.

For vicinal surfaces below the roughening temperature, the terrace-step-kink (TSK) model provides an adequate description. In this model, there is a creation energy for an isolated step and steps have repulsive interactions with each other due to the no-crossing rule and an elastic or dipole interaction if present.[7–9] When steps basically remain straight a 1D model may be sufficient. This can be conveniently described using the reduced free-energy density, defined as the surface free energy per unit area projected onto the low index facet plane. For a small miscut angle (and therefore low step density), the reduced free energy density assumes the form[7,8]

$$f(s)=f^{0}+\frac{\beta}{h}s+gs^{3}, \tag{1}$$

where $s=|\tan\phi|$ is the slope of the surface, β the creation energy of an isolated step, h the height of the step, and g the step interaction parameter which is a function of the step stiffness, the temperature, and any energetic interaction be-

a)Electronic mail: Williams@surface.umd.edu

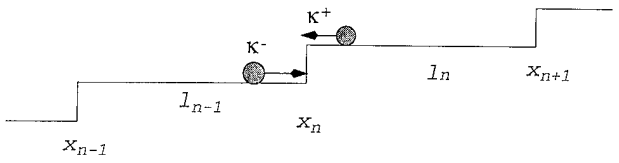


FIG. 1. Illustration of the labeling of steps and terraces and kinetics coefficients.

tween the steps.[10] We make a mean field approximation by assuming that the effective Hamiltonian or the free-energy functional for a nonuniform system has the form of Eq. (1) and depends only on the local surface slope. The simplest way to proceed is to define the local slope on each terrace as h/l_n, where $l_n=x_{n+1}-x_n$ is the width of the nth terrace. See Fig. 1 for the labeling of steps and terraces. Hence we can write

$$H=L_y\sum_n f(l_n)l_n=L_y\sum_n F(l_n), \tag{2}$$

where $F(l)=f(l)l=f^0l+\beta+gh^3/l^2$. Following Rettori and Villain,[5] we define the chemical potential μ_n at a step edge as the free energy change arising when the average step position moves due to attachment of an adatom. Therefore, from Eq. (2), we find

$$\mu_n=2ga^2h^3(1/l_{n-1}^3-1/l_n^3)+\mu_0 \tag{3}$$

by assuming a square lattice for the adatoms, where a is the lattice constant, and μ_0 is the chemical potential of the atom in the solid. Note that the step creation energy β drops out of the expression for the step chemical potential. This is because Eq. (2) is based on the assumption that the number of steps remains constant and thus only makes sense when the steps are in the same direction (i.e., there is no annihilation of steps of opposite signs).

Other cases are more complicated and introduce new physical considerations. For example, at the tops or bottoms of a sinusoidal profile where the average direction coincides with the facet direction (and hence both *up* and *down* steps are present), two-dimensional (2D) islands usually form on terraces. As the average positions of the step and antistep move closer, they have a greater chance to annihilate one another, therefore reducing the total length of the steps. Using h/l as the step density is certainly not correct for this case. Ozdemir and Zangwill[6] studied a model for this case which ignored the interaction term g/l^3 between steps and antisteps but still did not explicitly take into account any changes due to the creation energy of the steps. That might explain why they did not see the formation of true facets in their model. A more general 2D step model that explicitly takes into account this creation energy is currently being investigated.[11] However, here we concentrate both theoretically and experimentally on the simpler, but still challenging, problem of profiles consisting of steps of the same sign, and therefore Eq. (2) is sufficient. Except for the orientation along the facet direction for long wavelength perturbations at small amplitude, we expect that different discretization methods should yield results similar to those of the continuum model.

III. KINETICS OF THE STEP MODEL

Having defined the step chemical potential, we can proceed to study the relaxation process arising from different kinetic assumptions in the step model. Many different kinetics have been studied previously by different authors, but no standard terminology has emerged. In this article, we divide the possible kinetic processes into two families, depending on whether or not the adatoms on each terrace obey a local conservation condition.

A. Kinetics with local adatom conservation: Attachment/detachment limited and diffusion limited cases

One standard treatment going back to the Burton–Cabrera–Frank (BCF) model[12] assumes that the two neighboring step edges serve as boundary conditions for a diffusion equation describing adatom diffusion on a given terrace. The mass transport on each terrace obeys a local mass conservation law. Originally, it was assumed that steps acted as perfect sinks for adatoms. A more general treatment assuming a finite attachment probability of the adatoms to the step edges was given by Ozdemir and Zangwill.[6] This problem is usually solved under the quasistatic approximation. We discuss here two special limiting cases: (a) the attachment/detachment rate limited case corresponding to a very small attachment rate at the step edges compared with the adatom diffusion rate and (b) a diffusion limited mechanism (originally considered by BCF) corresponding to fast attachment rates and relatively slow diffusion.

For the attachment/detachment limited process, the rapid diffusion ensures that adatoms have a well-defined and spatially uniform chemical potential on each particular terrace. The first order kinetics theory predicts that

$$\frac{\partial x_n}{\partial t}=\kappa^+(\mu_n-\mu_n^t)+\kappa^-(\mu_n-\mu_{n-1}^t), \tag{4}$$

where μ_n^t is the chemical potential of an adatom on the nth terrace, and $\kappa^{+(-)}$ is the kinetic coefficient for adatom attachment from the terrace above (below) the step. In the quasistatic approximation, time derivatives in the adatom concentration field are neglected, so we require that the net flux onto any terrace from the two neighboring steps is zero. This fixes the chemical potential on the terraces as $\mu_n^t=(\kappa^+\mu_n+\kappa^-\mu_{n+1})/(\kappa^++\kappa^-)$. Thus from Eq. (3) and Eq. (4), we find the equation of the motion of steps under an attachment/detachment limited mechanism:[4]

$$\begin{aligned}\frac{\partial x_n}{\partial t}&=\frac{\kappa^+\kappa^-}{\kappa^++\kappa^-}(2\mu_n-\mu_{n-1}-\mu_{n+1})\\&=\frac{2gh^3a^2\kappa^+\kappa^-}{\kappa^++\kappa^-}\left(\frac{-1}{l_{n-2}^3}+\frac{3}{l_{n-1}^3}-\frac{3}{l_n^3}+\frac{1}{l_{n+1}^3}\right).\end{aligned} \tag{5}$$

In the opposite diffusion limited case, rapid adatom exchange at the step edges causes the adatom concentration to

assume the local equilibrium value $c_n^{eq} = c_0^{eq} \exp[(\mu_n - \mu_0)/k_BT] \approx c_0^{eq}(1 + (\mu_n - \mu_0)/k_BT)$, where c_0^{eq} and μ_0 are the equilibrium values for the completely relaxed surface. In the absence of evaporation and deposition, and under the quasistatic approximation, the diffusion field describing the adatom concentration on each terrace is a linear function of the terrace width. Thus one can relate the adatom flux or current to the known values c_n^{eq} at the steps. From local mass conservation we have[5]

$$\frac{\partial x}{\partial t} = \frac{D_s c_0^{eq} a^2}{k_B T}\left(\frac{\mu_n - \mu_{n-1}}{l_{n-1}} + \frac{\mu_n - \mu_{n+1}}{l_n}\right) \tag{6}$$

as the equation of the step motion under the diffusion limited mechanism.

B. Kinetics without local conservation: Attachment/detachment limited case

A second general class of mass transport mechanisms arises when adatoms can hop over the step edges quite easily. In such cases, local conservation is not important since contributions from many other terraces influence the adatom density on a given terrace. Moreover, the adatom density cannot have a singularity at the step edges. This picture is physically relevant to the attachment/detachment limited case with rapid and efficient surface diffusion. Because adatoms can freely diffuse to any place on the surface, their concentration and, therefore, the adatom chemical potential μ_n^t assumes a constant value for all terraces. In the absence of evaporation and deposition $\mu_n^t = \mu_0$. Therefore, from Eq. (4) we have[4]

$$\frac{\partial x_n}{\partial t} = 2ga^2h^3(\kappa^+ + \kappa^-)\left(\frac{1}{l_{n-1}^3} - \frac{1}{l_n^3}\right) \tag{7}$$

as the equation of step motion under the attachment/detachment limited mechanism without local conservation. In the literature, the term attachment/detachment limited does not consistently refer to cases either with or without local conservation, so it is necessary to specify the conservation condition to avoid confusion. Note that the formula for this case is the same as for the evaporation–condensation mechanism, where the constant chemical potential is determined by that of the vapor phase.

IV. DECAY OF A STEP BUNCH FOR THE STEP MODEL

Although standard lithographic techniques provide a method of preparing a sinusoidal perturbation involving steps of both signs, this is not the simplest or cleanest physical example with which to study step models because of the problems noted above with the theoretical understanding of step annihilation.[13] Therefore in this article, we restrict our study to systems with steps in the same direction, specifically, the decay of a step bunch under different kinetic mechanisms.

Because of the step–step repulsion [the g term in Eq. (2)], a step bunch will spontaneously spread out, thus decreasing its slope, to reduce the free energy. Since no steps are created or destroyed in the process, the step creation energy β is irrelevant here (note its absence in the various equations of motion). Step movement is associated with mass transport, and different transport mechanisms will yield different dynamical behaviors. The purpose of this section is to derive certain qualitative differences, specifically the scaling laws implied by the different mechanisms, and compare these with experiments.

First we study the motion of an isolated step bunch consisting of N steps [$(N-1)$ terraces] bounded by infinite flat terraces. In the attachment/detachment limited case, we assume that the adatom chemical potential on the two flat terraces equals the equilibrium value. The geometry is similar to that considered by Ozdemir and Zangwill,[6] except at the two ends. Most of their conclusions for sinusoidal profiles have direct analogs here. Most importantly, we obtain the scaling behavior through a separation of variables. Considering a "shape preserving" solution

$$l_n(t) = A_n \tau^{\beta}, \tag{8}$$

where $\tau = t + t_0$, and the A_n are constants depending only on the system size and kinetic coefficients. It is easy to show that $\beta = 1/4$ for the attachment/detachment limited mechanism (with and without local conservation) and $\beta = 1/5$ in the diffusion limited case.[6] The meaning of "shape preserving" here is that after the surface has acquired the shape preserving profile, the profile of the system at any later time can be obtained by multiplying each terrace width by a constant.

It is instructive to look at the continuum limit for this problem. As an example, we present the conserved attachment/detachment case, but the reasoning for the rest of the cases is similar. We obtain the continuum equation from our step equations by replacing differences involving the step indices by derivatives. In many cases, the resulting equations reduce to those derived in other ways by earlier workers, but this derivation may put them on firmer theoretical ground. As an example, for the attachment/detachment limited mechanism with local conservation, the equation of motion for the continuum model is

$$\frac{\partial z(x,t)}{\partial t} = -B_1 \frac{\partial}{\partial x}\left[\frac{1}{|z_x|}\frac{\partial}{\partial x}(|z_x| z_{xx})\right], \tag{9}$$

where $z(x,t)$ is the height of the surface, and z_x is the partial derivative of z with respect to the x variable. Using the scaling form $z(x,t) = Z(x/t^{1/4}) = Z(u)$, from Eq. (9) we have

$$uZ' = 4B_1 \frac{d}{du}\left[\frac{1}{Z'}\frac{d}{du}(Z'Z'')\right]. \tag{10}$$

The exponent 1/4 is consistent with the step model. We will solve a similar scaling equation for the isotropic continuum model later in this section. Here we just derive some scaling rules from general considerations. The symmetry of the problem requires $Z(-u) = -Z(u)$. Observe that if $Z_0(x/t^{1/4})$ is a solution of the problem, $\lambda Z_0(x/t^{1/4})$ is also a solution. Thus we can use the same scaling function Z_0 to describe bunches of different sizes, with the scale factor λ propor-

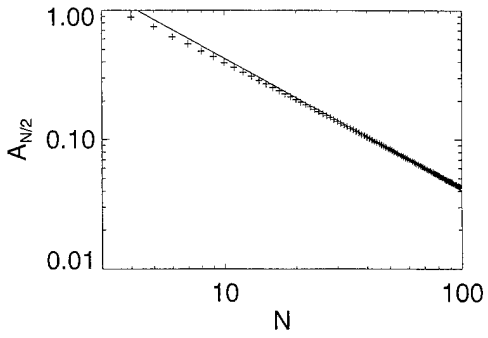


FIG. 2. Solutions of A_n, the prefactor governing the width of the terrace [Eq. (8)], at the center of the bunch ($n=N/2$) for the conserved attachment/detachment limited case as a function of different bunch size N. The solid line is a fit to N^{-1}.

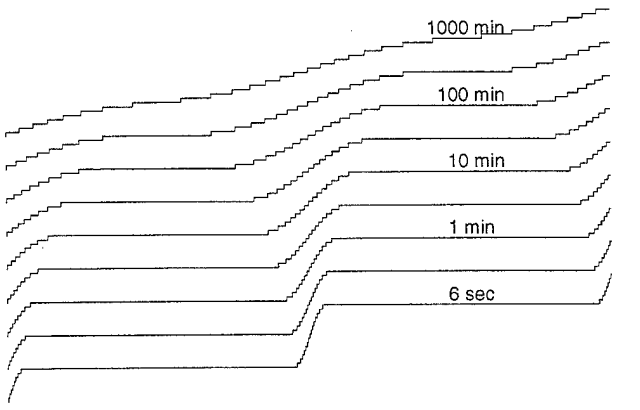


FIG. 3. Illustration of the decay of periodic step bunches from the numerical calculation using the locally conserved attachment/detachment limited mechanism. The number of steps in each bunch is 15.

tional to N [i.e., $2\lambda Z_0(\infty)=Nh$]. Then we have the following scaling law for the time evolution of the slope z_x at the center of the bunch:

$$z_x(0,t)=\frac{\partial}{\partial x}[\lambda Z_0(x/t^{1/4})]=\frac{hZ_0'(0)}{2Z_0(\infty)}Nt^{-1/4}\sim(t/N^4)^{-1/4}. \tag{11}$$

Similar relationships hold for the discrete step models. By numerically determining the A_n's for different bunch sizes N, we can now calculate the terrace width at the center of the bunch as a function of time from Eq. (8). Figure 2 plots $R_N=A_{N/2}$ (or $A_{(N+1)/2}$ if N is an odd number) versus the bunch size N. Notice that $1/R_N$ is just the prefactor for the power law decay of the slope of an isolated step bunch. The result shows $R_N\sim 1/N$ as would be expected; the agreement is better for large bunch sizes. Therefore for the attachment/detachment limited mechanism with local conservation, both the continuum model and the step model predict that the slope at the center of the step bunch scales as

$$s_0(t)\sim 1/l_{N/2}(t)\sim(\tau/N^4)^{-1/4}. \tag{12}$$

By similar approaches, we find that for all three cases of step motion, aside from a short initial transition, the decay of the slope at the center of an isolated step bunch obeys the scaling law

$$s_0(N,t)\approx m(t/N^{\alpha})^{-\beta}, \tag{13}$$

where m is a constant independent of the bunch size. This scaling law is exact for the continuum model (or the step models in the large bunch size limit) but is only an approximation for the step models with finite bunch sizes. For the attachment/detachment rate limited mechanism with local conservation, we find $\beta=1/4$ and $\alpha=4$. For the diffusion limited mechanism, we find $\beta=1/5$ and $\alpha=4$. For the attachment/detachment rate limited mechanism without local conservation, we find $\beta=1/4$ and $\alpha=2$.

In the experiments described below, the surfaces always have a finite miscut angles and step bunches are separated by finite distances. The complete analysis is only possible numerically, but for small modulations, we can linearize the step motion equations. Recently, Bonzel and Mullins[14] analyzed the continuum equations for a surface with small sinusoidal perturbation. Similar to this work, the systems they considered also do not involve annihilation of steps. As their results have shown, the step bunch should decay exponentially when the surface profile is in the linear region, with the time constant dependent on both the average miscut angle and periodicity of the modulation. Therefore in the decay of an initially steep structure, the initial power law decay of the step bunch should break down and crossover to exponential decay at long times. We can obtain the time constants of the decay from the linear expansion of the step motion equations [Eqs. (5)–(7)]. For a modulation with periodicity L and average terrace width l, the relaxation rate can be written as [using appropriate prefactors from Eqs. (5)–(7)]

$$1/\tau\sim l^{-4}(18-24\cos(kl)+6\cos(2kl))\approx 3k^4 \tag{14}$$

for the conserved attachment/detachment case, and

$$1/\tau\sim l^{-5}[18-24\cos(kl)+6\cos(2kl)]\approx 3k^4/l \tag{15}$$

for the diffusion limited case and

$$1/\tau\sim l^{-4}[6-6\cos(2kl)]\approx 3k^2/l^2 \tag{16}$$

for the nonconserved attachment/detachment case, where $k=2\pi/L$ is the wave number. At the long wavelength limit (i.e., when the separation of the initial step bunches is much larger than the average terrace width), these results produce the familiar L^4 scaling for the two conserved cases and L^2 scaling for the nonconserved case. A full numerical solution of the decay of periodic step bunches as shown in Fig. 3 confirms the crossover from a power law decay to an exponential decay when the slope of the bunch has decreased to about five times the average slope of the surface.

As a final note to the step models, we compare the step motion equations obtained above with equilibrium fluctuation results. Bartelt *et al.*[15] analyzed the fluctuations of steps on Si(111) which had been measured using reflection electron microscopy (REM).[16] It was shown that the dependence of relaxation (or creation) time on the wave number of the

fluctuations satisfies an approximate q^{-2} relation, with measured mobility $\Gamma \approx 5.5\times10^{7}$ Å^3/s.[15] This suggests that adatom attachment/detachment is the main mechanism for step motions.[15] However as was shown by Pimpinelli *et al.*,[17] this result alone cannot distinguish whether the adatom diffusion or the attachment/detachment is the limiting process. Under the diffusion limited assumption, Pimpinelli *et al.*[17] estimated that $D_s c_{\rm eq}^0 \sim 10^8$ s^{-1}. For the attachment/detachment limited assumption with local conservation, care must be taken in comparing the kinetic coefficients $\kappa^{\pm}$ with the mobility Γ as measured from the equilibrium fluctuation.[15] Following Bales and Zangwill[18] and Pimpinelli *et al.*,[19] a formal derivation of the mobility Γ can be obtained by extending the BCF model to two dimensions and doing the linear perturbation analysis under the quasistatic approximation.[11] The calculation is rather complicated and not particular instructive. Here we just quote the results. If the step motions are mainly caused by the exchanges of atoms between neighboring steps, then $2\kappa^+\kappa^-/(\kappa^+ + \kappa^-) \approx \Gamma/(a^2 k_B T)$. If exchanges of atoms between different positions on the same step (e.g., due to the diffusion on the terrace roughly parallel with the step edges) make important contributions to the fluctuations of the step position, then we have $(\kappa^+ + \kappa^-) \approx \Gamma/(a^2 k_B T)$. The first scenario is suitable for extremely long wavelength fluctuations, and the second one is suitable at shorter wavelengths. The crossover length can be determined by the ratio between the diffusion rate and the attachment rate, which is still an outstanding question. In the analysis of the experiments below, we assume there is no asymmetry in the adatom attachment rate and equilibrium fluctuation is caused by the second scenario $[\kappa^+ = \kappa^- = \Gamma/(2a^2 k_B T)]$.

To put matters into perspective, we apply Mullins' isotropic continuum model to the step bunch relaxation problem. Consider a 1D surface profile described by $z(x,t)$. Mullins predicts that the evolution of the surface profile obeys the following partial differential equation if the mass transport is through surface diffusion:

$$\frac{\partial z}{\partial t} = -B\frac{\partial^4 z}{\partial x^4}, \tag{17}$$

where

$$B = \frac{D_s c_{\rm eq} a^4 h^2 \gamma_c}{k_B T}. \tag{18}$$

D_s and $c_{\rm eq}$ have the same definition as in the diffusion limited case. γ_c is the surface tension which is assumed orientationally isotropic in the continuum model. Below the roughening transition, γ_c cannot be derived from the Gruber–Mullins form of free energy. Nonetheless, it is instructive to study this model and compare it with the experiments.

We study the evolution of a surface profile that connects two large terraces at different heights; therefore, the profile satisfies the boundary condition $z(x,\pm\infty) = \pm H/2$. We assume the following scaling form:

$$z(x,t) = \lambda Z[x/(Bt)^{\beta}]. \tag{19}$$

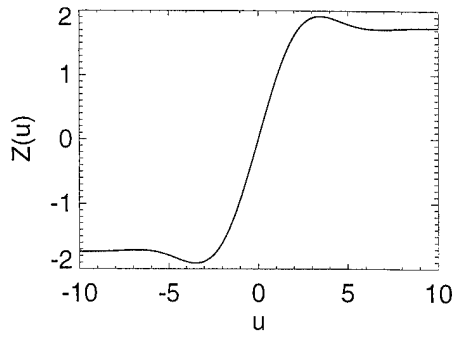


FIG. 4. Shape preserving profile determined from the solution of the scaling equation for the isotropic continuum model [Eq. (20)].

Denote $u = x/(Bt)^{\beta}$, then from Eq. (17) we have

$$\beta u Z'(u) = (Bt)^{1-4\beta} Z'''', \tag{20}$$

where $Z'(u)$ denotes the first order derivative of Z with respect to u. Thus a shape preserving profile must have $\beta = 1/4$. To determine $Z(u)$, it is more convenient to consider the derivative of $Z(u)$ first because it will vanish at plus and minus infinity. Writing $W(u) = Z'(u)$ leads to

$$W''''(u) - uW'(u)/4 - W(u)/4 = 0. \tag{21}$$

The assumption of a power series solution

$$W(u) = \sum_{n=0}^{\infty} a_n u^n \tag{22}$$

to Eq. (21) leads to the recursion relation

$$a_{n+4} = \frac{a_n}{4(n+2)(n+3)(n+4)}. \tag{23}$$

We can use the symmetry and boundary condition of the problem to find the linearly independent solutions. The final result is

$$W(u) = \sum_{n=0}^{\infty} \frac{(-1)^n \Gamma(3/4)(u^2/8)^n}{n!\,\Gamma(n/2+3/4)}. \tag{24}$$

The coefficients are chosen so that $W(0) = 1$. $Z(u)$ can also be obtained,

$$Z(u) = \sum_{n=0}^{\infty} \frac{(-1)^n \sqrt{8}\,\Gamma(3/4)(u/\sqrt{8})^{2n+1}}{n!\,\Gamma(n/2+3/4)(2n+1)}. \tag{25}$$

Figure 4 shows the form of $Z(u)$. Equation (19) predicts that $z_x(x,t) = \lambda (Bt)^{-1/4} W[x(Bt)^{-1/4}]$. $W(u)$ reaches its maximum at $x = 0$ and since $W(0) = 1$, we have $z_x(0,t) = \lambda (Bt)^{-1/4}$. Since λ is proportional to the height difference between the two terraces $[2\lambda Z(\infty) = H]$ and H is proportional to the number of steps in the step bunch, we have the following scaling relation for the maximum slope of the step bunch:

$$z_x(0,t) = \frac{h}{2Z(\infty)B^{1/4}} (t/N^4)^{-1/4}, \tag{26}$$

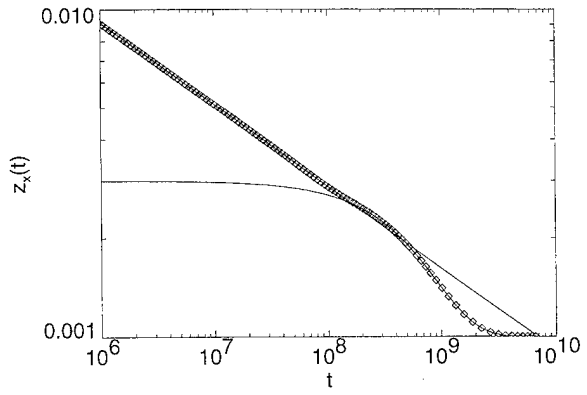


FIG. 5. The time evolution of the steepest slope of the surface for a system of $H/L=0.001$, where H is the height of the step bunch and L is the system size or periodicity, from numerically integrating the continuum equation. The two solid lines are for comparison with the power law $t^{-1/4}$ for the initial decay and exponential decay for long time limit.

where h is the single step height and $Z(\infty)\approx 1.733$. Comparing with Eq. (13), we see that the isotropic continuum model yields the same scaling exponents as the discrete step model using the attachment/detachment limited mechanism with local conservation.

On the other hand, predictions of other characteristics of the decay of the surface profile for the isotropic continuum model are quite different from the step model. As can be seen from Fig. 4, the regions that connect the flat terraces with the step bunch can have slopes opposite to that of the step bunch. Physically, if the direction of two large terraces coincides with the low index facet, slopes of opposite sign correspond to steps of opposite sign (steps and antisteps). Therefore in this case, the isotropic continuum model predicts the creation of step and antistep pairs. In contrast, in the step model it is assumed that there are no extra steps or antisteps created; i.e., the step creation energy is prohibitively high. Thus the total number of steps is conserved. This is a more realistic picture for surfaces well below the roughening temperature, as the experiment on Si(111) will show.

To complete the study of the continuum model, instead of an isolated step bunch, we study a periodic system with step bunches separated by a distance L. This is of interest because in this case the final relaxed surface will have a finite slope H/L. We perform a numerical integration of the partial differential equation, Eq. (17). Figure 5 shows the evolution of the slope at the center of the bunch (which corresponds to the maximum slope of the surface) for a system with $H/L=0.001$. The calculation shows the surface assumes the shape preserving profile in Eq. (25) very quickly and retains this profile until the slope decreases to within an order of magnitude of the final slope. The decay rate obeys a power law approximately with exponent ~ -0.25 as in Eq. (26). As the slope decreases and step bunches widen and overlap with one another, the power law breaks down. When the slope of the step bunch becomes comparable to the average slope, the profile of the surface is approximately sinusoidal relative to the average slope. The decay of the slope of the bunch can then be better fit to $H/L+C\exp(-t/\tau)$. This is consistent with the exponential decay rate of a sinusoid in Bonzel and Mullins's[14] continuum model.

V. EXPERIMENTS ON THE RELAXATION OF STEP BUNCHES ON Si(111)

Recently, experiments have been performed on the particular surface geometry discussed before.[20] In these experiments, we have investigated the thermal relaxation of step bunched surfaces. These surfaces consist of bunches of very closely spaced steps separated by terraces much wider than the equilibrium step separation. We use direct current heating to produce these metastable structures of a controlled size. The dramatic changes in surface morphology produced by the use of direct current heating on the Si(111) surface have been well documented, first by Latyshev *et al.*,[21] and since then by many others.[9,22–28] This step bunching effect on Si(111) is dependent on both the direction of the applied current relative to the step edges and the temperature of the surface. A uniformly stepped surface will become bunched when direct current is applied in the step up direction at temperatures around 1190 °C. The surface remains uniformly stepped, however, for step up current for temperatures near 945 and 1245 °C. For direct current applied in the step down direction, the opposite occurs. For the conditions we use to obtain step bunching, the number of steps within the bunches and the distances between bunches increases with increasing time of applied current. Thus the length of time during which direct current is applied to the sample is used to determine the size of the resulting structures produced on the surface.

The experiments were performed in ultrahigh vacuum (base pressure of 4×10^{-11} Torr) using a homemade scanning tunneling microscope (STM).[29] We use two methods of heating the sample, resistive heating and electron bombardment heating from a tungsten filament positioned behind the sample. In the case of resistive heating, an optical pyrometer was used to measure the sample temperature. For electron bombardment heating, the resistance of the sample was monitored to measure the sample temperature. We used a sample, $15\times3\times0.4$ mm^3 in size, of nominally flat, n-type Si(111) with a measured equilibrium step spacing of 1300 Å. The sample was placed into vacuum without any chemical precleaning and degassed. This was followed by a 1 min flash to 1275 °C using direct current applied in the step up direction.

We studied the thermal decay of three different starting surfaces. To prepare the initial step bunched surfaces, we used direct current, applied in the step down direction, at a temperature of ~1260 °C, to bunch samples for 2, 1, and 0.5 min. For these heating times, we obtained surfaces with structures from ~50 to 28 Å in height. The voltage drop across the sample was ~8 V and the current used ~6 A. The next step was then to relax these starting surfaces. To avoid the effects of a direct current on the decay of the surface structures, we used electron bombardment heating at a temperature of 930 °C (only a small current of 0.068 A was passed through the sample for temperature measurement).

For each starting surface, relaxations were done for two different times. For the starting surfaces which were direct current heated for 2, 1, and 0.5 min, relaxations were done for 30 and 120, 15 and 60, and 11 and 24 min, respectively. Times were chosen to produce a similar amount of relaxation on the different initial surfaces. After each relaxation, the sample was quenched and imaged at room temperature. The pressure during heating was less than 6×10^{-10} Torr.

Figures 6(a)–6(c) show 3D STM images of the three initial step bunched surfaces (with structures created by heating with direct current for 2, 1, and 0.5 min, respectively). Figures 7(a) and 7(b) show the initial surface of Fig. 6(a) after relaxations for 30 and 120 min, respectively. There are two things to note in these images. First, even after 120 min, the bunches have only spread out a small amount. We are still in an early stage of the relaxation. The second thing to note is the presence of the crossing steps[9] on the surface, whose numbers increase with increasing relaxation time. Both of these factors had an effect on our choice of the relaxation parameter used to measure the extent of decay. We considered several possibilities including the largest terrace width and the root-mean-square (rms) height of the surface. The problem with the largest terrace width as a relaxation measure was the initial presence of crossing steps between the bunches. Because the crossing steps are nearly parallel to the bunches, it is impossible to distinguish between crossing steps which are located close to a bunch and steps which are actually part of a widening bunch. The root-mean-square roughness of the surface was found not very sensitive to the decay of the surface structures in this early stage of the relaxation.

The relaxation measure which we found to be the most useful is the maximum slope of the bunches. This measure is more sensitive to the decay than rms roughness in this early stage of the relaxation, but insensitive to the effects of crossing steps. For each heating preparation, the maximum slope of the bunches was measured for 11 to 22 bunches on the surfaces. The slopes were then averaged and plotted as a function of time as shown in Fig. 8. The error bars on each point represent the standard deviation of the mean. As is apparent from the plot, the initial slope (and thus the step–step distance) of the bunches is different for each of the starting surfaces. This step–step spacing within the bunches decreases as the number of steps within the bunches increases, which is important in the analysis of the data. As shown above, the thermal relaxation of the surfaces is driven by the repulsive step–step interaction. Since the initial step–step separation determines the initial driving force on the surface, comparison of the relaxation curves for the different starting surfaces requires setting the zero of time for the experimental data so that all curves have the same slope at $t=0$.

In analyzing the experimental data, we considered predictions from the three different relaxation mechanisms discussed earlier in this article. The first two mechanisms assume a local conservation condition, in the two extremes of attachment/detachment limited (fast diffusion but slow attachment/detachment of adatoms to/from the step edges) and diffusion limited (fast attachment/detachment rate but relatively slow diffusion). The third mechanism does not require local adatom conservation and assumes that the attachment/detachment of adatoms to/from the step edges is the limiting factor in adatom motion. Since the numerical solutions [see Eq. (13)] predict that the decay of the slope of

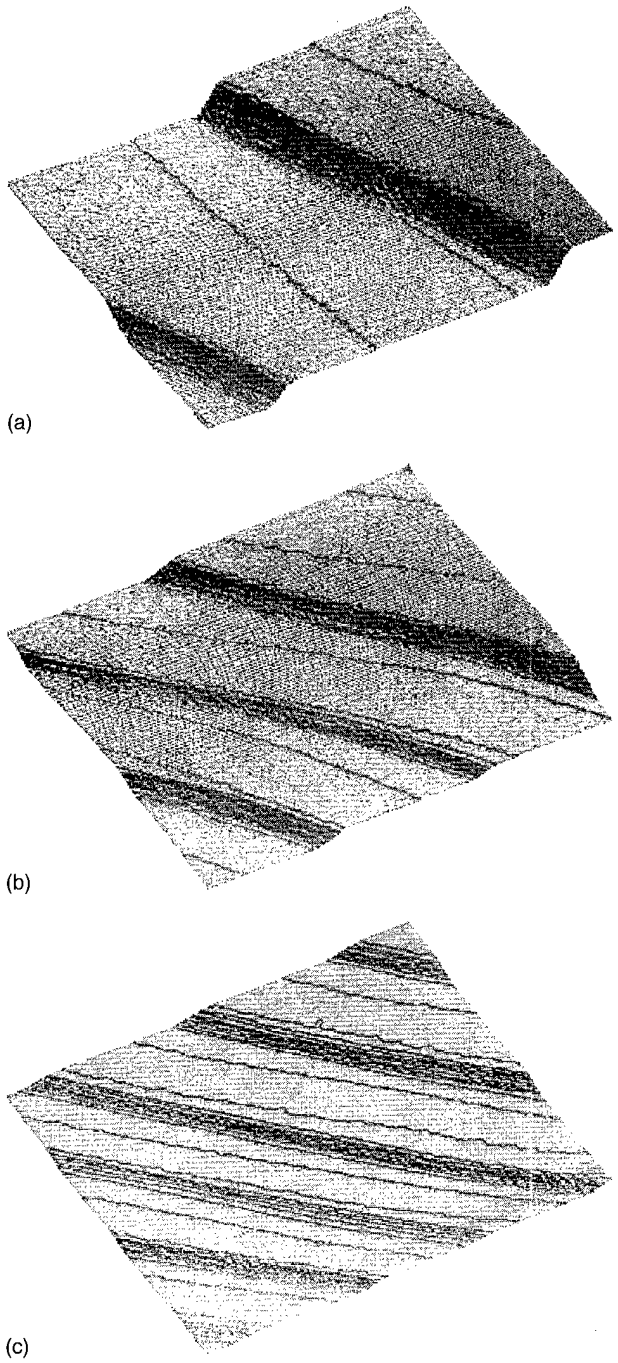


FIG. 6. 50 000 Å×50 000 Å STM images of the three different initial surfaces with (a) a bunch size of 16 steps, a bunch–bunch separation of 22 300 Å, and an initial step–step separation within the bunches of 124 Å (prepared by heating with dc in the step down direction for 2 min), (b) a bunch size of 12 steps and a bunch–bunch separation of 10 900 Å (prepared by heating with dc for 1 min), and (c) a bunch size of 9 steps and a bunch–bunch separation of 7300 Å (prepared by heating with dc for 0.5 min).

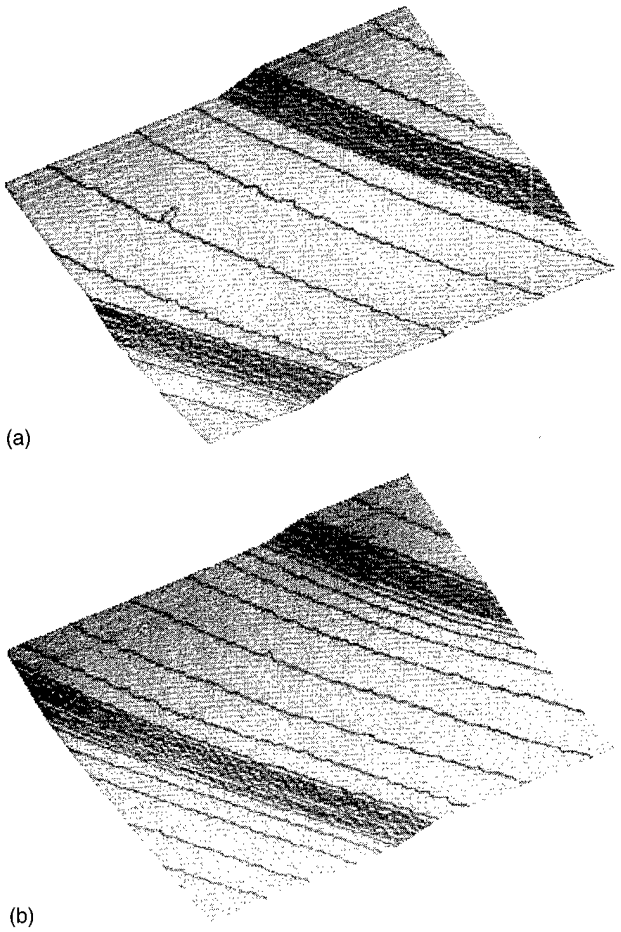


FIG. 7. 50 000 Å×50 000 Å STM images of surfaces with an average bunch size of 16 steps. Individual steps are a single interplanar spacing, 3.14 Å, high. The starting surface, shown in Fig. 6(a), was prepared by heating with dc in the step down direction for 2 min. This was followed by relaxations under indirect heating for (a) 30 and (b) 120 min.

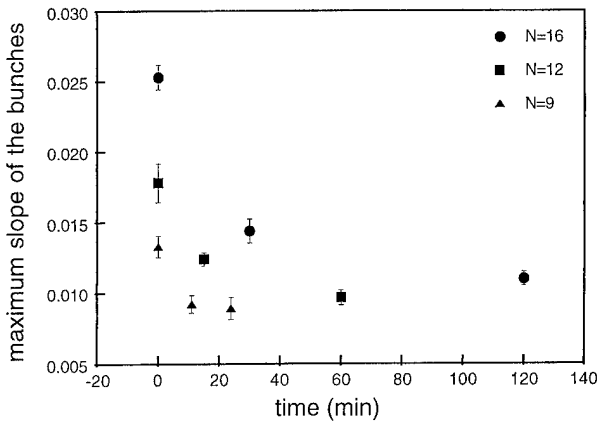


FIG. 8. Plot of maximum slope of the bunches vs time for the three experimental data sets. Each data point is an average of approximately 11–22 measurements, and the error bars represent the standard deviation of the mean.

the bunches follows a scaling relation with respect to the number of steps in the bunches, we first investigate if such a scaling relation exists in the experimental data. We offset the time axis of the relaxation curves so that all start out subject to the same initial driving force. We performed a three parameter fit to obtain the best values of α, and the time shifts for data sets two and three. We find that the experimental data does indeed scale with the number of steps in the bunches. A conservative error bar on α of ± 0.5 was obtained by varying α about the best value of 4.3, and finding the values at which two experimental data points moved outside their error bars from the scaled curve. The measured value of α is consistent with a conserved model of step motion for which α is predicted to be 4.

To make quantitative comparisons between the three theoretical mechanisms and the experimental results, we need values of the step–step interaction coefficient g, the step mobility Γ, and the diffusion coefficient for atoms on the terraces $D_s c_0^{eq}$. In general, the step–step interaction coefficient can be determined by direct observation of the equilibrium distribution of step–step spacings.[30] REM observations of steps on Si(111) at 900 °C show that the steps interact with a repulsive interaction which decays as the inverse square of separation, and a magnitude consistent with a stress-mediated interaction.[16,30] The value of the step interaction coefficient at 900 °C has been found to be $g=0.015$ eV/Å^2, with an uncertainty of approximately 30%. The kinetic parameters can be determined by direct observations of the equilibrium fluctuations of the steps.[31] REM observations of step fluctuations at 900 °C show that fluctuations decay with a time constant that depends on the square of their wavelength.[15,32] The result has been alternatively interpreted as due to step attachment/detachment limited kinetics with a step mobility of $\Gamma=5.5\times10^7$ Å^3/s,[15] and is due to diffusion-limited kinetics with a diffusion constant of $D_s c_0^{eq}=1\times10^8$ s^{-1}.[32] These parameters were used in numerical solutions of Eqs. (5)–(7), with $\Gamma=f(k_+,k_-)$ as discussed above, to make predictions on the characteristics of the decay of the slope (see Fig. 3 for the illustration of the change of profile). The numerical calculations were done for starting surfaces with bunch sizes matching those of the experimental system, $N=16$, 12 and 9, and an initial slope of 0.0253, the initial slope of all experimental data sets after the time axis was reset. The resulting calculated relaxation curves were then scaled to produce the curves shown in Fig. 9. The results show quantitative agreement between the measured rates and the predicted rates, well within the experimental uncertainties.

Next we look to the rate of decay of the slope to try and distinguish between the two mechanisms which require local adatom conservation. The relaxation of the slope shown in Fig. 10 follows a power law with an exponent close to -0.23 (-0.2 to -0.3 give reasonable fits). We cannot distinguish between the two conserved mechanisms; as in these two cases, the solutions described above yield time-scaling exponents of $-1/4$ and $-1/5$, respectively. As in the case of the analysis of the step fluctuations, the data do not allow these two cases to be distinguished.

As a final test of the utility of the mesoscale theory, we

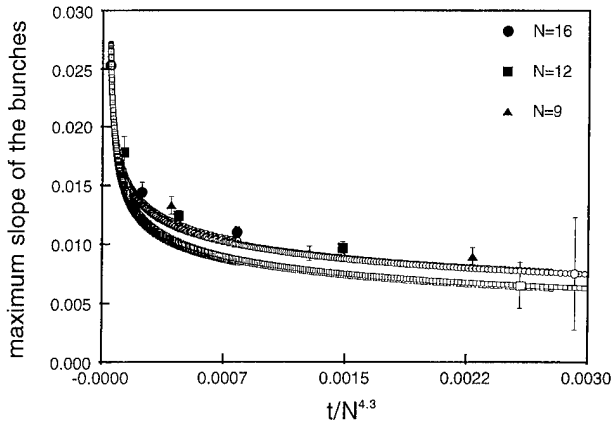


FIG. 9. Plot of scaled slope vs time for the experimental data and results of the numerical solutions. The three experimental data sets are plotted using different symbols, solid circles for $N=16$, solid squares for $N=12$, and solid triangles for $N=9$, and the error bars again represent the standard deviation of the mean. The closely spaced points show the results of the numerical solutions for the two conserved relaxation mechanisms [see Eqs. (5) and (6)]. For each mechanism, relaxation results were obtained for the three initial surface configurations matching the experiment (initial slope of 0.0253, the starting slope of all the data after appropriate time shifts) and then scaled using the experimental best α of 4.3. The hollow squares with dots in the centers denote the results for the conserved attachment/detachment limited case, and the hollow circles with dots in the centers denote results for the diffusion limited case. The error bar shown for each mechanism is the error bar for the specific time for which it is plotted and was obtained from the numerical solution results assuming an error in g of 30%, an error in Γ of 50%, and an order of magnitude variation in $D_s c_0^{eq}$. The error bar in each case is representative of the larger percentage variations in slope observed for the time shown in the figure. The experimental results are consistent with either of the two mechanisms which require adatom conservation, and inconsistent with the mechanism which does not require adatom conservation (not shown).

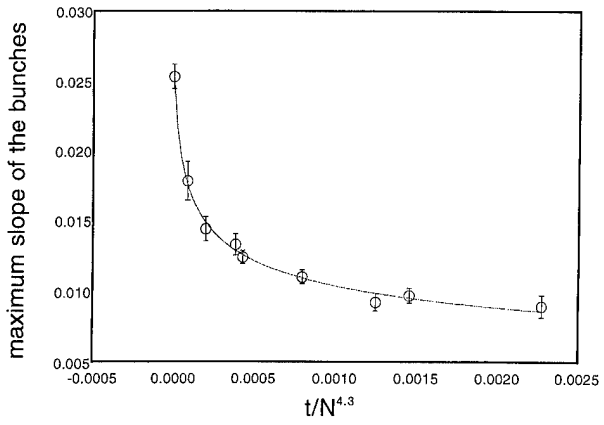


FIG. 10. Plot of the scaled experimental data with a fit of $t^{-0.23}$. We cannot distinguish between the two mechanisms which require local conservation (attachment/detachment limited and diffusion limited).

compare these experimental results to the previously discussed predictions of Mullins' original isotropic continuum theory applied to our surface geometry. The range of slopes investigated experimentally in this study were from 12 to 5 times the initial slope. In this range, we are in agreement with the continuum prediction of $t^{-1/4}$ (Fig. 5). However, as mentioned earlier, the surface profile which evolves in the continuum analysis contains the unphysical prediction of the creation of antisteps on the surface, which was never observed experimentally. In addition, we used our fit of the experimental data to calculate the prefactor $D_s c_0^{eq} a^4 h^2 \gamma_{isotropic}/k_B T$ ($D_s c_0^{eq}=10^8$ s^{-1}), obtaining an estimate of the effective surface tension $\gamma_{isotropic}$ of 0.015 eV/Å^2. This result effectively has no meaning since it cannot be identified with any physical parameter in our system, which is below the roughening temperature.

VI. SUMMARY

We have presented theoretical and experimental results which show the applicability of coarse-grained models of step motion to the prediction of the rate of decay of metastable structures of arbitrary shape. We have considered the specific case of a structure consisting of step bunches with steps of the same sign separated by wide flat terraces. Three distinct mechanisms of step motion were considered. Numerical solutions, as well as analytical solutions in the continuum limit, of the equations of step motion for this geometry show that in the large amplitude (or large slope) regime the decay has a power law dependence on time. A crossover to exponential decay occurs when the slope of the structures becomes comparable to (within about a factor of 5) the average slope of the surface, in agreement with previous studies of the small amplitude regime.[14] The decay of the structures is shown to obey a size-scaling relationship that depends on the initial size of the structure (number of steps in the bunch) independent of the spacing of the bunches. Numerical solutions of an isotropic continuum model (i.e., one in which there are no steps) give similar time and size scaling results. However, the prediction of the shape evolution of the profile in the isotropic continuum model is inconsistent with experimental observations, and the prefactor for the rate of decay cannot be predicted on any physical basis.

We have compared the theoretical results with the measured decay of metastable step bunches of three different sizes fabricated using direct current heating. The metastable structures consist of step bunches from 3 to 5 nm in height (step–step separations on the order of a couple hundred angstroms) separated from one another by terraces 7300 to 22000 Å wide. Their decay demonstrates size scaling ($\alpha=4.3 \pm 0.5$) as predicted by the theory for step motion governed by local conservation of mass (i.e., atoms motion via rapid exchange across steps is not the governing component of mass transport). The time scaling observed ($\beta=-0.2$ to -0.3) is consistent with either a mechanism in which the rate limiting process in step motion is diffusion on the terraces ($\beta=-0.2$) or attachment and detachment of atoms at the step edges ($\beta=-0.25$). The experimentally measured decay rate can be predicted quantitatively using values for the parameters in the theory which were determined from measurements of the equilibrium properties of the surface. The approach taken in this work, which considers appropriate coarse-graining of kinetic processes to remove the system-dependent details of

atomic processes, provides a useful and successful method of predicting surface mass transport. The rather simple system studied here, decay of a single-component quasi-one-dimensional metastable structure in the absence of external perturbations, provided an excellent test of this approach. Extension of this work to more complex systems including growth, chemical reaction, and full 2D structures is feasible, and will present interesting theoretical and experimental challenges.

ACKNOWLEDGMENTS

This work has been supported in part by the ONR (EF, MJ, EDW) and in part by the NSF-MRG (DJL, JDW, EDW).

[1]W. W. Mullins, J. Appl. Phys. **28**, 333 (1957).
[2]W. W. Mullins, Philos. Mag. **6**, 1313 (1961).
[3]W. Mullins, J. Appl. Phys. **30**, 77 (1959).
[4]P. Nozières, J. Phys. France **49**, 257 (1988).
[5]A. Rettori and J. Villain, J. Phys. France **49**, 257 (1988).
[6]M. Ozdemir and A. Zangwill, Phys. Rev. B **42**, 5013 (1990).
[7]E. E. Gruber and W. W. Mullins, J. Phys. Chem. Solids **28**, 875 (1967).
[8]C. Jayaprakash, C. Rottman, and W. F. Saam, Phys. Rev. A **30**, 6549 (1984).
[9]E. D. Williams, E. Fu, Y.-N. Yang, D. Kandel, and J. D. Weeks, Surf. Sci. **336**, L746 (1995).
[10]E. D. Williams, R. J. Phaneuf, J. Wei, N. C. Bartelt, and T. L. Einstein, Surf. Sci. **310**, 451 (1994).
[11]D.-J. Liu and J. D. Weeks (unpublished).
[12]W. K. Burton, N. Cabrera, and F. C. Frank, Proc. R. Soc. London, Ser. A **243**, 299 (1951).
[13]M. Keefe, C. Umbach, and J. Blakely, J. Phys. Chem. Solids **55**, 965 (1994).
[14]H. P. Bonzel and W. W. Mullins (unpublished).
[15]N. C. Bartelt, J. L. Goldberg, T. L. Einstein, E. D. Williams, J. C. Heyraud, and J. J. Métois, Phys. Rev. **48**, 15453 (1993).
[16]C. Alfonso, J. Bermond, J. Heyraud, and J. Métois, Surf. Sci. **262**, 371 (1992).
[17]A. Pimpinelli, J. Villain, D. E. Wolf, and J. J. Métois, Surf. Sci. **295**, 371 (1993).
[18]G. S. Bales and A. Zangwill, Phys. Rev. B **41**, 5500 (1990).
[19]A. Pimpinelli, I. Elkinani, A. Karma, C. Misbah, and J. Villain, J. Phys. Condensed Matter **6**, 2661 (1994).
[20]E. Fu, M. D. Johnson, E. D. Williams, D.-J. Liu, and J. D. Weeks, Phys. Rev. Lett. (in press).
[21]A. Latyshev, A. Aseev, A. Krasilnikov, and S. Stenin, Surf. Sci. **213**, 157 (1989).
[22]A. Latyshev, A. Aseev, A. Krasilnikov, and S. Stenin, Surf. Sci. **227**, 24 (1990).
[23]Y. Homma and M. Suzuki, Appl. Surf. Sci. **60/61**, 479 (1992).
[24]M. Suzuki, Y. Homma, Y. Kudoh, and R. Kaneko, Ultramicroscopy **42–44**, 940 (1992).
[25]M. Suzuki, Y. Homma, Y. Kudoh, and R. Kaneko, Appl. Surf. Sci. **60/61**, 460 (1992).
[26]H. Tokumoto, K. Miki, Y. Morita, T. Sato, M. Iwatsuki, M. Suzuki, and T. Fukuda, Ultramicroscopy **42–44**, 816 (1991).
[27]H. Tokumoto, K. Miki, H. Murakami, and K. Kajimura, J. Vac. Sci. Technol. B **9**, 699 (1991).
[28]M. Ramstad, R. J. Birgeneau, K. J. Blum, D. Y. Noh, B. O. Wells, and M. J. Young, Europhys. Lett. **24**, 653 (1993).
[29]Y.-N. Yang and E. Williams, Phys. Rev. Lett. **72**, 1862 (1994).
[30]N. Bartelt, T. Einstein, and E. Williams, Surf. Sci. **240**, L591 (1990).
[31]N. C. Bartelt, T. L. Einstein, and E. D. Williams, Surf. Sci. **312**, 411 (1994).
[32]A. Pimpinelli, J. Villain, D. E. Wolf, J. J. Métois, J. C. Heyraud, I. Elkinani, and G. Vimin, Surf. Sci. **295**, 143 (1993).

First-principles study of Si 2*p* core-level shifts at water and hydrogen covered Si(001)2×1 surfaces

Alfredo Pasquarello[a)]
Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA), IN-Ecublens, CH-1015 Lausanne, Switzerland and AT&T Bell Laboratories, Murray Hill, New Jersey 07974

Mark S. Hybertsen
AT&T Bell Laboratories, Murray Hill, New Jersey 07974

Roberto Car
Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA), IN-Ecublens, CH-1015 Lausanne, Switzerland and Department of Condensed Matter Physics, University of Geneva, CH-1211Geneva, Switzerland

(Received 22 January 1996; accepted 15 March 1996)

Using a first-principles approach, we study Si 2*p* core-level shifts at water and hydrogen covered Si(001)2×1 surfaces. After allowing for full relaxation of the surface structures, core-level shifts are calculated including core-hole relaxation effects. We find that dissociated water on the Si(001)2×1 surface induces a core-level shift of 1.1 eV to higher binding energies, in good agreement with experiment. In the case of the hydrogen terminated Si(001)2×1 surface, calculated surface shifts are small (about 0.2 eV) and comparable to shifts of subsurface Si atoms. © *1996 American Vacuum Society.*

The characterization of silicon surfaces is an essential prerequisite to understanding the early stages of the silicon oxidation process. In particular, a considerable amount of effort has been focused on the water covered Si(001) surface,[1–7] because of the importance of wet oxidation in silicon-based device technology. It is by now well established that H_2O is adsorbed dissociatively on Si(001)2×1 by saturating the free dangling bonds of the clean surface with –OH and –H groups.[1]

Core-level (Si 2*p*) photoelectron (PE) spectroscopy is used routinely on Si surfaces and interfaces in an attempt to determine the microscopic structures. This technique is sensitive to the local potential, providing a measure of the chemical environment (or oxidation state) of the silicon atoms. When H_2O is adsorbed on Si(001), the Si 2*p* spectra show the appearance of a peak shifted by about 0.8–1.0 eV to higher binding energies with respect to the bulk peak.[2–6] This peak is generally attributed to the surface silicon atoms to which OH groups are bonded. Because of the well defined nature of water covered Si(001) at saturation coverage, it is possible to calibrate the intensity of this peak.[6] In this way, estimates of the number of partially oxidized silicon atoms at the Si(001)/SiO_2 interface can be extracted from PE spectra.[6,8]

In previous work, we addressed Si 2*p* core-level shifts at the Si(001)–SiO_2 interface using a first-principles approach.[9] By making use of interface models, we showed that the suboxide peaks, which are about 1 eV apart in PE spectra, could be attributed to silicon atoms with a different number of oxygen neighbors. The effect of further neighbors turned out to be negligible. Calculated shifts were in good quantitative agreement with experiment provided core-hole relaxation effects, which were found to account for more than 50% of the shifts, were properly included.

In this article, we apply our approach to calculate Si 2*p* core-level shifts at water and hydrogen covered Si(001) surfaces. In the case of water covered Si(001), surface silicon atoms occur in an oxidation state Si^{+1}. According to the first-neighbor dependence that resulted from our study on the Si(001)–SiO_2 interface,[9] the core levels of these atoms are expected to shift by about 1 eV. Agreement between theory and experiment in this case, in which the structure is well known, would strengthen the picture which emerged from the study of Si–SiO_2 interface models. We complement this study by calculating core-level shifts at the hydrogen terminated Si(001)2×1 surface.[10–12]

The results in this work are obtained within the local density approximation to density functional theory (DF-LDA). The surface geometries were obtained by fully relaxing the atomic coordinates to minimize the total energy. This was done by using the Car–Parrinello method,[13] which provides the electronic structure as well as the forces that act on the ions. Only valence electrons are explicitly considered using pseudopotentials (PPs) to account for the core–valence interactions. A norm-conserving PP is used for Si,[14] whereas the O atoms are described by an ultrasoft Vanderbilt PP.[15] The Coulomb potential is used for H. The electronic states were expanded on a plane-wave basis set. Exchange and correlation were included using Perdew and Zunger's interpolation formulas.[16] A description of the method is given in Ref. 17.

In order to achieve good convergence in the electronic properties, we have used a plane-wave cutoff of 16 Ry for the wave functions and of 150 Ry for the augmented electron density. Our system contains a 2×2 surface unit of side $L=7.65$ Å (based on the theoretical equilibrium lattice constant of Si). The dimension of the cell in the direction orthogonal to the surface is 31.7 Å, containing ten layers of

a)Electronic mail: pasquarello@eldp.epfl.ch

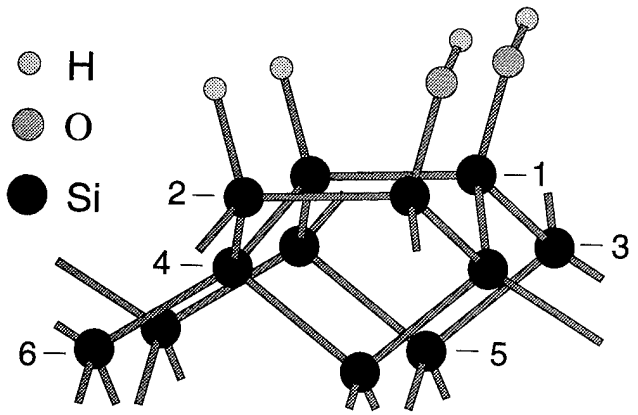


FIG. 1. Ball and stick model of the relaxed positions resulting from dissociative water adsorption on Si(001). The labels for the distinct Si positions correspond to Table I.

Si (≈13 Å), the lower extremities being saturated with hydrogen atoms. The Brillouin zone of our simulation cell was sampled using only the Γ point. In the minimization process, all the surface atoms as well as the first six Si layers were allowed to relax.

The Si $2p$ core-level shifts were calculated both within the initial-state approximation and including core-hole relaxation effects. The difficulty of using a PP approach in which core levels are not explicitly considered is overcome as follows. Initial-state shifts are obtained in first-order perturbation theory, by evaluating the expectation value of the self-consistent local potential on the Si $2p$ atomic orbital.[18,19] Final state effects are included by performing differences in total energy calculations according to the procedure of Ref. 20. Two separate calculations are performed. In a first step, the electronic ground state is determined. Then, the PP of a given atom is replaced by another PP which simulates the presence of a screened $2p$ hole in its core. Following these procedures, relative core-level shifts can be determined. Because of the use of a PP approach, the absolute values of the core-level binding energies are not obtained. A more detailed description of our procedure is given in Ref. 9. By comparing DF-LDA shifts with experiment for a series of molecules, an overall accuracy of 0.5 eV with respect to experiment has been established.[9,21] These tests span a wide range of core-level shifts (several eV). Relative shifts for chemical environments that are more similar are expected to be more accurate, e.g., recent work on the clean Si(001) surface[20] and on the H-terminated Si(111) surface.[19]

The surface geometry resulting from dissociative water adsorption on Si(001) is shown in Fig. 1. At the end of the relaxation process, we found a Si–Si dimer bond of 2.39 Å, a Si–H bond of 1.54 Å, and a Si–O bond of 1.66 Å. The Si–H and Si–O bonds form angles of 75° and 73° with the surface plane, respectively. The Si dimer is tilted by only 1.5°, with the OH-bonded Si atom higher. The O–H bond is roughly perpendicular to the dimer orientation and nearly parallel to the surface.[22] Overall, the structural parameters we found are very close to those obtained in a previous LDA calculation.[7]

TABLE I. Calculated Si $2p$ core-level shifts at the water adsorbed Si(001) surface. Shifts are calculated in the initial-state approximation and including core-hole relaxation (full shift). The shifts are given with respect to the bulk value, obtained by averaging the shifts between the sixth and the eighth Si monolayer. Shifts are given in eV. See Fig. 1 for the geometry.

	Initial state	Full	Expt.
Si(1) (bonded to –OH)	−0.52	−1.09	−1.00[a] −0.9[b]
Si(2) (bonded to –H)	−0.02	−0.08	−0.25[a] −0.3[b]
Si(3) (second layer)	0.07	0.11	
Si(4) (second layer)	0.02	0.03	
Si(5) (third layer)	0.06	0.09	
Si(6) (third layer)	−0.18	−0.18	

[a]Reference 4.
[b]Reference 6.

Calculated Si $2p$ core-level shifts for this surface structure are given in Table I. The shifts are given with respect to a reference value which corresponds to the average shift of atoms in the Si slab sufficiently distant (sixth to eighth monolayer) from the surface, but far enough from the hydrogen terminated back surface. The Si atoms that present the most significant shifts are those that are bonded to an OH group. In the initial-state approximation, we found a shift of −0.52 eV, which increased to −1.09 eV when core-hole relaxation was taken into account. For the other surface Si atoms, the ones bonded to a H atom, we calculated a full shift −0.08 eV. These calculated values are in good agreement with experimental data which give shifts ranging from −0.8 to −1.0 eV (Refs. 2–6) for the OH-bonded Si atom and of −0.25 eV (Ref. 4) [−0.3 (Ref. 6)] for the H-bonded one. The experimental values reported in Table I are taken from Refs. 4 and 6, which are the only ones providing shifts for both surface Si atoms.

In Table I, we also give calculated shifts for atoms in lower lying layers. We found scattered values ranging from −0.2 to +0.1 eV. Since all these values fall rather close to each other as well as to the bulk peak, it is rather difficult to distinguish their contributions separately in the PE spectra. The same remark holds for the surface Si atom bonded to H atoms. The experimental value reported in Table I is the result of a fitting procedure with a limited number of degrees of freedom; in our opinion, it should be taken as the effective contribution due to several Si atoms rather than to the single contribution due to the H-bonded Si atoms.

We complement this study by considering the hydrogen terminated Si(001)2×1 surface.[10–12] Figure 2 shows the relaxed structure of this surface. We found an equilibrium geometry defined by structural parameters in close agreement with other LDA calculations in the literature.[23,24] We found Si dimer and Si–H bonds of 2.38 and 1.54 Å, respectively. The Si–H bond forms an angle of 74° with the plane of the surface.

Calculated core-level shifts for the hydrogen terminated surface are given in Table II. Overall, the shifts are rather small. The H coordinated Si shows a shift close to that found

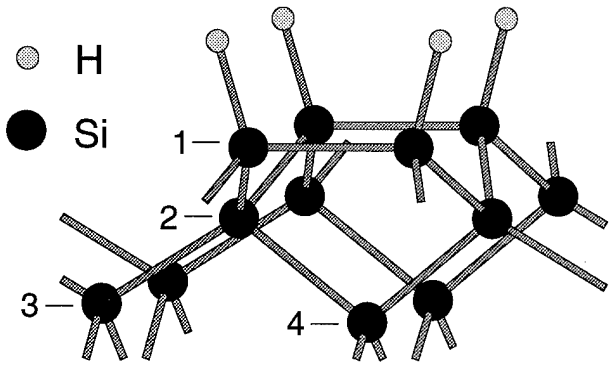


FIG. 2. Ball and stick model of the relaxed positions of the monohydride Si(001)2×1 surface. The labels for the distinct Si positions correspond to Table II.

for the dissociated water case. The shifts for subsurface Si atoms are also similar. The experimental data is taken from Ref. 25, and are again the result of a fit. Thus, the experimental value reflects the effective contribution of all Si atoms which present shifts differing from the bulk value. The overall comparison between theory and experiment is satisfactory.

In conclusion, we have studied Si $2p$ core-level shifts at water and hydrogen covered Si(001)2×1 surfaces. In the case of the water adsorbed surface, the surface Si atom, bonded to the OH group and nominally in a Si^{+1} oxidation state, presented a shift of −1.1 eV, in very good agreement with experiment. Core-hole relaxation was found to account for about 50% of this shift. These results support the interpretation given to Si $2p$ core-level shifts at the Si(001)–SiO_2, which assigned a −1 eV shift for every nearest-neighbor oxygen atom. In the case of the monohydride Si(001)2×1 surface, calculated shifts are small and agree with experiment.

TABLE II. Calculated Si $2p$ core-level shifts at the monohydride Si(001)2×1 surface. Shifts are calculated in the initial-state approximation and including core-hole relaxation (full shift). The shifts are given with respect to the bulk value, obtained by averaging the shifts between the sixth and the eighth Si monolayer. Shifts are given in eV. See Fig. 2 for the geometry.

	Initial state	Full	Expt.
Si(1) (dimer)	−0.08	−0.16	−0.3[a]
Si(2) (second layer)	0.00	0.00	
Si(3) (third layer)	−0.22	−0.24	
Si(4) (third layer)	0.07	0.08	

[a]Reference 25.

ACKNOWLEDGMENTS

Two of us (A. P. and R. C.) acknowledge support from the Swiss National Science Foundation under Grant No. 20-39528.93. The calculations were performed on the NEC-SX3 of the Swiss Center for Scientific Computing (CSCS) in Manno.

[1]Y. J. Chabal and S. B. Christman, Phys. Rev. B **29**, 6974 (1984).
[2]D. Schmeisser, F. J. Himpsel, and G. Hollinger, Phys. Rev. B **27**, 7813 (1983).
[3]W. Ranke and Y. R. Xing, Surf. Sci. **157**, 339 (1985).
[4]C. U. S. Larsson, A. S. Flodström, R. Nyholm, L. Incoccia, and F. Senf, J. Vac. Sci. Technol. A **5**, 3321 (1987).
[5]C. Poncey, F. Rochet, G. Dufour, H. Roulet, F. Sirotti, and G. Panaccione, Surf. Sci. **338**, 143 (1995).
[6]F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B **38**, 6084 (1988).
[7]A. Vittadini, A. Selloni, and M. Casarin, Phys. Rev. B **52**, 5885 (1995).
[8]Z. H. Lu, M. J. Graham, D. T. Jiang, and K. H. Tan, Appl. Phys. Lett. **63**, 2941 (1993).
[9]A. Pasquarello, M. S. Hybertsen, and R. Car, Phys. Rev. Lett. **74**, 1024 (1995); Phys. Rev. B **53**, 10 942 (1996).
[10]T. Sakurai and D. Hagstrum, Phys. Rev. B **14**, 1593 (1976).
[11]Y. J. Chabal and K. Raghavachari, Phys. Rev. Lett. **53**, 282 (1984).
[12]J. J. Boland, Phys. Rev. Lett. **65**, 3325 (1990).
[13]R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[14]G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
[15]D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
[16]J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[17]A. Pasquarello, K. Laasonen, R. Car, C. Lee, and D. Vanderbilt, Phys. Rev. Lett. **69**, 1982 (1992); K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, Phys. Rev. B **47**, 10142 (1993).
[18]Y. Miyamoto and A. Oshiyama, Phys. Rev. B **44**, 5931 (1991).
[19]X. Blase, A. J. R. da Silva, X. Zhu, and S. G. Louie, Phys. Rev. B **50**, 8102 (1994).
[20]E. Phelke and M. Scheffler, Phys. Rev. Lett. **71**, 2338 (1993).
[21]L. Pedocchi, N. Russo, and D. R. Salahub, Phys. Rev. B **47**, 12992 (1993).
[22]The structure we obtained corresponds to the one depicted in Fig. 2(a) of Ref. 7.
[23]J. E. Northrup, Phys. Rev. B **44**, 1419 (1991).
[24]A. Vittadini, A. Selloni, and M. Casarin, Surf. Sci. Lett. **289**, L625 (1993).
[25]K. Yamamoto and M. Hasegawa, J. Vac. Sci. Technol. B **12**, 2493 (1994).

New approach to preparing smooth Si(100) surfaces: Characterization by spectroellipsometry and validation of Si/SiO_2 interfaces properties in metal-oxide-semiconductor devices

D. Schmidt, H. Niimi, B. J. Hinds, D. E. Aspnes,[a] and G. Lucovsky
Departments of Physics, Materials Science and Engineering, and Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 23 February 1996; accepted 20 April 1996)

Wet chemical removal of thermally grown SiO_2 layers on Si(100) substrates has been studied as a function of the pH of the etching solutions in the range of -0.32–1.6 by adding controlled amounts of H_2SO_4 to a 1:30 HF:H_2O mixture. Characterization of the stripped Si(100) surfaces by spectroellipsometry showed that the *smoothest* surfaces were obtained at a 1:0.50:30 HF (49 wt %):H_2SO_4(98 wt %):H_2O etch with a pH of approximately 0.5. Electrical characterization of metal-oxide-semiconductor (MOS) capacitors fabricated on these surfaces with oxide layers prepared by remote plasma enhanced chemical vapor deposition showed (i) the lowest density of interface traps, D_{it}, (ii) the lowest tunneling currents, J_0, and that (iii) the highest breakdown fields, E_{BD}, occurred at the same pH value that produced the smoothest surfaces. In contrast, MOS capacitors fabricated with high-temperature thermally grown oxides were not significantly affected.
© *1996 American Vacuum Society.*

I. INTRODUCTION

In devices with submicron feature sizes scaled down for ultralarge scale integration (ULSI) circuit applications, the complementary scaling of gate oxide thicknesses to ~6 nm and less reduces significantly the influence of bulk oxide defect and defect precursors on electrical performance and reliability. This means that defect and defect-precursor bonding arrangements at the Si–SiO_2 interface will then play a dominant role in determining device performance and reliability. For devices with gate oxides less than ~6 nm thick, the roughness of the Si surface prior to either thermal oxidation, or oxide deposition by remote plasma enhanced chemical vapor deposition (RPECVD) is an important factor that must be considered in relating performance and/or reliability to the specific interface formation process. This means that considerable attention must be placed on all aspects of processing that occur prior to the sequence of steps that form the Si–SiO_2 interface.

Solutions containing HF, NH_4F, and H_2O are commonly used to remove sacrificial oxides and for other aspects of Si device processing, including interprocess cleaning. For Si(100) wafers that are used in ULSI devices, the two most often applied steps in preoxidation processing are (i) an RCA clean and (ii) a final etch in dilute HF to remove the sacrificial chemical oxides generated in the RCA clean. It has been demonstrated that these steps combine to roughen the Si(100) surface.[1–3] In marked contrast, Si(111) surfaces can be etched to a nearly atomically flat condition using HF:NH_4F:H_2O solutions with a pH~7. Solutions with this pH also remove thin oxide layers and leave the Si surface hydrophobic. However, improvements in Si(100) surface quality using similar pH values are at best marginal.[4]

The accessible range of pH values for HF:NH_4F:H_2O mixtures is from approximately 1.6 for 1:30 HF(49%):H_2O to 8 for NH_4F. However, with control of pH by adding strong acids (such as HCl, one can control the relative amounts of species (H_2F_2, HF, HF_2^-, F^-) in solution. At low values of pH, the H_2F_2 and HF species predominate, which may effect the etching process. In one such study, HCl (0.5 mol/ℓ) was added to various HF solutions to describe the kinetics of the SiO_2 etch rate as a function of solution species and the net effect on hydrogen passivation of the surface.[5] However, the effect on surface smoothness and the use of other acids with moderate oxidizing potentials were not studied over a range of pH values. As seen in other studies, the variation of pH has significant effects on surface smoothness in Si(111). The influence of the concentrations of these solution species with respect to etching different Si surfaces clearly needs additional study,[6,7] in particular to determine which species or combination of species promotes the smoothest Si(100) surfaces.

This article presents a study of the influence of HF:H_2SO_4:H_2O mixtures of pH values between -0.33 and 1.6 on the surface roughness of Si(100) and on the electrical properties of Si–SiO_2 interfaces formed on these surfaces. This work thereby extends etching studies on Si(100) to significantly lower pH values than could be obtained for the more conventional HF:NH_4F:H_2O mixtures as well as providing an oxidizing anionic species. The relative surface roughness for different etching solutions has been determined by measuring the pseudodielectric function $\langle\epsilon\rangle=\langle\epsilon_1\rangle+i\langle\epsilon_2\rangle$ by spectroscopic ellipsometry (SE) and comparing these results with reference data[8] obtained on atomically flat Si(111) surfaces prepared by an RCA clean followed by immersion in a 40% solution of NH_4F in H_2O. The electrical characteristics of metal-oxide-semiconductor (MOS) structures fabricated on these surfaces by thermal oxidation and RPECVD were determined and compared with

[a]Electronic mail: aspnes/@unity.ncsu.edu

the SE results. We found a good correlation between SE-determined smoothness and MOS performance for the RPECVD oxides, but none for the thermally grown oxides.

II. EXPERIMENTAL PROCEDURES

Two inch p-type Si(100) wafers with a resistivity of 2–6 Ω cm were initially subjected to a conventional hot RCA clean in which the sacrificial chemical oxides were removed in 1:10 dilute HF:H_2O. This was followed by a conventional wet thermal oxidation at 950 °C during which 100 nm of SiO_2 was grown. The thermally grown oxide was then etched off in different HF:H_2SO_4:H_2O mixtures with pH values between about −0.33 and 1.6 immediately before (i) optical characterization by SE, and (ii) either (a) RPECVD or (b) thermal oxidation to generate an oxide layer for a MOS device structure. For the oxide removal step, etching solutions with different pH values were obtained by adding 1–20 ml H_2SO_4(~98%) to a 5 ml:150 ml HF(49%):H_2O mixture. pH values are calculated from the system of equilibrium equations presented by Warren,[9] using K_{a2} (H_2SO_4)=0.012 and the appropriate mass balance. Direct measurement of pH with standard electrodes is not feasible due to their reaction with HF. The samples were taken out of the etching solution (at room temperature) after complete removal of the thermal oxide as evidenced by their surfaces becoming hydrophobic. Ellipsometric data were obtained in the spectral range of 206–450 nm. Care was taken to maintain each etched surface under a flowing dry N_2 atmosphere before and during SE measurement. These data were converted to $\langle\epsilon\rangle$ form. In another set of experiments, water and methanol rinses were performed after the etching step to remove any residual fluoride and/or carbon contamination. The etched Si(100) surfaces prepared in this manner were also studied by SE.

Electrical characterization was performed by capacitance–voltage ($C-V$) and current–voltage ($I-V$) measurements on MOS capacitors fabricated on these surfaces. For the first set of devices the wafers were transferred to an ultrahigh vacuum compatible multichamber system less than one minute after removing the thermally grown oxide layer as described above. Thin oxide layers (~10 nm) were prepared by a 300 °C two-step RPECVD process followed by a 30 s rapid thermal anneal at 900 °C in a 0.1 Torr oxygen ambient.[10] For the second set of devices, 7.5 nm thermal oxides were grown at 900 °C. The time between etching and loading of the wafers into the thermal oxidation system was approximately five minutes. As an additional control, $\langle\epsilon\rangle$ spectra were obtained between 412 and 755 nm from Si samples etched in the same solution as the plasma-processed wafers. The gate oxide layer thicknesses were also determined using ellipsometry, and cross checked using the data obtained from $C-V$ measurements. For the MOS capacitors, a 350-nm-thick metal electrode was formed by vapor depositing Al. Individual capacitors with areas of 6.3×10^{-4} cm^2 and 2.3×10^{-4} cm^2 were prepared for the $C-V$ and $I-V$ studies, respectively. Areas were defined by standard photolithography and wet-chemical-etching techniques. A 150-nm-thick Al film formed the backside contact to the Si wafer.

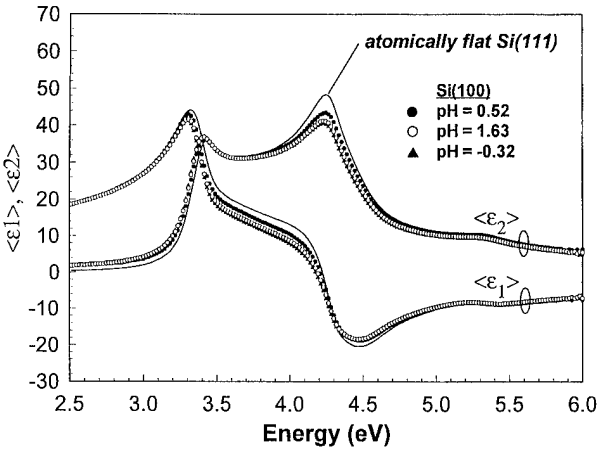


FIG. 1. Spectral dependence of the real and imaginary parts of the pseudodielectric function $\langle\epsilon\rangle=\langle\epsilon_1\rangle+i\langle\epsilon_2\rangle$ for three Si(100) surfaces after removing a thermally grown oxide in HF:H_2SO_4:H_2O etches with pH values of −0.32, 0.52, and 1.59. These are compared with the $\langle\epsilon\rangle$ spectrum of atomically flat H-terminated Si(111) (Ref. 8).

The Al depositions were followed by a postmetallization anneal at 400 °C for 30 min in N_2. Additional control capacitors were made on the same type of Si(100) wafers without any preoxidation processing; i.e., the wafers were not subjected to the RCA clean, the rinse in the 1:10 dilute HF:H_2O etch, or the thermal oxidation and oxide removal in HF:H_2SO_4:H_2O solutions.

III. EXPERIMENTAL RESULTS

Figure 1 shows $\langle\epsilon\rangle$ spectra of three Si(100) samples obtained after removing the 100 nm thermal oxide by etching in solutions with pH values of −0.32, 0.2, and 1.6. A reference spectrum[8] obtained from an atomically flat H-terminated Si(111) sample prepared by an RCA clean followed by etching in a 40% NH_4F solution is also shown for comparison. Differences between the Si(100) and Si(111) spectra are minimized for a pH of ~0.5. A simple measure of effectiveness is the height of the 4.25 eV E_2 peak in $\langle\epsilon_2\rangle$, which for thin oxide layers decreases by 6.4/nm from its atomically flat Si(111) value of 48.3. For rough surfaces, the corresponding decrease in peak height is 9.2/nm. The peak value of $\langle\epsilon_2\rangle$ is shown as a function of pH in Fig. 2. The data reveal that the highest value of $\langle\epsilon_2\rangle$ occurs for a pH of about 0.5. Decreases in $\langle\epsilon_2\rangle$ at the E_2 peak have generally been attributed to (i) surface contamination by hydrocarbons, etc., and/or (ii) surface roughness.[11,12] Surface contamination resulting from HF and H_2SO_4 solutions has been studied by Auger electron spectroscopy (AES) and x-ray photoelectron spectroscopy.[13] These studies showed the presence of relatively small amounts of carbon and/or fluoride contamination (in the percent range), which could be removed by a rinse with methanol and water. To establish the extent to which this contamination could contribute to these data, we also show in Fig. 2 the difference between as-etched and methanol/water-rinsed surfaces. The difference is consistently about 1, indicating that the dominant reduction is due to roughness. If $\langle\epsilon_2\rangle$ values were affected from organic im-

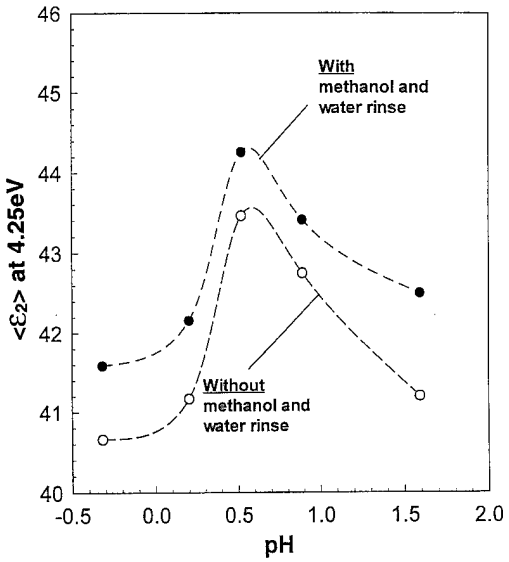


FIG. 2. Peak values of $\langle \epsilon_2 \rangle$ at 4.25 eV as a function of pH of etching solution with (●) and without (○) a methanol–water rinse.

purities, there would be an expected increase in $\langle \epsilon_2 \rangle$ with H_2SO_4 concentration due to oxidation of organic molecules. However, $\langle \epsilon_2 \rangle$ decreased at pH<0.5 (higher $[H_2SO_4]$) further supporting a surface roughness explanation for the observed $\langle \epsilon_2 \rangle$ values. We conclude that the *smoothest* Si(100) surfaces are obtained for a pH 0.5±0.1, and that surfaces prepared in both lower and higher pH mixtures are rougher. We note that the $\langle \epsilon_2 \rangle$ peak of 41.1 obtained by Utani and Adachi[12] with a 60:1 dilution of HF in H_2O (pH~1.8) is entirely consistent with the trend established in Fig. 2.

Figure 3 shows the calculated concentration of reactive species in the etching solution where the points a, b, c, d, and e correspond to pH values of 1.6, 0.9, 0.52, 0.2, and −0.32, respectively, for the data shown in Fig. 2. As can be seen, the

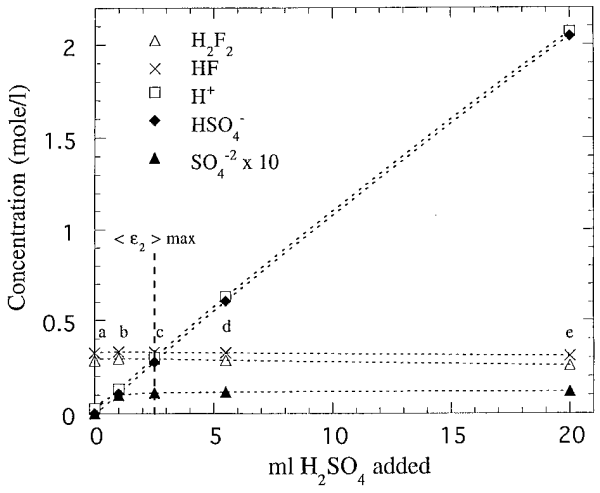


FIG. 3. Neutral and ionic species in solution calculated from Ref. 9 for etching solutions used in Fig. 2 as a function of added H_2SO_4. Points a, b, c, d, and e correspond to solution pH values of 1.6, 0.9, 0.52, 0.2, and −0.32, respectively, with point c resulting in the maximum $\langle \epsilon_2 \rangle$ of Fig. 2.

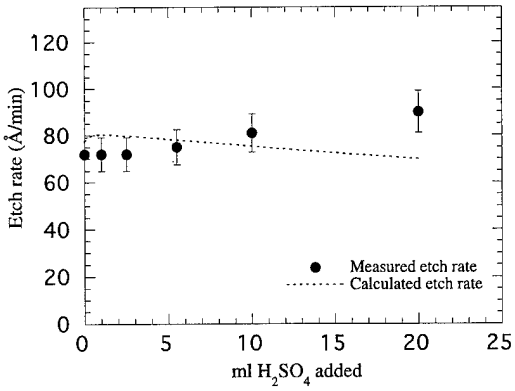


FIG. 4. Measured etch rates of sacrificial oxides in 155 ml 1:30 HF (49%) solutions as a function of added H_2SO_4 compared to calculated rates from Ref. 5 (dashed line) based on the concentrations of [HF] and $[H_2F_2]$.

concentration of HF and H_2F_2 remains nearly constant while HSO_4^- and SO_4^{2-} increase. At pH values below 0.9, only HSO_4^- increases. Since sulfate anions are known to be oxidizers, a possible mechanism for the observed increase in smoothness is the selective oxidation of Si facets to SiO_2 which is, thus, removed by HF. There are two important points that support a selective surface oxidation/etching model for surface smoothing. The first is that the same experiments using HCl instead of H_2SO_4 to decrease pH did not result in any improvement in surface roughness. Cl^- has significant etching power. Second, H_2SO_4 was seen to not dramatically effect the etching rate, thus, not to be an intimate part of the HF/SiO_2 etching process. Figure 4 shows the SiO_2 etching rate as a function of added H_2SO_4 compared to calculated etching rates based on concentrations of [HF] and $[H_2F_2]$. As can be seen, pH values as low as that which resulted in the smoothest Si(100) surface did not change the SiO_2 etch rate. This strongly suggests that the mechanism for Si surface smoothing is not the effect of the anion on the process of SiO_2 etching by HF. It should be noted that at very low pH values there is a moderate increase in the etch rate, however, the ionic strength of this solution is very high and deviations are not surprising. The effect of ionic strength is difficult to study since conventional methods of adding concentrate spectator ions (such as KCl) would have a detrimental effect on the cleaning process.

As further evidence of the trends previously discussed, Fig. 5(a) displays $\langle \epsilon_2 \rangle$ values at 2.81 eV for the control samples described above, i.e., for the Si(100) samples stripped in the solutions used to prepared these surfaces for plasma-deposited oxides. At 2.81 eV $\langle \epsilon_2 \rangle$ values for the smoothest surfaces are lowest, providing a confirming perspective of the results shown in Fig. 2.

The smoother surfaces as characterized by SE would be expected to translate to increased performance of MOS capacitors and is explored for both PECVD and thermally grown oxide gate dielectrics. The concentration of interface defects (D_{it}) at midgap, determined by conventional high frequency and quasistatic $C-V$ measurements averaged over at least 4 devices, is shown in Fig. 5(b). For the plasma-deposited oxides the D_{it} values show a minimum for pH

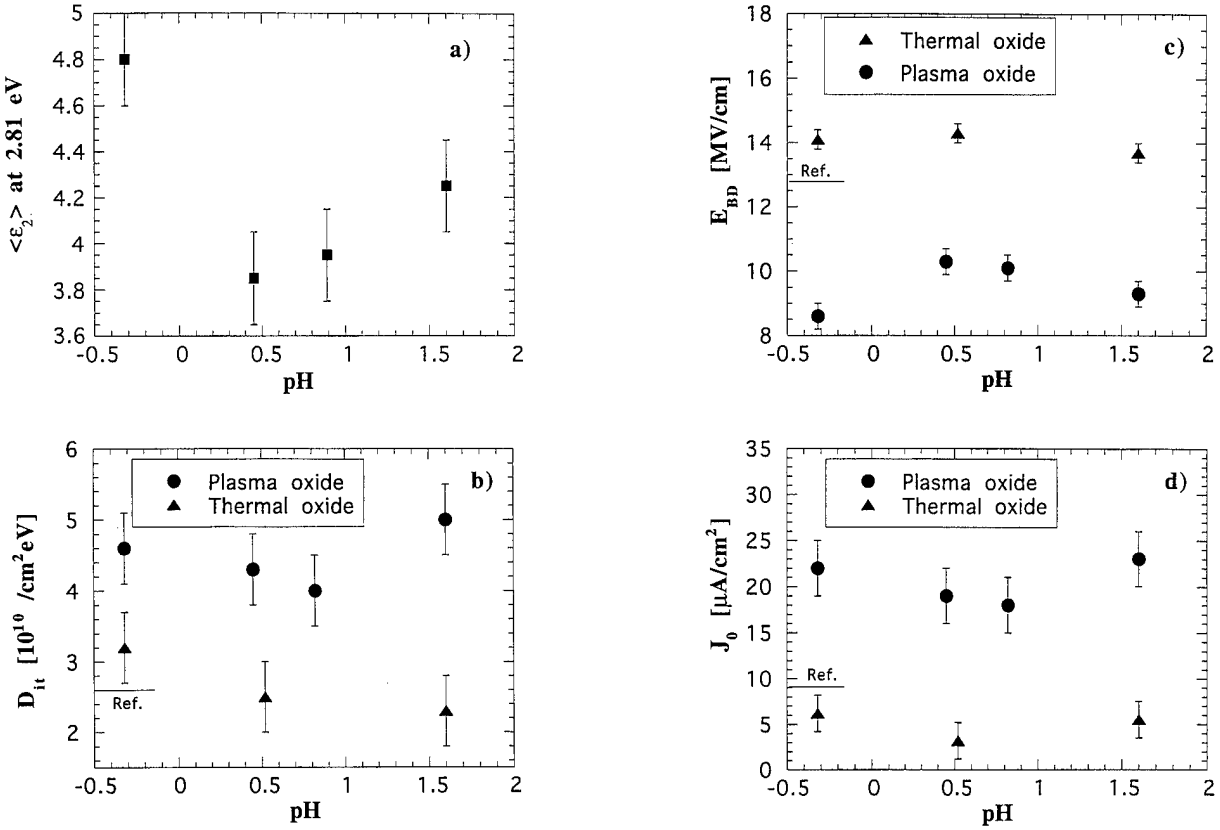


FIG. 5. Comparison of various surface, interface, and oxide data for etched surfaces and devices as a function of pH: (a) $\langle\epsilon_2\rangle$ values at 2.81 eV for wafers etched in the same solution as the substrates for the thermal and plasma oxides; (b), (c), and (d): midgap D_{it}, breakdown field E_{BD}, and tunneling current J_0 values, respectively, of MOS capacitors fabricated on these surfaces with plasma and thermal oxides. The lines in (b), (c), and (d) labeled reference are MOS capacitors with thermal oxides fabricated on surfaces not subjected to preoxidation etching.

values of about 0.5. For the thermal oxides the D_{it} values are lower by a factor about 0.5 with respect to the plasma-deposited oxides, and these D_{it} values display a different dependence on pH. For example, for the thermal oxides there is very little difference in D_{it} for pH values of 0.5 and 1.8, whereas D_{it} is increased slightly for a pH value of -0.32. Note further that the average D_{it} value for the reference thermal oxide devices without preoxidation processing falls in the middle of the range of D_{it} values for the thermal oxides that were etched. The breakdown field (E_{BD}) data in Fig. 5(c) show that MOS devices with plasma oxides display lower E_{BD} values than those for MOS devices with thermal oxides, and in addition display a pH dependence paralleling that of D_{it} in the sense that the best performance (highest E_{BD}) occurs at a pH of ~0.5. The devices with thermal oxides display a much weaker dependence on pH; however, paralleling their D_{it} behavior, the performance of devices subjected to pH values of 0.45–1.6 display improved performance with respect to devices formed on wafers with no preoxidation processing. Finally, as shown in Fig. 5(d) the lowest tunneling current density (J_0) for a bias of 5 V in the Fowler–Nordheim regime for devices with both thermal and plasma oxides is obtained at a pH of ~0.5. Note further that the dependences of D_{it} and J_0 on pH for the plasma-deposited oxides are essentially the same. Finally, the value of J_0 for the reference wafer is higher than those of the thermal oxides that had been subjected to preoxidation processing.

IV. DISCUSSION

Our ellipsometric studies of Si(100) surfaces etched in H_2SO_4:HF:H_2O solutions with pH values between -0.32 and 1.7 indicate that the smoothest surfaces are obtained with a 1:0.50:30 HF(49 wt %):H_2SO_4(98 wt %):H_2O etch with a pH of 0.52. However, the values of $\langle\epsilon_2\rangle$ at the E_2 peak for these surfaces were significantly smaller than that obtained from atomically flat H-terminated Si(111) surfaces prepared with an RCA clean followed by a final etch in a 40% NH_4F:H_2O mixture.[8] The best Si(100) surface had a peak $\langle\epsilon_2\rangle$ value of 44.3, which can be modeled to correspond to either 0.64 nm of oxide or 0.45 nm of surface roughness. These smaller peak values of $\langle\epsilon_2\rangle$ are here interpreted as showing that the Si(100) surfaces are rougher on an atomic scale than optimized Si(111) surfaces. $C-V$ and $I-V$ characterizations of MOS structures formed on these surfaces were consistent with the SE measurements in the sense that the best electrical performance was obtained when sacrificial oxides were removed before device fabrication using H_2SO_4:HF:H_2O mixtures with a pH~0.5, the same pH that gave the highest values for the E_2 peak of $\langle\epsilon_2\rangle$. As noted above, the pH value of 0.5 falls outside the range accessible with HF:NH_4F:H_2O solutions that are conventionally used to

prepare Si surfaces. Preliminary SE studies on Si(100) samples with doping concentrations other than those used above show similar trends in surface roughness with the *p*H of the etching solution.

More generally, we found no improvement relative to conventional etching in dilute 1:30 HF:H_2O when the *p*H was lowered by adding HCl instead of H_2SO_4. This result establishes that *p*H is not the sole factor in either promoting or reducing surface roughness on Si(100), but that specific anions can also play a role. The SiO_2 etching rate was not effected by H_2SiO_2 concentration indicating that the sulfate anion is not prevalent in the SiO_2/HF etching mechanism. Since H_2SO_4 is an oxidizer, there is the possible mechanism of selective Si surface oxidation to SiO_2 with subsequent removal by HF. Importantly, there is a report of the oxidizer H_2O_2 being added to the final step of a NH_4OH buffered HF etch (high *p*H) resulting in markedly smoother films as detected by scanning tunneling microscopy.[14] Strong oxidizing agents such as HNO_3 and $HClO_4$ would be expected to not be selective. The role of the sulfate anion and other mildly oxidizing or kinetically hindered oxo-acids as well as electrically biasing the potential of the substrate are being currently investigated.

Finally, the results presented for thermally grown oxides also demonstrate that even if Si(100) surface roughening does occur in predeposition processing, the effects of this roughening are largely *removed* during thermal oxidation. In contrast, correlations between preoxidation wafer processing and the performance of MOS devices involving thermally grown oxides have been observed for Si(111) indicating a quantitatively different relationship for surface smoothing effects occurring during thermal oxidation of this surface.[4] In contrast, the results for Si(100) plasma-deposited oxides parallel those for Si(111), i.e., the *p*H of the etch used to remove sacrificial oxides prior to plasma processing does have a significant effect on MOS device performance.[4] These differences in behavior for devices with thermally grown and plasma-deposited oxides for Si(100) and Si(111) substrates indicate that considerable care must be exercised in comparing the electrical performance of devices with thermally grown and/or deposited oxides, in particular when attempting to correlate trends in device performance with surface roughness arising from predeposition surface treatments.

ACKNOWLEDGMENTS

This research was supported by NSF, ONR, and SRC.

[1]T. Ohmi, M. Miyashita, M. Itano, T. Imaoka, and I. Kawanabe, IEEE Trans. Electron Devices **39**, 537 (1992).
[2]D. Graef, M. Brohl, S. Bauer-Mayer, A. Ehlert, P. Wagner, and A. Schnegg, Mater. Res. Soc. Symp. Proc. **315**, 23 (1993).
[3]M. M. Heyns *et al.*, Mater. Res. Soc. Symp. Proc. **315**, 35 (1993).
[4]T. Yasuda, Y. Ma, Y. L. Chen, G. Lucovsky, and D. Maher, J. Vac. Sci. Technol. A **11**, 945 (1993).
[5]S. Verhaverbeke, H. Bender, M. Meuris, P. W. Mertens, H. F. Schmidt, and M. M. Heyns, Mater. Res. Soc. Symp. Proc. **315**, 457 (1993).
[6]W. Vandervorst, H. Bender, M. Meuris, P. W. Mertens, H. F. Schmidt, and M. M. Heyns, Microelectron. Eng. **28**, 27 (1995).
[7]F. Tardif, T. Lardin, C. Paillet, J. P. Joly, B. Beneyton, P. Patruno, D. Levy, K. Barla, and W. Sievert, Microelectron. Eng. **28**, 121 (1995).
[8]T. Yasuda and D. E. Aspnes, Appl. Opt. **33**, 7435 (1994).
[9]L. J. Warren, Anal. Chim. Acta **53**, 199 (1971).
[10]G. Lucovsky, Y. Ma, S. V. Hattangady, V. Misra, J. J. Wortman, U. Emmerichs, C. Meyer, and H. Kurz, Jpn. J. Appl. Phys. **33**, 7061 (1994).
[11]D. E. Aspnes, J. Vac. Sci. Technol. **18**, 289 (1981).
[12]K. Utani and S. Adachi, Jpn. J. Appl. Phys. **32**, 3572 (1993).
[13]T. Bitzer, Ph.D. thesis, Technical University, Berlin, Germany (1993).
[14]T. Ohmi, T. Tsuga, and J. Takano, in *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993), p. 257.

Schwoebel barriers on Si(111) steps and kinks

S. Kodiyalam,[a] K. E. Khor, and S. Das Sarma
Department of Physics, University of Maryland, College Park, Maryland 20742-4111

(Received 29 February 1996; accepted 4 May 1996)

Motivated by our previous work using the Stillinger–Weber potential, which shows that the $[\bar{2}11]$ step on the relaxed 1×1 Si(111) has a Schwoebel barrier of 0.61 ± 0.07 eV, we calculate here the same barrier corresponding to two types of kinks on this step—one with rebonding between upper and lower terrace atoms (type B) and the other without (type A). From the binding energy of an adatom, without additional relaxation of other atoms, we find that the Schwoebel barrier must be less than 0.39 eV (0.62 eV) for the kink of type A (type B). Such a bound is argued to be a robust feature following from the presence of rebonding at the step edge or kink site. From the true adatom binding energy we determine the Schwoebel barrier to be 0.15 ± 0.07 eV (0.50 ± 0.07 eV) for the kink of Type A (B). The decrease in the Schwoebel barrier is roughly consistent with our previous estimates of its upper bound of 0.05 eV. However, as the true binding energy plots show discontinuities due to significant movement of atoms at the kink site, we speculate on the possibility of multi-atom processes having smaller Schwoebel barriers. © *1996 American Vacuum Society.*

I. INTRODUCTION

The Schwoebel barrier, which was introduced as the additional barrier for adatom diffusion over a step edge from the upper to lower terraces,[1] has been a subject of current interest for its influence on the growth of a singular (flat) surface.[2–6] It was pointed out by Villain[7] that in the presence of such a barrier, growth by step flows was stable only if the surface was sufficiently vicinal with possible instabilities setting in during the growth a flat surface. It is now accepted that it leads to a coarsening of the evolving surface morphology.[2–6] The present study is motivated by recent observations of another kind of instability on the high temperature 1×1 phase of Si(111)—the reversible step bunching instability during sublimation.[8,9] As these experiments can be reinterpreted in the presence of the Schwoebel barrier,[10] we summarize here our previous calculations of the same corresponding to straight (high symmetry) steps[10] and present results corresponding to kinked steps, both of which use the empirical Stillinger–Weber (SW) potential. The use of this potential here (and in our previous study[10]) has been motivated by the fact that features that follow from changes in coordination number (of the adatom probing the potential energy topography) are expected to survive even if the details of the empirical potential used change. We attempt to identify such features here.

Table I summarizes our previous results.[10] The straight (high symmetry) $[\bar{2}11]$ and $[\bar{1}\bar{1}2]$ and other step orientations are shown in Fig. 1. Note that the $[\bar{2}11]$ step shows a large Schwoebel barrier of 0.61 ± 0.07 eV. However, an analysis of experimental data on the electromigration of steps[9] using a diffusion equation showed that the upper bound on the Schwoebel barrier[10] (in a particular limit of the equation parameters) was very small (0.05 eV). Therefore, here we calculate the Schwoebel barrier corresponding to unit depth kinks on the $[\bar{2}11]$ step to determine if it continues to be large.

II. MOLECULAR DYNAMICS METHOD

Using the SW potential, diffusion barriers are determined by mapping the adatom potential energy a function of the (x,y) position of the adatom [in the (111) plane] for two types of kinks on the $[\bar{2}11]$ step—type A [see Fig. 2(a)], which has rebonding of atoms along the kink site and type B [see Fig. 2(b)], which has, at the kink site, an upper terrace atom rebonded to the lower terrace. Here, the adatom is nothing but an additional atom (not belonging to the bare 1×1 Si(111) surface that probes the potential energy topography). It must be noted that the rebonding in these configurations (and in our previous study on straight steps[10]) is similar to that on the Si(001) surface whose reconstruction energy predicted the SW potential (0.85 eV)[11] is consistent with *ab initio* calculations.[12] The energies (per step edge atom) of the straight $[\bar{2}11]$ and $[\bar{1}\bar{1}2]$ steps (configuration types A) are reproduced by the SW potential in the same order (0.72 eV and 0.62 eV, respectively[13]) as compared to tight binding calculations (0.70 eV and 0.38 eV, respectively[14]). Moreover, the kink energy of the type A configuration studied here (SW prediction: 0.22 eV)[13] is roughly consistent with that derived from experiments (0.23 eV)[15,16] on fluctuating steps on 1×1 Si(111). The type B configuration however has a very small (SW prediction:≈-1 meV)[13] kink energy. In our previous study on straight steps, configurations of type B were neglected since the step ($[\bar{1}12]$) that allowed for this structure would have very large step-step interactions inconsistent with experimental estimates.[13] However, the large interactions were due to the presence of one rebonding atom per lattice constant along the step edge. The configuration of type B is nevertheless considered here, since (besides being energetically favorable according to the SW potential), at low kink densities, i.e., for a nearly straight $[\bar{2}11]$ step, the number of such rebonding atoms per unit length along the

[a]Electronic mail: sanjay@faldo.umd.edu

TABLE I. V is the adatom potential energy. Shown here is its value at the relevant minimum and saddle points.

Global minimum in $V = -3.31 \pm 0.05$ eV			
Configuration	V at saddle point (eV)	⇒Diffusion barrier (eV)	⇒Schwoebel barrier (eV)
Si(111) surface	-2.34 ± 0.05	0.97 ± 0.07	...
$[\bar{2}11]$ step	-1.73 ± 0.05	1.58 ± 0.07	0.61 ± 0.07
$[1\bar{1}2]$	-2.18 ± 0.05	1.13 ± 0.07	0.16 ± 0.07

step edge would be correspondingly low and therefore the expected step-step interactions would be smaller. The adatom potential energy V has been computed as the difference in the minimum potential energy of the system with the adatom at infinity (non-interacting) and the same with the interacting adatom.

Standard molecular dynamics (MD) procedures of integrating Newton's law (with dissipation to reduce temperature) and the steepest descent equations have been used to determine the minimum potential energy of the system. These routines determined the adatom potential energy to an precision of 10^{-4} eV. The (x,y) coordinates of the adatom are fixed during the integration process. The system consisted [as before (Ref. 10)] of six bi-layers of Si(111) in an MD cell, the bottom three layers of which are fixed at bulk lattice coordinates throughout the calculation. The system size along the x axis (ℓ_x) (which was parallel the $[\bar{2}11]$ step edge) was $5\frac{1}{2}a_1$ and along the y axis (ℓ_y) was $3\frac{2}{3}a_2$. Periodic boundary conditions along the x axis identified the points $(0,0,h)$ (h is the step height) and $(0,\ell_y,0)$ and the same along the y axis identified the points $(0,k,0)$ (k is the kink depth) and $(\ell_x,0,0)$. These boundary conditions made it possible to have exactly one vicinal step with one kink (and no antikinks) in the MD cell. As before,[10] the atoms on the (x,y) boundaries were however held fixed during the computation of V to prevent the entire configuration from shearing, particularly when the adatom is moved away from a deep minimum. The kinks were roughly in the middle of the cell consisting of movable atoms [see Figs. 3(a) and 4(a)]. As the $(x-y)$ size of this cell is larger than that used previously[10] in studying the system size dependence of V on the Si(111) surface and since it is more square in shape, we expect the error in V due to finite size effects to be smaller than before[10] (0.01 eV). It must be additionally noted that with the system size used here (with the adatom absent), the kink energy was within 6 meV (4 meV) of that calculated previously[13] for the kink of type A (type B).

The MD procedures began with the initial configurations for each (x,y) position of the adatom corresponding to the relaxed adatom-free structures. The z coordinate of the adatom was then varied in small steps in a wide range to roughly determine the point (z_0) at which its potential energy is the smallest. Initializing this z coordinate at z_0, the integration procedures were followed, first with only the adatom relax-

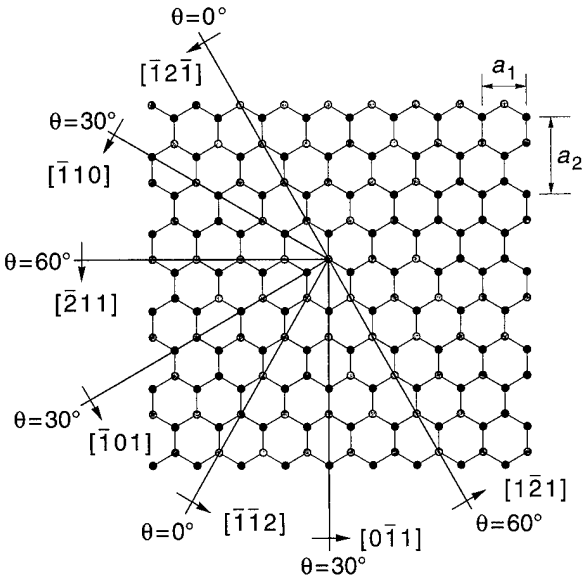


FIG. 1. One bilayer of the Si(111) surface consisting of the upper monolayer (gray) and lower monolayer (black). The figure shows the threefold and reflection symmetry of this surface: Steps running along directions with equal θ are identical. (Borrowed from Ref. 13).

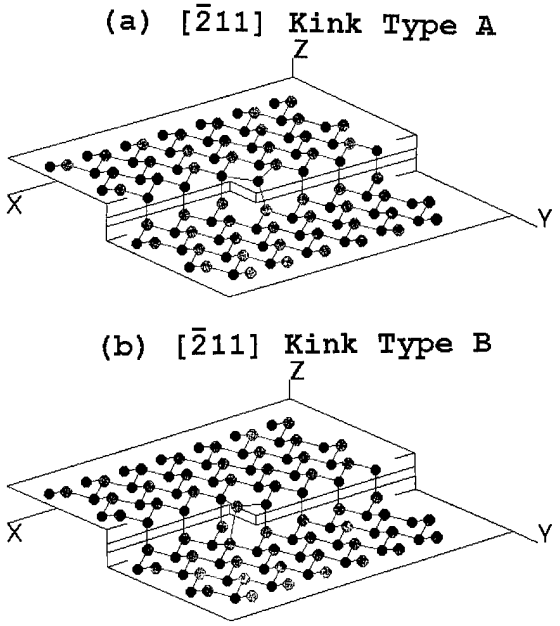


FIG. 2. The two kinks studied here: type A has atoms rebonded along the kink site whereas type B has, at the kink site, an upper terrace atom rebonded to the lower terrace. On both terraces the upper monolayer is shown in grey and the lower monolayer in black. (Borrowed from Ref. 13).

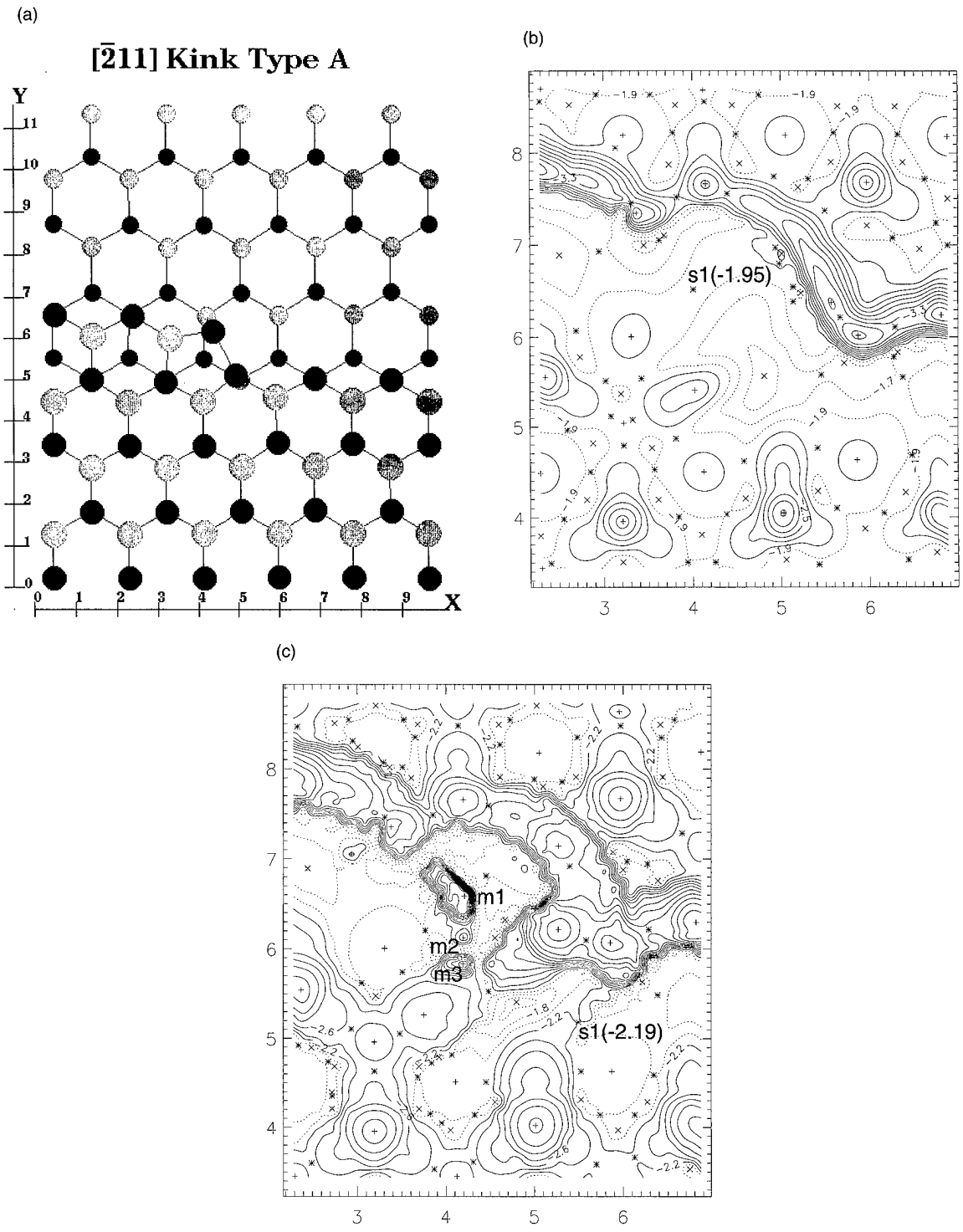


FIG. 3. Shown in (a) is the top view of the kink with rebonding along the kink site with the upper terrace atoms larger than those on the lower terrace. (b) shows the corresponding adatom potential energy derived from the local geometry (V_{lg}) whereas (c) shows the true adatom potential energy (V). In both plots, contours are separated by 0.2 eV with the minima, saddle points, and maxima marked (and sometimes labeled) by +(m), *(s), and ×(M), respectively. Figures in parenthesis are corresponding values in eV. Contours in (b) $\geqslant -1.9$ eV and those in (c) $\geqslant -2.0$ eV are marked with dashed lines.

ing while other atoms remain fixed. The minimum of the potential energy reached this way (V_{lg}) depends only on the local geometry of atoms around the adatom. The atoms that were held fixed are then allowed to relax together with the adatom to recompute the minimum which is now the true adatom potential energy (V). With the kink roughly in the middle of the region explored [see Figs. 3(b), 3(c), 4(b), 4(c)], V and V_{lg} are computed on a rectangular grid with the spacing between points being $a_1/16$ ($a_2/30$) along the $x(y)$ axis for a total length of $2\frac{1}{2}a_1$ ($1\frac{2}{3}a_2$). An interpolation

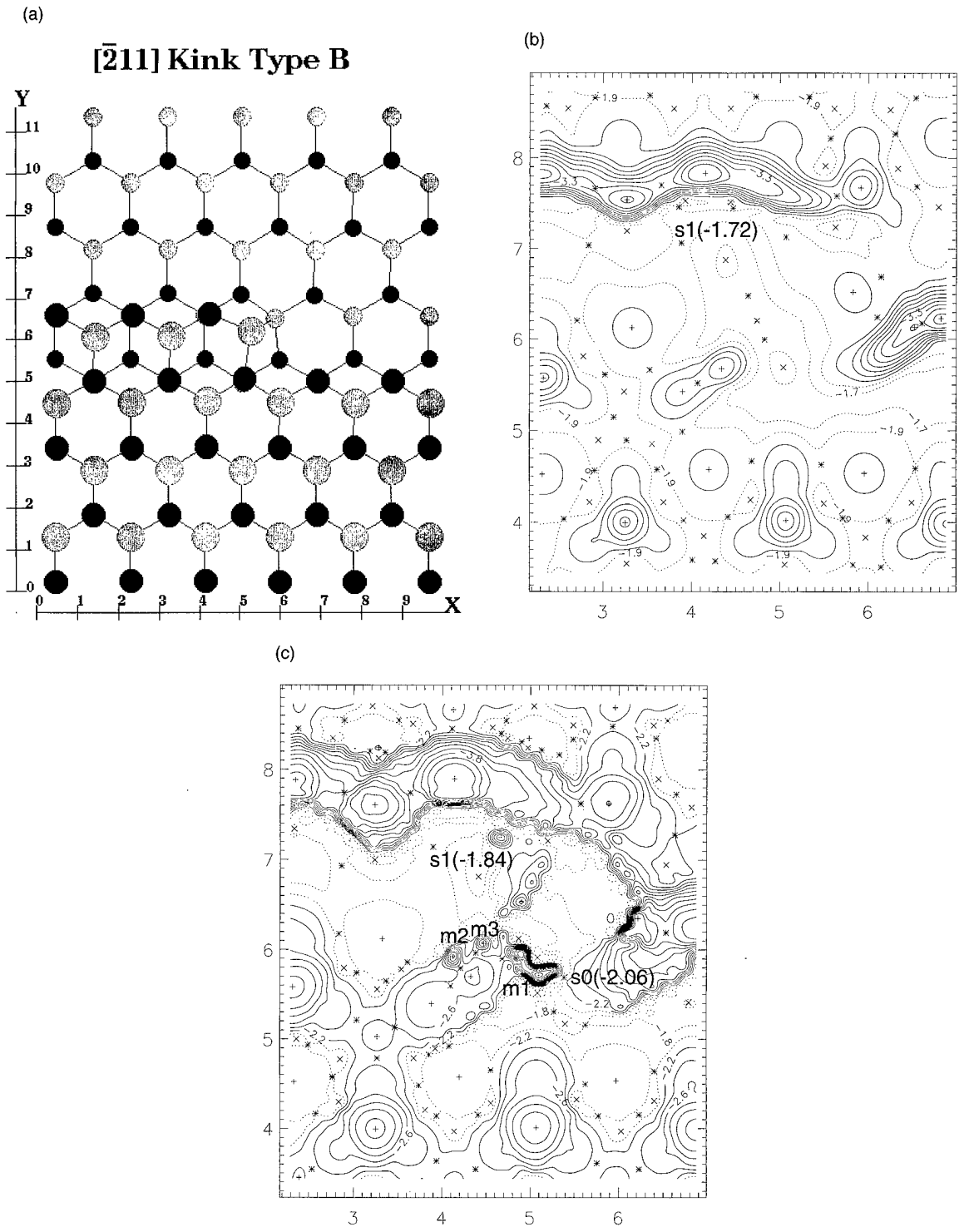


FIG. 4. Shown in (a) is the top view of the kink with an atom of the upper terrace (larger atoms) rebonded to the lower terrace. (b) shows the corresponding adatom potential energy derived from the local geometry (V_{lg}) whereas (c) shows the true adatom potential energy (V). In both plots, contours are separated by 0.2 eV with the minima, saddle points, and maxima marked (and sometimes labeled) by +(m), *(s), and ×(M), respectively. Figures in parenthesis are corresponding values in eV. Contours in (b) ≥ -1.9 eV and those in (c) ≥ -2.0 eV are marked with dashed lines.

scheme applying periodic boundary conditions along the x and y axes similar to that used in the simulations, was used to construct the contour plots of Fig. 3. From our previous study (Ref. 10) the error in V due to (the same) finite grid size and similar interpolation scheme was estimated to be ± 0.05 eV. We therefore assume here that this error remains the same. Being much larger than the errors due to finite size effects, it is assumed to be the error bar in V. Barrier values being differences are therefore estimated to have an error bar of ± 0.07 eV.

III. DISCUSSION

It must be noted that the SW potential was constructed to reproduce only the bulk diamond and liquid properties of Silicon.[17] No additional tuning to the (111) surface/step properties has been carried out. Specifically, this potential does not reproduce the correct energetics, for the (111) surface configurations with adatoms[18] as compared to *ab initio* calculations.[19–21] However, as mentioned previously, it satisfactorily predicts energy differences when rebonding occurs—as was the case with the Si(001) surface,[11] the relative energies of the straight $[\bar{2}11]$ and $[\bar{1}\bar{1}2]$ steps[11,14] and the energy of $[\bar{2}11]$ kink of type A.[13,15,16] These observations suggest that the energetics predicted by this potential is reliable when atoms undergo a change in their coordination number. We use this idea in an attempt to identify the robust features of this study—those that follow from changes in adatom coordination number. As the functional form of the SW potential does not provide an independent measure of atomic coordination, here we assume [as before (Ref. 10)] that the potential energy (V_{lg}) obtained from relaxing only the adatom, over the relaxed (but fixed) adatom free configurations, to be a good measure of the coordination number of the adatom. Although very similar to the true adatom potential energy (V) it helps to identify features that follow directly from the local geometry of atoms around the adatom. Hence, *assuming* (for barrier calculations on the Si(111) surface/step/kink configurations) that the SW potential is more reliable when it predicts an absence of additional relaxation (due to the presence of the adatom) which if present changes the potential energy topography significantly, it follows [as argued previously (Ref. 10)] that features that follow from a strong correlation between V_{lg} and V are robust, i.e., would survive changes in details of the empirical potential used and are expected to be reproduced in more satisfactory *ab initio* or tight binding calculations. Specifically, we had assumed previously (Ref. 10) that if the saddle point determining the diffusion barrier in V_{lg} was nearly at the same position as that in V then the diffusion/Schwoebel barrier is a robust feature. Even if the position of the relevant saddle points differed, there is however a bound on the true Schwoebel barrier that follows purely from V_{lg} [this is the difference between V_{lg} at barrier determining saddle point on the step or kink configurations and V at the diffusion barrier determining saddle point on the free Si(111) surface]. It is a strict upper bound on the Schwoebel barrier (that follows from V on the step or kink configuration) since any relaxation that occurs during the computation of V can only reduce the relevant saddle point energy. As this bound is completely independent of the details in V, we now argue that it must be a robust feature.

From the contour plots of V_{lg} corresponding to the kinks studied here [Figs. 3(b) and 4(b)] as well as the same corresponding to straight steps studied previously[10] [Figs. 5(b) and 5(c)] a common trend emerges—the bound on the Schwoebel barrier (or equivalently the energy of the barrier determining saddle point in V_{lg}) is small (or more negative) when atoms at the step edge or kink site are rebonded. It can be seen (from Figs. 3 and 5) that this saddle point [$s1$ in Fig. 3(b), $s1$ in Fig. 4(b), $s3$ in Fig. 5(b) and $s6$ in Fig. 5(d)] is in the neighborhood of an upper terrace atom that has moved from its bulk terminated position in a direction away from the saddle point due to the presence of rebonding at the step edge or kink site. The magnitude of this displacement (in the $x-y$ plane) is found to be correlated to the saddle point energy which is more negative if the displacement is large. The displacement (and the relevant saddle point energy) for the $[\bar{2}11]$ kinks of types A,B, the $[\bar{2}11]$ step and the $[\bar{1}\bar{1}2]$ step is 0.99 Å (−1.95 eV), 0.21 Å (−1.72 eV), 0.00 Å (−1.63 eV) and 0.72 Å (−1.89 eV), respectively. Therefore the bound on the corresponding Schwoebel barrier is 0.39 eV, 0.62 eV, 0.71 eV, and 0.45 eV, respectively.

The true adatom potential energy V, corresponding to the kink configurations studied here, is shown in Figs. 3(c) and 4(c). These plots show discontinuities near the kink site indicated by the presence of minima, some of which are labeled in the figures. Studying the final atomic configuration when the adatom is in these regions shows that the discontinuities are due to large scale rearrangements of atoms leading to a loss in the identity of the adatom. In other words, in these configurations, the adatom seems to occupy a lattice position after dislodging another atom which now appears to be the adatom. Diffusion in these regions cannot therefore be viewed as a *single* atom process, thereby making the contour plots in these regions less meaningful. The saddle point determining the Schwoebel barrier in case of the $[\bar{2}11]$ kink of type A[$s1$ in Fig. 3(c)] is however not in the proximity of such regions. Therefore the corresponding Schwoebel barrier of 0.15 ± 0.07 eV is assumed to be relevant for processes involving only a single adatom. In the case of the kink of type B, it appears that $s0$ [in Fig. 4(c)] is the relevant saddle point that determines the Schwoebel barrier. However, this point is close to a minimum that indicates a discontinuity in V. The Schwoebel barrier determined by this point (0.28 ± 0.07 eV) is therefore considered as corresponding to a multi atom, and not a single atom, process. Hence, discounting this point, the Schwoebel barrier [determined by $s1$ in Fig. 4(c)] is 0.50 ± 0.07 eV.

In our previous study of barriers over straight steps,[10] discontinuities in V, of the kind that are seen here, were not observed. This was because the temperature used during the simulation was very small—$\approx 2\times10^{-4}$ eV. Here we used a larger temperature ($\approx 3\times10^{-3}$ eV) which resulted, in the presence of the adatom, the large scale movement of atoms around the kink site. The experimentally relevant temperature[9] however continues to be much larger—$\approx$0.1 eV. At these temperatures we expect multi-atom processes to have smaller Schwoebel barriers—this is supported by the observation here that such a process occurs near the $[\bar{2}11]$ kink of type B with the Schwoebel barrier being smaller than that for the single atom process. Carrying out simulations at high temperatures with a *single* probe atom cannot yield useful information since such potential energy plots are expected to suffer from severe discontinuities due to such multi-atom processes. The general decrease in the barrier

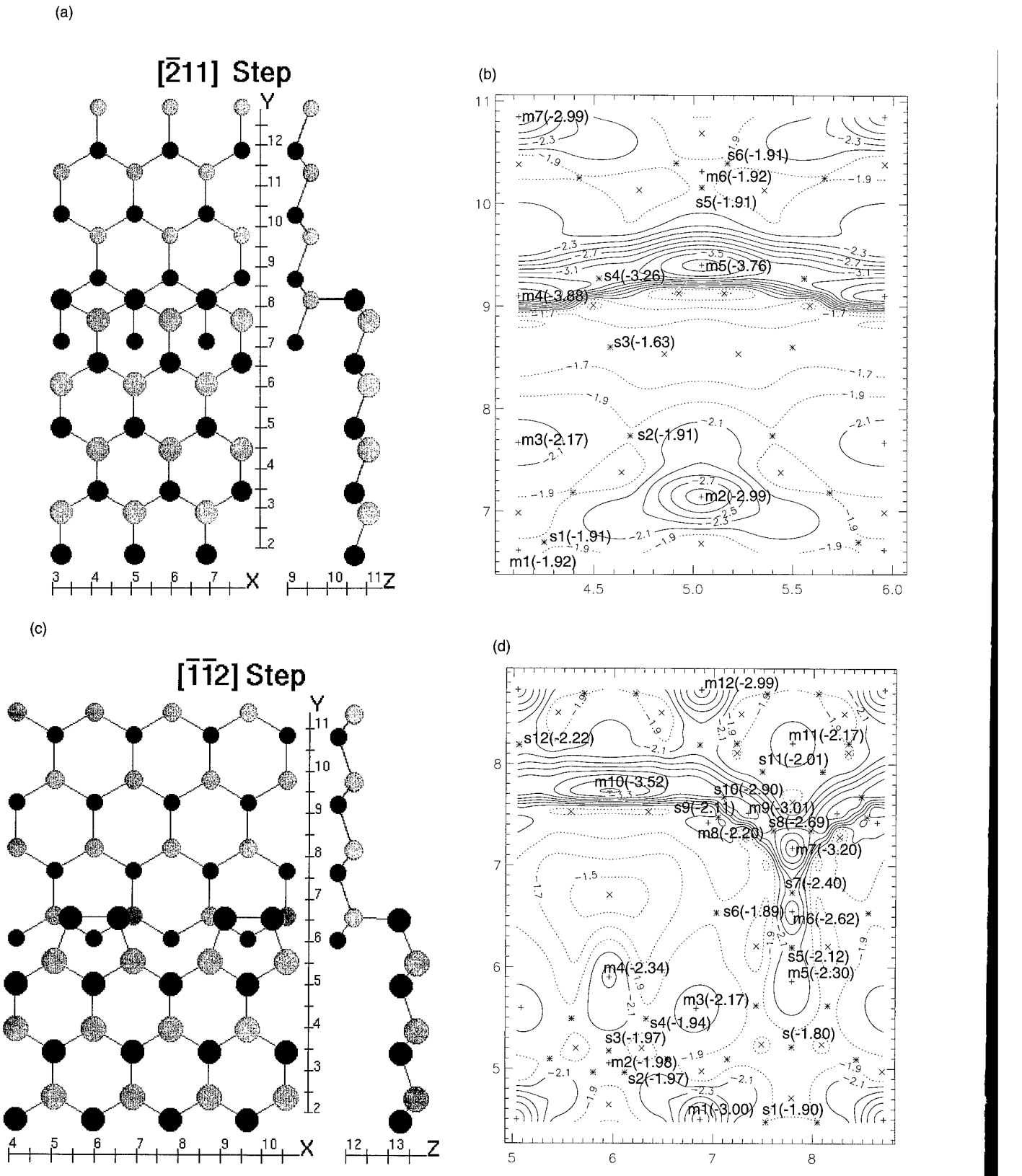


FIG. 5. In the step configurations [(a) and (d)] the upper terrace atoms are shown larger than those on the lower terrace and within each terrace the upper monolayer is in grey and the lower in black. (b) and (d) show the adatom potential energy derived from the local geometry (V_{lg}) corresponding to (a) and (c), respectively. In these plots contours are 0.2 eV apart with those ≥ -1.9 eV shown by dashed lines. The minima, saddle points, and maxima are marked (labeled) by +(m), *(s), and ×(M), respectively, with the figure in parenthesis being their corresponding value in eV.

values is however consistent with the small upper bound on the Schwoebel barrier (0.05 eV) developed previously[10] from an analysis of experimental data on the electromigration of steps.[9]

IV. CONCLUSION

Schwoebel barriers, calculated using the empirical Stillinger–Weber potential, for unit depth kinks along the $[\bar{2}11]$ step are smaller in magnitude as compared to that calculated previously[10] for the straight $[\bar{2}11]$ step. This decrease can be expected directly form the adatom potential energy plots (V_{lg}) that follow purely from the local geometry of atoms around the adatom in these atomic configurations. These plots as well as similar plots for the straight $[\bar{2}11]$ and $[\bar{1}12]$ steps calculated previously[10] show that the upper bound on the true Schwoebel barrier calculated using V_{lg} is correlated to the the displacement of an atom on the upper terrace of these configurations that is near the relevant saddle point in V_{lg}. The true adatom potential energy plots (V) however show discontinuities due to large scale movement of atoms near the kink sites resulting sometimes in smaller barriers than when such movements do not occur. We therefore speculate that multi-atom processes occurring in the configurations studied here and previously[10] could have smaller Schwoebel barriers.

ACKNOWLEDGMENTS

This work has been supported by the NSF-MRG and the U.S. ONR.

[1]R. L. Schwoebel and E. J. Shipsey, J. Appl. Phys **37**, 3682 (1966).
[2]M. D. Johnson, C. Orme, A. W. Hunt, D. Graff, J. Sudijono, L. M. Sander, and B. G. Orr, Phys. Rev. Lett. **72**, 116 (1994).
[3]M. Siegert and M. Plischke, Phys. Rev. Lett. **73**, 1517 (1994).
[4]A. W. Hunt, C. Orme, D. R. M. Williams, B. G. Orr, and L. M. Sander, Europhys. Lett. **27**, 611 (1994).
[5]I. Elkinani and J. Villain, J. Phys. I (France) **4**, 949 (1994).
[6]C. J. Lanczycki and S. Das Sarma, Phys. Rev. Lett. **76**, 780 (1996).
[7]J. Villain, J. Phys. I (France) **1**, 19 (1991).
[8]A. V. Latyshev, A. L. Aseev, A. B. Krasilnikov, and S. I. Stenin, Surf. Sci. **213**, 157 (1989).
[9]E. D. Williams, E. Fu, Y. N. Yang, D. Kandel, and J. D. Weeks, Surf. Sci. **336**, L746 (1995).
[10]S. Kodiyalam, K. E. Khor, and S. Das Sarma, Phys. Rev. B **53**, 9913 (1996).
[11]K. E. Khor and S. Das Sarma, Phys. Rev. B **36**, 7733 (1987).
[12]M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 5668 (1982).
[13]S. Kodiyalam, K. E. Khor, N. C. Bartelt, E. D. Williams, and S. Das Sarma, Phys. Rev. B **51**, 5200 (1995).
[14]D. J. Chadi and J. R. Chelikovsky, Phys. Rev. B **24**, 4892 (1981).
[15]C. Alfonso, J. M. Bermond, J. C. Heyraud, and J. J. Metois, Surf. Sci. **262**, 371 (1992).
[16]E. D. Williams, R. J. Phaneuf, Jian Wei, N. C. Bartelt, and T. L. Einstein, Surf. Sci. **294**, 219 (1993).
[17]F. H. Stillinger and T. A. Weber, Phys. Rev. B **31**, 5262 (1985).
[18]X. P. Li, P. B. Allen, and J. Q. Broughton, Phys. Rev. B **38**, 3331 (1988).
[19]J. E. Northrup and M. L. Cohen, Phys. Rev. B **29**, 1966 (1984).
[20]J. E. Northrup, Phys. Rev. Lett **57**, 154 (1986).
[21]K. E. Khor and S. Das Sarma, Phys. Rev. B **36**, 7733 (1987).

An inquiry concerning the principles of Si 2*p* core-level photoemission shift assignments at the Si/SiO_2 interface

F. R. McFeely[a)]
T. J. Watson Research Center, IBM, Yorktown Heights, New York 10598

K. Z. Zhang and Mark M. Banaszak Holl[b)]
Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055

Sunghee Lee and John E. Bender IV
Chemistry Department, Brown University, Providence, Rhode Island 02906

(Received 9 February 1996; accepted 20 April 1996)

It has long been held that high resolution x-ray photoemission spectra of the Si 2*p* core level at $SiSiO_2$ interfaces provide important structural information, and that any credible interfacial model must be able to account for these data. To this end it has traditionally been assumed that there existed a one-to-one relationship between chemically shifted spectral features and interfacial silicon atoms in specific formal oxidation states. A series of new measurements of Si 2*p* core level binding energies for cluster-derived Si/Si oxide interfaces appear to stand in direct contradiction to this critical assumption. In this article we present a critique of various responses to the challenge posed by the new observations. Particular attention is given to the logical consequences of either maintaining or rejecting the one-to-one relationship between shifted features and formal silicon oxidation states and to the challenges each of these responses must meet if they are to prevail.
© *1996 American Vacuum Society.*

I. INTRODUCTION

The structure and reactivity of the $Si(100)/SiO_2$ interface has been a subject of continuing interest and controversy over the past few years.[1] While a variety of experimental methods have been brought to bear on these problems, high resolution x-ray photoemission studies of the 2*p* core level binding energies of the interfacial atoms have played a particularly important role.[2–9] The technique offered the promise of simple to interpret, reliable information concerning the structure of the silicon-oxygen network at the interface. Prior to 1993 it was generally held that in the spectra of Si/SiO_2 interfaces, chemically shifted Si 2*p* core level peaks of 0.9, 1.8, and 2.5 eV were the result of interfacial Si atoms bound to one, two and three oxygen atoms.[10] The binding energy shift exhibited by fully formed SiO_2 exhibits a small dependence on overlayer thickness, which is generally held to originate from varying image charge stabilization of the core hole state.[2,3,5,611] Aside from this effect, no details of the structure other than the number of oxygen atoms to which the Si atom in question is directly bound was generally held to have any measurable effect on the spectra.[12] This set of empirical assumptions has been heavily relied upon for the construction of structural models.[2,5,6,8,13] In addition, the assignment scheme, and the structural models derived from it, are routinely used when attempting to optimize processes for semiconductor device development. For these reasons, a reliable, clearly delineated assignment scheme is critically important.

Recently, however, both the simplicity and reliability of the assignment rules have been called into question. New measurements have been made of Si 2*p* core level shifts for a class of cluster derived Si/SiO_2 interfaces.[14] These interfaces, formed by the low temperature attachment of cluster molecules with pre-formed Si-O bonds, constituted the first class of silicon/silicon oxide interfaces for which aspects of the local structure could legitimately claimed to be known *a priori*. This permitted the first examination of the assignment rules for the Si 2*p* core level chemical shifts which was not intimately entangled with the problem of structure determination.[15] From the standpoint of the simple clear cut assignment scheme described above, the results were, to say the least, rather disturbing. In this article we will review the key observations made on the cluster interfaces, along with some new supporting evidence from a new class of molecules, the silatranes.[16] We will then attempt to systematize the various positions which have, implicitly or explicitly, been taken with respect to the need to modify the rules and principles by which the photoemission spectra are assigned and models of Si/SiO_2 interfaces are constructed.

II. THE DISQUIETING CLUSTER INTERFACE SPECTRA

In 1993 Banaszak Holl and McFeely published a paper containing the Si 2*p* photoemission spectrum of $H_8Si_8O_{12}$ spherosiloxane clusters chemisorbed on a Si(100)-2x1 surface (Fig. 1).[14(a),17] Analogous spectra from the chemisorbed clusters $H_{12}Si_{12}O_{18}$ and $H_{14}Si_{14}O_{21}$ soon followed.[14(b)] Each of these spectra contained no less than three features which were to varying degrees incompatible with the then universally accepted rules for the assignment of Si 2*p* spectra from silicon/silicon oxide interfaces. For the sake of convenience in the forthcoming discussion, we enumerate these formally.

[a)]Electronic mail: mcfeely@watson.ibm.com
[b)]Electronic mail: mbanasza@umich.edu

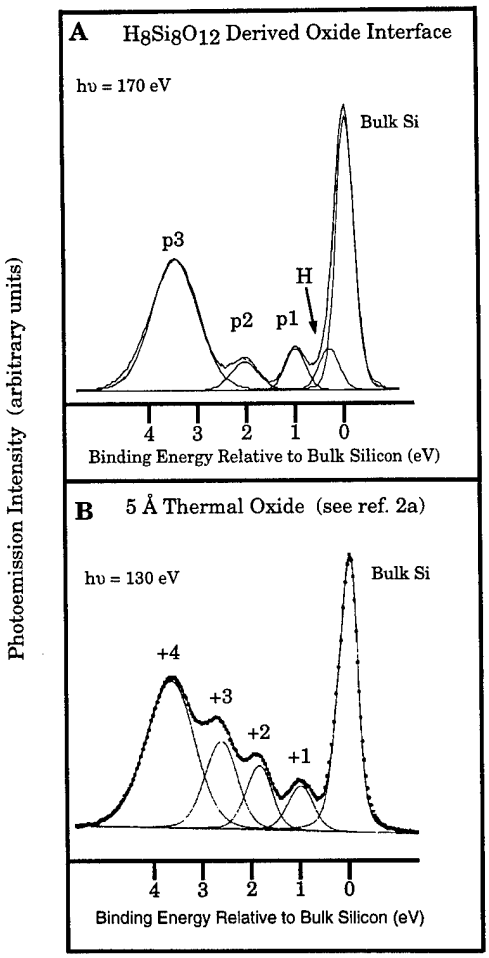


FIG. 1. Panel A: cluster derived silicon/silicon oxide interface. Panel B: thermal silicon/silicon oxide interface with formal oxidation states assigned to features using the orthodox assignment model.

C1: The binding energy of a surface Si to which a cluster vertex is bound, i.e., O_3Si-*Si*, is shifted by 1.0 eV with respect to the bulk, despite the fact that it is a formal Si^0 species without any direct bond to an oxygen atom.

C2: The cluster vertex Si effecting the attachment to the surface, i.e., O_3*Si*-Si, is shifted by 2.1 eV, a value almost exactly midway between those held to be characteristic of Si^{+2} (1.8 eV) and Si^{+3} (2.5 eV) at $Si(100)/SiO_2$ interfaces.

C3: $HSiO_3$ moieties are chemically shifted 3.6 eV–3.7 eV, far removed from the 2.8–2.9 eV predicted on the basis of the simple additivity of chemical shifts.

III. RESPONSES TO THE NEW DATA

As the results presented above do not fit neatly into the traditional assignment scheme, they naturally suggest that some modification of the assignment principles may be in order. The experimental observations, however, are, to stress a point, observations only. They do not transparently reveal the underlying physics responsible for the anomalous shifts. As such, they themselves allow considerable latitude in the assessment of the nature and extent of such modifications. In this section, we consider some of the positions which have been taken explicitly or implicitly in the literature with regard to the appropriate modifications of the traditional assignment procedures. In so doing we will review a series of positions ranging from the orthodox, to the most aggressively revisionist, orthodoxy in this case referring to positions substantially upholding the pre-1993 assignment scheme. Emphasis will be given to a discussion of the logical and evidentiary foundations of the positions in question and to any conclusions necessitated with regard to the cluster data. As in all matters not firmly settled by an overwhelming preponderance of experimental evidence, the reader will doubtless note the possibility fully self-consistent positions lying in between those we have explicitly set out. Those specifically elucidated (orthodoxy, revisionism, and accommodationism) should most properly be considered as somewhat arbitrary guideposts along a continuum. In the spirit of forthrightness, we alert the reader to the fact that we are not disinterested parties in this debate. We are on record as advocates of revisionism, and continue (variously) to hold positions on this end of the spectrum. We will however attempt to draw the reader's attention to the weak points in our own position as well as those in opposing views.

A. Orthodoxy

As orthodoxy implies subscribing to a well established set of ideas, we designate as orthodox those asserting that the pre-1993 assignment principles are in no need of revision in light of the evidence cited above, and apply universally to the silicon/oxygen system. While we are aware of no one who has made this exact affirmation in print, authors who assign their spectra by means of the pre-1993 standard principles without explicit acknowledgment of divergent possibilities suggested by the cluster results, are by clear implication upholders of the orthodox position. To summarize, the tenets of the orthodox assignment scheme are

O1: The binding of an oxygen atom to a Si atom at the interface produces a shift of approximately 0.9 eV in the Si 2*p* binding energy. The effect of oxygen binding is additive, each additional oxygen inducing a similar incremental shift, so that Si_2SiO_2 and $SiSiO_3$ configurations should give rise to shifts at 1.8 and 2.5 eV on the Si(100) surface , respectively.

O2: No configurations other than those stated above give rise to a spectroscopically measurable chemical shift. The Si 2*p* binding energy is determined completely by the identity of the four nearest neighbors of the central atom, with the exception of *O3* below.

O3: A small decrease in the dielectric screening of the core hole final state occurs as the emitting Si atom moves away from the bulk Si into the oxide. The primary observational consequence of this is a gradual increase in the SiO_2 chemical shift (to a maximum added shift of 0.4 eV) with overlayer thickness in the 0–10 Å regime.

Some comment on this particular choice of propositions as representative of the orthodox position is in order. Proposition *O3* is clearly the least central to orthodoxy, and there are indeed some workers who have (at least in the past) assigned spectra using generally orthodox methods without adopting this principle. We have included it because it pro-

vides an apparently simple means to a ccount for observation *C2*. This attempt raises a number of subtle points which we wish to address below, hence its inclusion. Absent the *O3* principle, the orthodox assignment principles seem simply at a loss to account for this discrepancy. Principle *O2*, it may be noted, would be regarded as a rather extreme statement by some workers who have in the past interpreted the photoemission data by methods which we would characterize, on balance, as orthodox. This is particularly true of those who have argued for the importance of bond angle variations in inducing small chemical shifts. These bond angle effect arguments, however, have largely been applied to the "Si^{+4}" peak, which is not our primary subject in this article. Furthermore as we have discussed elsewhere, the spherosiloxane cluster studies directly contradict the claims that these effects are important, as do recent theoretical calculations upholding the orthodox viewpoint.[20] Moreover, it is our view, based upon a survey of recent publications,[2(b),4,9] that the notion of bond angle fluctuations as a significant effect to be incorporated into the analysis of the photoemission data has not found widespread acceptance.

The orthodox assignment scheme rests on several foundations. First, the assignment of a 1 eV shift to a surface silicon bound to three silicons and one oxygen atom has firm underpinnings in studies of water chemisorption.[2(c),18] We know of no systems containing *known* Si_2SiO_2 and $SiSiO_3$ moieties yielding Si 2*p* chemical shifts in the 1.8 eV and 2.5 eV range. (Clearly, the thermal oxide interfaces themselves cannot be used for this purpose.) In fact, the only cases we know of in which a moiety $SiSiO_3$ is known to be present are the various cluster interfaces where shifts of 2.0–2.1 eV, not 2.5 eV are measured. Nevertheless, the fairly regular spacing of the interfacial oxide peaks at 0.9, 1.8 and 2.5 eV are clearly reminiscent of the regular spacings observed for halide chemisorption on Si(111) surfaces.[19] As these monolayer systems are considerably simpler from the standpoint of structural chemistry, the di- and tri-halide origin of the peaks shifted by multiples of a basic increment are quite generally accepted. The ascription of the 1.8 and 2.5 eV peaks to Si_2SiO_2 and $SiSiO_3$ moieties has therefore been based upon the principle of the regular additivity of shifts, in analogy with the halide systems. This idea has recently been bolstered by theoretical calculations from Pasquarello, Hybertson, and Car.[20] In this calculation, the binding energy shifts are insensitive to the details of the second nearest neighbors,[21] but no explanation for the discrepancy posed by the cluster interface spectra was given.

The challenge to the orthodox position is to account for the results of the cluster studies. As this assignment scheme asserts that spectroscopically detectable binding energy shifts are dependent upon the first coordination sphere only, the key structural moieties in the cluster interfaces should clearly fall within their purview. Observation *C1*, that a formally Si^0 species should give rise to a 1 eV shifted core level, is absolutely incompatible with the orthodox picture. Therefore, to espouse orthodoxy requires, for consistency's sake, the simultaneous assertion that this experimental observation is either some sort of misleading artifact or that the assignment is wrong in some fundamental way. This prospect will be considered in conjunction with our discussion of revisionism below.

Orthodoxy must also face the challenge posed by observation *C2*. This is more troubling than it appears at first sight. The orthodox position is placed in a curious position here. The cluster spectra provide a model system which is apparently ideal for determining the binding energy of an $SiSiO_3$ moiety experimentally.[14] Moreover, it is, to the best of our knowledge, the *only* such model. Nevertheless, the 2.1 eV shift exhibited by this moiety is rejected in favor of the value 2.5 eV. This latter value is evidently to be preferred based upon its appearance in the thermal oxide interface spectra, and the presumed linear variation of binding energy shift with formal oxidation state, a condition which, if this latter value does not actually quite satisfy, it at least approximates more closely. The difference between the 2.5 eV peak observed in the thermal oxide and the 2.1 eV shift observed for the $SiSiO_3$ moiety in the cluster interface spectra is not a trivial one. A shift of 2.1 eV is actually closer to the 1.8 eV shift proposed for Si_2SiO_2 fragments than to the 2.5 eV shift.

The orthodox assignment scheme offers one means by which this observation could be rationalized, principle *O3*. In the cluster interface spectra, the O_3Si- moiety is bound directly to the substrate, while in every plausible model of the interface these fragments are on the SiO_2 side of the interfacial region, removed by up to 4 Å or so from the silicon substrate.[2,5,6,8,22,23] According to *O3* this could in principle be accompanied by some upward shift in binding energy.[24] Invoking *O3*, however, to rationalize this discrepancy creates new problems of its own. The O_3SiSi feature of the cluster interfaces and the Si^{+1} feature of the thermal interface spectra (or the HO-*Si* feature from water chemisorption, for that matter) both occur at the bulk Si side of the interfacial region. Most interface models, whether based upon orthodox or upon radically revised interpretations of photoemission data, and tridimite-based models,[25] agree in placing the Si_2SiO_2 moieties there. If the difference between the 2.5 eV of the thermal oxide $Si^{+3}$and the 2.1 eV in the clusters is attributed to a 0.4 eV "*O3* effect," and the Si^{+2} and Si^{+1} are so situated geometrically as to have the same *O3* contribution as the cluster Si^{+3}, then we must select the cluster Si^{+3} as the best species to use in obtaining an undistorted picture of the Si^{+1}-Si^{+3} binding energy progression. The resulting series, 0.9–1.0, 1.8, 2.0–2.1, does not seem very regular at all, and appears at clear variance with the theoretical assertion[20] that the initial state contributions to the shifts should vary linearly. Suddenly, from a self evident regularity, simply accounted for, one seems forced to conclude that if one invokes *O3* as the sole cause of the discrepancy in binding energy between Si^{+3} species in the thermal and cluster interfaces, the approximately regular spacing of peaks in the oxide interface obtains only because a large incremental decrease in initial state shift in adding a third oxygen atom is fortuitously compensated by a decrease in core hole relaxation. In so doing however, proponents of orthodoxy would

be compelled to relinquish at least some of their claim that their position is bolstered by theory. Pasquarello, Hybertson, and Car state unequivocally that the initial state contributions to the Si $2p$ binding energies at the Si/SiO_2 interface vary linearly with formal oxidation state.[20] Thus, experiment and theory currently seem at an impasse regarding this point. In conclusion, orthodoxy, which recall is defined as asserting universal applicability to the Si/O interface system, can possibly survive observation *C2*, but not without some damage to the plausibility of its foundations.

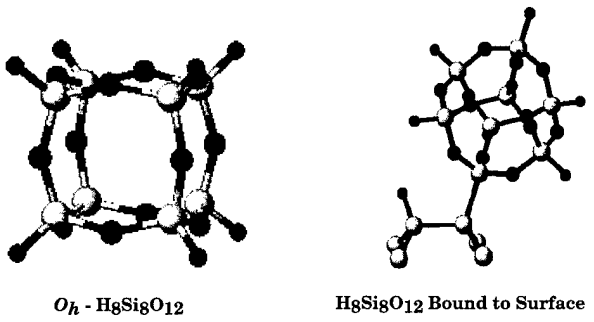


FIG. 2. Cluster molecule $H_8Si_8O_{12}$ and a schematic view of the cluster bound to a silicon surface.

B. Revisionism

Revisionism, as the name implies, refers to the doctrine that the orthodox assignment rules *O1–O3* must be modified in light of experimental observations *C1–C3* to produce a new set of assignment rules, of equally universal application within the Si/Si oxide interface systems.[26] Implicit in any revisionism are two fundamental assertions.

R1: The cluster interfaces are not pathological or any sort of special case. They are typical examples of Si-O bonding and as such their spectra are assignable on the basis of the same rules as any other Si-O system.

R2: The experiments on the cluster derived interfaces have been performed correctly, and the assignments given for the spectral features is as given by the authors.

Point *R1*, which serves to distinguish revisionism from the accommodationist position discussed below, is straightforward and in essential agreement with the orthodox view. In examining the structural moieties giving rise to *C1* and *C2*, one must go out to at least 3rd nearest neighbors before the structure reaches any anomaly (a hydrogen atom). As the orthodox position is that, to all extents and purposes, only the first coordination sphere matters, the clusters should not be anomalous. While not endorsing precisely this reasoning, advocates of revisionism would agree that absent some specific, convincing physical reason why the cluster interfaces should be anomalous, they should not be so regarded.

Principle *R2* is of course the foundation of revisionism. It is clear from the preceding discussion that if orthodoxy is to prevail, the assignments given in the cluster interface spectra must be fundamentally wrong ones, and it must be possible to construct new assignments based upon orthodox principles which successfully explain the data. We examine here what we consider the two most pertinent challenges to the revisionist interpretation of the cluster interface spectra, the contamination hypothesis and the partial dissociation hypothesis.

The contamination hypothesis deals with *C1* by asserting that the 1.0 eV shifted feature in the cluster and silatrane spectra are in every case an artifact caused by the contamination of the surface by water. The spectra are thus held to really be that of clusters and co-adsorbed water, with the latter responsible for the 1.0 eV feature. To be sure, the Si(100)-2x1 surface is highly reactive towards water, and careful procedures had to be adopted to obtain the spectra reported. We do not however believe that the water hypothesis can be sustained. To begin, note that the water hypothesis must assert that ALL of the intensity in Fig. 1, peak $p1$, is due to water. This makes for difficulties in understanding the intensity relationships among peaks $p3$, $p2$, $p1$, and H in the $H_8Si_8O_{12}$ spectrum of Fig. 1 and in the analogous intensity relationships for spectra of the other clusters. Banaszak Holl and McFeely noted that peaks $p3$–H should stand in the ratio of 7:1:1:1 for single vertex attachment to the surface. This single vertex intensity relationship is approximately satisfied by all the clusters. We have repeated the $H_8Si_8O_{12}$ result in particular many times over the past three years, using different synthetic lots of source cluster, always with this same result. It strains our credulity to believe we would not have seen considerably more variability in this spectrum if it were the result of contamination. Even if this point is waived, however, problems remain. If peak p1 is ascribed to water, then its intensity with respect to peaks p3 and p2 implies that water adsorbs in approximate 1:1 proportion to all the clusters, as well as showing an identical 1:1 behavior in the chemically very different silatrane case. If true, this would be an exceedingly curious phenomenon. The clusters and H_2O each efficiently inhibit the chemisorption of the other. A water saturated surface is inert towards clusters, and a cluster saturated surface is inert towards water chemisorption. One supposes therefore that in the water contamination scenario, what must be going on is a race for surface sites between the two species. The vapor pressure difference between $H_8Si_8O_{12}$ and $H_{14}Si_{14}O_{21}$ appears to be on the order of a factor of 100 at the very least, which during the exposure process leads to a similar reduction in the flux of cluster arriving at the surface. If the $H_8Si_8O_{12}$ is adsorbed in 1:1 proportion, one must conclude that the race for sites between background water and the $H_{14}Si_{14}O_{21}$ clusters should be overwhelmingly in favor of the former, and that not more than a percent or so of surface sites should wind up being occupied by the clusters. This expectation is clearly belied by the experiments.[14]

In spite of these difficulties with the water hypothesis, we have nevertheless undertaken further experiments to ascertain its likelihood. To do this we have prepared surfaces containing only about half-saturation coverage of $H_8Si_8O_{12}$, a representative spectrum of which is shown in Fig. 3. The half-saturated surface still clearly shows surface dimer peaks indicative of reactive, clean surface sites. The fact that we can measure these spectra, which require substantial addi-

tional sample manipulations, is strong evidence that we can adsorb the clusters cleanly. To verify that the surface shown in Fig. 3 is still reactive towards water, the sample was removed to the reaction chamber and deliberately dosed with water. As shown by the dotted curve in Fig. 3, the expected result was obtained The surface dimer site peak is completely removed and additional water induced intensity grows in at about 1 eV.

As a final piece of evidence, we have for the silatrane case been able to demonstrate that for a closely related isoelectronic molecule, water-free depositions can be effected. To this end we employed azasilatrane, $HSi[(NH)CH_2CH_2]_3N$, which is identical to silatrane except that each oxygen is replace by an NH group.[27] The absence of oxygen within this cluster made possible the use the O 1*s* photoemission intensity to rule out significant contamination by water.

In conclusion, we find the hypothesis of water contamination unsustainable. Even if one allows for the strange constancy of the magnitude of the putative contamination from cluster to cluster, it offers no insight into why azasilatrane, alone amongst these clusters should be immune from contamination effects. Finally, we note that the water hypothesis, by itself does not by itself solve all of the difficulties faced by the orthodox position, as it offers no explanation for observation *C2*, which, as we have seen, is reconcilable with the orthodox doctrine only awkwardly, if at all.

The second possibility we shall consider is the hypothesis of partial decomposition. The partial decomposition structure most consistent with the orthodoxy assignment scheme involves breaking one silicon-oxygen bond at the surface attachment vertex and rebonding it to another surface silicon. This rebonding would provide an Si-O moiety at the surface with the $\sim$1.0 eV Si 2*p* binding energy shift necessary to account for *C1*. This sort of partial decomposition scenario is difficult to subscribe to on a number of grounds. First of all, this process involves the spontaneous breakup of a chelated site to form two radical sites separated in space, one at the cluster vertex and one resulting from the breaking of the dimer bond accompanying the rebonding of the oxygen atom at the surface. While surface reconstruction might account (somehow) for the quenching of the surface radical, the cluster radical seems unavoidable. This process would therefore be expected to be extremely endothermic. Further compensating dissociations are virtually impossible to square either with the preservation of high symmetry evidenced by the sharpness of the valence band structure, or the absence of additional resolvable core level features.[14] The partial dissociation theory also requires that five different cluster complexes, which have both subtle and substantial differences in the chelate formed at the vertex Si, all partially dissociate in the exact same manner. This is difficult to accept. Note that upon heating, which we believe *does* cause the clusters to react further with the surface, the sharp valence band features rapidly give way to a spectrum reminiscent of amorphous SiO_2 and new, resolvable features are apparent in the Si 2*p* core-level spectrum.[14(b),14(c)] Even if these difficulties are waived, the partial decomposition model must account for

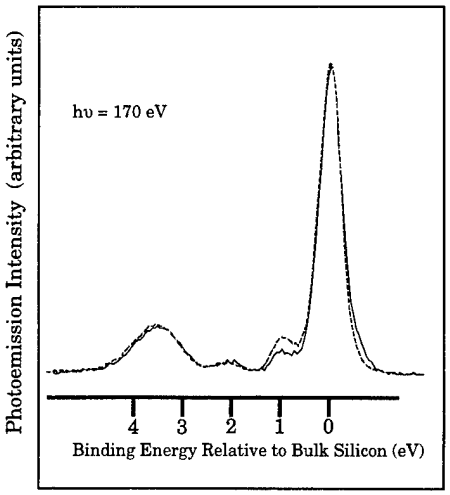


FIG. 3. The solid line shows a Si(100)-2×1 surface that has been partially dosed with $H_8Si_8O_{12}$. Note the presence of a shoulder to lower binding energy than the bulk, indicative of surface dimers. Subsequent dosing with water causes the disappearance of the dimer featuers and the growth of an Si–OH feature at 1.0 eV as expected [Ref. 2(c)].

the observed chemical behavior of the cluster interfaces. The interfaces produced by $H_8Si_8O_{12}$ show no spectral change when exposed to thousands of langmuirs of water. Even short exposure to atmosphere produces no discernible changes.[14(b)] This observation is incompatible with the existence of highly reactive free radicals on the surface.

Another unfortunate consequence of the partial decomposition hypothesis is that it makes observation *C2* even more problematic for the orthodox assignment model than it was before. The disruption of the cluster vertex upon partial decomposition changes this silicon atom from a formal Si^{+3} to a Si^{+2}. This leaves one with the predicament that somehow a single oxygen atom in duces 0.9 eV of chemical shift while the addition of a second can induce an additional 1.2 eV. If the silicon cluster vertex site is a free radical, one would also expect some multiplet splitting in the core level spectrum, which is not in evidence. We could go on, considering the effects of shifting hydrogen atoms, but to the best of our understanding these shifts only serve to patch up (poorly) one problem, leaving worse ones in their wake. In summary, neither we nor anyone else has succeeded in constructing any alternative structural model or contamination hypothesis which plausibly explains the cluster interface spectra. We hope shortly to offer infrared evidence to further bolster this conclusion.[28,29]

Having upheld, at least for now, the experimental foundations of revisionism we proceed to the problem of exactly how it is to be implemented. This is a difficult problem, *as in our opinion the fundamental lesson of the cluster studies is that reliable model systems of known structure are absolutely required to make unimpeachable assignments of interface core level features*. Simple rules of thumb can take one only so far, and subtleties can be missed. An excellent case in point is provided by observation *C3*, that $HSiO_3$ species are shifted by 3.6–3.7 eV. The principle of regular additivity of

chemical shifts, which is at least consistent with overall shift trends, would take the 2.5 eV shift of interfacial SiO_3, and add to it the 0.3 eV shift induced by H on bare Si, and predict this feature to come in the 2.7–2.9 eV range. Numerous assignments have been made on this basis in the literature.[30] This is moreover a case in which the germaneness of the cluster studies to the interface cannot be challenged, as recently it has been shown that hydrogen atom attachment to ultrathin SiO_2 interfaces produces a feature at exactly the 3.6 eV predicted by the cluster studies.[31]

The most resolutely conservative position for a revisionist to take is that only the binding energies associated with O_3Si–Si, with the second Si being a bulk Si atom, have been measured, and that this and only this should be incorporated into the analysis of interface spectra. While certainly safe, such a diffident attitude is not the sort which is conducive to progress in understanding Si/SiO_2 interfaces. For the sake of intellectual consistency would one not also be obliged show equivalent reticence in attributing the 1.8 eV peak to Si_2SiO_2 or the 2.5 eV peak to oxide matrix $SiSiO_3$, as these assignments are based merely upon plausible analogies? Clearly, this is carrying rigor to the point of intellectual rigor mortis. It is also inconsistent with revisionist principle *R1*, that the cluster systems are typical ones. As such the same latitude should be granted to the revisionist as to the orthodox in the use of carefully chosen plausibility arguments to construct physical models of the interface worthy of scrutiny.

With this perspective, while as revisionists we disagree with orthodox principal *O2*, we accept the idea that the principal of regular additivity of shifts, an implicit underpinning of orthodoxy, is a useful rule of thumb to which recourse can legitimately be made, although not regarded in itself as a source for definitive conclusions or a substitute for the direct evidence of appropriate model systems. Adopting this stance necessitates a revision of the orthodox assignments on the grounds that they are insufficiently specified by not including the effects of second nearest neighbors,[32] which the cluster results show can be significant.[14] Orthodoxy requires that any silicon atom without a direct bond to oxygen exhibit a negligible Si 2*p* chemical shift. Cluster results show that if such an atom is bound to an $-SiO_3$ group, so as to place 3 oxygen atoms in the second coordination sphere, a 1 eV shift results. Similarly, we can consider the 2.5 eV peak in the thermal interface spectra. All parties to this debate, whether rigorously orthodox or radically revisionist, associate this peak with SiO_3 groups, more specifically with O_3Si-SiO_3 structures. As the Si_3Si-SiO_3 binding energy falls at 2.1 eV, it is reasonable to ascribe the bulk of the binding energy difference to the effect of additional oxygen atoms in the second coordination sphere, afforded by the mutual linking of the SiO_3 moieties. (As orthodox principle *O3* is fully compatible with revisionism one can allow some sm all final state contribution to this difference as well.)

Given these two cases, and a toleration for the additivity principle as a guideline, does it make sense to maintain that there should be no difference in binding energy between the two formal Si^{+1} fragments Si_3SiO and $Si_2Si(O)(SiO_3)$?[33] We think not. Taking the same sort of approach as was originally used to construct the orthodox as signment rules, one could propose a linear scaling of the effect of $-SiO_3$ group attachment between the two above cases. Thus the addition of an SiO_3 group to an otherwise Si+1 species would produce an extra shift of about 0.8 eV and to a Si^{+2} an additional 0.6 eV. The adoption of this set of auxiliary assignment rules, and retaining the remainder of the orthodox, one would arrive at a species of "minimal revision" which would be able to address most of the configurations likely to be found in a model of the $Si(100)/SiO_2$ interface. Even this Minimal Revisionist Model has dramatic consequences for photoemission based models of the $Si(100)/SiO_2$ interface. If it is correct, then many proposed models cannot be right, as they attrib ute all of the observed core level photoemission intensity at 0.9–1.0 eV to $Si_2Si(O)(SiO_3)$ species, while the considerations above would indicate that 1.7 eV or so is a more appropriate position for the Si 2*p* core level peak arising from this moiety.

As the explanation of the cluster results constitutes a challenge for the orthodox position, an accounting for the thermal oxide interfaces is a challenge for revisionism. At first thought, it might be expected that significant second nearest neighbor effects would be incompatible with the discrete peak structure which is experimentally observed. To address this challenge, Banaszak Holl, Lee, and McFeely put forth a more comprehensive and more radical revisionist scheme by adopting an effective electronegativity model.[34] This scheme predicts qualitatively the same revisionist assignment principles as given above, and also gives a systematic means of evaluating chemical shifts of other sorts of configurations, such as the formal Si^0 fragment $Si_2Si(SiO_2)_2$ (to which it assigns a shift of approximately 1 eV), employing a methodology which has had past successes.[35] It must be emphasized clearly that the further such a scheme strays from moieties which have been directly observed in well characterized model systems, the more it becomes speculative. However, applying revisionist photoemission assignment principles, they wer e able to construct a model of the $Si(100)/SiO_2$ interface that clearly reproduced the photoemission data. Interested readers are referred to this paper for details. Given that we have no cluster data for the other crystallographic faces of Si (the clusters do not even adsorb on the (111) surface to a significant extent), we have been hesitant to pursue further model building at this time. The structural model proposed based upon revisionist assignment principles for the $Si(100)/SiO_2$ interface has been strenuously challenged by Pasquarello, Hybertson, and Car, on the basis of theoretical calculations, however, as these authors apparently offer no explanation for the cluster observations, it is difficult to regard this matter as settled.

C. Accommodationism

Between the diametrically opposing views of Orthodoxy and Revisionism lies the doctrine of Accommodationism. This position denies the catholic claims made by both of the foregoing, and asserts that while the cluster interface spectra

have been correctly interpreted, they constitute a special case of some sort, that inferences from the cluster derived interfaces to the thermal interfaces are unw arranted, and that for the latter case the orthodox assignment doctrine holds good.

Accommodationism currently has a peculiar status. There is a chance, in our opinion, that this position may ultimately prove to be substantially correct. However, to date, no one has elaborated an argument along this lines. The necessary elaboration, which is crucial, is a plausible reason grounded in fundamental physical terms to regard the two cases as different. At this time, it is not clear to us what form such an argument would take.

IV. CONCLUSIONS

In this article we have argued that based upon the most straightforward interpretation of the various cluster interface spectra, the orthodox scheme for the assignment of Si $2p$ XPS spectra from the Si/SiO_2 interface must be revised in some way. Furthermore, we have shown that it is no simple matter to devise an alternative explanation for the cluster interface spectra.

The most resolute and uncompromising defense of orthodoxy is provided by the theoretical work of Pasquarello, Hybertson, and Car.[20] These calculations represent pioneering efforts to apply modern local density functional theory to these problems, and they find no evidence for any of the second nearest neighbor effects we claim to have measured. In our opinion, these calculations constitute at this time the best foundation for the orthodox position, far better than the small body of direct experimental observations, or an appeal to simplicity which can no longer be easily sustained. Since we advocate revisionism, we are logically committed to the position that that these calculations are insufficiently accurate in their treatment of the effects of second nearest neighbor atoms. It is an important challenge to the revisionist community to provide cogent reasons why this should be so. While aware of this problem, we frankly admit lacking the necessary theoretical expertise to address it. In any case, whether one presumptively accepts their results as true or not, we believe it is clear at the present time that these sort of calculations have not been subjected to the exhaustive testing necessary to render their conclusions unimpeachable. There are thus insufficient grounds to warrant the wholesale relegation of the cluster experiments to the status of certainly wrong results unworthy of serious consideration. The challenge posed by the cluster interface spectra must be met, and not merely dismissed, by the supporters of the orthodox doctrine. The calculations by the mselves, however, do not shed any obvious light on the nature of any alternative explanation of the cluster spectra.

Unfortunately the theoretical and experimental approaches have thus far tended to talk past one another somewhat. Where one concludes the balance of evidence lies at this time depends upon whether one believes the calculations or the cluster experiments are more trustworthy. The one concrete suggestion we can make to the theorists is that they employ their molecular dynamics approach to investigate the stability of the clusters on the surface. Perhaps such an approach could reveal a type of partial decomposition which has not occurred to us, or add weight to our assertion that they remain intact. For our part, we accept the challenge of bringing additional experimental evidence to bear on this fascinating problem.

ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation (DMR-9596208) and IBM (to M.M.B.H). J.E.B. IV thanks the National Science Foundation for a graduate fellowship.

[1]For lead references see: (a) *The Physics and Chemistry of SiO_2 and the Si-SiO_2 Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993); (b) *The Physics and Technology of Amorphous SiO_2*, edited by R. A. B. Devine (Plenum, New York, 1988); (c) *SiO_2 and Its Interfaces*, edited by S. T. Pantelides and G. Lucovsky, Mater. Res. Soc. Symp. Proc. 1990; (d) *The Physics and Technology of Amorphous Silicon Dioxide,* edited by C. R. Helms and B. E. Deal (Plenum, New York, 1989).

[2](a) F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B **38**, 6084 (1988); (b) F. J. Himpsel, D. A. Lapiano-Smith, J. F. Morar, and J. Bevk, in *The Physics and Chemistry of SiO_2 and the Si-SiO_2 Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993), p. 237; (c) F. J. Himpsel, F. R. McFeely, J. F. Morar, A. Taleb-Ibrahimi, and J. A. Yarmoff, in *Proceedings of the 1988 Enrico Fermi School on Photoemission and Absorption Spectroscopy of Solids and Interfaces with Synchrotron Radiation* (North Holland, Varenna, 1988); (d) G. Hollinger, E. Bergignat, H. Chermette, F. Himpsel, D. Lohez, M. Lanoo, and M. Bensoussan, Philos. Mag. B **55**, 735 (1987).

[3]F. G. Bell and L. Ley, Phys. Rev. B **37**, 8383 (1988).

[4](a) Z. H. Lu, S. P. Tay, T. Miller, and T.-C. Chiang, J. Appl. Phys. **77**, 4110 (1995); (b) Z. H. Lu, M. J. Graham, S. P. Tay, D. T. Jiang, and K. H. Tan, J. Vac. Sci. Technol. B **13**, 1626 (1995); (c) Z. H. Lu, S. P. Tay, R. Cao, and P. Pianetta, Appl. Phys. Lett. **67**, 2836 (1995); (d) Z. H. Lu, M. J. Graham, D. T. Jiang, and K. H. Tan, Appl. Phys. Lett. **63**, 2941 (1993); (e) Y. Tao, Z. H. Lu, M. J. Graham, and S. P. Tay, J. Vac. Sci. Technol. B, **12**, 2500 (1994).

[5]T. Hattori, Crit. Rev. Solid State Mater. Sci. **20**, 339 (1995).

[6]F.J. Grunthaner and P.J. Grunthaner, Mater. Sci. Rep. **1**, 65 (1986).

[7]A. Ishizaka, S. Iwata, and Y. Kamigaki, Surf. Sci. **84**, 355 (1979).

[8]M. Niwano, H. Katakura, Y. Takeda, Y. Takakuwa, N. Miyamoto, A. Hiraiwa, and K. Yagi, J. Vac. Sci. Technol. A **9**, 195 (1991).

[9]J. R. Engstrom, D. J. Bonser, and T. Engel, Surf. Sci. **268**, 238 (1992).

[10]Si $2p$ core-level shifts are referenced to bulk Si defined as 0.0. This allows direct comparison between different data sets by removing the effects of band bending.

[11]Films greater than 10 Å are known to exhibit charging effects, see Ref. 4(e).

[12]Various dissenting views have been presented, mostly concerning the "Si^{+4}" feature which is not the primary concern of this article; for example, see Refs. 5, 6, and 8. These positions do not appear to have been widely adopted.

[13]A. Ourmazd, P. H. Fuoss, J. Bevk, and J. F. Morar, Appl. Surf. Sci. **41/42**, 365 (1989).

[14](a) M. M. Banaszak Holl and F. R McFeely, Phys. Rev. Lett. **72**, 2441 (1993); (b) S. Lee, S. Makan, M. M. Banaszak Holl, and F. R. McFeely, J. Am. Chem. Soc. **116**, 11819 (1994); (c) S. Lee, M. M. Banaszak Holl, W. H. Hung, and F. R. McFeely, in *Modular Chemistry*, edited by J. Michl (in press).

[15]For example, the cluster system provided the first set of interfaces for which a known number of Si-O rings of differing size were present.

[16](a) K. Z. Zhang, J. Bender, S. Lee, M. M. Banaszak Holl, and F. R. McFeely (unpublished).

[17]Spectra are presented as Si $2p_{3/2}$ core-levels with the Si $2p_{1/2}$ component mathematically removed using standard procedures (Refs. 2a, 2c, 14). A spin-orbit splitting of 0.6 eV and a branching ratio of 0.5 was experimentally determined to be optimal for the 170 eV photon energy employed. All spectra were fit equally well using both the Si $2p_{3/2}$ and

$Si_{1/2}$ components on spectra from which the Si $2p_{1/2}$ component had not been removed. The fits were quite robust with respect to small changes in the values of the spin-orbit splitting and branching ratio, which are sometimes varied by small amounts (Ref. 19d).

[18](a) H. Ibach, W. Wagner, and D. Bruchmann, Solid State Commun. **42**, 457 (1982); (b) E. M. Oellig, R. Butz, H. Wagner, and H. Ibach, Solid State Commun. **51**, 7 (1984); (c) Y. Chabal and S. B. Christman, Phys. Rev. B **29**, 6974 (1984).

[19](a) T. D. Durbin, W. C. Simpson, V. Chakarian, D. K. Shuh, P. R. Varekamp, C. W. Lo, and J. A. Jarmoff, Surf. Sci. **316**, 257 (1994); (b) L. J. Whitman, S. A. Joyce, J. A. Yarmoff, F. R. McFeely, and L. J. Terminello, Surf. Sci. **232**, 297 (1990); (c) R. D. Schnell, D. Rieger, A. Bogen, F. J. Himpsel, K. Wandelt, and W. Steinmann, Phys. Rev. B **32**, 8057 (1985); (d) F. R. McFeely, J. F. Morar, N. D. Shinn, G. Landgren, and F. J. Himpsel, Phys. Rev. B **30**, 764 (1984).

[20](a) A. Pasquarello, M. S. Hybertson, and R. Car, Phys. Rev. Lett. **74**, 1024 (1995); (b) A. Pasquarello, M. S. Hybertson, and R. Car, Phys. Rev. B (in press).

[21]Other high-level calculations, discounted by Pasquarello *et al.* on the basis of the structural models employed, have noted the existence of effects on the Si 2*p* binding energy that did not relate to the first neighbor coordiation sphere: (a) H. Kageshima and M. Tabe, in *Control of Semiconductor Interfaces*, edited by I. Ohdomari, M. Oshima, and A Hiraki (Elsevier, New York, 1994), p. 227; (b) Y. Miyamoto and O. Oshiyama, Phys. Rev. B **44**, 5931 (1991).

[22]A. Ourmazd, P. H. Fuoss, J. Bevk, and J. F. Morar, Appl. Surf. Sci. **41/42**, 365 (1989).

[23]A notable exception is K. Ohishi and T. Hattori, Jpn. J. Appl. Phys. **33**, L675 (1994) and Ref. 5 in which a structural model is presented with $-SiO_3$ units in direct contact with the bulk silicon. These fragments are assigned to the observed binding energy shift of 2.5 eV, in strict adherence to tenet *O1* of the orthodox model. Grunthaner and Grunthaner (Ref. 6) have also assigned $-SiO_3$ units in direct contact with the bulk silicon at 2.5 eV.

[24]It is worth noting that if the inherent assumptions and conclusions of Grunthaner *et al.* and Hattori *et al.* are correct (see Refs. 5, 6, 23), there is no *O*3 effect for Si^{+3} moieties, and the argument presented below is moot. However, if *O*3 is removed from consideration, there is no apparent solution to the discrepancy posed by *C*2 (*vida supra*).

[25]A. Ourmazd, D. W. Taylor, J. A. Rentschler, and J. Bevk, Phys. Rev. Lett. **59**, 213 (1987).

[26]At this time, we restrict our comments to areas in which we believe we have sound experimental data which challenges the orthodox assignments. However, as noted earlier, similar formal oxidation state based theories are used for the assignment of silicon/halogen spectra. In any truly "universal" theory for Si 2*p* core-level shift assigments, the basic principles of the t heory should be equally applicable to the halogen case. For example, the explanation provided for the observed 1 eV shift for Si_3SiSiO_3 fragments should also be applicable to the prediction of the binding energy shift of $Si_3SiSiCl_3$, see Ref. 19(a) for an initial discussion of this problem. A detailed discussion of this problem can be found in W.C. Simpson, J.A. Yarmoff, W.-H. Hung, and F.R. McFeely, Surf. Sci. Lett. (in press).

[27]J. Bender, K. Z. Zhang, M. M. Banaszak Holl, and F. R. McFeely (to be published).

[28]B. Bent, M. M. Banaszak Holl, Y. Chabal, S. Christman, J. Eng, G. Flynn, S. Lee, S. Mantl, F. R. McFeely, K. Radermacher, and G. Williams (experiments in progress).

[29]J. Bender, Z. Deng, J. N. Greeley, and M. M. Banaszak Holl (experiments in progress).

[30](a) H. Ogawa, N. Terada, K. Sugiyama, K. Moriki, N. Miyata, T. Aoyama, R. Sugino, T. Ito, and T. Hattori, Appl. Surf. Sci. **56–58**, 836 (1992); (b) T. Hattori, and H. Ogawa, Appl. Phys. Lett. **61**, 577 (1992); (c) T. Hattori, Thin Solid Films **206**, 1 (1991); (d) P. J. Grunthaner, M. H. Hecht, F. J. Grunthaner, and N. M. Johnson, J. Appl. Phys. **61**, 629 (1987).

[31]S. Lee, M. M. Banaszak Holl, W. H. Hung, and F. R. McFeely, Appl. Phys. Lett. **68**, 1081 (1996).

[32]Several groups have in the past made assignments for the Si/SiO_2 interface that relied on effects beyond the first neighbors. See Refs. 5, 6, and 8.

[33]$Si_2Si(O)(SiO_3)$ is a common fragment "Si^{+1}" fragment found in structural models based on the orthodox assignment scheme, see Refs. 2, 5, and 6.

[34]M. M. Banaszak Holl, S. Lee, and F. R. McFeely, Appl. Phys. Lett. **65**, 1097 (1994).

[35](a) S. Hasegawa, L. He, T. Inokuma, and Y. Kurata Phys. Rev. B **46**, 12478 (1992); (b) T. Hattori, T. Igarashi, M. Ohi, and H. Yamagishi, Jpn. J. Appl. Phys. **8**, L1436 (1989); (c) M. Takakura, T. Ogura, T. Hayashi, and M. Hirose, Jpn. J. Appl. Phys. **27**, L2213 (1988); (d) G. Lucovsky, J. Phys. (Paris) Colloq. **C4**, 741 (1981); (e) B. J. Lindberg, K. Hamrin, G. Johansson, U. Gelius, A. Fahlman, C. Nordling, and K. Siegbahn, Phys. Scripta. **1**, 286 (1970); (f) U. Gelius, P. F. Heden, J. Hedman, B. J. Lindberg, R. Manne, R. Nordberg, C. Nordling, and K. Siegbahn, Phys. Scripta. **2**, 70 (1970); (g) P. Kelfve, B. Blomster, H. Siegbahn, K. Siegbahn, E Sanhueza, and O. Goscinski, Phys. Scri. **21**, 75 (1979).

Defect properties of Si-, O-, N-, and H-atoms at Si—SiO_2 interfaces

G. Lucovsky,[a] Z. Jing, and D. R. Lee
Departments of Physics, Materials Science and Engineering, and Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 12 February 1996; accepted 20 April 1996)

A chemical bonding model for defects and defect precursors at Si—SiO_2 interfaces is presented. Bonding geometries of neutral and charged Si-, O-, N-, and H-atoms are described in terms of the valence shell electron pair repulsion model, thereby identifying a qualitative distinction between charged (a) Si-atoms, and charged (b) O-, N-, or H-atoms. Threefold coordinated Si- atoms are stable in three charge states (+, 0, and −), and can be active as trapping and/or recombination centers depending on their charge state. In contrast, there is a direct relationship between charge state and bonding coordination for O-, N-, and H-atoms, and as such their roles in defect generation processes are qualitatively different. Reaction mechanisms based on these differences in local bonding are (i) discussed for the generation of defects comprised of threefold coordinated Si-atoms, and/or positively charged H-, O-, and N-atoms, with coordinations of two, three, and four, respectively, and (ii) compared with experiments. © *1996 American Vacuum Society.*

I. INTRODUCTION

As device dimensions shrink in the deep submicron for ultra large scale integrated (ULSI) circuit applications, proportional decreases in gate dielectric thicknesses of field effect transistors (FETs) to about 3–3.5 nm are required. In this regime of ultra thin dielectrics, interfacial defects and defect precursors play dominant roles in device performance and reliability. For example, there is a large body of experimental data for thin oxide dielectrics which has demonstrated that: (i) H-atom transport from polycrystalline-Si gate electrodes to the Si—SiO_2 interface under stress-bias conditions generates interfacial traps, D_{it}, and positively charged defects, Q_{ss},[1,2] and (ii) N-atom incorporation at Si—SiO_2 interfaces improves device reliability,[3–5] while interfacial N—H bonding leads to significantly higher defect densities (each of these comparisons is with respect to non-nitrided interfaces).[5]

This article presents a model based on bonding geometries of neutral and charged Si-, O-, N-, and H-atoms at Si—SiO_2 interfaces that is derived from an extension of the valence shell detection pair reulsion (VSEPR) model to amorphous solids and their iterfaces, providing a framework for understanding the experimental observations. Section II reviews bonding arrangements of neutral and charged Si-, O-, N-, and H-atoms in terms of the VSEPR model.[6] Section III discusses electronically active defects at Si—SiO_2 interfaces. Section IV presents reaction equations for defect generation at Si—SiO_2 interfaces in terms of the chemical bonding geometries of the neutral and charged Si-, O-, N-, and H-atoms discussed above. Section V summarizes the main points of the paper.

II. CHEMICAL BONDING OF SI-, O-, N-, AND H-ATOMS

The VSEPR model of chemical bonding describes the bonding geometries of inorganic molecules.[6] The model is based on a set of empirical rules that describe the way that atoms combine to form molecules, and has been successful in predicting the stereo-chemistry of different classes of inorganic molecules. It is readily extendable to molecular ions through a consideration of isoelectronic analogs; e.g., a positive charged O-atom has the same number of electrons available for bonding as a neutral N-atom and this is used as a basis for defining the local bonding geometry of the charged O-atom with respect its nearest-neighbor atoms in forming molecular ions.

Following the language of Ref. 6, the VSEPR rules are:

(1) The preferred arrangement of a given number of electron pairs in the valence shell of an atom is that which maximizes their distance apart.
(2) A non-bonding pair of electrons takes up more room on the surface of an atom than a bonding pair.
(3) The size of a bonding electron pair decreases with increasing electronegativity of the ligand.
(4) The two electron pairs of a double bond (or the three electron pairs of a triple bond) take up more room that does the one electron pair of a single bond.

The VSEPR rules are not generally applicable to crystalline solids due to long range interactions and constraints imposed by periodicity. However, based on a large experimental data base for *good* glass-forming materials like SiO_2, As_2S_3, GeS_2, etc., these rules also provide a qualitative guideline for describing local bonding environments in non-crystalline solids. In particular, the local bonding arrangements of the constituent elements in melt-quenched glasses such as SiO_2, GeO_2, As_2S_3, As_2Se_3, GeS_2, and $GeSe_2$, and their alloys, as determined from infrared and Raman spectroscopies, and/or direct structural probles such as x-ray diffraction and EXAFS[7,8] are consistent with the bonding arrangements predicted from this application of the VSEPR rules to network structures. The idealized structures of these glasses are chemically ordered random covalent networks,[8]

[a]Electronic mail: gerry_lucovsky@ncsu.edu

and the disorder that distinguishes these materials from crystalline solids is mostly at the bond-angle and dihedral-angle level, and not at the coordination geometry level. For example, bond- and dihedral-angles are fixed and repeated periodically in a crystal lattice, whereas in network glasses there are statistical distributions of these angles centered around *preferred*-angles predicted from application of VSEPR rules. The concept of a chemically ordered random covalent network can also be applied to *amorphous* materials prepared by thin film deposition, including silicon dioxide, SiO_2, silicon nitride, Si_3N_4, and alloys of SiO_2 and Si_3N_4 (silicon oxynitrides).[9,10] The materials system of interest in here is the interface between a crystalline silicon substrate and non-crystalline SiO_2 dielectric, and as such the VSEPR bonding rules are assumed to apply on the SiO_2 *side of the interface*; i.e., at the metallurgical interface and into the SiO_2 layer. In addition to the bonding of Si- and O-atoms at such an interface, the local bonding arrangements of H- and N-atoms must also be addressed to maintain contact with device processing technologies.

The bonding geometries of neutral and charged Si-, O-, N-, and H-atoms discussed in Sections II A and II B are restricted to environments that are expected to occur at (i) as-grown or as-deposited interfaces, and (ii) interfaces subjected to conventional levels of electrical-stress bias. Qualitatively different defects can be produced at interfaces subjected to intense x-ray, γ-ray, and particle fluxes which dislodge atoms from their normal bonding positions. Examples include the O-atom hole center in SiO_2 glasses,[11] and two-fold coordinated neutral N-atoms in Si_3N_4 thin films.[12]

A. Bonding geometries of neutral Si-, O-, N-, and H-atoms

The bonding geometries of neutral Si-, O-, N-, and H-atoms in σ-bonds in thin film oxide and nitride dielectrics are consistent with VSEPR Rules (1) and (2) (see Fig. 1). Consider first SiO_2: each Si atom makes fours bonds with its four oxygen neighbors which are arranged in a tetrahedral geometry to comply with Rule (1). Each oxygen atom has two silicon neighbors and two non-bonding pairs, which are also arranged in a distorted tetrahedral geometry, also as required by Rule (1). This distortion derives from (i) the two O-atom non-bonding pairs taking up more space than the two O-atom bonding pairs, and (ii) a back-donation of electrons from the non-bonding oxygen orbitals into anti-bonding orbitals on the Si-atoms, the called $p\pi$—$d\pi$ bonding interaction.[6] Studies of SiO_2 have indicated a large distribution of Si—O—Si bond angles ($\pm 30°$) around an average bond angle of $\sim 145°$ consistent with the proposed extension of VSEPR Rules (1), (2), and (4). The average angle of 145° is indicative of some degree $p\pi$—$d\pi$ bonding, which is also evident from the magnitude of the average Si—O bond length which is about 0.15 Å less than anticipated on the basis of the sum of the appropriate Si- and O-atom bonding radii.[6]

The bonding in Si_3N_4 parallels that of SiO_2 with π-bonding interactions also being evident in a shorter (0.15–0.20 Å) than expected Si–N bond length.[6] The bonding geometry of Si-atoms is tetrahedral in accord with the VSEPR model, whereas the bonding geometry of the four orbitals on the N-atoms is markedly non-tetrahedral, indicating significant π-bonding, and requiring an extension of VSEPR rules. In particular, back donation of electon density from the non-bonding p-orbital of nitrogen into anti-bonding orbitals of silicon with a sd^3 character is the $p\pi$—dπ bonding interaction that forces the N-atom into a planar bonding arrangement with three σ-bonds to Si-neighbors directed at vertices of an equilateral triangle, and the non-bonding p_z-orbital perpendicular to that bonding plane.[6] This planar bonding geometry for nitrogen and its three silicon neighbors has been verified by Raman scattering in non-crystalline Si_3N_4, (Ref. 13), and also occurs in both crystalline forms of Si_3N_4. The formation of relatively strong π-bonds between N-atoms and three Si-atom neighbors is quantitatively different than for the bonding of O-atoms to two Si-atom neighbors, and will be shown to be a significant factor in accounting for improving device reliability by N-atom incorporation at Si—SiO_2 interfaces.

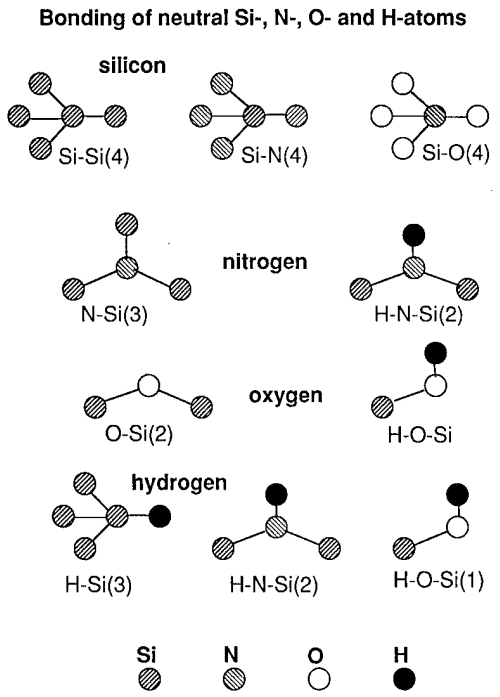


FIG. 1. Bonding arrangements of neutral Si-, O-, N-, and H-atoms in Si—SiO_2 heterostructures.

Hydrogen atoms can also be incorporated into amorphous SiO_2 and Si_3N_4 networks, bonding to both atomic constituents in arrangements consistent with their neutral atom valence of one. The bonding arrangements deduced from infrared (ir) studies are consistent with the VSEPR model: Si—H, O—H, N—H and N—H_2 (Ref. 14) (see Fig. 2). One aspect of the incorporation of H-atoms in O—H and N—H groups in SiO_2 and Si_3N_4, respectively, is in H-bond formation,[6] which shows up in ir line-shapes in the bond-stretching modes.[14] H-bonding interactions fall outside of the VSEPR model, but represent an example of three-center

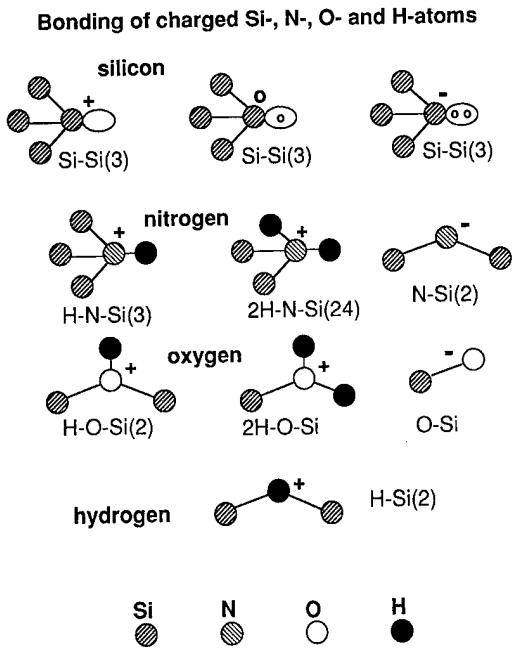


FIG. 2. Bonding arrangements of charged Si-, O-, N-, and H-atoms in Si—SiO_2 heterostructures.

bonding for neutral atoms. This serves to provide a framework for understanding other three-center bonds involving positively charged H-atoms (see Sec. II B), e.g., $\equiv Si—H^+—Si \equiv$.[15–17] The "$\equiv Si—H^+—Si\equiv$" notation indicates that the Si- atom is bonded to the H^+-atom as well as three other atoms of the network structure.

B. Bonding geometries of charged Si-, O-, N-, and H-atoms

The bonding geometries of charged O- and N-atoms can be addressed in the context of isoelectronic relationships: for example, O^+ is isoelectronic with N^0; i.e., both atomic species have five electrons available for bonding. In a parallel manner, N^+ is isoelectronic with C^0, with both atoms having four electrons available for bonding. Depending up the competition between repulsive interactions between non-bonding and bonding orbitals, and π-bonding interactions, the local bonding of O^+ to three neighbors will be a distorted tetrahedron, or in a planar geometry for three bonds with the fourth bonding direction occupied by a non-bonding pair perpendicular to the bonding plane. The local bonding of N^+ to four atoms will be tetrahedral; however, if the four neighbors to the N-atom are markedly different (e.g., 2 Si- and 2 H-atoms), it will be a distorted tetrahedron. Threefold coordinated, positively charged O-groups are designated as *oxonium*, and fourfold coordinated, positively charged N-groups are designated as *ammonium*.[6] H^+ cannot be discussed in terms of isoelectronic species simply because the H-atom contains only one electron, and after it is removed to form H^+ there are no electrons present to mimic the electronic structure of another atom whose atomic number is reduced by one. However, H^+ can bond to two Si-neighbors in two-atom, three-center bonding arrangement that parallels the three-center bonds between B- and the bridging H-atoms in B_2H_6;[6,15–17] hence the isoelectronic relationship is between two bonding groups: Si—H^+—Si and B—H—B, rather than between two atoms.

Si^+ has three electrons available for bonding in a pyramidal geometry to three neighbors. However, in contrast to H-, O-, and N-atoms where coordination and charge state are linked, threefold coordinated Si-atom can support three charge states, (+, 0, or −) according to the electron occupancy (0, 1, or 2, respectively) of the so-called lone-pair orbital,[18] and therefore will be treated separately after the discussion of negative charge states for H-, O-, and N-atoms.

The bonding geometries of H^-, O^-, and N^- can also be described in terms of isoelectronic relationships. H^- is isoelectronic to the rare gas atom He^0, and does not form a stable bond with Si, but may play a role as an intermediate and transient excited state in solid state reactions involving muon capture.[16] In contrast, O^- and N^- can form stable bonding arrangements with coordinations reduced by one with respect to the neutral bonding. O^- is isoelectronic with F^0, and can be singly bonded with the six-remaining electrons occupying three non-bonding pair orbitals. If O^- coordinates in this way, the bond will be weak and highly reactive due to the fact that the six electrons contained in three non-bonding orbitals produce a significant destabilizing repulsive effect. Neutral Si- and negatively-charged O-atoms are not known to form a stable terminal bond, either in molecules or in the solid state. The bonding of N^- is expected to parallel that of O^0, and be twofold coordinated with two non-bonding pairs. This type of bonding has not as yet been identified in combination with Si-atom neighbors.

Threefold coordinated group IV atoms such as silicon must be treated separately. Because of their neutral valence band occupation of four, they are stable in three charge states with the same number of nearest-neighbor bonds.[18] This is because the two charged atoms, Si^- and Si^+, are isoelectronic, respectively, with the group III and group IV elements, B and P, and therefore they display local bonding geometries similar to those of P- and B-atoms. Finally, the threefold coordinated neutral Si^0 center is the singly-occupied Si-dangling bond which has received much attention as a defect in amorphous silicon and amorphous silicon alloys,[18] and as an E′-center in SiO_2 and Si_3N_4 glasses and thin films.[19,20]

The application the VSEPR model to the Si-, N-, O-, and H-atoms thereby identifies possible local bonding geometries for these atoms in the body of the oxide, and at the Si—SiO_2 interface. In the discussions that follow the focus is on formation of two types of defects which are of importance in NMOS devices, interfacial traps and fixed positive charge. A parallel set of defect interfactions occurs in PMOS structures, but will not be addressed specifically in this paper.

III. DEFECT BONDING AT Si—SiO_2 INTERFACES AND Si—SiO_2 HETEROSTRUCTURES

Because of its importance in Si-based microelectronics, the Si—SiO_2 materials system has received much attention.

For details, the reader is referred to papers in four topical conferences held between 1976 and 1992.[21–24] Defect studies have proceeded into two different ways: they have been made on: (i) defects in capacitor and FET structures, before and after current and voltage stress-biasing, and (ii) defects induced in SiO_2 thin films and bulk glasses as well, or in Si—SiO_2 heterostructures by exposure to intense UV-radiation, x-rays, γ-rays, and/or charged or neutral particles (e.g., electrons and neutrons, respectively). The defect concentrations of interest in electronic devices are typically in the range from about 10^{10} cm^{-2} to 10^{12} cm^{-2}, whereas defect levels induced by radiation, and/or charged and neutral particles are usually substantially higher, up to 10^{14} cm^{-2}. After irradiation, defect levels are high enough so that they can be studied by spectroscopic techniques, including optical absorption and ESR, and as noted above, there may be qualitative differences in defect bonding geometries in such materials so that they are not necessarily described in terms of the VSEPR model.

Recently, the focus has shifted away from bulk oxide defects to interfacial and near-interfacial defects since these are the ones that determine performance and reliability in devices with ultra thin gate dielectrics. Two types of interfacial defects dominate in n-channel, or NMOS structures: (i) interface trapping states, denoted as D_{it}, and (ii) fixed positive charge, denoted as Q_{ss}.

The remainder of this paper focuses on these defects in the context of VSEPR bonding descriptions. The distinguishing aspects of this approach is based on *intrinsic* differences between bonding defects on group V and VI atoms, such as N and O, respectively, and the group IV atom Si. For example, within the framework of the VSEPR model, different charge states of N- and O-atoms are stable for different numbers of bonding neighbors, whilst, Si-atoms bonded to three neighbors are stable in three different charge states. Bonding configurations of N- and O-atoms with different charge states can contribute to shifts in flatband voltages of capacitors, threshold voltages of FETs, and charge carrier scattering in channel layers of FETs. Changing the charge state of these centers, either in a defect creation or neutralization process requires permanent trapping (effectively recombination) of a charged carrier, and a change in bonding coordination that can be accomplished by acquiring or losing a H-atom bonding partner. On the other hand threefold coordinated Si-atoms can act as trapping sites for charged carriers of either sign, and this does not require a change in bonding coordination.

IV. DEFECT REACTIONS AT Si—SiO_2 INTERFACES

Representative defect-generation reactions that can take place at, or in the immediate vicinity of Si—SiO_2 interfaces are described by the nine equations presented below. The changes of coordination of O-, N-, and H-atoms with changes in charge state in these reactions are consistent with the VSEPR model. In order for these reactions to take place, the energy of the resulting defect state must fall within the forbidden bandgap of SiO_2; however, in several instances, e.g., examples (i), (iii), and (iv) there is experimental data suggesting the reactions do indeed occur.[1] Nine representative reactions are:

(i) $\equiv Si—H+H \rightarrow \equiv Si^{o}+H_2$: generation of D_{it} centers ($\equiv Si^{o}$ is a Si dangling bond);

(ii) $\equiv Si—Si\equiv +H+h^{+} \rightarrow \equiv Si—H^{+}—S\equiv$: generation of Q_{ss} (H_2^{+});

(iii) $\equiv Si—O—Si\equiv +H+h^{+} \rightarrow H—O^{+}—(Si\equiv)_2$: generation of Q_{ss} (O_3^{+});

(iv) $\equiv Si—H+\equiv Si—O—Si\equiv +h^{+} \rightarrow$ $\equiv Si^{o}+H—O^{+}—(Si\equiv)_2$: generation of D_{it} and $Q_{ss}(O_3^{+})$;

(v) $H—N—(Si\equiv)_2+H+h^{+} \rightarrow H_2—N^{+}—(Si\equiv)_2)$: generation of $Q_{ss}(O_3^{+})$;

(vi) $\equiv Si—H+H—N—(Si\equiv)_2+h^{+} \rightarrow$ $\equiv Si^{o}+H_2—N^{+}—(Si\equiv)_2$: generation of D_{it} and $Q_{ss}(N_4^{+})$;

(vii) $H+N—(Si\equiv)_3+h^{+}-/no/->H—N^{+}—(Si\equiv)_3$: $N\equiv(Si\equiv)_3$ *does not act a defect precursor*;

(viii) $H+H—O—Si\equiv +h^{+} \rightarrow H_2—O^{+}—Si\equiv$: generation of $Q_{ss}(O_3^{+})$; and

(ix) $H—O—Si\equiv +H—Si\equiv +h^{+} \rightarrow H_2—O^{+}—Si\equiv + {}^{o}Si\equiv$: generation of D_{it} and $Q_{ss}(O_3^{+})$.

The notations O_3^{+} and N_4^{+} are shorthand descriptions for defect centers involving oxonium (e.g., $H—O^{+}—(Si\equiv)_2$ and ammonium (e.g., $H_2—N^{+}—(Si\equiv)_2$) bonding arrangements, respectively. A defect reaction in which the coordination of an atom is changed by virtue of the atom going from a neutral to a charged state will requires a network relaxation to accommodate the change in bonding geometry that accompanies that change in coordination. This can be a significant factor in either driving the defect reaction to the right as in (v) and (vi), or inhibiting the reaction in (vii) as will be discussed later on.

Figures 3(a) and 3(b) indicate schematically the defect generation reactions in (iii) and (v). The first reaction, (i) $Si—H+H \rightarrow \equiv Si^{0}+H_2$, demonstrates that H-atoms transported to the Si—SiO_2 interface can extract H-atoms from Si—H bonding sites thereby creating a dangling bond that acts as interfacial trap, D_{it}.[1] The driving force for this reaction is the large difference between the bond energies of Si—H, 3.4 eV, and H—H, 4.5 eV.[6] The Si—H bond that is attacked can be at the interface having at least one Si-atom neighbor, or in the oxide having three O-atom neighbors. The electrical activity of the dangling bond created by the H-atom abstraction reaction in depends on its position within the forbidden gap of Si and/or SiO_2. The energy of the neutral dangling bond in hydrogenated amorphous silicon relative the conduction band edge has been shown to be a function of the chemistry of its nearest neighbors, moving toward the conduction band edge as the electronegativity sum of the three nearest neighbors increases.[25] Since trapping kinetics depend on trap depth, the release times of Si-dangling bond traps will then vary significantly with their distance from the Si—SiO_2 interface.

The second reaction, (ii) $Si—Si\equiv +H+h^{+}$

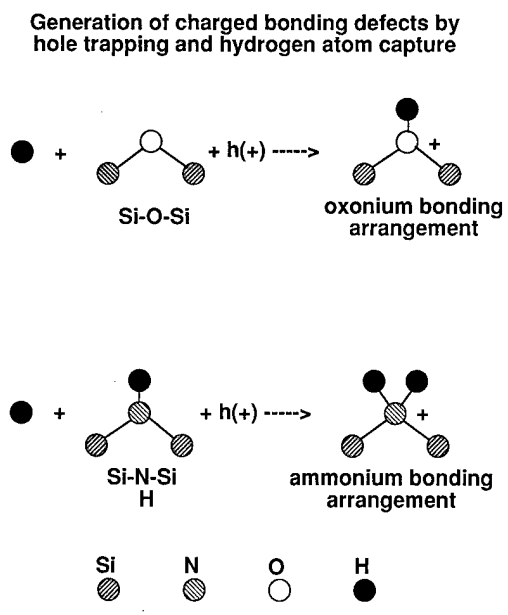


FIG. 3. Schematic representation of defect generation reactions for (a) oxonium centers (O_3^+), and (b) ammonium centers (N_4^+).

$\rightarrow$Si—H^+—Si$\equiv$, is newly proposed in this paper; note that $\equiv$Si—H^+—Si$\equiv$ is isoelectronic to B—H—B bridge in B_2H_6 molecules. Once created, the $\equiv$Si—H^+—Si$\equiv$ defect can be active in two ways: (i) it can contribute to fixed charge and thereby shift flatband and threshold voltages in devices, or (ii) it can serve as a recombination center for electrons, creating a Si-atom dangling bond in the process of trapping an electron as in Eq. (1):

$$\equiv Si—H^+—Si\equiv + e^- \rightarrow Si—H + \equiv Si^0. \quad (1)$$

The third reaction, (iii) $\equiv$ Si—O—Si$\equiv$+H + $h^+ \rightarrow$H—O +—(Si$\equiv$)$_2$, is similar to (ii) in the sense that it (i) involves the creation of a positively charged defect [see Fig. 3(a)], and (ii) requires concurrent H-atom attachment and hole trapping for creation. However, in this case de-activation of the H—O^+—(Si$\equiv$)$_2$ center by electron trapping returns the Si—O—Si bond group to its original neutral state freeing a H-atom without the creation of a Si-atom dangling bond as in Eq. (2);

$$H—O^+—(Si)_2 + e^- \rightarrow \equiv Si—O—Si\equiv + H. \quad (2)$$

The fourth reaction has been discussed in Ref. 26; if the defect pair are near-neighbors then trapping of an electron will return the local bonding to its original state, as in Eq. (3):

$$\equiv Si^0 + H—O^+—(Si\equiv)_2 + e^- \rightarrow \equiv Si—H + \equiv Si—O—Si\equiv. \quad (3)$$

In this case the H-atom motion will be displacive rather than diffusive, as would be required if the two bonding groups were more distant network neighbors. Reactions (v), see Fig. 3(b), and (vi), respectively, are analogs of reactions (iii) and (iv), in the sense that the neutral N—H group is isoelectronic with a neutral O-atom.

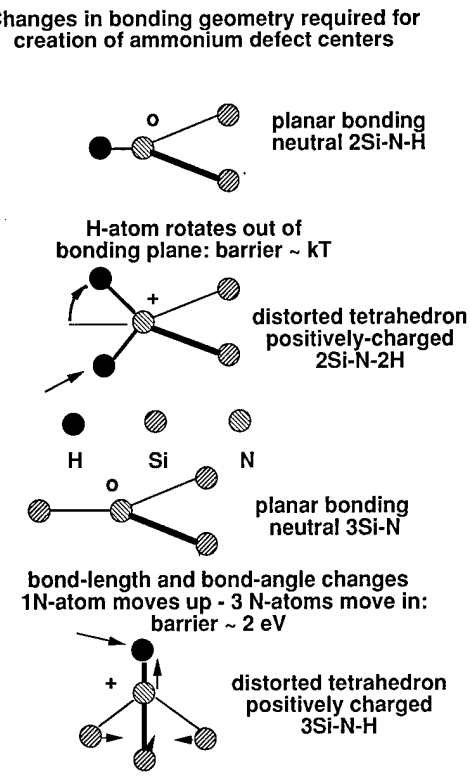


FIG. 4. Differences in defect precursor reactions between the bonding of N-atoms to (a) one H- and two Si-atoms, and (b) three Si-atoms.

There is another point regarding increases in coordination that take place when oxonium and ammonium centers are created. Following VSEPR rules, the geometry of an oxonium center is generally pyramidal so that relatively small changes in bond angles are required when the coordination is increased from two to three. These changes will not lead to significant energy barriers as estimated from empirical bond-bending force constants. Similarly, for creation of the ammonium center, the H-atom of the originally bonded N—H group must rotate out of the bonding plane and in combination with hole trapping and attachment of a second H-atom participate in the creation of a distorted tetrahedral arrangement (see Fig. 4). The energy barrier for this bond angle change is at most 0.02 to 0.05 eV, as estimated from empirically determined bond-bending force constants and will not represent a significant limitation to room temperature reaction kinetics. Reverse reactions for (iv) and (v) are given in Eqs. (4a) and (4b), respectively:

$$H_2—N^+—(Si\equiv)_2 + e^- \rightarrow H—N—(Si\equiv)_2 + H, \quad (4a)$$

and

$$\equiv Si^0 + H_2—N^+—(Si\equiv)_2 + e^- \rightarrow \equiv Si—H + H—N—(Si\equiv)_2. \quad (4b)$$

Figure 4 indicates a significant difference between the defect precursor behavior of N-atoms bonded only to Si-atoms, and to both Si- and H-atoms as in arrangement that includes an N—H bond. What is illustrated in Fig. 4 is the basis for comment in reaction (vii) that N—(Si$\equiv$)$_3$ *does not act as a*

defect precursor. The bonding of a N-atom to three Si-atoms is unique in the stereo-chemistry of N-atoms to three neighbors of the same chemical species. With the exception of a similar bonding between one N-atom and three Ge-atoms, the bonding arrangements of N-atoms bonded to three similar atoms or groups are pyramidal, e.g., as in ammonia, NH_3 and $N{\equiv}(CH_3)_3$.[6] However, for symmetric bonding to three Si- (or Ge-atoms) as in Si_3N_4 or silane-derived molecules, such as $N{-}(SiH_3)_3$, the bonding arrangement is planar. As noted above, these planar arrangements for Si (and Ge) are stabilized by $p\pi$-$d\pi$ bonding interactions.[6] Experiments have shown that unlike the incorporation of N—H bonding groups, the incorporation of N-atoms bonded only Si-atoms at Si—SiO_2 interfaces (i) does not degrade initial device performance, but instead (ii) improves reliability with respect to *un*-nitrided interfaces.[5] This aspect of defect/defect precursor behavior is also supported by *ab initio* calculations which demonstrate that there is an approximately 2 eV barrier to the conversion of a planar $N{-}Si_3$ group to a tetrahedral ammonium center by hole trapping and H-atom attachment.[27]

Finally, an additional intrinsic defect precursor is the the H—O—Si≡ configuration [see examples (viii) and (ix)]. The hydroxyl (H—O) group can be converted to a charged defect by hole trapping as in examples (viii) and (ix). Neutralization can occur by electron capture, as for example in Eqs. (5) and (6):

$$H_2{-}O^+{-}Si{\equiv}+e^-\rightarrow H{-}O{-}Si{\equiv}+H, \tag{5}$$

and

$$H_2{-}O^+{-}Si{\equiv}+{}^0Si+e^-\rightarrow H{-}O{-}Si{\equiv}+H{-}Si{\equiv}. \tag{6}$$

The reactions involving creation of O_3^+ and N_4^+ are supported by *ab initio* calculations which have been presented in Ref. 26 which demonstrated:

(i) The stability of the positively charged oxonium and ammonium centers O_3^+ and N_4^+, respectively.
(ii) The instability of neutral oxonium and ammonium centers, O_3^+ and N_4^0, respectively.
(iii) The absence of significant barriers (i.e., $E_b{>}kT$) to defect generation processes in (ii) to (v).

The *ab initio* calculations on small molecules have shown that the O—H, and N—H bond energies for the O_3^+ and N_4^+ are 5.7 eV and 5.0 eV, respectively, whereas the same bonds are *de-stabilized* by energies of 3.4 and 1.5 eV, respectively for the O_3^0 and N_4^0 centers (for details, see Ref. 26). In addition other studies (see, Refs. 15 and 16) have specifically addressed bond-centered hydrogen in crystalline Si. Based on the calculations presented in Refs. 15 and 16, and an empirical relationship presented in Ref. 17, the binding energy of a $Si{-}H^+$ bond is estimated be between 2.5±0.5 eV; i.e., greater than the Si-Si bond energy of 2.34 eV, but less than the Si—H bond energy of 3.35 eV.[6] For example, an upper limit can be estimated from the expression given in Ref. 17 for the bond dissociation energy, E_d, of a Si—H^+ bond in a three center bonding arrangement,

$$E_d = k\mu_{H-Si}(\Delta I)/R, \tag{8}$$

where k is an empirical constant equal to $=2.44\times10^{18}$ C^{-1} μ_{H-Si} is the bond moment, ΔI is an effective ionization energy, and R is the bond length. If we assume that the bond moment is determined by one positive charge at a distance R from the Si-atom, and take $\Delta I = 7.8$ eV, following the procedure given in Refs. 6 and 17, then E_d is approximately equal to 3.0 eV, the upper estimate given above.

V. DISCUSSION

This article has made several points relative to defects and defect precursors at, or in the immediate vicinity of Si—SiO_2 interfaces. The first is based on the application of the VSEPR model to amorphous materials and their interfaces and deals with a distinction between bonding defects in which the number of nearest-neighbor bonds is independent of their charge state, and bonding defects in which the number of nearest-neighbors changes with their charge state. The threefold coordinated Si atom represents the first type of defect. It can exist in three different charge states according to its electron occupancy without any change in bonding coordination. The positions of the three trapping levels relative the band edges of the crystalline Si substrate, and/or the SiO_2 dielectric, will be different for different nearest-, and more distant-neighbor bonding arrangements. For example, for the neutral centers with an electron occupancy of one, the more electronegative these neighbors, the higher into the gap is the energy state of the defect. This means that the release times for Si-atom dangling bonds that have trapped charge carriers will vary considerably depending on their distance from the Si—SiO_2 interface, since moving the Si-atom dangling bond into the oxide increases effective electronegativity of the atoms backed to that dangling bond.

In contrast, the *over-* and *under*-coordinated bonding arrangements with respect to the twofold and threefold coordinations for neutral O- and N-atoms, are not stable unless a change in charge state accompanies the change in coordination. In this context, the over- and under-coordinated bonding arrangements are charged, either positively or negatively, respectively. The three charge states of these two atoms can then be represented by O_1^-, O_2^0 and O_3^+, and N_2^-, N_3^0, and N_4^+. The inter-conversion between these charge states requires capture of a hole or electron and this is represented in the following symbolic reactions:

$$O_2^0+e^-\rightarrow O_1^-;O_2^0+h^+\rightarrow O_3^+;O_3^++e^-\rightarrow O_2^0;O_1^-+h^+\rightarrow O_2^0. \tag{9}$$

Each of the hole capture reactions reactions is accompanied by capture of a neutral bonding partner, and each of the electron capture reactions, by release of a neutral bonding partner; in the examples presented above these are H-atoms. A similar set of symbolic reactions can also be written for the N-atoms, where the coordination of the N-atom in the expression below includes at least one N—H bonding pair:

$$N_3^0+e^-\rightarrow N_2^-;N_3^0+h^+\rightarrow N_4^+;N_4^++e^-\rightarrow N_3^0;N_2^-+h^+\rightarrow N_3^0. \tag{10}$$

In addition to O- and N-atoms, H-atoms are stable in the neutral and positively-charged states and the relevant symbolic reaction is given below:

$$H_1^0 + h^+ \rightarrow H_2^+ ; H_2^+ + e^- \rightarrow H_1^0. \quad (11)$$

By virtue of the changes in coordination that accompany changes in charge state, these types of defects are active in two different ways: in their charged states (i) they change surface potentials of the crystalline Si at the channel in an FET, and thereby change the threshold voltage, and the carrier mobility; or (ii) they act as terminal traps for charged carriers. In their neutral states O- and N-atoms can trap free carriers of either sign; however, these trapping events require changes in bonding coordination, so that H-atom participation (or its equivalent) is required, and the trapping is once more terminal.

There are significant differences between the defect precursor behaviors of N-atoms bonded to one H-atom and two Si-atoms, N-atoms bonded to three Si-atoms. The planar geometry H–N–$(Si\equiv)_2$ centers can convert to a tetrahedral H_2–N^+-$(Si\equiv)_2$ by hole trapping and attachment of a H-atom with at most a small reaction barrier associated with a rotation of the originally bonded H-atom out of the original bonding plane. However, the conversion of a planar N–$(Si\equiv)_3$ group to a tetrahedral H–N^+-$(Si\equiv)_3$ arrangements presents a significant energy barrier, ~2 eV, due to the bond-length changes required. It has been demonstrated that N-atom incorporation at Si—SiO_2 interfaces improves device reliability with respect to interfaces which contain either Si—O or SiN—H bonds.[5]

The defect generation processes (i) and (iii) through (ix) can also take place in bulk oxides, nitrides or oxynitride alloys, or at the interfaces between dielectrics layers in stacked dielectric structure such an oxide-nitride-oxide (ONO) composite. Experiments performed on ONO sandwich structures indicated significant fixed positive charge ($\sim 10^{12}$ cm^{-2}) at the internal O—N and N—O interfaces prior to a 30 s 900 °C RTA,[28] and a significant reduction by more than a factor of 10 after the anneal. This reduction is attributed to release of H-atoms from bonding configurations involving N—H groups.

The chemical bonding framework for characterizing defects, and defect precursors for the Si—SiO_2 interface, can be applied to other semiconductor-dielectric interfaces including interfaces between SiC, Ge, (Si,Ge) alloys and oxides and/or nitrides. These extensions must take into account differences in bonding between (i) Si and O, and (ii) C or Ge and O, that may change the qualitative aspects of some of the types of reactions described above. For example, one factor to be considered for Ge and O is that there are two different bonding arrangements between Ge and O, octahedral and tetrahedral as exemplified by the rutile- and quartz-like oxides (both of which have the same composition GeO_2).[6] A second factor is the tendency of C and O to from multiple bonds as in the π-bonding in both CO_2 and CO; this carries over to C and N as well, where the gaseous species is CN.

The defect generation processes in (iv), (vi), and (ix) that result in the charged H_2^+, O_3^+, and N_4^+ centers, and where the H-atom is donated from a near-neighbor Si—H bond thereby creating a Si-dangling bond defect partner, can be also viewed as an extension of the valence alternation pair (VAP) model originally proposed for defects in chalcogenide glasses.[29] In its original form the VAP model included pairs of negatively and positively charged defects which had coordinations in their charged states that were different from the neutral bonded atoms. As in the discussion above, the bonding coordination of the negatively-charged centers was one lower, and the bonding coordination of the positively charged center was one higher than neutral bonding, so that in amorphous selenium, the defect generation reaction could be written as

$$2Se_2^0 \rightarrow Se_3^+ + Se_1^-, \quad (12)$$

where, the subscripts again indicate the number of nearest-neighbors, and superscripts the charge state. The defect generation reactions in (iv), (vi), and (ix) are similar to those in Eq. (12), except that the *valence alternation pair defect* that is produced consists a neutral Si-atom dangling bond, Si_3^0, and a positively charged O- or N-atom, O_3^+ or N_4^+, respectively, so that the pair is charged rather than neutral as Eq. (11).

ACKNOWLEDGMENTS

This work has been supported by ONR, the NSF, and SRC.

[1]D. J. DiMaria, Microelect. Eng. **28**, 63 (1995); E. Cartier and J. H. Stathis, *ibid.*, 3 (1995).
[2]R. E. Stahlbush, E. Cartier, and D. A. Buchanan, Micorelectron. Eng. **28**, 15 (1995).
[3]G. Q. Lo and D. L. Kwong, IEEE Electron Device Lett. **12**, 175 (1991).
[4]M. Green, D. Brasen, K. W. Evans-Lutterodt, L. C. Feldman, K. Krisch, W. Lennard, H.-T. Tang, L. Manchanda, and M.-T. Tang, Appl. Phys. Lett. **65**, 848 (1994).
[5]D. R. Lee, G. Lucovsky, M. S. Denker, and C. Magee, J. Vac. Sci. Technol. A **13**, 607 (1995); D. R. Lee, C. Parker, J. R. Hauser, and G. Lucovsky, J. Vac. Sci. Technol. B **13**, 1788 (1995).
[6]F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry*, 3rd ed. (Interscience, New York, 1972).
[7]R. J. Bell and P. Dean, Discuss. Faraday Soc. **50**, 55 (1970).
[8]G. Lucovsky, F. L. Galeener, R. H. Geils, and R. C. Keezer, in *The Structure of Non-Crystalline Materials*, edited by P. H. Gaskell (Taylor and Francis, London, 1977), p. 231.
[9]D. V. Tsu, G. Lucovsky, M. J. Mantini, and S. S. Chao, J. Vac. Sci. Technol. A **5**, 1998 (1987).
[10]Y. Ma and G. Lucovsky, J. Vac. Sci. Technol. B **12**, 2504 (1994).
[11]D. L. Griscom, in *International Topical Conference on SiO_2 and Its Interfaces*, edited by S. T. Pantelides (Pergamon, New York, 1976), p. 232.
[12]T. Shimizu, *Extended Abstracts of the 1995 International Conf. on Solid State Devices and Materials* (The Japan Society of Applied Physics, Tokyo, 1995), p. 133.
[13]F. L. Galeener, W. Stutius, and G. T. McKinley, in *The Physics of MOS Insulators*, edited by G. Lucovsky, S. T. Pantelides, and F. L. Galeener (Pergamon, New York, 1980), p. 77.
[14]G. Lucovsky, J. Non-Cryst. Solids **141**, 241 (1992).

[15]D. J. Chadi and C. H. Park, Phys. Rev. B **52**, 8877 (1995).
[16]C. G. Van de Walle, Phys. Rev. B **49**, 4579 (1994); C. G. Van de Walle, and N. H. Nickel, *ibid.* **51**, 2636 (1995).
[17]J. E. Huheey, *Inorganic Chemistry,* 2nd ed. (Harper and Row, New York, 1978).
[18]D. C. Allen and J. D. Joannopoulos, in *The Physics of Hydrogenated Amorphous Silicon II*, edited by J. D. Joannopoulos and G. Lucovsky (Springer, Berlin, 1984), p. 1.
[19]D. L. Griscom and W. B. Fowler, in *International Topical Conference on SiO_2 and Its Interfaces,* edited by S. T. Pantelides (Pergamon, New York, 1976), p. 97.
[20]W. L. Warren, P. M. Lenahan, and S. E. Curry, Phys. Rev. Lett. **65**, 207 (1990).
[21]*International Topical Conference on SiO_2 and Its Interfaces,* edited by S. T. Pantelides (Pergamon, New York, 1976).
[22]*Physics of MOS Insulators,* edited by G. Lucovsky, S. T. Pantelides, and F. L. Galeener (Pergamon, New York, 1980).
[23]*The Physics and Chemistry of SiO_2 and the Si—SiO_2 Interface,* edited by C. R. Helms and B. E. Deal (Plenum, New York, 1988).
[24]*The Physics and Chemistry of SiO_2 and the Si—SiO_2 Interface II,* edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993).
[25]G. Lucovsky and S. Y. Lin, AIP Conf. Proc. **120**, 55 (1984).
[26]Z. Jing, G. Lucovsky, and J. L. Whitten, J. Vac. Sci. Technol. B **13** (1995).
[27]Z. Jing (unpublished).
[28]Y. Ma, T. Yasuda, S. Habermehl, and G. Lucovsky, J. Vac. Sci. Technol. B **11**, 1533 (1993).
[29]M. Kastner, D. Adler, and H. Fritzsche, Phys. Rev. Lett. **37**, 1504 (1976).

Oxidation of silicon (100): Experimental data versus a unified chemical model

P. Thanikasalam, T. K. Whidden, and D. K. Ferry
Nanostructures Research Group, Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-5706

(Received 22 January 1996; accepted 25 February 1996)

It has been observed in many studies of oxidation kinetics that silicon dioxide growth in the thin regime (<30 nm) is faster than the predictions of the linear-parabolic Deal–Grove relationship. We have developed a conceptual and mathematical model for the thermal oxidation of silicon which is based on an initial dissociative chemisorption of molecular oxygen on the 2×1 silicon (100) surface followed by subsequent diffusion and dissociative reactions of molecular oxygen at the Si/SiO_2 interface. This model accounts for current experimental observations on the structural modification of the reconstructed surface to a 1×1 superlattice on limited exposure to molecular oxygen and provides a mechanistic rationale for the self-limiting mechanism of the oxide film at ~0.6 nm during low-temperature oxidation processes. The rate-equation model, consistent with the proposed reactions at the Si/SiO_2 interface, has been refined to give an excellent fit to experimental data within all thickness regimes, but especially in the initial rapid growth regime where the growth rate is proportional to $(P_{O_2})^{1/2}$. The rate equation reduces to a linear dependence on oxygen partial pressure in the thicker regime, where it predicts classic Deal–Grove behavior. We present the development of the model for oxidations performed between 780–1100 °C consistent with observed oxide growth and film properties. The activation energies for the reaction-controlled regime and the diffusion-controlled regime are consistent with literature data. We observe a sharp transition in a characteristic length parameter at ~950 °C which may be possibly due to the experimentally observed change in oxide film density. A detailed analysis and explanation is presented. © *1996 American Vacuum Society.*

I. INTRODUCTION

Silicon dioxide is, arguably, the most important component in the fabrication of a metal–oxide–semiconductor (MOS) device. Despite extensive research over the last thirty years, accurate and quantitative models of the initial oxidation of the silicon surface have proven illusive.[1–12] Models of the oxidation process in the thick oxide regime have been successfully developed,[3,4,8] but fail when applied for the prediction of oxidation rates at thicknesses less than approximately 20 nm. These failures may be ascribed to an incomplete understanding of the growth mechanisms in the very early stages of oxidation. Very large scale (VLSI) and ultralarge scale integration (ULSI) device fabrication requires the formation of silicon dioxide films ranging in thickness from 3 nm to 1000 nm, and as a consequence there is an immediate need for models capable of predicting oxidation rates in the thin regime.

The thermal oxidation of silicon has usually been modeled within the framework originally formulated by Deal and Grove[3] in the mid-1960's. They modeled the oxidation rate in terms arising from two processes: the chemical reaction of oxygen with silicon at the oxide/substrate interface and the diffusion of oxygen through the growing oxide film. The combination of rate relationships for these processes yielded the classic Deal–Grove "linear-parabolic" rate law. This model does not, however, accurately predict the growth curves observed for oxides of thicknesses below 20 nm. Within this "thin oxide" regime, significantly higher oxidation rates are observed than predicted by the Deal–Grove model. Massoud and Plummer[4] have proposed a modification of the original Deal–Grove model which incorporates two additional exponential terms, one in order to obtain a better fit in the initial regime of oxidation and the other for the intermediate oxidation phase. Their model improves upon Deal–Grove rate predictions for the thin oxide regime, but fails at oxide thicknesses of less than 4 nm. A variety of other models have been reported in the literature,[5–15] but none have successfully yielded accurate rate predictions in the thin oxide regime. Our model is based upon kinetic relationships involving the interfacial production of a substoichiometric oxide, SiO, in addition to the final product SiO_2. In this article, we compare the predictions of an oxidation rate model that we have developed[16] with the available experimental data for both the thin and thick oxide regimes.

II. OXIDATION MECHANISM

Our model is based on the assumption that the oxidation proceeds via the dissociative chemisorption of molecular oxygen, initially at the silicon surface and subsequently at the oxide/silicon interface.[16] We assume that the starting Si(100) surface has a 2×1 reconstruction with one dangling bond per silicon atom and a number of unique adsorption sites for the oxygen molecule. The geometries and energetics

of the possible adsorption sites on the Si(100) 2×1 surface have been discussed by Oshiyama and Miyamoto[17] and by Zheng.[5]

The driving force for the initial reaction between oxygen and the silicon surface is orbital mixing between the oxygen molecule and the silicon surface. The highest occupied molecular orbitals (HOMOs) of the oxygen molecule consist of two half-filled, orthogonal and degenerate orbitals of antibonding symmetry. These may accept electron density from the highest energy filled orbitals at the silicon surface, presumably the dangling bonds associated with the 2×1 reconstruction. The incorporation of electron density from the surface silicon into the oxygen π^* orbital weakens the oxygen-oxygen bond, leading to dissociation of the oxygen molecule and the insertion of an oxygen atom into neighboring silicon-silicon backbonds. Depending on the temperature, three stages of oxidation at the surface may exist. At low temperatures, there is insufficient energy available for dissociation of the silicon surface dimers and oxygen atoms may bridge between dimer rows, yielding half monolayer coverage of the surface. Such coverage has been reported for low temperature exposure of reconstructed Si(100) to molecular oxygen.[18] At higher temperatures (but insufficient for sustained oxidation of the silicon), dissociation of the silicon dimers leads to a 1×1 oxygen superlattice on the surface, consistent with literature.[19] Finally, the oxide films thickness is expected to self-limit at approximately 0.5–0.6 nm at temperatures below those required to drive oxygen diffusion through the surface oxide, due to inaccessibility of silicon-silicon bonds to incoming molecular oxygen.[16] This thickness is consistent with self-limiting oxide films observed on Si(100). At temperatures and pressures sufficient for sustained oxidation of the silicon substrate (i.e., conditions supporting significant diffusion of molecular oxygen through the growing oxide), our model suggests a layer-by-layer growth of the film.[16]

III. OXIDATION KINETICS

Adopting a Langmuir kinetic formalism,[20] the chemical reaction sequence for the oxidation of silicon may be written as

$$\mathrm{Si}+\tfrac{1}{2}\mathrm{O_2}\underset{k_i}{\overset{k_1}{\Leftrightarrow}}\mathrm{SiO}, \tag{1}$$

$$\mathrm{SiO}+\tfrac{1}{2}\mathrm{O_2}\overset{k_2}{\rightarrow}\mathrm{SiO_2}. \tag{2}$$

Based on chemical kinetics and Fick's law of diffusion the expression for the total growth rate is then

$$N\frac{dx}{dt}=k_3\frac{D_{\mathrm{eff}}C^*}{D_{\mathrm{eff}}+k_3x}+\alpha C^*e^{-x/a}. \tag{3}$$

Thus:

$$\frac{dx}{dt}=\frac{k_3D_{\mathrm{eff}}C^*}{N[D_{\mathrm{eff}}+k_3x]}+\frac{\alpha C^*}{N}e^{-x/a}, \tag{4}$$

where k_3 is the linear reaction rate at the interface, D_{eff} is the diffusion coefficient, C^* is the equilibrium concentration of the oxygen molecules in the gas phase, N is the number of oxygen molecules in a unit volume of the oxide, α is a pre-exponential fitting parameter and x is the oxide thickness. We interpret the parameter a which appears in the exponential term as a characteristic length associated with the onset of diffusional limits on the transport of oxygen to the oxide/silicon interface.

Expression (4) for dx/dt may be compared with existing experimental data on the oxidation rate of silicon. In the thin limit (<20 nm), our model successfully predicts the observed oxide growth rates with significantly better accuracy than previously reported models. Beyond oxide thicknesses of ca. 20 nm, our relationship gives rates that are comparable with those predicted by the Deal–Grove model and with experimental observations.

In the thin regime $[x \ll (D_{\mathrm{eff}}/k_3)]$, the Deal–Grove model reduces to

$$x\cong\frac{k_3C^*}{N}t. \tag{5}$$

In our model, for $x \ll (D_{\mathrm{eff}}/k_3)$, Eq. (4) reduces to

$$\frac{k_3D_{\mathrm{eff}}C^*}{N}dt\cong\frac{k_3D_{\mathrm{eff}}}{\alpha}e^{-x/a}dx, \tag{6}$$

and

$$x\cong\frac{\alpha C^*}{N}t. \tag{7}$$

Equation (7) is found to match the observed growth rates within the thin oxide regime if

$$\frac{k_3}{\alpha}<1. \tag{8}$$

When the fitting parameter α in equation (4) is set to zero, the relationship reduces to the Deal–Grove form where the growth can be characterized by the two Deal–Grove parameters k_3 and D. The model has been fit to the available experimental data in the thick oxide regime and the corresponding values k_3 and D are obtained (see Fig. 5). Using these activation energies for k_3 and D in an Arrhenius relationship, values of k_3 and D are obtained for different temperatures and the value of a is varied to fit the experimental data in the thin oxide regime, while the value of α is kept constant. The only variable fitting parameter in the thin regime is a. The value of α is kept constant at 4.0×10^6 μm/min.

IV. RESULTS AND DISCUSSIONS

We have chosen data taken from Massoud, Plummer, and Irene,[4] Irene and Van de Meulen,[21] Kamigaki and Itoh,[22] and Chao, Lee, and Lei[23] as experimental values of oxidation rates with which to develop and test our model. Massoud's samples[4] had ca. 1.0 nm of oxide prior to thermal oxidation whereas Irene[8,11,21] reported the presence of 0.3–0.6 nm of oxide before the wafers were introduced into the oxidation

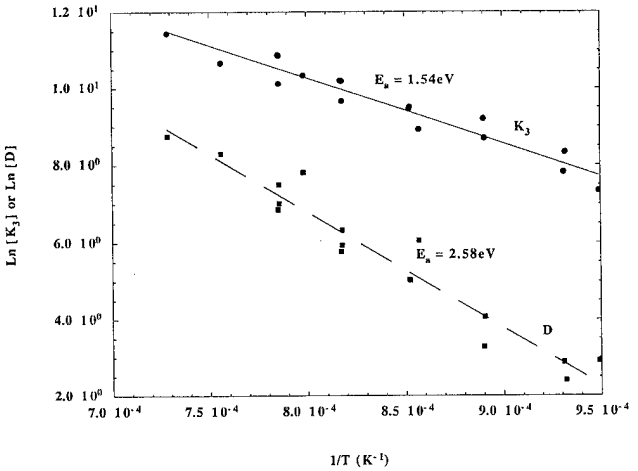


FIG. 1. The diffusion coefficient D and the linear reaction-coefficient k_3 plotted as a function of temperature. The activation energies are given next to the plots.

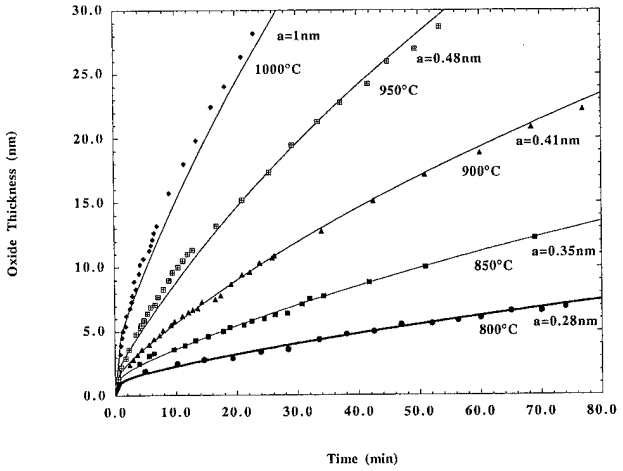


FIG. 2. The data of Massoud, Plummer, and Irene (Ref. 4) versus our model. The corresponding values of the parameter a which give the best fit are given against the oxidation temperature.

chamber. The samples of Kamigaki and Itoh[22] reportedly had a 0.6 nm thick oxide layer while the ellipsometry of Chao, Lee, and Lei[23] detected a 1.6–2.0 nm oxide film prior to oxidation. Massoud, Irene, and Chao used an *in situ* ellipsometer whereas Kamigaki's samples were taken out of the oxidation furnace in order to perform the oxide thickness measurements by ellipsometer. The reported initial oxide thicknesses agreed quite well with the initial oxide thicknesses found in fitting the Deal–Grove model to determine the k_3 and D values. All the samples were lightly doped. All the experiments compared in this paper were performed at atmospheric pressure. Following the first evaluation, equation (4) was used to fit each set of data, but with k_3 and D evaluated from Fig. 1. The only variable was the parameter a.

Figure 1 shows the linear reaction coefficients k_3 and the diffusion coefficients D for all the data on a single plot. The associated activation energies for k_3 and D are 1.54 eV and 2.58 eV, respectively. The activation energy for k_3 is close to the Si-Si bond energy of 1.83 eV,[24] consistent with a mechanism in which the rate controlling step of the interfacial reaction is the dissociation of Si-Si back bonds. The higher activation energy observed for the diffusion coefficient is consistent with decreased growth rate observed in the thick oxide regime, where the rate controlling step in the oxidation is the diffusion of molecular oxygen rather than the interfacial reaction. It has been previously reported[3] that the activation energy associated with the bulk diffusion of oxygen in fused silica is 1.17 eV. This is far smaller than the value derived from our model. It may be that the structural nature of the silicon dioxide formed in thermal oxidations of silicon differs significantly from that of the silica studied in these earlier reports, with concommittent differences in the characteristics of oxygen transport through the lattice. This could lead to the observed differences in activation energies. Certainly, lattice matching considerations would suggest that thin films of silicon dioxide on silicon, at least near the oxide/substrate interface, are subject to significant compressive stress. This could induce the formation of metastable structure in the oxide film having diffusion characteristics for oxygen that differ significantly from those of unstressed silicon dioxide.

Figure 2 shows the oxide thickness as a function of oxidation time for experimental data of Massoud, Plummer, and Irene[4] and fits to our model. The value of the characteristic length, a, varies between 0.28 nm and 1.0 nm over the temperatures 800–1000 °C. In Fig. 3, we show the comparison between our model and Irene's experimental data. Figure 4 shows only the growth curves for those experiments performed at atmospheric pressure. In Fig. 5 we show the comparison between our model and data of Chao, Lee, and Lei. The fit is excellent for all temperature ranges in both sets of data. The range of the characteristic length a is 0.3–1.0 nm

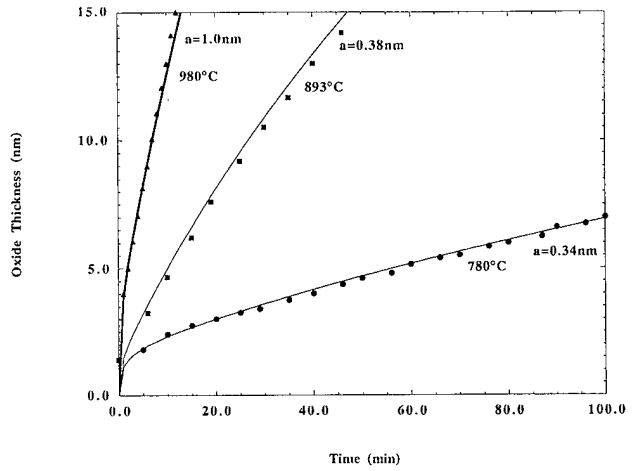


FIG. 3. The data of Irene and van der Meulen (Ref. 21) versus our model. The corresponding values of the parameter a which give the best fit are given against the oxidation temperature.

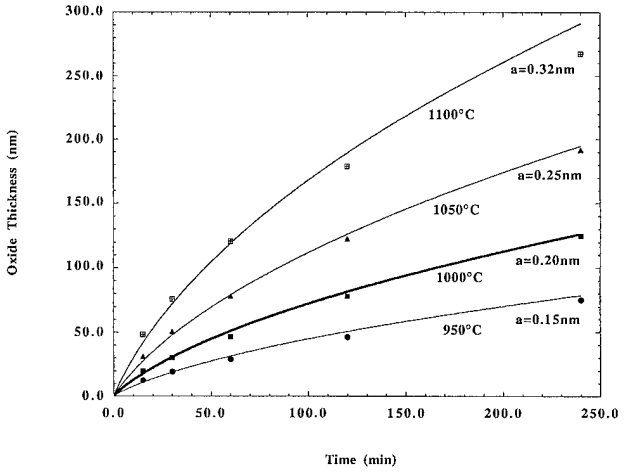


FIG. 4. The data of Chao, Lee, and Lei (Ref. 23) versus our model. The corresponding values of the parameter a which give the best fit are given against the oxidation temperature.

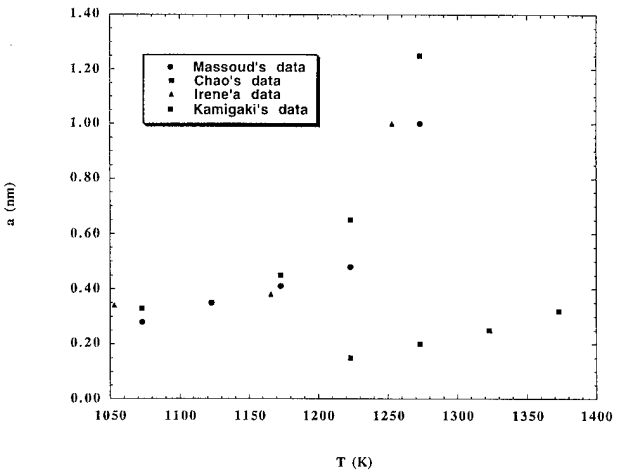


FIG. 6. Variation of parameter a with oxidation temperature, where α is kept constant at ca. 10^6 μm/min.

for data of Irene and van der Muelen[21] and 0.33–1.25 for data of Ref. 23, consistent with that observed in the more complete data set of Ref. 4. Kamigaki's experiments were done within several different pressure regimes. The range of a values found for Kamigaki and Itoh[22] experiments were from 0.15 to 0.30 nm for temperature ranges of 950–1100 °C, a distinctly different range than in any other data set. The origin of this anomalous range for Kamigaki's data is not presently clear. Kamigaki and Itoh's experiments were performed using nitrogen to dilute the oxygen gas. It may be that the molecular nitrogen blocks some of the interfacial adsorption sites, resulting in atypical reaction rates at the interface. This may give rise to ranges of a that are significantly different from those observed in the other studies compared in this article. Minimal data points in the experiments of Refs. 22 and 23 prevented us from making a distinction between reaction and diffusion controlled regimes for computation of the activation energies.

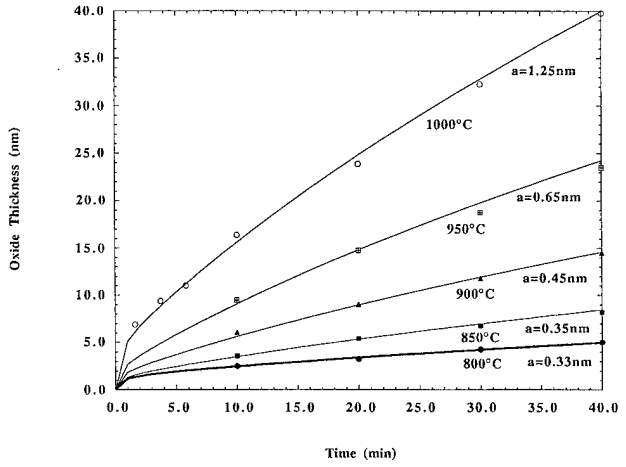


FIG. 5. The data of Kamigaki and Itoh (Ref. 22) versus our model. The corresponding values of the parameter a which give the best fit are given against the oxidation temperature.

In Fig. 6, we plot the values of the characteristic length a as a function of the growth temperature. The average value of a for the temperatures up to 950 °C is ca 0.35 nm, approximately the height of the Si-O moiety at the Si/SiO_2 interface. The rise in a above 950 °C presumably signals modified interfacial processes or a structural change in the oxide film or a combination of the two. The parameter a may be characteristic of an initial surface oxide formation analogous to the self-limiting oxide formation observed at much lower temperatures. At low temperatures, a is ca. 0.33 nm, consistent with an initial film formed by monolayer coverage of oxygen atoms on the silicon surface. At increased oxidation temperatures oxygen molecules may diffuse further through the initially formed oxide layer to the point at which no unoxidized silicon back bonds are available for reaction, yielding somewhat larger values for a. Alternatively or additionally, the variations observed in a may arise from structural differences within the oxide film which affect diffusion processes for molecular oxygen. It has been reported by Irene, Tierney, and Angiello[25] that there is a change in the density of the oxide film at ca. 950 °C. They have observed a 3% change in density between films formed at 600 °C and those formed at 1150 °C. Based on the values reported by Irene, Tierney, and Angiello[25] 2.28–2.20 g/cm^3, comparisons with available data[26] indicate the formation of either tridymite or cristobalite forms of silica. The change in the value of a may thus be related to variations in the diffusion of oxygen through the growing oxide film which in turn depends on the structural porosity of the silicon lattice. Such a structural variation would not necessarily need to be present throughout the oxide film but could be localized to regions near the oxide/substrate interface. Indeed, given the amorphous character of thermal oxides, it is most likely that regions possessing crystalline order are localized near the interface. The available data do not permit us to differentiate between pos-

sible forms of silica in this system or whether there is actually a structural change occurring at 950 °C.

The fitting parameter, α, is kept constant so that $\alpha/k_3 > 1$. This parameter describes the rate at which the oxygen molecules penetrate the Si 2×1 reconstructed surface to oxidize the upper most layers of the substrate. A typical value of α in the temperature range of 800–1000 °C is $\sim 10^6$ μm/min, whereas the linear reaction rate coefficient, k_3, is $\sim 10^3$ μm/min. We expect α to have a stronger dependence on gas phase oxygen pressure rather than the oxidation temperature. Apparently, in this initial, extremely fast growth regime, the diffusivity of oxygen does not play a major role.

V. CONCLUSIONS

We have compared relevant experimental data for the dry oxidation of silicon with a model invoking the dissociative chemisorption of oxygen at the interface between a growing silicon dioxide film and the silicon substrate. This relationship has been tested against experimental data taken from the literature and found to give an excellent fit. The relationship exhibits a dependence on the value of a characteristic length, a, which is observed to vary with temperature. The value of a is believed to be related to the thickness of an initial oxide film produced by an extremely fast reaction of oxygen with the 2×1 reconstructed silicon surface. Alternatively, variations in a may arise from variations in the structural characteristics of all or part of the silicon dioxide films grown at varying temperatures. Further experimental evidence is required before conclusive remarks could be made on the structural evolution of the initial oxidation phase.

ACKNOWLEDGMENT

This work was supported by the Advanced Research Projects Agency.

[1]S. Wolf and R. N. Tauber, in *Silicon Processing for the VLSI Era* (Lattice Press, Sunset Beach, CA, 1986), Vol. 1.
[2]S. M. Hu, J. Appl. Phys. **55**, 4095 (1984).
[3]B. E. Deal and A. S. Grove, J. Appl. Phys. **36**, 3770 (1965), and references therein.
[4]H. Z. Massoud, J. D. Plummer, and E. A. Irene, J. Elec. Chem. Soc. **132**, 2693 (1985).
[5]X. M. Zheng and P. V. Smith, Surf. Sci. **232**, 6 (1990).
[6]L. N. Lie, R. R. Razouk, and B. E. Deal, J. Elec. Chem. Soc. **129**, 2828 (1982).
[7]Y. J. van der Meulen, J. Elec. Chem. Soc. **119**, 530 (1972).
[8]E. A. Irene, J. Elec. Chem. Soc. **125**, 1708 (1978).
[9]J. Blanc, Appl. Phys. Lett. **33**, 424 (1978).
[10]A. G. Revesz and R. J. Evans, J. Phys. Chem. Solids **30**, 551 (1969).
[11]E. A. Irene, J. Appl. Phys. **54**, 5416 (1983).
[12]J. M. Gibson and D. W. Dong, J. Elec. Chem. Soc. **127**, 2722 (1980).
[13]A. Fargeix, G. Ghibaudo, and G. Kamarinos, J. Appl. Phys. **54**, 2878 (1983).
[14]S. A. Ajuira, P. U. Kenkare, A. Nghiem, and T. C. Mele, J. Appl. Phys. **76**, 4618 (1994).
[15]B. J. Brstik and P. J. McMarr, Phys. Rev. B **48**, 17 972 (1993).
[16]T. K. Whidden, P. Thanikasalam, M. J. Rack, and D. K. Ferry, J. Vac. Sci. Technol. B **13**, 1618 (1995), and references therein.
[17]A. Oshiyama and Y. Miyamoto, in *New Horizons in Low-Dimensional Electron Systems* (Kluwer Academic, Amsterdam, 1992).
[18]H. Yaguchi, K. Fujita, S. Fukatsu, Y. Shiraki, R. Ito, T. Igarashi, and T. Hattori, Surf. Sci. **275**, 395 (1992).
[19]T. Ohmi, M. Morita, A. Teramoto, K. Makihara, and K. S. Tseng, Appl. Phys. Lett. **60**, 2126 (1992).
[20]M. A. Morris, M. Bowker, and D. A. King, in *Comprehensive Chemical Kinetics*, edited by H. Bamford, C. F. H. Tipper, and R. G. Compton (Elsevier, New York, 1984), and references therein.
[21]E. A. Irene and Y. J. van der Meulen, J. Elec. Chem. Soc. **123**, 1380 (1976).
[22]Y. Kamigaki and Y. Itoh, J. Appl. Phys. **48**, 2891 (1977).
[23]T. S. Chao, C. L. Lee, and T. F. Lei, J. Elec. Chem. Soc. **138**, 1756 (1991).
[24]L. Pauling, *The Nature of Chemical Bond*, 3rd ed. (Cornell University, Ithaca, NY, 1960).
[25]E. A. Irene, E. Tierney, and J. Angiello, J. Elec. Chem. Soc. **129**, 2594 (1982).
[26]A. F. Wells, *Structural Inorganic Chemistry* (Clarendon, London, 1975), p. 805.

Kinetics of field-induced oxidation of hydrogen-terminated Si(111) by means of a scanning force microscope

L. Ley,[a] T. Teuschler, K. Mahr, S. Miyazaki,[b] and M. Hundhausen
Institut für Technische Physik, Universität Erlangen-Nürnberg, Erwin-Rommel-Strasse 1, D-91058 Erlangen, Germany

(Received 22 January 1996; accepted 24 March 1996)

Using an atomic force microscope (AFM) with a highly doped Si tip (0.01–0.02 Ω cm) we induced the oxidation of hydrogen-terminated Si(111) surfaces by scanning the surface with an applied bias U_{ox} between sample tip (negative) and substrate (positive) in air of 60% relative humidity. The resulting oxide height was measured relative to the unoxidized substrate immediately afterwards using the same setup in the AFM mode. Typical oxide thicknesses of some nm were achieved. p-type as well as n-type samples were investigated and the oxidation time at each point was varied by varying the scan speed of the tip during oxidation. Some of the salient results are as follows: (i) There is a definite threshold voltage U_{th} such that no oxidation takes place below U_{th}; (ii) U_{th} depends on doping and is lowest (2.7 V) for n^+- and highest (5.4 V) for p^+-type material; (iii) U_{th} is independent of scan speed and thus oxidation time for speeds between 0.25 and 7 μm/s. Assuming that the oxidation time is inversely proportional to the tip scan speed during oxidation we tried to fit the data to current models of oxide growth (parabolic growth law, linear-quadratic growth law, and Mott–Cabrera mechanism). None of them fits the experimental data. Instead, we find that the results are well described by a power law of the form $Z=(U_{ox}-U_{th})\,\alpha_0\,(t_{ox}/t_0)^{\gamma}$ where Z and t are the oxide thickness and the oxidation time, respectively, and α_0 is a factor that depends on experimental conditions such as doping, tip shape, and humidity. The exponent γ has a value of 1/4 to within experimental uncertainty. © *1996 American Vacuum Society.*

I. INTRODUCTION

Electric fields are known to enhance the oxidation rate of Si since the pioneering work of Jorgensen.[1] Dagata *et al.*[2] were the first to apply the field-induced oxidation (FIO) to create an oxide pattern on Si with the help of a scanning tunneling microscope and since then a number of related reports have appeared in the literature.[3–6] The potential of these oxides as etch masks for device fabrication on the nanometer scale has been emphasized by Snow and Campbell[6] and the general field of scanning probe microscopy (SPM) based techniques to create nanometer-scale structures on different surfaces has been reviewed in Ref. 7. We recently demonstrated that oxide patterns written by a scanning probe microscope on H-terminated Si(100) exhibit an increase in friction force compared to the unoxidized Si surface that scales with the oxide thickness.[8]

In this article we report an investigation of the kinetics of field-induced oxidation of hydrogen-terminated Si(111) surfaces using a scanning force microscope (SFM) with a conducting tip. The advantage of this approach lies in the fact that the same probe can be used to oxidize and detect the resulting oxide thickness in consecutive scans without the need to remove the sample from the microscope.

II. EXPERIMENTAL DETAILS

The SFM was custom built at our laboratory employing rectangular 0.01–0.02 Ω cm n^+-type Si cantilevers (0.2 N/m) with integrated tips (apex radius ~20 nm) which could be electrically biased without metal coating. Oxidation by SFM was performed in air with ~60% relative humidity and at room temperature. The SFM was operated in contact mode at a constant loading force $F_{\perp}$ of ~20 nN.

We investigated oxide growth on p^+-type samples with a resistivity of 0.005–0.01 Ω cm, moderately doped 5.6–10.4 Ω cm p-type samples, and 0.01–0.02 Ω cm n^+-type material. Before oxidation, oxide-free H-terminated Si(111) surfaces [Si(111):H] were prepared on Czochralski (Cz) grown wafers. To this end, the Si(111) surfaces were chemically treated under clean room conditions using procedures after Ref. 9, finished by a 3 min dip in 40% aqueous NH_4F solution and a 3 s rinse in ultrapure water (18 MΩ cm). As a result, hydrophobic, H-terminated, and atomically stepped surfaces with flat terraces are formed which are well stabilized against oxide formation in air ambient.[9,10] A native oxide thickness well below 0.1 nm is expected to form within 3 h, the time scale of our experiments.[10] In contrast, in most cases atomic scale roughness remains on Si(100) after chemical treatment.

As judged by IR-absorption spectra of our samples their dominant surface termination is due to Si—H bonds with less than 10% contributions from either $Si–H_2$ or $Si–H_3$ configurations. Surface contamination fell short of the XPS detection limit (~0.7 at. %) and chemically shifted Si $2p$ components due to initial oxidation were absent.

For local FIO we used the method of Ref. 3, applying a positive voltage U_{ox} to the sample with respect to the tip at selected positions while scanning the sample surface. Subse-

[a]Electronic mail: lothar.ley@physik.uni-erlangen.de
[b]Permanent address: Department of Electrical Engineering, Hiroshima University, Higashi-Hiroshima 724, Japan.

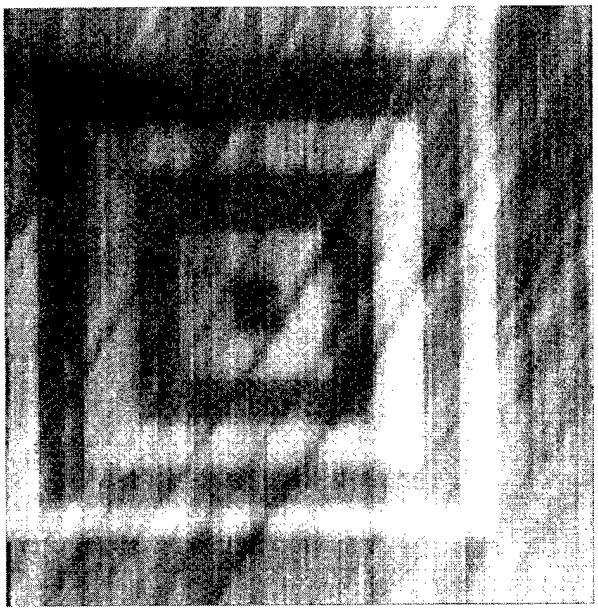


FIG. 1. Oxide pattern formed by field-induced oxidation using a conducting-probe scanning force microscope [p-type Si(111):H, ρ=5.6–10.4 Ω cm, U_{ox}=10 V, v_{ox}=2 μm/s, $F_{\perp}$=20 nN, image size: 1.2×1.2 μm^2].

quently, the resulting oxide-height z relative to the substrate was measured by the same SFM tip.

III. RESULTS

In Fig. 1 an example of the resulting topography after oxidation is shown for p-type material. In the bright areas a voltage U_{ox} of 10 V was applied during the previous patterning scan whereas in the dark areas U_{ox} was set to 0 V. During subsequent imaging tip and sample were grounded. The scan speed was 2 μm/s during oxidation as well as for imaging. The image size is 1.2×1.2 μm^2. Bright areas show protrusions formed by the FIO process, exhibiting an oxide height z of 3.8±0.5 nm above the unoxidized substrate. We find no increase of the topographical roughness over the oxidized areas compared to the unstructured surface. On thermally and naturally oxidized Si(111) surfaces, i.e., not patterned by FIO, atomic steps and terraces have been reported by Homma, Suzuki, and Yabumoto.[11]

The oxide height for a given voltage and scan speed depends on the doping of the substrate in the way depicted in Fig. 2 where we have plotted z vs U_{ox} for three differently doped substrates. The writing speed v_{ox} was kept fixed at 2 μm/s in all cases. It is apparent that the oxidation requires a certain threshold voltage U_{th} below which no oxidation is observed. This threshold voltage depends not only on the type of doping (n or p type) but also on the dopant concentration. We measure values for U_{th} of 2.7, 4.2, and 5.4 V for our n^+-, p-, and p^+-type material, respectively. The uncertainty in U_{th} is estimated at ±0.1 V. Above threshold the oxidation thickness increases linearly with U_{ox} in such a way that the slope dz/dU_{ox} increases from n^+- over p- to p^+-type material. It thus appears that in the material (p^+)

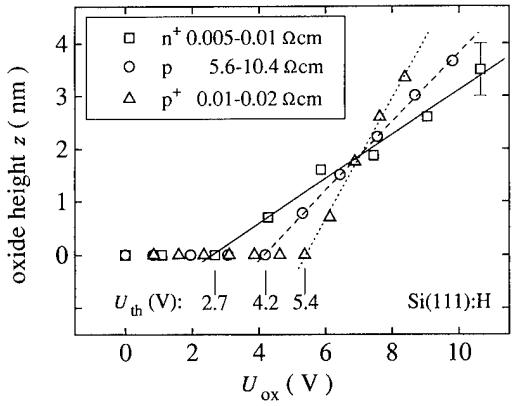


FIG. 2. Oxide height z as a function of the sample voltage U_{ox} applied for oxidation for three different doping levels (v_{ox}=2 μm/s, $F_{\perp}$=20 nN).

that requires the highest potential to initiate oxidation the oxidation rate also depends most sensitively on the field strength.

In Fig. 3(a) the oxide height is plotted versus U_{ox} for different scan speeds during oxidation. The sample is lightly p type. Again, no FIO takes place below U_{th} and this thresh-

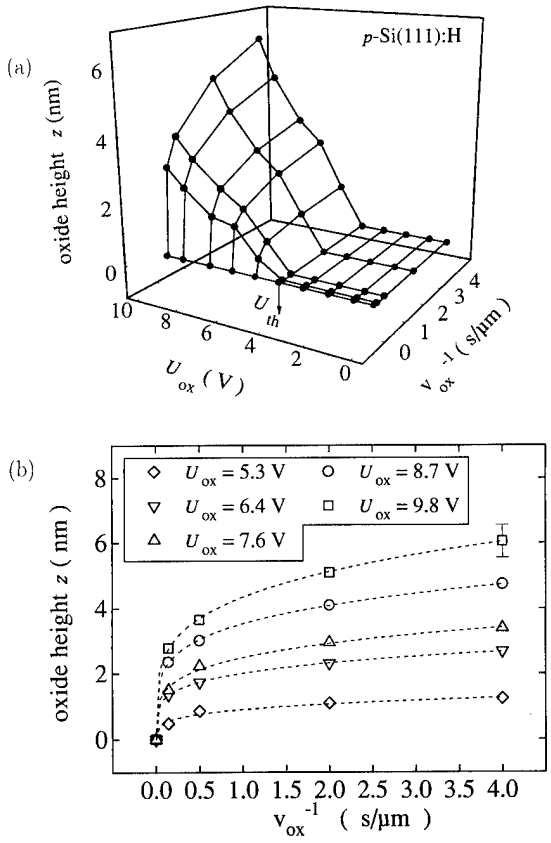


FIG. 3. (a) Oxide height z as a function of the sample voltage U_{ox} at different scan speeds v_{ox} during oxidation. (b) z as a function of the inverse scan speed v_{ox}^{-1} during oxidation at different voltages U_{ox} [p-type Si(111):H, ρ=5.6–10.4 Ω cm, $F_{\perp}$=20 nN].

old voltage is furthermore independent of the scan speed while above U_{th} the slope dz/dU_{ox} increases with decreasing v_{ox}. The data of Fig. 3(a) are replotted in Fig. 3(b) as z versus the inverse scan speed v_{ox}^{-1} during oxidation for a set of different voltages U_{ox}. Both higher v_{ox}^{-1} and higher U_{ox} lead to thicker oxide films.

IV. DISCUSSION

Before discussing these results in the light of models for the oxidation kinetics and other experiments it is necessary to establish the relationship between our experimental parameters oxide height z and scan speed during oxidation v_{ox} with the commonly employed parameters oxide thickness Z and oxidation time t. Clearly, the oxidation time at each point is inversely proportional to v_{ox} even if the proportionality factor is not known exactly. On the other hand, there is ample evidence[4,12,13] that the oxide height z above the original silicon surface is proportional to the total oxide thickness with a proportionality factor $z/Z \approx 0.44$.[13] Consequently, we use in what follows the parameters z and v_{ox}^{-1} synonymously with Z and t.

There is an extensive literature on the functional relationship between oxide thickness and oxidation time. The point of departure is in most cases the Deal–Grove model which combines the diffusion limited rate yielding $Z \sim \sqrt{t}$ appropriate for thick oxides with the initial linear regime $Z \sim t$ reflecting the interface reaction. This is the so-called "linear-quadratic" growth model $AZ^2+BZ=t-\tau$, where A and B are parameters and the integration constant τ takes an initially present oxide into account.[14] An attempt to fit our data with this model failed. An extension of the Deal–Grove model that considers a spatial variation of the diffusion coefficient in growth direction on account of the strain field likely to be present in the oxide was proposed by Blanc.[15] It yields a functional relationship between oxidation time and oxide thickness of the form $t-\tau=AZ^2+BZ-C\ln(1+DZ)$, where C and D are additional parameters describing the variation of the oxygen diffusivity through the oxide. This model fits the highly precise thermal oxidation data of Hopper, Clarke, and Young[16] exceedingly well. It nevertheless fails to describe our data despite the fact that Hopper and co-workers' data also cover the thickness regime below 10 nm. A major difference between Hopper and co-workers' data and ours is oxidation temperature: around 830 °C in their case and room temperature here. Needless to say, our data cannot also be fitted to the Cabrera–Mott model,[17] $Z^2 \sim Et$, where E is now proportional to the electric field across the oxide which we take to be U_{ox}/z.

We did, however, succeed in a two-parameter fit of the whole set of data to an empirical kinetic expression,

$$z=\alpha(v_0/v_{ox})^{\gamma}, \tag{1}$$

with parameters α and γ and an arbitrary constant v_0 which was chosen to be $v_0=1$ μm/s. The result is represented in Fig. 3(b) as dashed lines. The parameters α and γ obtained for a given value of U_{ox} are plotted in Fig. 4 versus U_{ox}. It is

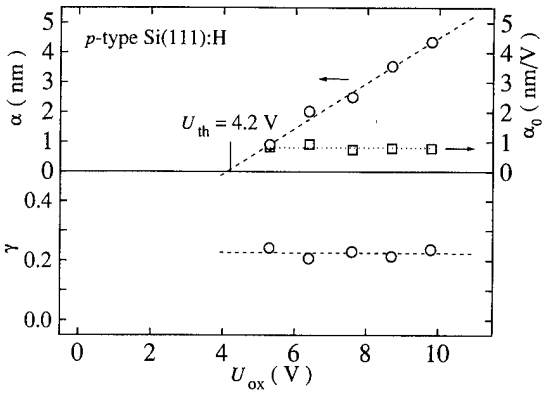


FIG. 4. Parameters α, γ, and α_0 from a fit of Eqs. (1) and (2) to the data of Fig. 3 vs the sample voltage U_{ox} [p-type Si(111):H, ρ=5.6–10.4 Ω cm, $F_\perp$=20 nN].

obvious from this plot that the exponent is independent of U_{ox} within experimental error: $\gamma=0.23\pm0.02$, while $\alpha=\alpha_0(U_{ox}-U_{th})$. We thus rewrite Eq. (1) as

$$z=\alpha_0(U_{ox}-U_{th})(v_0/v_{ox})^{\gamma} \quad (U_{ox}>U_{th}). \tag{2}$$

The normalized prefactor α_0 is about 0.78 nm/V for the p-type sample (Fig. 3) and we find a dependence on doping as well as on experimental conditions, such as tip shape and humidity. In order to substantiate Eq. (2) we replot the data of Figs. 2 and 3 in Fig. 5. Here, z vs $(U_{ox}-U_{th})v_{ox}^{-1/4}$ is shown, yielding straight lines for the respective complete sets of data. To check general applicability, we also fit Eq. (2) to data for FIO on n-type (2–6 Ω cm) Si(111) covered with a thin protective oxide, taken from Ref. 4 and shown in Fig. 5 as crosses. Here, too, we obtain $\gamma \approx 1/4$ while α_0 is smaller than for our samples.

Two examples of silicon oxidation not induced by an electric field that also obey the power law kinetics with an exponent $\gamma=1/4$ are given in Figs. 6 and 7. The data are taken from Refs. 18 and 19. The upper panels show in each case the oxide thickness Z versus oxidation time together with the power law (dashed line), whereas the lower panels give the

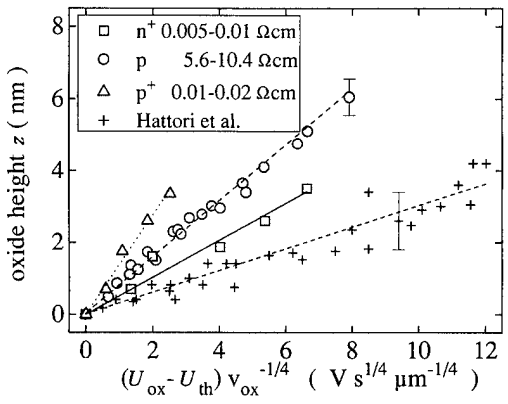


FIG. 5. Oxide height z as a function of $(U_{ox}-U_{th})v_{ox}^{-1/4}$ for the data of Figs. 2 and 3 [Si(111):H, $F_\perp$=20 nN], and data from Ref. 4.

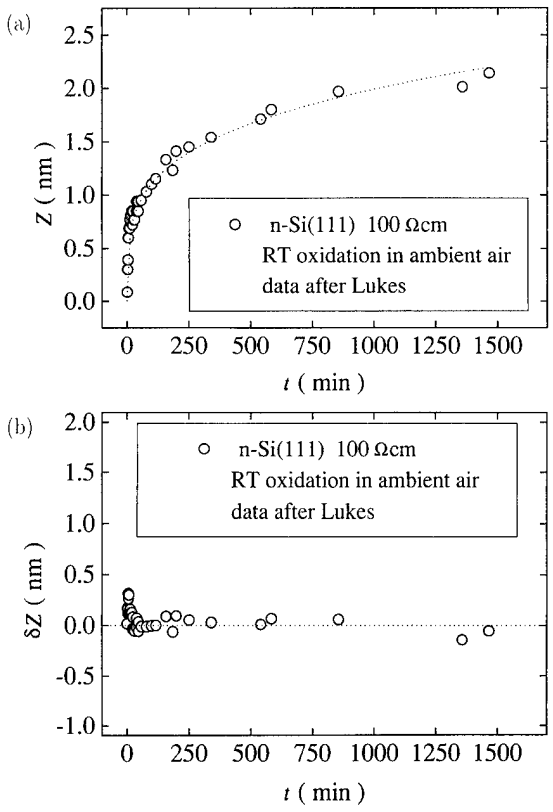


FIG. 6. (a) Fit of the oxidation data of Lukeš (Ref. 18) to a power law with exponent 1/4. (b) Residues of the fitting procedure.

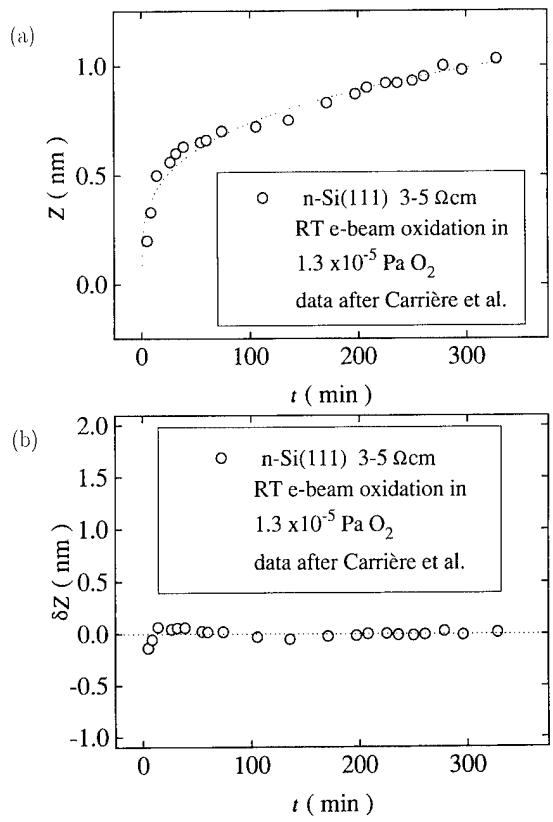


FIG. 7. Same as Fig. 6 for the oxidation data of Carrière and co-workers (Ref. 19).

residues. We have chosen these two examples because they both deal with RT oxidation in a thickness regime that is comparable to ours. At the same time the oxidation procedures differ in an important aspect. Carrière and co-workers[19] (Fig. 7) oxidize under ultrahigh-vacuum conditions in a partial pressure of 1.3×10^{-5} Pa O_2 and in the presence of electron bombardment. This must therefore be considered a dry oxidation process. In contrast, Lukeš[18] oxidizes under ambient conditions which implies a certain amount of humidity. Humidity is also present in our case and it appears to be a necessary condition for FIO as demonstrated recently by Kramer, van den Berg, and Schönberger who find oxidation rates that scale with relative air humidity in an experiment very similar to ours.[20] The results of Figs. 6 and 7 and our results therefore indicate that neither the electric field nor the presence of water vapor is essential for the observed kinetics. A description of oxidation kinetics by a simple power law is in fact not uncommon. Reisman *et al.*[21] have demonstrated that all dry oxidation data of Si published up to 1987 can be described by such a power law, albeit with exponents ranging generally between 0.57 and 0.90. Only in one case was an exponent as low as 0.24 observed which would agree with the data presented so far.

At the present state, the exact microscopic mechanisms underlying the observed doping dependence and the empirical growth relation remain speculative. Nevertheless, some aspects should be pointed out. A doped-silicon tip is used for sample oxidation which is initially covered with a thin native oxide which alternatively might also be worn away by contact during scanning. While a voltage drop over this tip oxide may partly account for U_{th} it cannot explain its variation with substrate doping. This behavior, however, as well as the variation of dz/dU_{ox} with doping, might be related to carrier depletion near the apex region of the tip[22] which we expect to depend on the doping of the tip as well as on that of the sample.

Further, before layer growth of the oxide commences it is necessary to reactivate the stabilized Si(111):H surface for initial oxidation. Becker *et al.* demonstrated *in vacuo* electron-stimulated desorption of hydrogen from n^+-type (0.007 Ω cm) Si(111):H by bombardment with electrons above 2 eV from a scanning tunneling microscope tip.[23] Schwartzkopff *et al.* in a similar experiment find ~4.5 V as a H-desorption threshold from p-type (~1 Ω cm) Si(111):H.[24] In both cases the electron current was in the nA range, while in our experiment it remained below 10 pA. This implies in case of Fig. 1 that an electron dose of less than 5 pC per scan line and/or the presence of the intense electric field between tip and sample, which itself is also dependent on oxide thickness, was sufficient for oxidation.

The rate limiting mechanism for oxide layer growth leading to the observed kinetics will be difficult to deduce from measurements of z versus oxidation time alone. Power-law

kinetics for Si oxidation have been derived assuming very different underlying mechanisms.[25,26]

V. SUMMARY

In conclusion, we have presented experiments on the doping dependence of SFM-based FIO of Si(111):H, as well as on its kinetics. FIO is observed only for voltages exceeding a doping-dependent threshold above which oxidation kinetics follows a power law incompatible with several classical models for oxide growth developed in the regime of thicker oxids. Further work should address differently doped tips, different tip materials, as well as a controlled variation of environmental conditions such as humidity, temperature, and gas partial pressures.

ACKNOWLEDGMENT

Part of this work was supported by the Deutsche Forschungsgemeinschaft under SFB 292.

[1]P. J. Jorgensen, J. Chem. Phys. **37**, 874 (1962).
[2]J. A. Dagata, J. Schneir, H. H. Harary, C. J. Evans, M. T. Postek, and J. Bennet, Appl. Phys. Lett. **56**, 2001 (1990).
[3]H. C. Day and D. R. Allee, Appl. Phys. Lett. **62**, 2691 (1993).
[4]T. Hattori, Y. Ejiri, K. Saito, and M. Yasutake, J. Vac. Sci. Technol. A **12**, 2586 (1994).
[5]D. Wang, L. Tsau, and K. L. Wang, Appl. Phys. Lett. **65**, 1415 (1994).
[6]E. S. Snow and P. M. Campbell, Appl. Phys. Lett. **64**, 1932 (1994).
[7]See for example, R. Wiesendanger, Appl. Surf. Sci. **54**, 271 (1992).
[8]T. Teuschler, K. Mahr, S. Miyazaki, M. Hundhausen, and L. Ley, Appl. Phys. Lett. **66**, 2499 (1995).
[9]G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari, Appl. Phys. Lett. **56**, 656 (1990).
[10]T. Yasaka, K. Kanda, K. Sawara, S. Miyazaki, and M. Hirose, Jpn. J. Appl. Phys. **30**, 3567 (1991).
[11]Y. Homma, M. Suzuki, and N. Yabumoto, J. Vac. Sci. Technol. A **10**, 2055 (1992).
[12]L. Tsau, D. Wang, and K. L. Wang, Appl. Phys. Lett. **64**, 2133 (1994).
[13]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).
[14]B. E. Deal and A. S. Grove, J. Appl. Phys. **36**, 3770 (1965).
[15]J. Blanc, Philos. Mag. B **55**, 685 (1987).
[16]M. A. Hopper, R. A. Clarke, and L. Young, J. Electrochem. Soc. **122**, 1216 (1975).
[17]N. Cabrera and N. F. Mott, Rep. Prog. Phys. **12**, 163 (1948).
[18]F. Lukeš, Surf. Sci. **30**, 91 (1972).
[19]B. Carrière, J. P. Deville, and A. El Maachi, Philos. Mag. B **55**, 721 (1987).
[20]N. Kramer, M. R. van den Berg, and C. Schönberger, presented at IVC-13/ICSS-9, Yokohama, 1995.
[21]A. Reisman, E. H. Nicollian, C. K. Williams, and C. J. Merz, J. Electron. Mater. **16**, 45 (1987).
[22]P. De Wolf, J. Snauwaert, T. Clarysse, W. Vandervorst, and L. Hellemanns, Appl. Phys. Lett. **66**, 1530 (1995).
[23]R. S. Becker, G. S. Higashi, Y. J. Chabal, and A. J. Becker, Phys. Rev. Lett. **65**, 1719 (1990).
[24]M. Schwartzkopff, P. Radojkovic, M. Enachescu, and E. Hartmann, Verhandl. DPG (VI) **30**, 1546 (1995).
[25]E. H. Nicollian and A. Reismann, J. Electron. Mater. **17**, 263 (1988).
[26]L. Li and J. Peeters, J. Appl. Phys. **74**, 639 (1993).

Photoinduced surface reactions of reverse-biased *n*-type porous Si

Eric J. Lee,[a] Theodore W. Bitner, Anselm P. Hall, and Michael J. Sailor[b]
Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0358

(Received 24 January 1996; accepted 24 March 1996)

Light induced surface reactions of reverse-biased *n*-type porous Si with H_2O, CH_3OH or CF_3COOH are studied by diffuse reflectance and transmission Fourier transform infrared spectroscopy. The goal of this work is to identify reagents which will react with the irradiated, electron deficient surface but not the non-irradiated, neutral surface of reverse-biased porous Si. The porous Si is the anode in an electrochemical cell with the desired reagent present as the electrolyte solution. H_2O/0.5M NaCl is found to react with the reverse-biased Si in both the light and the dark to form the native oxide. CH_3OH/0.1M $NaBF_4$ generates a small amount of methoxide-modified Si in both the light and the dark. CF_3COOH/1M CF_3COONa reacts with the surface when the porous Si is optically excited to create a surface-bound trifluoroacetic ester species. CF_3COOH/1M CF_3COONa does not react with reverse biased Si in the dark. The Si surface can be photolithographically patterned with the trifluoroacetic ester species. © *1996 American Vacuum Society.*

I. INTRODUCTION

In this article we describe a method for using the electronic structure of *n*-type Si to photopattern Si surfaces. The *n*-type Si is brought under positive potential control in an electrochemical cell. This is the reverse-biased condition where the bands are bent such that upon optical excitation, electrons are driven into the bulk of the semiconductor, changing the state of charge of the Si surface. We have investigated three common oxygen-containing nucleophiles—water, alcohols and carboxylic acids—to identify reagents which will react with the irradiated, positively charged Si surface but not the non-irradiated neutral Si, Scheme 1. The ability of a nucleophile to discriminate between the irradiated and non-irradiated surfaces will allow us to photopattern the species onto the Si surface.

There are few examples of light induced reactions with Si in the literature. Photoinduced electron transfer involving semiconductors and solution redox couples is an active area of investigation.[1] Irradiation of *n*-type Si has been shown to polymerize and dope polypyrrole films[2] and excitation of porous Si is known to enhance styrene polymerization.[3] Reactions at Si electrodes in solar cells have been studied in an effort to minimize corrosion.[4]

The Si used in this work is *n*-type Si (100) which has been etched in an HF acid solution to produce a photoluminescent porous layer. The surface of as-etched porous Si is H-terminated.[5] Some work on the surface chemistry of porous Si has been published, including electrochemical alkoxidation,[6] and modification of porous Si by chemical means.[7] An earlier communication of our work on photoelectrochemical modification of porous Si has been published.[8]

There are two advantages in using porous Si to study Si surface chemistry. Porous Si has a large surface area, aiding spectroscopic identification of surface species. A second advantage is that porous Si is photoluminescent[9] with an intensity which depends on the nature of the surface composition,[7] allowing for easy identification of patterned areas.

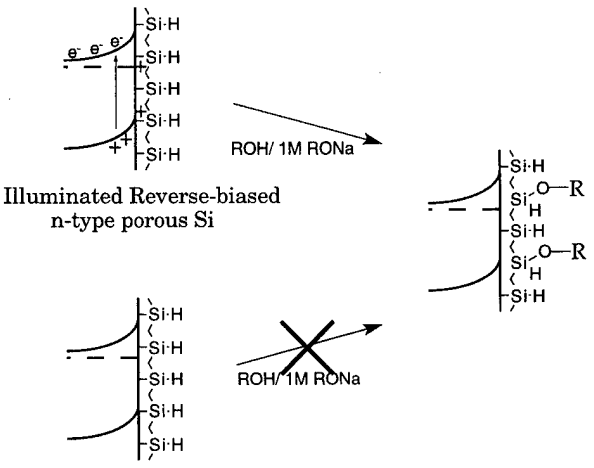
II. EXPERIMENT

A. Materials

The Si was (100) *n*-type 1 Ω·cm resistivity, polished on one side, purchased from International Wafer Service and used as received.

B. Apparatus

The light source was a tungsten filament Kodak ELH bulb controlled by a Variac. Light intensity was monitored by a calibrated photodiode. The potentiostat/galvanostat was a Princeton Applied Research Model 363. Fourier transform



[a]Electronic mail: e5lee@chem.ucsd.edu
[b]Author to whom correspondence should be addressed.

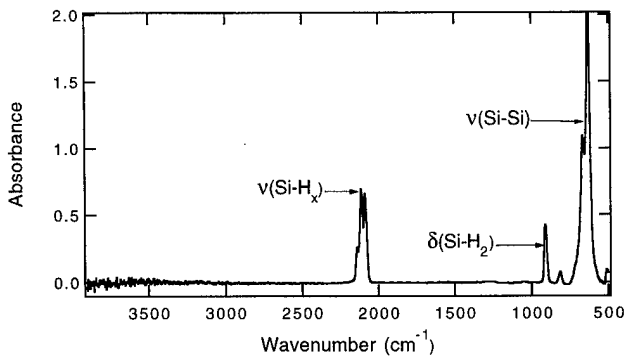


FIG. 1. Transmission FTIR spectrum of porous Si.

infrared (FTIR) spectra were obtained on a Nicolet Model 550 with an interchangeable diffuse reflectance apparatus by Spectra-Tech.

The electrochemical cell was custom made from Teflon and consists of two pieces, a top with a cylindrical bore of 1.2 cm diameter which acts as the solution bath, and a bottom piece with edges cut to fit the FTIR spectrometer transmission-mode sample holder. The Si was cut into squares approximately 2 cm on a side and secured between two O-rings by screwing together the bottom and top of the Teflon cell. Kalrez O-rings were found to withstand the carboxylic acid solutions much better than Viton. Both O-ring materials withstand HF. The counter electrode was a Pt wire coil.

C. Porous Si formation

The Si working electrode was galvanostatically etched in 1:1 49% HF(aq):CH_3CH_2OH at 88.5 mA/cm^2 for 2 min while under ~30 mW/cm^2 white light illumination. The resulting porous layer was washed thoroughly with CH_2Cl_2 and dried under N_2 stream. The porous Si sample was handled in air before and after derivatization.

The etch conditions were chosen to produce a thin, photoluminescent layer with large diameter pores. The current density of 88.5 mA/cm^2, near the electrochemical polishing regime, makes large diameter pores. An etch time of less than 2 min results in porous layers that are not photoluminescent while longer etch times produce thicker porous layers of which only a fraction is electrochemically accessible. The pore diameters have not yet been determined and the porous layer is estimated to be several microns thick.

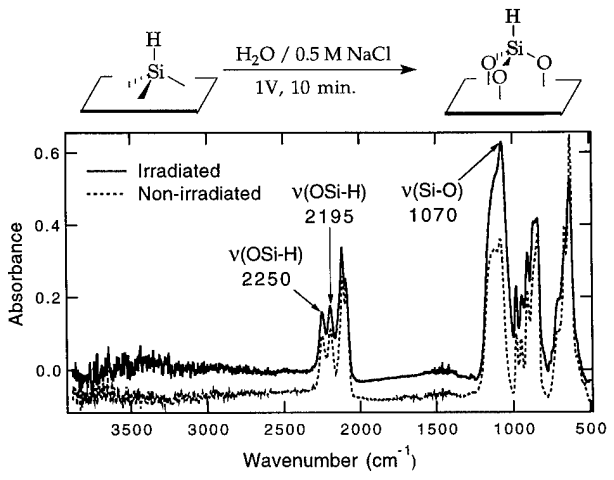


FIG. 2. FTIR spectra following electrolysis of porous Si in H_2O/0.5M NaCl at 1 V for 10 min while irradiated and not irradiated. Both samples develop oxide.

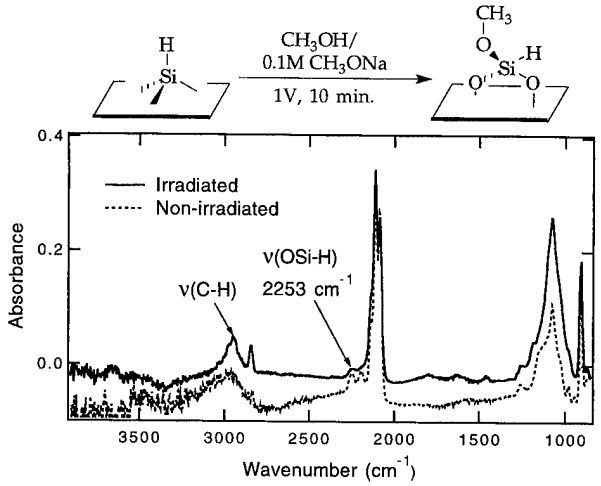


FIG. 3. FTIR spectra following electrolysis of porous Si in CH_3OH/0.1M $NaBF_4$ at 1 V for 10 min while irradiated and not irradiated. Both samples develop a small amount of the surface methoxide.

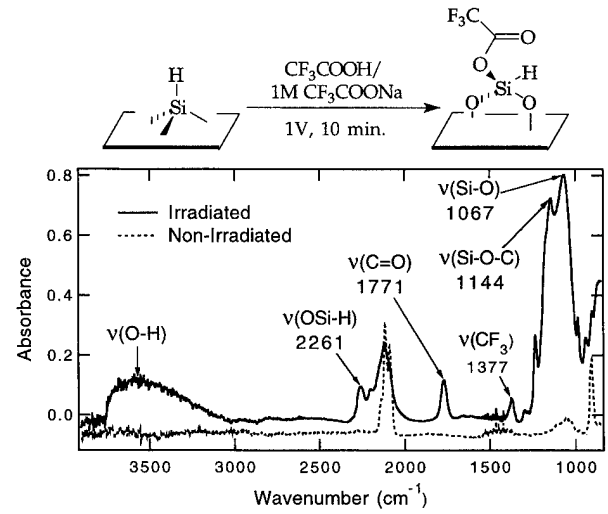


FIG. 4. FTIR spectrum following electrolysis of porous Si in CF_3COOH/1M CF_3COONa at 1 V for 10 min with irradiation indicates the surface has been modified with a trifluoroacetic ester species. In the dark, no reaction occurs.

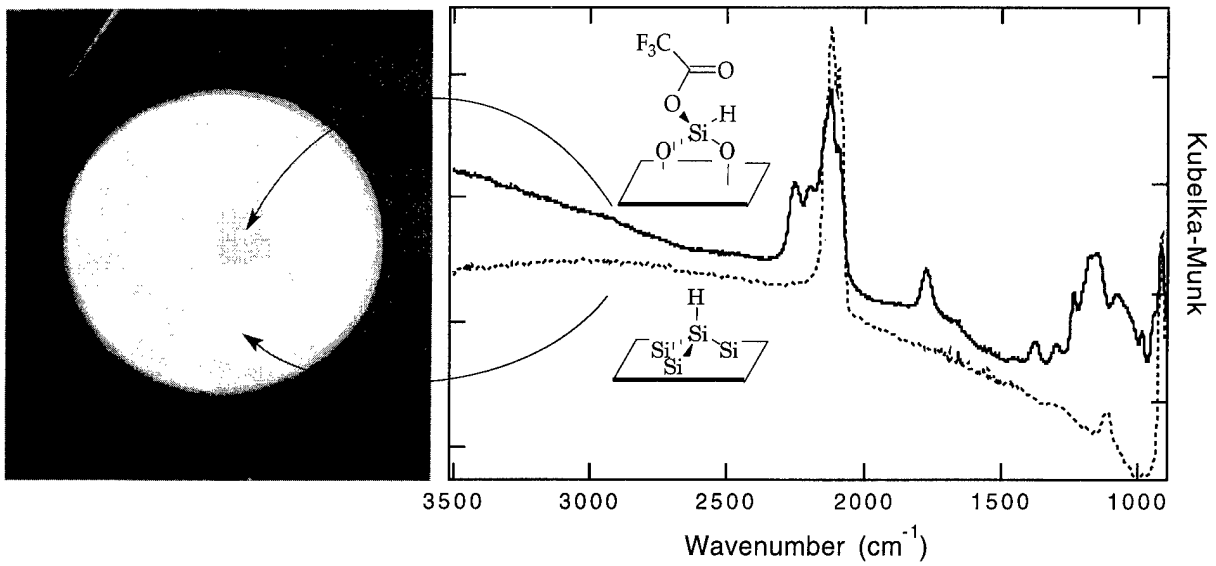


FIG. 5. Left: Photograph of the photoluminescence from a porous Si sample photopatterned with an ester-modified square region. Photoluminescence from the ester-modified region is less intense than that from the H passivated surface. Right: Diffuse reflectance FTIR spectra confirm that the patterned square is ester-modified porous Si, while the area outside the square retains the initial hydride surface.

D. Si Derivatization

Reagent solutions of H_2O/0.5M NaCl, CH_3OH/0.1M $NaBF_4$, CH_3OH/0.1M CH_3ONa and CF_3COOH/1M CF_3COONa were prepared. All chemicals except H_2O were purchased from Aldrich or Acros and used as received. H_2O was purified by a Barnstead Nanopure filtration system to a resistivity of 18M Ω·cm. The reagents were used as electrolyte solutions in the same Teflon electrochemical cell used to etch the porous Si (after thoroughly rinsing with CH_2Cl_2 to remove trace HF). Derivatization of the Si with the reagent solutions proceeded at constant potential of 1 V (Si biased positive) and a light intensity of 20 mW/cm^2 for 10–20 min. The derivatized sample was washed thoroughly with CH_3CN, CH_2Cl_2 and dried under N_2 stream. FTIR spectra were collected in a N_2-purged chamber. The modified surfaces were stable in air for at least the 30 min time period of the experiment. The surface-bound alkoxide or ester groups could be removed by washing with water or other alcohols, but not with CH_2Cl_2.

Photopatterned samples were generated by imaging a square mask placed in front of the light source through a Vivitar 70–180 mm zoom camera lens onto the Si surface. The patterned areas were squares of roughly 0.2 cm on a side, large enough to be sampled by front-face diffuse reflectance IR spectroscopy. Diffuse reflectance FTIR spectra are reported in Kubelka-Munk units.

III. RESULTS AND DISCUSSION

The porous Si spectrum, Fig. 1, is dominated by the ν(Si-H$_x$) stretching modes at 2089 cm^{-1} for ν(Si-H), 2116 cm^{-1} for ν(Si-H$_2$) and 2139 cm^{-1} for ν(Si-H$_3$).[10,11] A δ(Si-H$_2$) scissor mode absorption is present at 908 cm^{-1}. An absorption band at 627 cm^{-1} has been assigned to a combination of ν(Si-H) and ν(Si-Si) stretches. No absorption bands that can be assigned to Si-F stretches were detected.

The porous Si was oxidized by photoelectrolysis in H_2O/0.5M NaCl for 10 min at 1 V applied bias. The aqueous solution did not penetrate the porous layer well, but sufficient contact was made to generate a photocurrent within the cell. The photocurrent increased over time as oxidation of the Si surface rendered it more hydrophilic. FTIR spectroscopy identified the surface after reaction as the Si oxide, Fig. 2. Absorbances due to stretches of hydrides on Si backbonded to O appear at 2250 cm^{-1} and 2195 cm^{-1}. ν(Si-O) stretching modes at 1070 cm^{-1} and 840 cm^{-1} also appear upon derivatization. The ν(Si-Si) absorption band at 627 cm^{-1} is found to decrease upon oxidation.

Electrolysis in the dark generates a small dark current and surface oxidation almost as extensive as with the illuminated sample, Fig. 2. Apparently, H_2O/0.5M NaCl is too reactive with Si to discriminate between illuminated and non-illuminated Si.

Electrochemical derivatization of porous Si in CH_3OH/0.1M $NaBF_4$ in both the light and the dark results in a slight oxidation of the surface, and minimal deposition of surface methoxide, Fig. 3. Both the irradiated and non-irradiated samples have absorptions in the ν(Si-O) and ν(C-H) regions. The ν(C-H) absorptions exhibit the characteristic double peaks of -CH_3 symmetric and anti-symmetric stretching modes. The small absorptions at 2253 cm^{-1} are attributed to ν(OSi-H) and occur at the same frequency as the chemically prepared alkoxide-modified porous Si.[7] Chazalviel and co-workers studied electrolysis of p-type porous Si in CH_3OH,[6] and concluded that formation of Si-OCH_3 bonds activates that Si atom to further oxidation, leading to dissolution of surface Si (similar to the attack on Si by HF) rather than chemical modification.

Porous Si held at a +1 V under 20 mW/cm^2 white light illumination in CF_3COOH/1M CF_3COONa reacts to produce

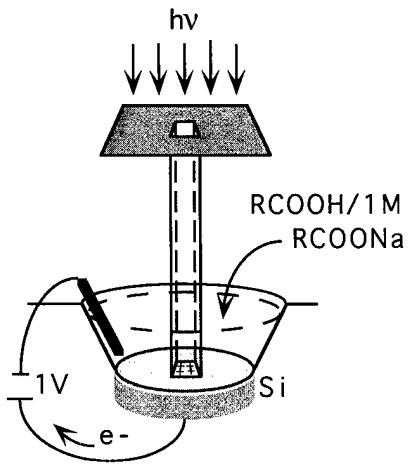
a Si ester modified surface, Fig. 4. The ν(OSi-H) stretching frequency is shifted to 2261 cm^{-1} due to the back-bonded trifluoroacetic ester species which has a greater inductive effect than either the simple oxide or the methoxide.[12] We also observed the appearance and growth of an absorption at 1771 cm^{-1} which we assigned to a carbonyl stretching mode. Shifts in the carbonyl stretching frequency are useful fingerprints for changes in its chemical environment. The ν(C=O) stretch of the trifluoroacetic acid starting material is at 1785 cm^{-1}. A model compound for the surface-bound species, $(CH_3)_3SiOOCCF_3$ (Aldrich), has a carbonyl absorption at 1772 cm^{-1}. Absorptions due to the CF_3 group are present at 1377 cm^{-1}. A ν(Si-O-C) asymmetric stretch is observed at 1144 cm^{-1}. A broad absorption around 3500 cm^{-1} is due to ν(O-H) stretches from surface-bound hydroxide and adsorbed moisture, and is an indication of the hydrophilicity of the surface. The presence of the ν(Si-O-C) absorptions, the presence and position of the ν(C=O) absorption and the position of the ν(OSi-H) absorption all indicate that an ester species has bonded to the surface. Binding in a bidentate carboxylate mode (where both O atoms of the CO_2 co-ordinate to the substrate) as is observed on many metal oxides is ruled out because such a species would have three $\nu(CO_2^-)$ stretches between 1600 cm^{-1} and 1400 cm^{-1},[13] none of which are observed.

Samples derivatized in the dark show no sign of surface modification, Fig. 4. A slight amount of oxidation due to handling is seen by the small ν(Si-O) band at 1070 cm^{-1}, but no ν(C=O) or ν(OSi-H) stretches are present.

As another control experiment, the counter electrode and Si were shorted together and illuminated. The photocurrent alone was sufficient to esterify the Si, though at a much lower rate than with an applied bias.

All three reagents, H_2O, CH_3OH, and CF_3COOH, react with porous Si upon photoelectrolysis. Water is reactive enough that the dark current is sufficient to oxidize Si. Electrolysis in CH_3OH in the light or the dark leads to dissolution of Si by the methoxide, leaving behind only a small amount of methoxide-modified Si. Trifluoroacetic acid reacts with the reverse-biased Si surface when the surface is sufficiently positively charged by irradiation but does not react with the non-irradiated Si.

The necessity for optical excitation of the Si to initiate reaction with carboxylic acids allows the Si surface to be photopatterned with the ester. The apparatus for photoelectrochemical patterning is shown in Scheme 2. A square region in the middle of the porous Si is illuminated during the electrolysis in CF_3COOH/1M CF_3COONa. Figure 5 shows a photograph of the photoluminescence from the circular porous Si sample patterned with an ester-modified square region. The ester-modified porous Si has a lower photoluminescence quantum yield because the ester provides additional surface sites for non-radiative recombination. Diffuse reflectance FTIR of the patterned surface identifies the square area as that containing the surface-bound ester, while the area outside the square has only the $\nu(Si\text{-}H_x)$ absorptions of the starting porous Si, Fig. 5.



IV. CONCLUSIONS

H_2O, CH_3OH, and CF_3COOH electrolyte solutions are found to react with the surface of reverse-biased n-type porous Si. H_2O was reactive in both the light and the dark, producing the simple oxide. CH_3OH produced only slight methoxide-modification in either the light or the dark. CF_3COOH was unreactive with the surface in the dark but optical excitation of the reverse-biased n-type Si rendered the Si surface susceptible to nucleophilic attack by CF_3COO^-, producing the trifluoroacetic ester-modified surface. This discrimination in reactivity between irradiated and nonirradiated areas allowed the porous Si to be photopatterned with the surface-bound trifluoroacetic ester species.

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation. One author (M. J. S.) acknowledges support from A. P. Sloan and Camille Dreyfus Foundation fellowships.

[1]M. Gratzel, *Heterogeneous Photochemical Electron Transfer* (Chemical Rubber, Boca Raton, FL, 1989).
[2](a) O. Inganas, T. Skotheim, and I. Lundstrom, Phys. Scr. **25**, 863 (1981); (b) O. Inganas and I. Lundstrom, J. Electrochem. Soc. **131**, 1129 (1984); (c) H. Yoneyama and M. Kitayama, Chem. Lett. 657 (1986).
[3]J. L. Heinrich, A. Lee, and M. J. Sailor, in *Microcrystalline and Nanocrystalline Semiconductors*, edited by R. W. Collins, C. -C. Tsai, M. Hirose, F. Koch, and L. Brus (Materials Research Society, Pittsburgh, 1995), Vol. 358, p. 605.
[4](a) A. J. Bard and M. S. Wrighton, J. Electrochem. Soc. **124**, 1706 (1977); (b) B. J. Tufts, A. Kumar, A. Bansal, and N.S. Lewis, J. Phys. Chem. **96**, 4581 (1992).
[5]P. Gupta, A. C. Dillon, A. S. Bracker, and S. M. George, Surf. Sci. **245**, 360 (1991).
[6]M. Warntjes, C. Vieillard, F. Ozanam, and J.-N. Chazalviel, J. Electrochem. Soc. **42**, 4138 (1995).
[7](a) E. J. Lee, J. S. Ha, and M. J. Sailor, in *Microcrystalline and Nanocrystalline Semiconductors*, edited by R. W. Collins, C. -C. Tsai, M. Hi-

rose, F. Koch, and L. Brus (Materials Research Society, Pittsburgh, 1995), Vol. 358, p. 387; (b) J. A. Glass, E. A. Wovchko, and J. T. Yates, Surf. Sci. **338**, 125 (1995).
[8]E. J. Lee, J. S. Ha, and M. J. Sailor, J. Am. Chem Soc. **117**, 8295 (1995).
[9](a) L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990); (b) S. Schuppler *et al.*, Phys. Rev. B **52**, 4910 (1995).
[10]J. M. Lavine, S. P. Sawan, Y. T. Shieh, and A. J. Bellezza, Appl. Phys. Lett. **62**, 1099 (1993).
[11]Y. J. Chabal, J. Mol. Struct. **292**, 65 (1993).
[12]D. R. Anderson, *Analysis of Silicones*, edited by A. L. Smith (Wiley, New York, 1974), pp. 247–286.
[13]J. E. Crowell, J. G. Chen, and J. T. Yates, J. Chem. Phys. **85**, 3111 (1986).

Stressing and high field transport studies on device-grade SiO_2 by ballistic electron emission spectroscopy

R. Ludeke,[a] H. J. Wen, and E. Cartier
IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598

(Received 22 January 1996; accepted 20 April 1996)

The tip of a scanning tunneling microscope was used to inject hot electrons across the gate and into the oxide of a metal-oxide-semiconductor structure. This method, known as ballistic electron emission microscopy (BEEM), allows an arbitrary choice of the energy of the injected electrons, which may be further accelerated by the application of a gate bias. The high current densities and choice of energy make BEEM an attractive method to study hot electron transport and breakdown phenomena in dielectrics. The studies reported here were made on $Pd/SiO_2/Si(100)$ structures with a SiO_2 layer thickness of 3.8 nm. Monte Carlo techniques were used to calculate the spreading of the electron beam as it traverses the oxide. A strong dependence of the spreading on the kinetic energy and oxide thickness were observed. Using the calculated beam spreads to determine current densities and injected charge densities, the charge-to-breakdown (Q_{bd}) was measured for several breakdown sequences. The Q_{bd}'s consistently exceeded by several orders of magnitude the values obtained by conventional Fowler–Nordheim (FN) tunnel injection under high field conditions. Most of the time, breakdowns could not be achieved for 3.8 nm oxides. It is concluded that impurity/defects still control all observed breakdowns; an intrinsic limit—although claimed to have been reached—has not yet been established. Changes in the BEEM spectra with injected electron charge are interpreted in terms of a three stage process to breakdown: (I) electron trap creation and filling at the SiO_2–Si interface, (II) prebreakdown believed to occur through thinning of the oxide that starts at the oxide-Si interface, and (III) oxide punch-through, characterized by an injection threshold $\lesssim 1$ V, close to that for Si. The role of hot hole injection into the SiO_2 valence band was also assessed and deemed a negligible factor in the degradation process under the zero or low oxide biases used in the experiment reported here. © *1996 American Vacuum Society.*

I. INTRODUCTION

From a design point of view, future generations of integrated circuits (IC) can be scaled well below the presently manufactured 0.25 μm technology, a projection that requires gate oxides ultimately reaching thicknesses below 2 nm.[1–3] However, the reliability of thin gate oxides is one of the most important problems in metal–oxide–semiconductor (MOS) technology. Hence, the ability of manufacturers to mass produce such aggressively scaled ICs will depend on the successful solution of a number of materials and processing problems, among which is the oxide integrity for gate oxides in the 4 nm range and below. This problem has been designated a potential show stopper.[4] In this thickness range, interface roughness, deviations from stoichiometry, and impurities in the oxide and at the SiO_2–Si interface are believed to be especially critical. Their roles in affecting device reliability and eventually oxide breakdowns are presently not known, and attempts at gaining further understanding of these issues are undertaken at a number of laboratories.[5–19] A fundamental physical description of oxide degradation and failure are of crucial importance in formulating models that predict device reliability and lifetimes from burn-in and accelerated aging studies.

Oxide breakdown studies are based on failure analysis of electrically stressed MOS capacitor structures or MOS field effect transistors (FETs). Conventionally, energetic (hot) electrons are injected into the conduction band of the SiO_2 by means of field assisted injection or Fowler–Nordheim (FN) tunneling. Following injection, the electrons gain energy through field acceleration until they reach the opposite interface. The injection may occur in a constant voltage or a constant current mode. The time is measured until device failure occurs (time to breakdown). This time, which is dependent on the fluence or total injected charge, exhibits a complex dependence on field and oxide thickness, which can be rationalized in terms of electron heating in the oxide film.[9] The total charge injected at failure is referred to as charge to breakdown, Q_{bd}. An example of the dependence of Q_{bd} on oxide field and thickness is shown in Fig. 1.[16] These Q_{bd}'s represent state-of-the-art values on high quality oxides. Their values have increased over the years as oxide growth and substrate cleaning methods have improved, implying that the origin of breakdowns are still associated with defects or impurities. It is general practice, however, that for the best oxides, such as those shown in Fig. 1, the breakdown is addressed as the *intrinsic* mode,[11,14,20] because breakdown distributions are narrow and quite distinct from early fails or low-field break-down. In a strict sense, however, *intrinsic* would imply an oxide free of defects, with the Q_{bd}'s reflecting the practical upper limit of oxide reliability.

A fundamental and little disputed aspect of oxide breakdown is the role of the hot electrons in generating charge

[a]Electronic mail: ludeke@watson.ibm.com

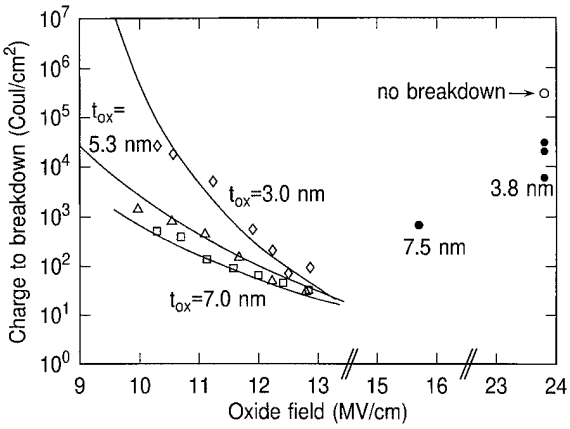


FIG. 1. Normalized charge to breakdown as a function of average electric field for indicated oxide thicknesses at room temperature. These results were obtained through field stressing (injection by FN tunneling at the polysilicon gate). Open data points are experimental results, solid lines are model calculations-adapted from Ref. 16. The solid data points represent oxide failures achieved with BEEM for 3.8 nm and 7.5 nm oxides. The energies of the STM injected electrons were converted to the equivalent field necessary to achieve those values.

traps at the interfaces. Microscopic details of such traps are either lacking or highly speculative, with the exception of traps generated at the cathode due to the presence of hydrogen. The latter is released at the anode through impact with energetic electrons (kinetic energy $\gtrsim$2 eV). The hydrogen diffuses to the cathode where it combines with another hydrogen atom and escapes. The release of the second hydrogen is believed to leave behind an interface state or an electron trap.[7,9] For thicker oxides under high oxide fields, conditions for which some electrons in the oxide reach energies exceeding 9 eV (the SiO_2 band gap energy), electron-hole pairs may be created in the oxide by impact ionization, with the holes being swept to the injecting cathode where they can be trapped by oxygen vacancies, for example. Electron-hole recombination can then create interface states and electron traps. This mechanism is important only for oxide thicknesses $\gtrsim$10 nm, and will not be further discussed, as these conditions are not encountered in the present experiments. Hole injection into the oxide may also occur at the Si–SiO_2 interface through electron-hole pairs generated by impact ionization of the hot electrons reaching the anode (Si substrate). Although deemed a low efficiency process,[21] some breakdown models rely exclusively on hole injection as the primary trap creation mechanism at the cathode, and the principal mechanism leading to progressive oxide degradation emanating from the cathode.[14,22] Other models deemphasize the role of hole trapping and attribute the start of oxide degradation to electron trapping at the anode.[12,13,15]

In this article, we present results of stressing thin SiO_2 layers by directly injecting hot electrons into the conduction band of SiO_2 with the tip of a scanning tunneling microscope (STM). This method obviates the need for field acceleration to energize the electrons in the SiO_2 conduction band, although this is still an option that allows additional experimental flexibility. The STM mediated injection provides unprecedented spatial resolution and a choice of the location of stressing. The results reported here were obtained for 3.8 nm device grade oxides. The principal conclusions of this study are: (i) in general, the oxides were difficult to break down in spite of injecting fluences exceeding by several orders of magnitude the reported Q_{bd} values of conventionally (FN) stressed oxides (such as those in Fig. 1) for comparable anode energies; (ii) when breakdowns were observed, the Q_{bd}'s far exceeded the best-of-breed published results; (iii) strong spectral evidence suggests that breakdown emanates from the anode following a clearly resolved trapping/charging stage and a prebreakdown phase; and (iv) hole injection appears to be inconsequential in breakdown processes initiated by STM-mediated hot electron injection.

II. EXPERIMENTAL DETAILS

A. Ballistic electron emission microscopy/ spectroscopy

Ballistic electron emission microscopy (BEEM) is a STM based microscopy that differs from conventional STM by the presence of a thin metallic or other conducting film overlying the substrate—the SiO_2–Si structure in the present case. The sole purpose of this conductive layer is to provide a potential reference plane (generally at ground potential) relative to which the STM tip is biased. The electrons injected into this layer by the STM tip have an energy equal to eV_T, where V_T is the STM tip bias. The thickness of the metal layer is comparable or preferably less than the electron mean free path, so that the electrons can traverse the layer without scattering (ballistically). If their energy is larger than the potential step at the interface—in the present case the difference eV_0 between the Fermi level in the metal and the bottom of the SiO_2 conduction band—the electrons have a finite probability of getting injected into the conduction band, traverse the oxide and of being collected as a collector current I_c in the Si substrate. This probability is dependent on the overlap of conduction band density of states at the interfaces, the transmission probability T_{ox} across the oxide, as well as transmission probabilities of a quantum mechanical origin. These issues have been detailed elsewhere.[23,24] Suffices to say here that T_{ox} is strongly dependent on the scattering rates of electrons with longitudinal optical phonons (energy dispersive) and acoustical phonons (momentum dispersive), rates which are strongly dependent on the kinetic energy of the electrons.[21,25] The latter is determined by a combination of the energy at the point of injection ($eV_T - eV_0$) and any gain in energy from field acceleration in the oxide due to the presence of an oxide potential V_{ox} applied by means of a bias potential V_b. An energy diagram for a BEEM experiment is illustrated in Fig. 2 for conditions of positive (accelerating) and negative (retarding) oxide potentials.

In the spectroscopy mode, the raster scan of the STM is stopped and the collector current I_c is measured as V_T is ramped over a range that includes the barrier step potential $V_0 \cdot I_c$ becomes finite once V_T exceeds V_0 for $V_{ox} > 0$ or exceeds $V_0 - V_{ox}$ for $V_{ox} \lesssim 0$, conditions illustrated in Fig. 2. The STM is generally operated at a constant tunneling cur-

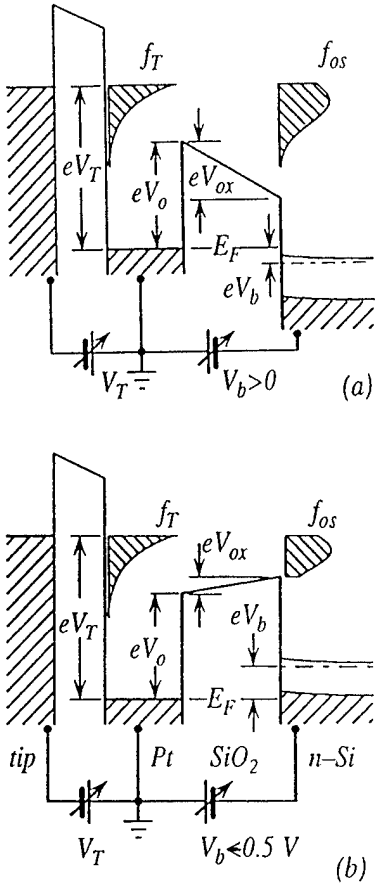


FIG. 2. Schematic energy band diagram for a BEEM experiment on an MOS structure for positive oxide bias V_{ox} (a), and for negative V_{ox} (b). V_{ox} is controlled by means of the gate bias V_b. For Pd/SiO_2/Si, $V_{ox}<0$ when $V_b<-0.5$ V. The energy of the electrons injected by the STM tip is determined by the tip bias V_T. Their energy distribution is represented by the exponential function f_T. The distribution of the electrons reaching the SiO_2–Si interface is shown by the hatched distributions f_{os}. eV_0 is the energy difference between the Fermi level E_F of the Pd layer and the bottom of the SiO_2 conduction band.

rent $I_T \cdot I_c$ can also be used for image contrast in the scanning mode, thereby generating an electron transmission image of the MOS structure, commonly referred to as a BEEM image,[26] that is recorded simultaneously with the STM topographic image.

In stressing studies using BEEM, electrons at a given energy (eV_T) are injected locally (well within a nm) for a given time, with the collector current determining the total fluence. Then an I_c versus V_T spectrum is taken to assess the consequences of the stressing in terms of changes in the local transmission characteristics, including threshold shifts δV_0. This process is then repeated until either the stressed area moves out of range due to instrument drifts, or the oxide breaks down as evidenced by a low value of V_0. Alternatively, spectra over a broad energy range are repeatedly taken at the same spot until either of the two previously mentioned events occur. As a point of reference, an I_c of 50 pA collected over an area of 5 nm^2 (realistic for oxides dominated by ballistic transport, i.e., oxide thickness $t_{ox} \lesssim 5$ nm) corresponds to a charge injection of 10^3 C/cm^2 s! In our experiments, we can achieve injection energies of 9 eV. To achieve this value by conventional FN injection, one would need to apply a field of ~24 MV/cm to a 5 nm oxide. Still larger effective energies can be achieved with BEEM by applying an accelerating oxide bias.

B. Sample preparation

Device-grade amorphous SiO_2 layers were thermally grown in the range 750–800 °C at 1 atm of dry oxygen for 30–50 min. The substrates were 125-mm-diam Si(100) wafers doped to the low 10^{17} range. No additional treatments were performed after the oxidation. Oxide thicknesses for the studies reported here are 3.8 and 7.5 nm. Approximately 8×15 mm^2 samples were cleaved from the wafers and introduced into an ultra high vacuum (UHV) preparation chamber. Arrays of Pd dots, 0.2 mm in diameter, were deposited by thermal evaporation of Pd through a shadow mask, with the substrate held at ~30 K. The finished sample was allowed to warm up to room temperature and was subsequently transferred under UHV into the STM chamber. The grounding contact needed to bias the tip was carefully positioned onto a selected Pd dot by means of three orthogonally mounted inchworms™. Deposition at low temperatures improves the surface morphology of the thin (~3.5 nm) Pd films, resulting in a morphology that consisted of a nodular structure typically 8 nm in diameter that protruded $\lesssim 1$ nm above the valleys. An STM image is shown in Fig. 3(a). A smooth surface morphology is needed to minimize BEEM image contrast arising from the surface topography of the metal, which has previously been observed to be strongly reflected in the BEEM image.[27,28] The corresponding BEEM image of the 3.8 nm Pd/3.8 nm SiO_2/Si(100) structure is shown in Fig. 3(b). The surface topography is clearly replicated in the BEEM image. However, the nodular structure of the topographic image appears rather flat in the BEEM image, that is, there is little evidence that thickness differences in the metal film affect the net transmission. The reason for this is that the mean free paths in the metal film are larger than the thickness of the metal film. However, near the grain boundaries the scattering drastically increases, resulting in the electrons being scattered outside the acceptance cone for transmission across the metal-SiO_2 interface.[29]

The change from the previously used Pt gates[23] to the presently used Pd metallurgy revealed a somewhat surprising enhancement in the net transmission characteristics across the MOS device. This enhancement is shown in Fig. 4, where we compare the BEEM spectrum of a 3.3 nm Pt film with that of a Pd film of comparable thickness, with both films deposited on 3.8 nm oxides. It should be noted that a tunneling current of 10 nA was used with the Pt gates,[23] whereas for Pd, this current was typically in the 1–5 nA range. The enhanced transmission for Pd not only provides enhanced statistics, but makes stressing studies much easier by allowing larger currents to be injected into the oxide. An added advantage of lower tunneling currents lies in the improved collimated nature of the injected current. This is a direct result from the increased tunnel gap during injection at

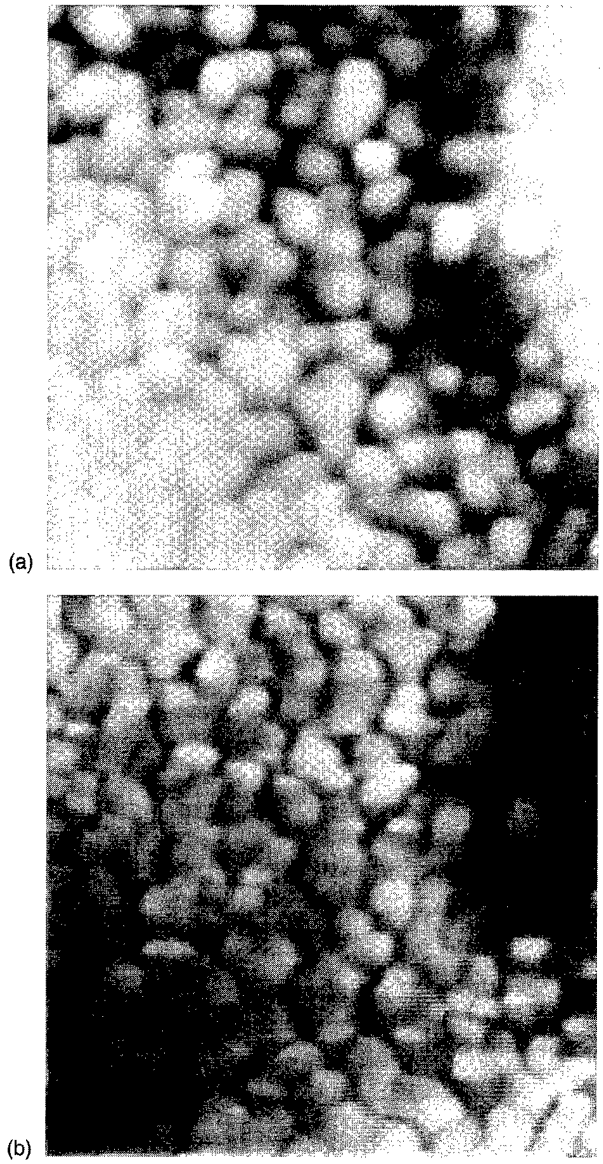


FIG. 3. 100×100 nm^2 STM topograph (a) and BEEM image (b) of a 3 nm Pd/3.8 nm SiO_2Si(100) structure. Black-white range over the image is 3 nm in topography, and 100 pA in BEEM. $V_T=-6$ V, $I_T=2$ nA, $V_b=0$ V. The average peak to valley in the topography image is less than 1 nm.

lower currents.[30] A more collimated electron beam is subject to fewer scattering events, which enhances the transmission across the MOS structure. This in turn results in crisper spectral features, as shown in Fig. 5. As the data indicates, under optimal conditions useful spectra can be readily obtained with tunnel currents of $\leq$0.5 nA.

C. Monte Carlo (MC) simulations

Details of the (MC) simulations, which are based on the code developed by Fischetti *et al.*,[25] have been given elsewhere.[23,24] In the present application, MC simulations were used to estimate the spreading of the electron beam as it traverses the SiO_2 layer. Knowledge of the spreading is needed in order to estimate current and charge densities during electrical stressing of the SiO_2. To calculate the spreading in the oxide, normal injection across the Pd–SiO_2 interface at one single point was assumed (delta function). The potential in the oxide included image force lowering and the applied oxide bias. Electrons returning to the metal were discarded. The position of the transmitted electrons was recorded and the normalized local current density $j(r)/I_{in}$, was calculated after sufficient statistics had been acquired (typically 10^5 electrons). The distance r is measured away from the injection point, and I_{in} is the injection current. In Fig. 6, typical results of the simulations are summarized for different combinations of the oxide thickness, the oxide bias, and the tip bias. The two distributions for $V_T=5.5$ (open sym-

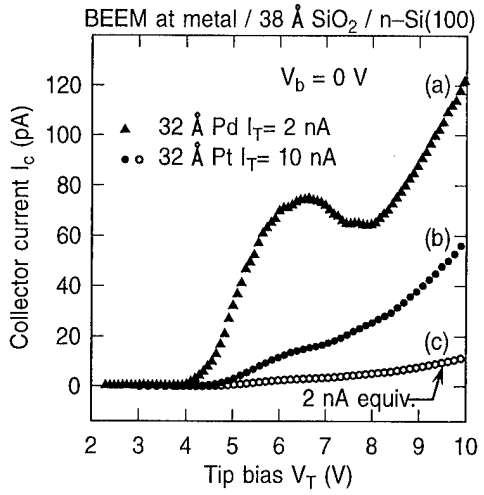


FIG. 4. Comparison of BEEM spectra for similar MOS structures except for the gate metal. Pd gates allow substantially larger transmission than the previously used Pt gates. Spectrum (a) is for a Pd gate, (b) for a Pt gate and (c) is curve (b) normalized to a 2 nA tunnel current.

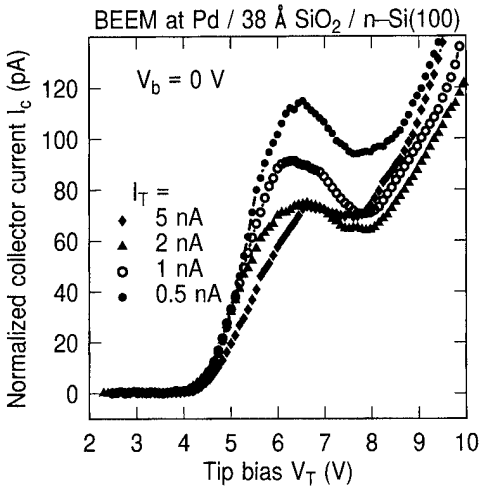


FIG. 5. Normalized collector currents for a Pd/SiO_2/Si(100) structure for a range of indicated tunnel currents. The curves have been normalized to the $I_T=2$ nA spectrum, to which the ordinate scale refers. The relative intensities scale inversely with I_T, which is attributed to the higher collimation (and transmission) for the smaller I_T's, for which the STM tip-surface distance is greater.

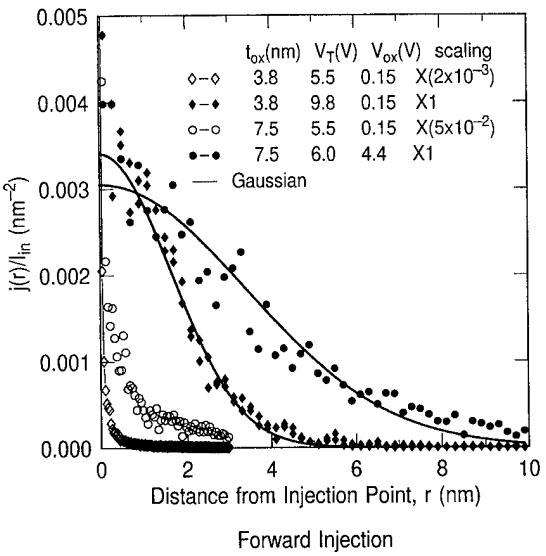


FIG. 6. Spatial distribution of the normalized current densities at the Si/SiO_2 interface after electron transport through 3.8 and 7.5-nm-thick SiO_2 films as calculated by MC simulations for various bias conditions. The figure serves to illustrate the strong dependence of the beam spreading on the electron energy in the oxide. At low electron energies (open symbols, $V_T=5.5$ V) little broadening is predicted. Strong broadening will occur if electrons are either injected at high energy (3.8 nm, $V_T=9.8$ V full diamonds), or if the electrons are heated inside the oxide by a strong electric field (7.5 nm, $V_{ox}=4.4$ V full dots).

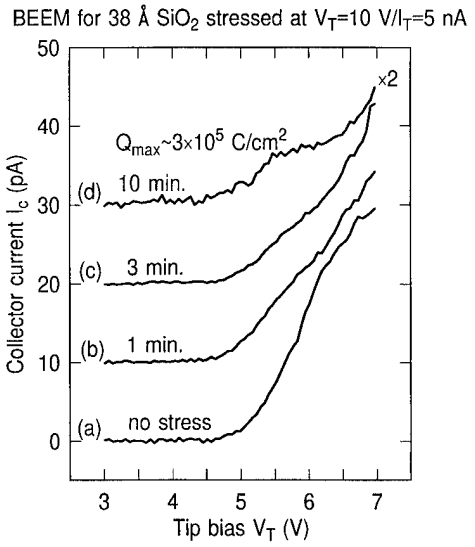


FIG. 7. Spectra of threshold region for BEEM-stressed Pd/3.8 nm SiO_2/Si structure. The kinetic energy of the injected electrons is ~6 eV. The total injected charge was $\sim 3\times10^5$ C/cm^2. This charge is plotted in Fig. 1. No oxide breakdown occurred. The stressing location is indicated in Fig. 8.

bols) were done at a small accelerating bias. At this tip bias, the electrons travel at an energy of 1.5 eV through the oxide, where phonon scattering is small and mostly via longitudinal optical phonons (forward scattering). This results in a streaming type motion in the presence of a small accelerating field and leads to very little beam broadening by the transport through the oxide. In these cases, the beam size on the substrate side is expected to be comparable to the size of the injection beam. The two distributions were scaled to the same heights in such a way that the additional broadening in the thicker oxide can be clearly seen. The peak at $r=0$ overshoots the y axis of the plot by one order of magnitude, such that actual broadening is considerably smaller than suggested at a first glance.

Most interestingly, considerably broader distributions are obtained if the electrons are either injected at high energies (for 3.8 nm film, solid diamonds) or accelerated by a strong electric field (7.5 nm film, solid circles) in the oxide. The broadening is caused by a rapid increase of the acoustic phonon scattering (Umklapp dominated, isotropic, quasi-elastic scattering) with increasing kinetic energy of the electrons. This scattering dominates at energies above 3 eV and leads to strong backscattering that both reduces the overall transmission efficiency and quite significantly spreads the beam. In the 3.8 nm film, the electrons were injected at a tip bias of 9.8 V, which corresponds to an initial electron energy in the oxide of about 5.8 eV. Alternatively, electrons were injected at an energy of 1.5 eV into a 7.5 nm oxide and subsequently heated by a strong electric field, such that their average energy reaches values of 4–5 eV. Again, significant beam spreading is predicted and the increased oxide thickness leads to additional broadening. Under conditions where the electrons reach energies in excess of 3–4 eV, the beam size on the substrate side is predicted to be largely determined by the spreading in the oxide. This has to be taken into account when calculating local current densities and charge to breakdown.

In the present determination of current and charge densities, it was assumed that $r=2.3$ nm. For this value 74% of the current injected into the oxide is included in the density determinations. The remaining current in the tail of the distribution contributes little to the density. This leads to an effective area of 17 nm^2 for the beam spread at the SiO_2–Si interface for the 3.8 nm oxide at a kinetic energy of 5.8 eV.

III. STRESS MEASUREMENTS

As pointed out before, the mere fact of taking a BEEM spectrum, such as those of Figs. 4 and 5, involves the injection of large charges into the oxide. A conservative, lower estimate of this charge is 300 C/cm^2 over a time span (about 1 s) that corresponds to a voltage change $\Delta V_T \approx 0.5$ V during a typical spectral sweep. For electrons injected at $V_T=10$ V, with $V_b=0$ V, their kinetic energy in the oxide is ≈6 eV, a value achieved with FN injection only in the presence of an oxide field of ~24 MV/cm for a 3.8 nm oxide. At such large energies or equivalent oxide fields, breakdown should be observed frequently if the local oxide characteristics were similar to those of the quality oxides represented in Fig. 1. However, breakdowns in 3.8 nm oxides were seldom observed during routine spectral scans with $1 \leq V_T \leq 10$ V and $V_b=0$ V, and never observed on oxides ≳7 nm thick unless a positive bias was applied. An extreme case of stressing without breakdown is illustrated in Fig. 7, which depicts a series of spectral scans taken at the same location of an MOS structure

with t_{ox}=3.8 nm. A spectrum was taken after each dosage of injected charge. Except for an almost imperceptible increase in the threshold position, indicative of some charging at the SiO_2–Si interface, no spectral changes are observed after 3 min of stressing ($Q_{in}\sim10^5$ C/cm^2). After 10 min of stressing the spectrum weakened and shifted to slightly higher energies. The weakening of the spectrum is indicative of some additional scattering in the oxide, possibly due to the generation of interface defects. These may be precursors to breakdown (see following section), but the evidence of a conventional BEEM spectrum and threshold indicates that the oxide is still largely intact. Additional stressing was not possible because of instrument drift, which, although extremely small at this particular time (~2.5 nm/h), caused the STM tip to drift out of the stressed area (~1 nm^2) over the duration of the stressing time. $Q_{inj}\sim3\times10^5$ C/cm^2 after completion of the stressing. This value, marked on Fig. 1 at an equivalent field of 24 MV/cm, is totally incompatible with the rest of the Q_{bd}'s for conventionally stressed oxides, strongly suggesting again that the Q_{bd}'s of Fig. 1 are not intrinsic but defect controlled. The stressing did not cause any observable damage in the Pd film. Topography and BEEM images of the stressed area are shown in Fig. 8(a) and 8(b), respectively. The stressed area has been marked with a cross. The BEEM image barely shows an area of slightly weakened transmission at the stress point.

Stress experiments that resulted in breakdown were successful rather infrequently (of the order of 10%), which suggest that the sites where breakdowns occur are quite dispersed, of rather low density and most likely controlled by impurities and/or defects. A breakdown sequence is illustrated in Fig. 9, which consisted of applying a V_T=10 V (kinetic energy ~6 in the oxide) for predetermined time intervals, each of which was followed by a spectral scan. After stressing for 10 s (curve *b*), a threshold shift to higher energies is clearly observed. A partial breakdown is observed after stressing for 20 s, with the threshold moving down to about 1 V and the collector current increasing by orders of magnitude (the dashed curves represent a 100× demagnification). The low threshold is a clear manifestation that the oxide barrier has disappeared and that a direct contact exists to the Si substrate. The total injected charge to breakdown corresponds to $\sim6\times10^3$ C/cm^2. Another breakdown sequence is shown in Fig. 10. The oxide was repeatedly stressed at the same point of a 3.8 nm oxide with V_T=10 V and V_b=0 V. Spectral scans were taken after each stress application. Only a fraction of the scans are shown in the figure, starting with the first scan (a) on a virgin section of the sample and ending with spectrum (g), which represents the locally broken down oxide. The sequence shows a weakening of the transmitted collector current that is accompanied by threshold shifts to higher energies. After injecting $\sim7\times10^3$ C/cm^2 the intensities again increase, as shown in spectrum (d), and the thresholds move progressively to lower energies ultimately reaching the breakdown stage (curve *g*) after $\sim3\times10^4$ C/cm^2 have been injected. Following such breakdown events, the surface of the MOS structure is heavily damaged. The damaged area is of the order of 10 nm across, and generally consists of a crater with surrounding rough hillocks of dimensions of 5–10 nm. It should be emphasized that the damage resulting from the breakdown is not due to a tip crash, but appears to be the result of the large currents injected locally and/or discharge of injected trapped charge in the stressed area. The morphology in the prebreakdown phase was not investigated.

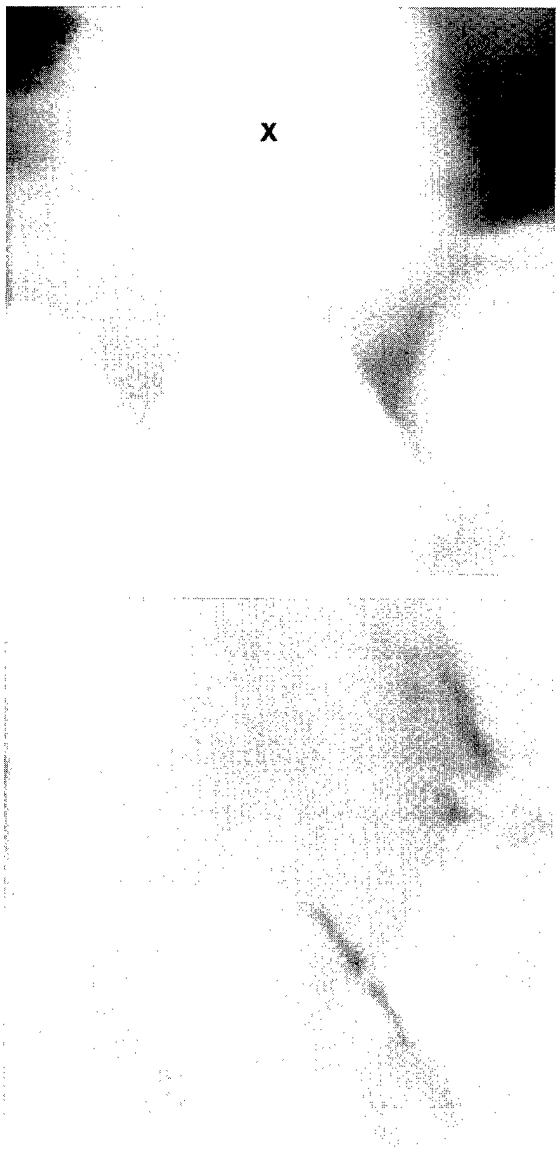


FIG. 8. 12.5×12.5 nm^2 STM topograph (a) and BEEM image (b) of a 3 nm Pd/3.8 nm SiO_2/Si(100) structure. Black-white range over the image is 1.9 nm in topography, and 30 pA in BEEM. V_T=−5 V, I_T=5 nA, V_b=0 V. The location of the BEEM stressing area is marked by an x.

IV. DISCUSSION AND CONCLUSIONS

A key feature of the stressing and breakdown studies with BEEM is the initial shift of the threshold V_0 to higher values. This shift is plotted as a function of injected charge in Fig. 11 for the data of Fig. 10. The increase δV_0, together with the attenuation of the intensity of I_0, depicted in Fig. 10, is

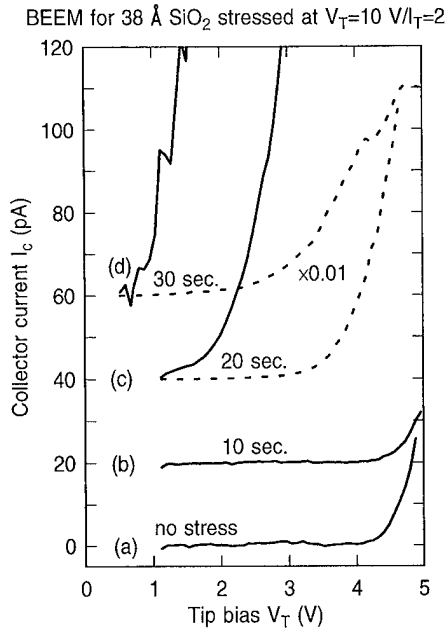


FIG. 9. A successful breakdown sequence for a 3.8 nm oxide stressed with 6 eV (kinetic energy) electrons. The shift to higher V_0 between curve (a) and (b) is due to negative charge at the SiO_2-Si interface. $V_b=0$ V. Total injected charge to breakdown (defined by reaching $V_0 \sim 1$ V, as in curve c) is $\sim 6\times 10^3$ C/cm^2.

equivalent to applying a negative bias to the oxide, as demonstrated in a previous publication.[23] The responsible field for this shift is due to electrons trapped in the vicinity of the SiO_2–Si interface. This view is consistent with previous observations based on FN injection.[10,12,13,15] However, one could argue that in the present case the injected electrons are

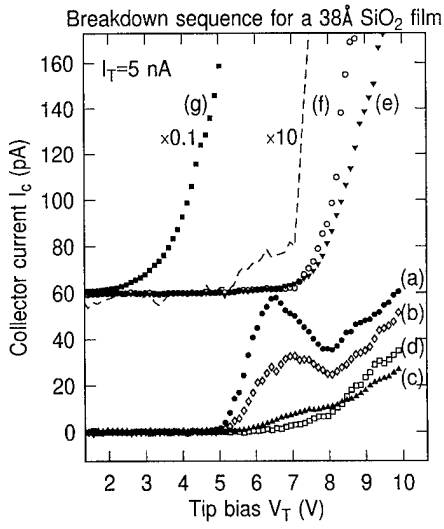


FIG. 10. Oxide breakdown sequence from no stress (curve a) to fully broken down oxide region or punch through (curve g) for a 3.8 nm oxide. The oxide was stressed between spectra at $V_T=10$ V for 10–30 s periods. Not every spectrum and stress sequence is shown. $V_b=0$ V. Of interest is a second threshold in the stressed area beginning with spectrum (d) and clearly seen in the magnified spectrum (f). The threshold is ~2 V above V_0 and appears to be related to trap generation reported in Ref. 9.

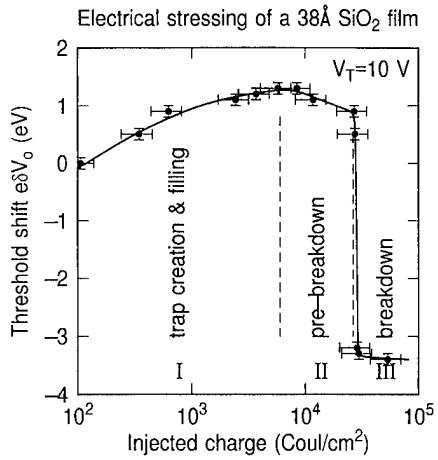


FIG. 11. Shifts in the fundamental threshold V_0 as a function of injected charge taken from the data of Fig. 10. Three phases of stressing are depicted. Prebreakdown commences after a maximum has been reached and coincides with the appearance of a second threshold, as discussed in Fig. 10.

already sufficiently energetic to cause damage at the injecting Pd electrode, a scenario that cannot entirely be excluded. However, any trapped charge in such hypothetical states would be screened by the metal and contribute a negligible amount to the threshold shift. A simple calculation, with the metallic screening treated by the classical image force,[31] gives the following relationship between $e\,\delta V_0$ and the distance d of the charge from the metal–SiO_2 interface:

$$e\,\delta V_0=\frac{e^2}{4\pi\epsilon_0\epsilon_r}\left[\frac{1}{r}-\left(\frac{1}{r^2+4d^2}\right)^{1/2}\right], \tag{1}$$

where r is the distance between the trapped charge and the injected electron as it crosses the plane parallel to the interface that includes the trap. ϵ_0 and ϵ_r are, respectively, the permittivity of vacuum and the dielectric constant. Clearly, $e\,\delta V_0$ approaches zero as d approaches the metal–SiO_2 interface, so that the trapped charge must reside in the vicinity of the SiO_2–Si interface to produce a sizeable δV_0. For a shift of 1 eV or more, with the negatively charged trap 3.5 nm from the Pd–SiO_2 interface, the electron must pass within 0.35 nm of the trap. Alternatively, if the charge were uniformly distributed, a 1 eV shift corresponds to an interface charge of 5.7×10^{12} electrons/cm^2. In a realistic situation there will be a number of filled traps, as evidenced by the increasing shift with dosage, and the net shift will arise from the sum of contributions from each individual charge. During stressing, Fig. 6 indicates that the electron beam has spread over an area of radius ~2 nm with tails of the distribution extending somewhat farther. At the threshold, however, the beam has spread very little and is thus more sensitive to local distributions. Fluctuations in thresholds are indeed observed even for unstressed areas a few nanometers apart, corroborating the highly localized resolution of BEEM, particularly for thin oxides. Consequently, the shifts in V_0 represent interface trap charging on a truly local (<1 nm) scale, but the trap creation from the energetic electrons occurs over an area of ~20 nm^2 for a 3.8 nm thick oxide.

Trap creation and charging also occur for $V_T < 10$ V, as can be ascertained form curve (a) of Fig. 10 which represents a spectrum for a previously unstressed area. Yet its intensity for $V_T \gtrsim 8$ V is uncharacteristically low compared to those of Figs. 4 and 5. The lower intensity implies an ongoing charging as the data is taken. This phenomenon is not observed in spectra taken randomly over the sample and suggests the presence of a weak spot at the point of injection that ultimately results in breakdown. Upon further stressing a maximum in δV_0 is reached, as shown in Fig. 11, which indicates that a maximum was reached in the density of negative charge trapped near the SiO_2–Si interface. A negative charge in the absence of an applied bias leads to a local energy profile similar to that depicted in Fig. 2(b). Further injection leads to a gradual decrease of V_0, which may result from an increased charge leakage due to the deterioration of the oxide or from either positive charges generated at the SiO_2–Si interface (anode), or electron trapping at the cathode.[9] The maximum in V_0 coincides with the onset of larger currents for $V_T \gtrsim 7$ eV, as seen in spectra (d)–(f). The enhanced current is due to an ongoing modification of the oxide, the simplest explanation being the destruction of the oxide that gradually progresses into the oxide starting from the SiO_2–Si interface. The local thinning explains the dramatic increases in I_c as the stressing progresses, curves (e) and (f). The $\times 10$ enhanced curve (f), shown by the dashed line, clearly reveals a remnant of a conventional BEEM spectrum with a threshold at 4.8 eV. The reason for the weakened threshold region in this spectrum is attributed to added scattering at defect sites in the progressively deteriorating oxide. However, once V_T exceeds 7 V, the oxide breakdown proceeds rapidly as suggested by the rapid rise of I_c. Upon completion of spectrum (f) punch-through of the oxide resulted, as indicated by the low $V_0 \approx 1$ V, and the large currents.

The onset of the current enhancement in what we refer to as the prebreakdown region in Fig. 11, and represented in Fig. 10 by curves (c)–(f), is approximately 2 eV above V_0. Perhaps coincidentally, this energy also corresponds to the threshold for hydrogen release and trap formation at the anode of MOS capacitor structures.[7] The hydrogen then diffuses to the cathode (injecting electrode) to produce electron traps. As discussed above, our measurements are not sensitive to such traps at the injecting interface, and it is not clear at present how the hydrogen release would affect the oxide breakdown scenario depicted here. Nevertheless, if present, electron trapping at the cathode would increase the accelerating field in the oxide, thereby accounting for both a decrease in V_0 in the prebreakdown regime, and an enhancement of oxide deterioration by increasing the electron energy as they reach the anode. The implication that most of the events that lead to breakdown are initiated at the SiO_2–Si interface (anode) is also consistent with the MC simulations. These indicate that beam spreading due to acoustic phonon scattering is accentuated in the oxide away from the injecting interface. Since this scattering occurs with little loss/gain in energy, the effectiveness in damaging the oxide is enhanced as the electrons sample a greater volume, which increases the probability that an electron encounters a weak spots where the breakdown may be initiated. This breakdown scenario starting at the oxide–Si interface is consistent with those proposed by several other workers,[12,13,15,32] but inconsistent with oxide degradation attributed to hot hole injection into the oxide at the SiO_2–Si interface.[14] The hot holes are created by impact ionization of the hot electrons as they enter the Si. The hole injection model requires the high fields present in conventional FN injection to accelerate the holes toward the cathode. This situation does not apply in the present case for a 3.8 nm oxide with no applied bias. In fact, the electron traps initially created by the stressing produce a retarding field for the holes, keeping them from being injected into the valence band of the SiO_2.

The Q_{bd} values obtained from various breakdown events have been plotted in Fig. 1, as well as the total charge injected in one experiment that did not reach breakdown. The injection energies of the BEEM experiments were converted to an equivalent electric field that would achieve the same energies under FN injection.[33] The stressing current at maximum tip bias was converted to a charge density by multiplying I_c by the time and dividing the product by the area of the electron beam reaching the SiO_2–Si interface (~ 17 nm^2), as determined from the MC simulations depicted in Fig. 6. The Q_{bd}'s thus obtained fall consistently above the curves obtained from conventional stress experiments.[16] These were interpreted in terms of an intrinsic breakdown limit, a conclusion that is challenged by the present results. Even the new points obtained here are not intrinsic, as much higher charge dosages did not routinely cause breakdowns in the 3.8 nm oxides. Consequently we conclude that even the higher observed breakdown dosages are still limited by extrinsic effects, either defects or impurities at the interface. This conclusion offers hope that in the future, device lifetimes and device reliability can be considerably improved. Our observations are consistent with an increase in oxide robustness observed over the years as capacitor areas and device dimensions shrink.[34–36] It is generally acknowledged that these trends are directly attributable to a decreasing probability of finding defects/impurities in the oxide as the device areas decrease.

It is a reasonable question to ask if Q_{bd}'s obtained with BEEM can be compared to those obtained under FN injection. The conversion of the former from an energy scale to an effective electric field scale assumes that only the electron energy is important and that field effects in the FN stress mode contribute negligibly to the breakdown. Previous studies have shown that the principal trap, defect and interface state generation mechanisms are dependent solely on the kinetic energy of the hot electrons injected into the oxide, with no evidence of a field dependence other than to accelerated the electrons to the required energies for defect generation.[7,9,37] This is not to say that the field is unimportant but rather that thus far there is little evidence of the field playing a primary role in determining the breakdown mechanism. For the electric field stressed modes, such as FN, defect formation can also occur from holes generated by impact

ionization in both the SiO_2 and the Si (anode) with subsequent injection into the SiO_2 valence band. The field is thus important to accelerate the holes towards the cathode where they can create traps and interface states. These processes also depend on the energy of the holes.[38] However, hole generation depends on the energy gained by the electrons, with a minimum oxide potential of 8 V being required for hole injection to occur from the Si (anode).[38] This potential threshold is higher than the gate voltages used by Schuegraf and Hu[16] to obtain the data of Fig. 1 with the exception that the data for the 7 nm oxide beyond 11 MV/cm correspond to a gate voltage exceeding 8 V. However, no change in slope is apparent in the data of the 7 nm oxide, strongly suggesting that hole injection for thin oxides and for the given field ranges is not a principal contributor to breakdown. It is also worth noting that the BEEM-stressed Q_{bd} for the 7.5 nm oxide in Fig. 1 was obtained with an applied field of ~6 MV/cm in addition to a kinetic energy of 4.2 eV at the point of injection. From these discussions, it can be surmised that the oxide field has a yet to be identified role on Q_{bd}'s for thin oxides (aside from accelerating the electrons in FN stressing). Consequently, the comparison of FN-stressed and BEEM-stressed data is a reasonable undertaking with any additional, but still unknown, field enhanced contributions in the FN case being compensated by the higher energies sustained by the BEEM-injected electrons throughout their passage across the oxide.

V. CONCLUSIONS

In conclusion, we have demonstrated that spatially highly resolved stress studies using the STM can give new insights into breakdown phenomena in ultrathin oxides. Unlike conventional stressing by means of FN injection, the present technique allows an arbitrary choice of injected electron energies, thereby greatly simplifying the injection and oxide transport processes. The latter have also been modeled by MC simulations, from which the spreading of the electron beam as it traverses the oxide was calculated. Stress induced breakdowns under conditions that far exceed those for conventionally stressed state-of-the-art oxides imply that breakdowns are still controlled by defects/impurities in the oxide, and that an intrinsic breakdown limit has not yet been reached. When observed, the breakdown scenario suggests a three-fold sequence of initial trap creation and filling, followed by a prebreakdown phase that appeared to be associated with a progressive oxide deterioration starting at the SiO_2–Si interface, and finally a catastrophic punch through, characterized by a very low barrier to electron injection. The data furthermore suggests that hole transport in the oxide is unlikely to occur for the low oxide biases used in the present experiments.

ACKNOWLEDGMENTS

Fruitful discussions with D. J. DiMaria and M. V. Fischetti are gratefully acknowledged, as are the growth efforts of the samples by C. D'Emic. The experimental aspects have greatly benefited from the technical assistance by M. Prikas. This research was partially supported by the Office of Naval Research, under Contract No. N00014-95-C-0087.

[1]C. Hu, Proc. IEEE **81**, 682 (1993).
[2]C. A. Mead, Analog Integrated Circuits and Signal Processing **6**, 9 (1994).
[3]Y. Taur, Y.-J. Mii, D. J. Frank, H.-S. Wong, D. A. Buchanan, S. J. Wind, S. A. Rishton, G. A. Sai-Halasz, and E. J. Nowak, IBM J. Res. Develop. **39**, 245 (1995).
[4]Semiconductor Technology Workshop, Working Group Reports, Semiconductor Industry Association, San Jose, 1993, p. 27.
[5]J. C. Lee, I.-C. Chen, and C. Hu, IEEE Trans. Electron Devices **35**, 2268 (1988).
[6]P. Olivo, T. N. Nguyen, and B. Riccó, IEEE Trans. Electron Devices **35**, 2259 (1988).
[7]D. J. DiMaria and J. W. Stasiak, J. Appl. Phys. **65**, 2342 (1989).
[8]P. Olivo, T. N. Nguyen, and B. Riccó, IEEE Trans. Electron Devices **38**, 527 (1991).
[9]D. J. DiMaria, E. Cartier, and D. Arnold, J. Appl. Phys. **73**, 3367 (1993).
[10]P. P. Apte, T. Kubota, and K. C. Saraswat, J. Electrochem. Soc. **140**, 770 (1993).
[11]M. Nafría, J. Suñé, and X. Aymerich, IEEE Trans. Electron Devices **40**, 1662 (1993).
[12]P. P. Apte and K. C. Saraswat, IEEE Electron Device Lett. **14**, 512 (1993).
[13]L. K. Han, M. Bhat, D. Wristers, J. Fulford, and D. L. Kwong, Proceedings of the IEEE International Electron Device Meeting, 1994, p. 617.
[14]K. F. Schuegraf and C. Hu, IEEE Trans. Electron Devices **41**, 761 (1994).
[15]S.-H. Lee, B.-J. Cho, J.-C. Kim, and S.-H. Choi, Proceedings of the IEEE International Electron Device Meeting 1994, p. 605.
[16]K. F. Schuegraf, D. Park, and C. Hu, Proceedings of the IEEE International Electron Device Meeting, 1994, p. 609.
[17]M. Depas, B. Vermeire, P. W. Mertens, M. Meuris, and M. M. Heyns, Semicond. Sci. Technol. **10**, 753 (1995).
[18]C.-H. Lin, J. Cable, and J. C. S. Woo, IEEE Trans. Electron Devices **42**, 1329 (1995).
[19]E. Hasegawa, A. Ishitani, K. Akimoto, M. Tsukiji, and N. Ohta, J. Electrochem. Soc. **142**, 273 (1995).
[20]D. J. DiMaria, D. Arnold, and E. Cartier, Appl. Phys. Lett. **61**, 2329 (1992).
[21]D. Arnold, E. Cartier, and D. J. DiMaria, Phys. Rev. B **49**, 10278 (1994).
[22]I.-C. Chen, S. Holland, and C. Hu, IEEE Trans. Electron Devices **32**, 413 (1985).
[23]R. Ludeke, A. Bauer, and E. Cartier, J. Vac. Sci. Technol. B **13**, 1830 (1995); Appl. Phys. Lett. **66**, 730 (1995).
[24]R. Ludeke, E. Cartier, and H. J. Wen, J. Electrochem. Soc. (to be published).
[25]M. V. Fischetti, D. J. DiMaria, S. D. Brorson, T. N. Theis, and J. R. Kirtley, Phys. Rev. B **31**, 8124 (1985).
[26]W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988).
[27]R. Ludeke and A. Bauer, Phys. Rev. Lett. **71**, 1760 (1993); J. Vac. Sci. Technol. A **12**, 1910 (1994); Phys. Scr. T **55**, 90 (1994).
[28]R. Ludeke, J. Vac. Sci. Technol. A **11**, 786 (1993).
[29]L. D. Bell, W. J. Kaiser, M. H. Hecht, and L. C. Davis, in *Scanning Tunneling Microscopy*, edited by J. E. Stroscio and W. J. Kaiser (Academic, San Diego, CA, 1993), p. 307.
[30]N. D. Lang, A. Yacoby, and Y. Imry, Phys. Rev. Lett. **63**, 1499 (1989).
[31]M. V. Fischetti, S. E. Laux, and E. Crabbé, J. Appl. Phys. **78**, 1058 (1995).
[32]D. R. Wolters and A. T. A. Zegers-Van Duynhoven, in *The Physics and Technology of Amorphous SiO_2*, edited by A. B. Devine (Plenum, New York, 1988), p. 391.
[33]For FN injection and quasi-ballistic transport, the equivalent oxide field F necessary to accelerate an electron to reach an energy E relative to the SiO_2 conduction band at the anode is: $F \approx (qV_0+E)/qw$, where $V_0=3.15$ V was assumed for a polysilicon gate, w is the oxide thickness and q is the electron charge.
[34]J. Suñé, I. Placencia, N. Barniol, E. Farrés, F. Martín, and X. Aymerich, Thin Solid Films **185**, 347 (1990).
[35]R. P. Vollertsen and W. G. Kleppmann, Proc. IEEE **4**, 75 (1991).
[36]D. J. Dumin, J. R. Maddux, R. S. Scott, and R. Subramoniam, IEEE Trans. Electron Devices **41**, 1570 (1994).
[37]E. Cartier and D. J. DiMaria, Microelectron. Eng. **22**, 207 (1993).
[38]D. J. DiMaria, E. Cartier, and D. A. Buchanan, J. Appl. Phys. (in press).

Ballistic-electron emission microscopy studies of charge trapping in SiO_2

B. Kaczer and J. P. Pelz
Department of Physics, The Ohio State University, Columbus, Ohio 43210

(Received 22 January 1996; accepted 13 April 1996)

We have used ballistic-electron emission microscopy (BEEM) to locally inject and microscopically characterize trapped charge in metal-oxide-semiconductor structures made with Pt metal electrodes, moderately thick (~25 nm) SiO_2 films and *p*-type Si substrates. Electrons injected into the oxide conduction band produce a local *suppression* in the transmitted BEEM current across the oxide. This suppression depends on the voltage V_{ox} applied across the oxide during injection, indicating that its source is *within* the oxide film, and does not simply result from hot-electron damage at the metal-oxide interface. This suppression is accompanied by a local increase in the BEEM *I-V* curve threshold voltage, consistent with a build-up of trapped BEEM electrons in the oxide. For a given trapped charge density we find that the threshold varies systematically with V_{ox} applied during *I-V* curve measurement. We use these variations to make estimates of the trapped charge density and approximate depth in the oxide film, and find them consistent with previous macroscopic observations. We further study the lateral extent of the suppression and the lifetime of this effect. We also measure the height of the Pt-SiO_2 barrier and observe its lowering due to the image force effect. © *1996 American Vacuum Society.*

I. INTRODUCTION

SiO_2 is a material of crucial importance for the overwhelming majority of today's active electronic devices. It is chemically stable, has excellent passivating and dielectric properties and its quality can be consistently controlled during production.[1] The properties of SiO_2 are well known from many macroscopic measurements, including *I-V* and *C-V* measurements, internal photoemission, etc.[2,3] In these measurements the microscopic properties are averaged over large areas. With the ongoing miniaturization of electronic devices the atomic-scale features of the materials used may soon be comparable with the size of the device features themselves, and will affect the performance of the devices. It is therefore critical to investigate the properties of these materials at the microscopic level. Cuberes *et al.*[4] first used the technique of ballistic-electron emission microscopy (BEEM) as a new, microscopic method to study electronic properties of very thin CaF_2 *insulating* films. Ludeke *et al.*[5–7] later showed that it was possible to use BEEM to perform nanometer scale investigations of thin (<10 nm) SiO_2 films sandwiched in metal-oxide-semiconductor (MOS) structures. Ludeke *et al.* studied hot electron transport processes in the SiO_2 films and observed microscopic irregularities in the electronic properties.

Here we report BEEM experiments performed on MOS structures made with thicker (~25 nm) SiO_2 films, in which *charge trapping effects* are found to be very significant (initial measurements are described elsewhere[8]). We have observed a local, *time-dependent suppression* in the BEEM current through the SiO_2 film, which is accompanied by a local increase in the BEEM threshold voltage. Our measurements demonstrate that both of these effects are related to a local build-up of charge in the oxide bulk, due to trapping of some of the injected electrons. We have developed a method to use BEEM to measure the magnitude of the trapped charge and its mean depth in the oxide, essentially by measuring shifts in BEEM threshold voltages as a function of the applied electric field in the oxide. Good agreement is found between our nanometer-scale measurements and previous macroscopic measurements for similar injection conditions. We have also investigated the lateral extent of the current suppression caused by injection of charge at one point, and have determined some lower limits on the lifetime of this effect. We also measure the height of the Pt-SiO_2 barrier and observe its lowering due to the image force effect.

II. EXPERIMENT

A. BEEM

The BEEM technique, which is an extension of scanning tunneling microscopy (STM), was first used by Kaiser and Bell[9] to investigate the microscopic properties of metal-semiconductor Schottky barriers. The configuration of BEEM for investigation of MOS structures is shown in Fig. 1(a). When the STM tip is brought close to the metal film with a negative tip bias [corresponding to $V_T>0$ V with polarities defined in Fig. 1(a)], a tunnel current I_T is produced consisting of electrons which tunnel from filled states in the tip into empty states in the metal film. Since the tunnel barrier height is smaller for electrons with higher energies, the flux of hot electrons in the metal film is peaked in energy near the Fermi energy of the tip [as shown in Fig. 1(b)] and is also highly spatially localized and forward-focused.[10] A fraction of these hot electrons is able to traverse the metal layer and reach the metal-oxide interface *ballistically*, i.e., without losing energy.[6,9] These ballistic electrons can enter the oxide provided they have enough energy (controlled by V_T) to surmount or tunnel through the barrier at the metal-oxide interface. This barrier has nominal height V_0 when the voltage drop across the oxide $V_{ox}=0$ V, but the barrier will

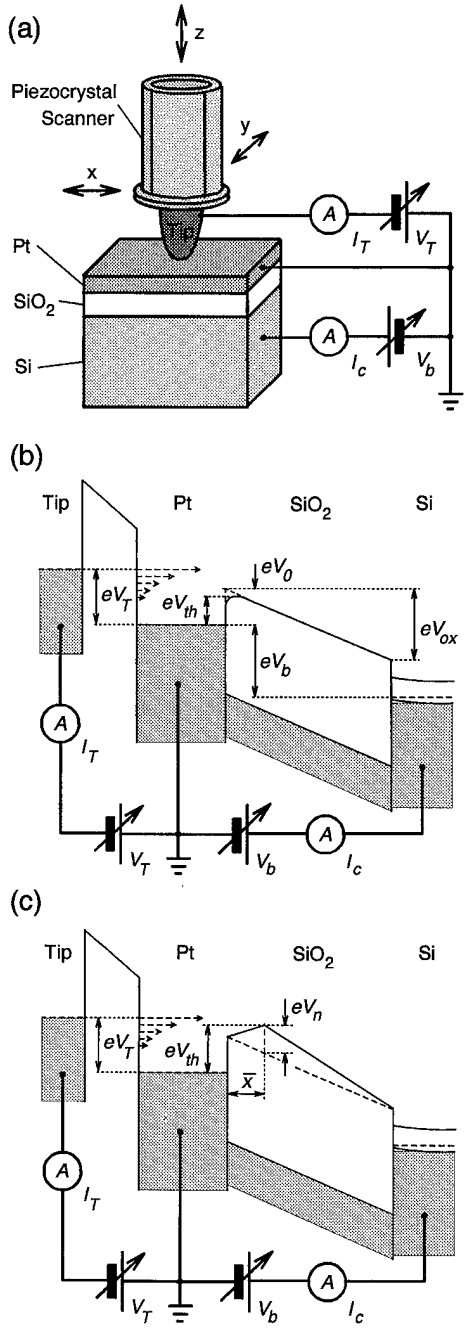


FIG. 1. (a) Schematic configuration of BEEM for investigation of MOS structures. The STM tip with the sample constitute a three terminal device. (b) Energy diagram of the experiment. The horizontal axis corresponds to the z axis in (a). Some of the hot electrons from the tip (current I_T, dashed arrows) can ballistically cross the metal layer and be injected into the oxide, provided they have sufficient energy (controlled by V_T). (c) Charge build-up in the oxide causes increase of the barrier height which the ballistic electrons have to overcome to be collected as current I_c in the substrate. See the text for definition of the remaining symbols.

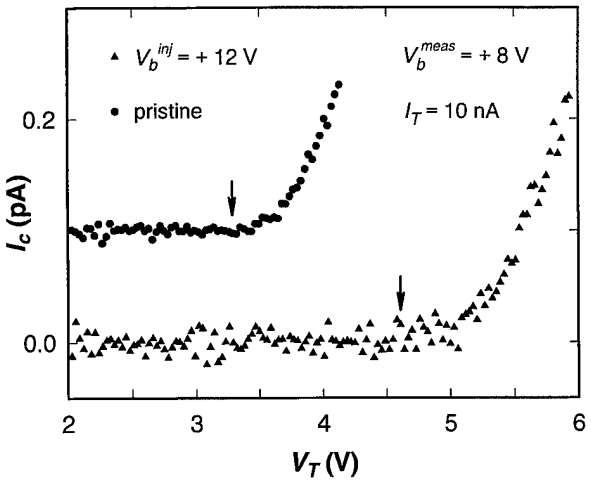


FIG. 2. An average of "short" I_c-V_T curves taken over pristine area (circles, 100 curves averaged) and over the center of injected region (triangles, 10 curves averaged). The method used to create the injected region is described in the text. The threshold voltage V_{th}=3.3 V for the pristine area while $V_{\rm th}$ for the injected region is increased to 4.6 V. Both I_c-V_T curves were measured at $V_b^{\rm meas}$=+8 V, I_T=10 nA.

be lowered below V_0 under a "forward" oxide bias ($V_{\rm ox} > 0$ V)[11] by so-called "image force lowering,"[3] or raised above V_0 under "reverse" oxide bias ($V_{\rm ox}<0$ V). Once in the oxide, injected electrons may be accelerated by an electric field present in the oxide, and/or be scattered by phonons, defects, or trapped charge. Some injected electrons will be scattered back into the metal film,[5] some may become trapped in the oxide,[12,13] and some may conduct across the oxide into the Si substrate and be measured as the external collector current I_c. Electrons which become trapped in the oxide will alter the electric field in the oxide and modify the barrier for subsequent electrons, as shown in Fig. 1(c). The model describing this will be discussed in more detail later in the article.

B. Modes of collecting data

We have used several different modes of collecting data in order to study different aspects of charge injection and charge trapping in oxide films. To monitor the barrier height at the metal-oxide interface, we measure "I_c-V_T curves"[9] as illustrated in Fig. 2. Since injected electrons must have enough energy to surmount this barrier, the BEEM threshold voltage $V_{\rm th}$ (indicated by the arrows in Fig. 2) correspond to the barrier height at a particular tip location.[5,9] We extract $V_{\rm th}$ from measured I_c-V_T curves by fitting the above-threshold region to the functional form $I_c \propto (V_T - V_{\rm th})^{2.5}$, which reasonably approximates the fitting model used in Refs. 5, 6, and 7. Due to noise in I_c ($\sim$25 fA rms in a 20 Hz bandwidth) the typical uncertainty in $V_{\rm th}$ is of the order of $\pm$0.1 V. In many cases we measure "full" I_c-V_T curves, with V_T covering a typical range of 2–10 V.[8] However, if we are only interested in extracting the threshold voltage $V_{\rm th}$, we often measure truncated or "short" I_c-V_T curves (as shown in Fig. 2), in which V_T is only increased until sufficient current I_c is detectable to determine an accurate value of $V_{\rm th}$. We do this to minimize the oxide charging which takes place during measurement of the I_c-V_T curve. This will be discussed in more detail later in the article.

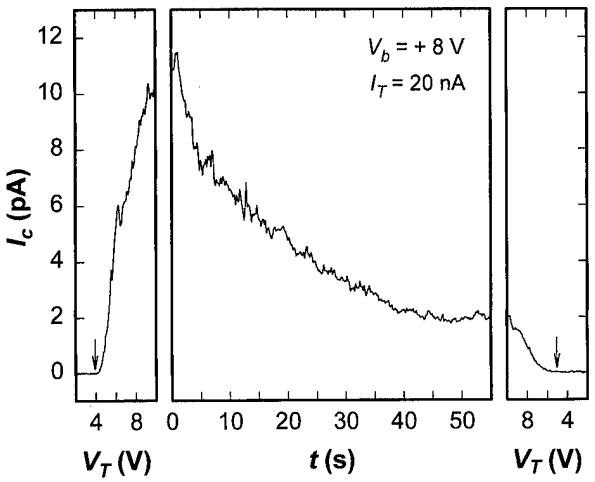


FIG. 3. Combination I_c-V_T/I_c-t curve. Left panel: An initial I_c-V_T curve is taken as V_T is ramped up to V_T^{hold} = 10 V. Central panel: V_T, I_T, and V_b are held constant to measure an I_c-t curve. Gradual suppression of I_c with increasing time is clearly visible. Right panel: A final I_c-V_T curve is measured by ramping V_T back down. Also note the increase in the threshold voltage (indicated by the arrows) from the initial to the final I_c-V_T curves.

In a second mode of collecting data, we directly address the time dependence of charging effects by measuring "combination I_c-V_T/I_c-t curves," as shown in Fig. 3. Each such I_c-V_T/I_c-t curve was measured by recording I_c as V_T is first ramped up to an above threshold value V_T^{hold} (to measure an initial I_c-V_T curve), then held at this voltage for a specified time (typically 10–60 s) to measure an I_c-t curve, then ramped back down to its initial (below-threshold) value to measure a final I_c-V_T curve. During the I_c-t curve, any measured changes in I_c are directly attributable to a build-up of trapped charge and/or possible interface damage in the part of the sample directly under the tip.

In a third data mode, a "BEEM image"[9] is measured to monitor the lateral variations in the BEEM current I_c through the oxide film. In this mode, I_c is measured as a function of tip position as the tip is scanned across the surface, while V_T and I_T are held constant. This yields simultaneously both the top surface topography and an image of the lateral distribution of I_c (provided, of course, that $V_T > V_{\text{th}}$). Such pairs of topographic and BEEM images are depicted in Fig. 4. Measurement of lateral images can be also combined with the measurement of I_c-V_T or combination I_c-V_T/I_c-t curves. The tip is stopped at predetermined locations during a scan and an I_c-V_T or combination I_c-V_T/I_c-t curve is measured. Unfortunately, these interruptions in scanning can cause disruptions in the continuity of the lateral images, due to sample drift and piezocrystal creep. This is particularly severe if the I_c-V_T curves are measured on a rectangular "grid" of sample locations aligned along the normal x- and y-scan directions, since several curves are measured along a single x-scan line. To distribute such disruptions more evenly, we measure the I_c-V_T curves on a rectangular grid which is *rotated* slightly with respect to the normal x- and y-scan directions, in such a way that at most one I_c-V_T curve is measured on a particular x-scan line.

C. Sample preparation and experimental set-up

Samples were prepared from p-Si(100) wafers (resistivity 3–20 Ω cm) with ~25 nm of high quality oxide (provided courtesy of National Semiconductor, Inc.). Samples of size ~5×20 mm^2 were rinsed sequentially in acetone, methanol and deionized water, mounted on sample blocks and introduced into our UHV system, which were then degassed while the system was baked (~180 °C, ~12 h). Then ~5 nm of Pt film was deposited in the shape of several 1 mm diameter dots by means of e-gun evaporation[14] and a shadow mask. The samples were then transferred in UHV into an adjacent STM chamber without any post-metallization anneal. The measured dot was gently contacted with a 0.1 mm diameter Au wire positioned with a mechanical manipulator inside the STM chamber.

As discussed elsewhere,[4–7] large tip voltages ($V_T > 4$ V) are required to surmount the large energy barrier at the metal-oxide interface and therefore UHV is essential for these BEEM measurements. For different materials[15,16] such large voltages tend to seriously alter the tip-surface interface if tunneling is done in air, but cause no observable damage if the tunneling is done in UHV, where V_T is routinely ramped up to 10 V without any apparent damage to the surface.[5–8] The reason for this is thought to be related to the presence of water vapor. From our own experience, tunneling in air on Pt surfaces damages or alters the Pt interface and/or the tip if V_T is increased above ~3 V (holes "drilled" into the metal layer were sometimes observed), even though stable tunneling is possible at lower voltages.

III. RESULTS

A. Dependence of suppression on oxide bias

Much of this work was motivated by the observation of current I_c through the oxide decaying with time (as illustrated in Fig. 3), even when V_T is held constant. This effect is important for two reasons. From a fundamental viewpoint, we would like to understand the origin of the suppression—whether it is due to charge build-up in the oxide or simply due to damage to the metal-oxide interface. However, there is a more practical reason for investigating the suppression effect. As it was shown by Ludeke *et al.*[5–7] a full I_c-V_T curve can carry important information about the dynamics of hot electrons in the conduction band of the oxide. However, since time is always an implicit variable during the measurement of an I_c-V_T curve, any explicit time-dependence of I_c must be understood and taken into account to properly interpret the I_c-V_T curve.

One possible cause for suppression of I_c is damage of the metal-oxide interface by energetic BEEM electrons without any build up of trapped charge, as was observed by Fernandez and Hallen on certain Au/Si interfaces.[17] To show that the suppression of I_c has its origin in the bulk of the oxide (and therefore not at the interface) we demonstrate that the

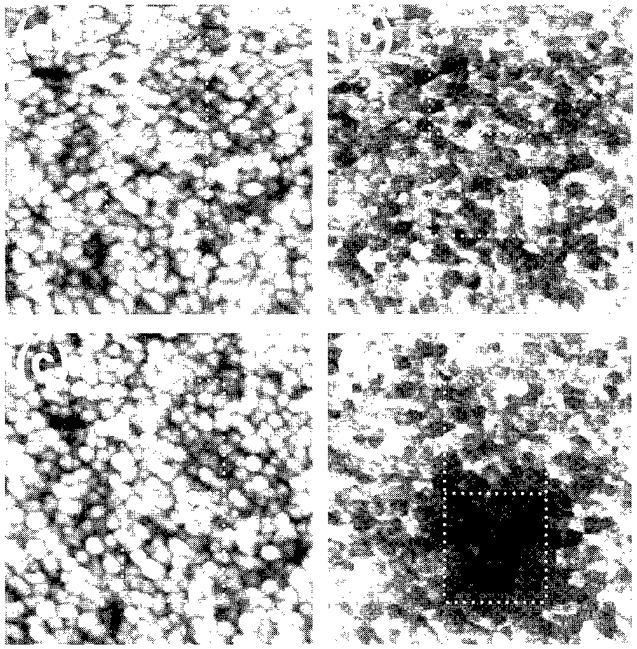


FIG. 4. Dependence of suppression of I_c on oxide bias V_b. (a) ~150×150 nm^2 topographic scan before injection, measured with V_T=5.4 V, $I_T \cong 10$ nA and V_b^{meas}=+8 V. The Pt topography (full gray-scale range 2.5 nm) shows a nodular character with average nodule size 6.5±2.0 nm. (b) Corresponding BEEM image (gray-scale range 0–5 pA). (c) Topography and (d) BEEM image after injection (see the text), over the two regions outlined by dashed lines. Injection in the upper (lower) region done with reverse (forward) oxide bias of V_b^{inj}=−4 V (V_b^{inj}=+12 V). Significant suppression is observed only for the forward-bias injected region.

suppression of I_c depends strongly on the oxide bias voltage V_b. Since the electric field created by V_b does not penetrate significantly into the metal film, the average kinetic energy and flux of the hot electrons reaching the metal-oxide interface should not be significantly affected by V_b.[18] If hot-electron damage to this interface were a major cause of the suppression, then it should *not* be possible to "turn off" the suppression simply by changing V_b from forward to reverse bias.

To see this dependence of suppression on V_b, a pristine 150×150 nm^2 area was first "pre-scanned" to image the lateral distribution of I_c [Figs. 4(a) and 4(b)]. Then, two nominally identical "injection" scans were performed in two adjacent 50×50 nm^2 regions using a 5×5 grid of I_c-V_T/I_c-t curves (V_T^{hold} = 5.4 V for ~60 s, I_T=20 nA) for each area. The two regions are indicated by dashed lines in Fig. 4. The only difference between the two areas was that a reverse bias V_b^{inj}=−4 V was applied for the upper region, while a forward bias V_b^{inj}=+12 V was applied for the lower region. For forward oxide bias, the average integrated collector charge density was $Q_c \cong 125$ C cm^{-2}, while for reverse oxide bias it was negligible, as the BEEM electrons do not have sufficient energy to surmount both the ~4 eV barrier at the metal oxide interface and the ~4 V reverse bias across the oxide film. The whole area was then imaged again. This is shown in Figs. 4(c) and 4(d). The strong suppression of I_c for V_b^{inj} = +12 V is clearly visible while the suppression for the V_b^{inj}=−4 V area is negligible. There is also no apparent damage to the metal surface. This shows that source of the suppression has to be within the oxide bulk as the only varied parameter—the electric field—is effectively screened at the metal-oxide interface.

B. Threshold shift observed for suppressed region

If the suppression is caused by electrons trapped in the bulk of the oxide near the suppressed region, then the electric field from this trapped charge should also *increase* the energy barrier at the metal-oxide interface, which in turn should produce an increase in the BEEM threshold voltage V_{th}. The two I_c-V_T curves in Fig. 3 exhibit this effect. The threshold voltage V_{th}=3.3 V for the pristine region while the threshold voltage V_{th} is shifted to 4.6 V for the I_c-V_T taken in the center of the V_b^{inj}=+12 V injected region. Both I_c-V_T curves were measured at V_b^{meas}=+8 V. These threshold shifts and their dependence on V_b can be used to determine the approximate magnitude of the trapped charge and its depth in the oxide. This will be addressed later in the article.

C. Suppression possible even for injection under reverse oxide bias

In Fig. 4(d) no substantial suppression is visible for the area injected under a V_b^{inj}=−4 V reverse oxide bias when injection was done with a tip bias V_T=5.4 V. This value of V_T is about 1.5 V higher then the expected barrier height V_0 and therefore we expect some electrons should still enter the oxide even under these reverse bias conditions. Furthermore, if the electrons have sufficient energy, they can completely traverse the oxide and be collected in the Si substrate even under reverse oxide bias, as was observed in Ref. 5. Hence we expect that oxide charging should still take place under reverse bias conditions, provided the energy eV_T of electrons during injection is sufficiently high. To test this, we have measured the suppression resulting from injection at V_b^{inj}=−4 V as a function of tip voltage V_T^{hold} during injection. A pristine area was again pre-imaged, similarly to Sec. III A. Then an injection was performed over a 50×50 nm^2 area at V_b^{inj}=−4 V in a similar manner to that described in Sec. III A, except that the maximum tip voltage V_T^{hold} was adjusted to different values. The whole area was then imaged again and the ratio of the average I_c over the injected region after ($\langle I_c^f \rangle$) and before ($\langle I_c^i \rangle$) the injection was determined. Figure 5 shows that the suppression of I_c is almost negligible at V_T^{hold}=5.4 V (about 1.5 V above V_0), as discussed in Sec. III A. Interestingly, the suppression increases substantially for a rather small increase in V_T^{hold} of about 0.5 eV. At this point, it is unclear whether this sharp turn-on is related to a particular "trap creation activation energy"[12] or simply represents the fact the total charge injected into oxide increases with V_T^{hold}. Also note that under these conditions the injected electrons do not have sufficient energy to traverse the whole length of the oxide, so the transmitted BEEM current I_c during injection remains negligible.

D. Imaging suppression after injection at one point

We have also investigated the lateral extent of the suppression caused by injection of charge at a single point. To estimate this the following experiment was performed. A prescan was first taken to get an image of I_c. This is shown in Fig. 6(b) together with the corresponding topography [Fig. 6(a)]. Then charge was injected in the center of that region by taking the a combination I_c-V_T/I_c-t curve shown in Fig. 3 ($V_t^{\text{hold}} = -10$ V). The area was than imaged again as is shown in Fig. 6(c). Again, the region of suppressed I_c in the center resulting from the injection is clearly visible and has approximate diameter of 20–30 nm. In order to obtain an average profile of the suppression this was repeated several times on different sample locations. The average profile is depicted in Fig. 7. The FWHM of this suppression profile is approximately 25 nm. The knowledge of this also turns out to be important for determining V_{th} on pristine areas of the sample, since two I_c-V_T curves must not be taken so close together that the later one would be affected by the trapped charge produced by the first I_c-V_T curve. We also point out that we do not know the *direct* relationship between the magnitude of suppression and the local amount of trapped charge. Therefore the size of the suppression profile gives us only an *estimate* of the extent of the trapped charge distribution after injection at one point.

We also addressed the issue of the stability of the effect of suppression. If the lifetime of the effect were comparable with the randomly varying delay between the injection and the measuring of V_{th}, then this could influence the measurements. To investigate this, the injected area [which was imaged immediately after injection in Fig. 6(c)] was imaged one more time ~14 hours after injection. The result is shown in Fig. 6(d). We see that the suppressed region appears roughly the same as in Fig. 3(c). In order not to loose the injected region due the sample drift and piezocrystal creep we took advantage of the nodular character of the top metal surface to actively lock the tip onto a local extreme and track this extreme until the later image was to be taken. During this tracking, V_T was held well below V_0 (V_T=1 V) to avoid additional charging of the oxide. While there was no precise quantitative answer, we concluded that the lifetime of the suppression is at least several hours.

E. Influence of I_c-V_T spacing

In order to lower the noise in the I_c-V_T curves taken to obtain the threshold voltage V_{th} it is common to average together several I_c-V_T curves. However, as shown above, injection of charge into the conduction band of the oxide (which is a necessary consequence of taking an I_c-V_T curve) may leave behind trapped charge which could influence the result obtained from an I_c-V_T curve measured in its close proximity. We therefore devised a simple experiment in which the spacing d between the positions of the short I_c-V_T curves in the 5×5 grid was varied. The V_{th} extracted from the average of these 25 I_c-V_T curves is plotted in Fig. 8. We see that for $d \geqslant 25$ nm the averaged V_{th} is independent of spacing while for $d<25$ nm it increases at smaller d. This

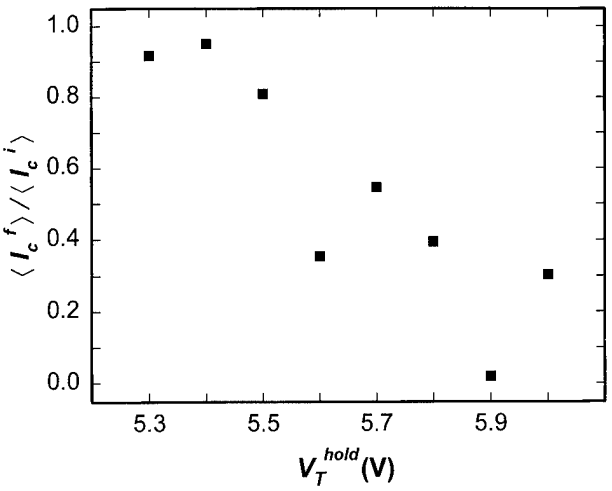


FIG. 5. Suppression of I_c vs tip bias V_T^{hold} under reverse oxide bias. Significant suppression of I_c is possible even for injection under a reverse oxide bias of $V_b^{\text{inj}}=-4$ V, provided the energy of electrons during injection (controlled by V_T^{hold}) is high enough. Almost no suppression is apparent for $V_T^{\text{hold}}=5.4$ V, but it increases substantially when V_T^{hold} is increased by ~0.5 V to ~5.9 V.

demonstrates that if the I_c-V_T curves are taken too close together ($\lesssim$25 nm) the charge left behind from the an I_c-V_T curve start to influence subsequent nearby I_c-V_T curves, which results in an increased apparent value for V_{th}. Also, a similar result is obtained when two I_c-V_T curves are taken subsequently at the same point (not shown). While the first yielded $V_{\text{th}} = 3.30$ V (V_b=+8 V, $I_t = 10$ nA), the second yielded $V_{\text{th}} = 3.65$ V.

F. Determining the magnitude and depth of the trapped charge

We have found that BEEM measurements can be used to estimate the amount and depth of the trapped charge in the oxide film. To do this, we consider a simple model for the trapped charge. In this model the trapped charge is represented by a uniform sheet charge at the distance $\bar{x}$ from the metal-oxide interface. This trapped charge of assumed electron density n_{trap} gives rise to additional voltage V_n, as is shown in Fig. 1(c). In this case, the expression for V_{th} (which corresponds to the maximum barrier height) is

$$V_{\text{th}}=\begin{cases} V_o+V_n-\bar{x}E_{\text{ox}} & \text{if } V_n \geqslant \bar{x}E_{\text{ox}} \\ V_o-\sqrt{\dfrac{e}{4\pi\varepsilon_d}\left(E_{\text{ox}}-\dfrac{V_n}{\bar{x}}\right)} & \text{if } V_n<\bar{x}E_{\text{ox}} \end{cases}, \quad (1)$$

with V_n given by[19]

$$V_n \equiv \frac{e\, n_{\text{trap}}}{\varepsilon_s}\frac{\bar{x}}{L}(L-\bar{x}).$$

Here $E_{\text{ox}} = (V_{\text{ox}}/L)$ is the applied electric field across the oxide without trapped charge, $V_{\text{ox}} \cong V_b - 0.6$ V (in the range of measured V_b),[11] $L \cong 25$ nm is the oxide film thickness, V_0 is the intrinsic barrier height at the metal-oxide interface [see Fig. 1(c)], $\epsilon_s \cong (3.9\ \epsilon_0)$ and $\epsilon_d \cong (2.15\ \epsilon_0)$ are the static

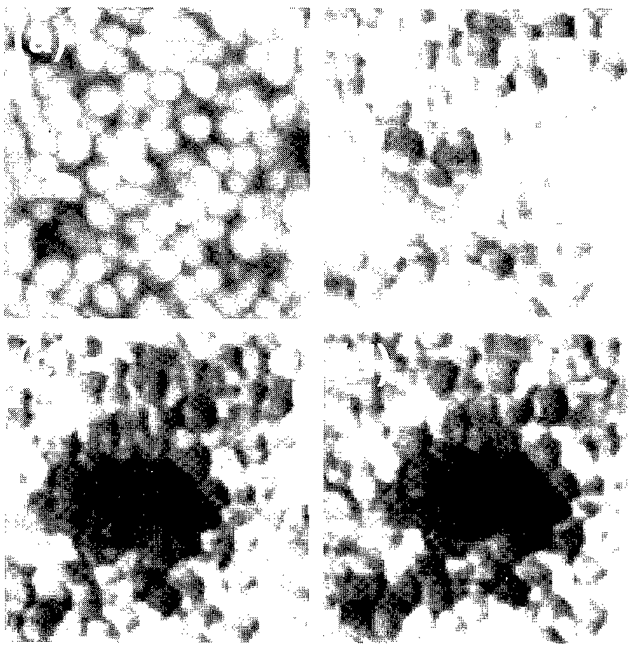


FIG. 6. Suppression of collector current I_c caused by injection at one point. (a) ~75x75 nm^2 topographic scan before injection, measured with V_T=5.4 V, $I_T \cong 10$ nA and $V_b^{\text{meas}}=+8$ V. (b) Corresponding BEEM image (gray-scale range 0–5 pA). (c) BEEM image after injection in the center of the region (Fig. 3). Approximate size of the suppressed region is 20–30 nm. (d) Same region after ~14 h. Suppression is still visible.

and optical dielectric constants of the oxide film, respectively, $\epsilon_0=8.85\times10^{-14}$ F cm^{-1}, and e is the elementary charge. For $V_n < \bar{x}E_{\text{ox}}$, the electric field due to the trapped negative charge increases V_{th} by reducing the amount of "image force lowering"[3] of the barrier below V_0, while for $V_n \geqslant \bar{x}E_{\text{ox}}$ the total electric field near the metal-oxide interface is actually reversed and the barrier is increased above V_0. Threshold voltages V_{th} were measured as a function of V_b for both pristine areas and in the center of regions injected in fashion similar to those discussed in Sec. III A (with $V_b^{\text{inj}}=+12$ V). Note that a "fresh" injected region was prepared for each new measurement. By measuring the I_c-V_T curves in the center of the injected region (~50 nm across) we ensure that the lateral extension of the trapped charge is sufficient to be reasonably approximated by an infinite sheet of charge. The model was fitted to both sets of measured threshold voltages by adjusting $\bar{x}$, n_{trap} and V_0 (note that for pristine regions we assume $n_{\text{trap}}=0$). The resulting fit is shown in Fig. 9. The model fits the measured threshold voltages for both pristine and injected areas very well, with fitted parameter values $\bar{x}\cong3.9$ nm, $n_{\text{trap}} \cong 1.3\times10^{13}$ cm^{-2}, $V_0 \cong 3.8$ V. We note that the approximate depth of the charge and its magnitude are in very reasonable agreement with previous macroscopic measurements for similar injection conditions.[13] The value of 3.8 V for V_0 (the Pt-SiO_2 barrier height at $V_{\text{ox}}=0$) corresponds well to the value measured by Ludeke *et al.* (~3.9 V) at $V_{\text{ox}}=0$ V.[5] Ludeke *et al.*, however, observed a similar threshold voltage even under strong forward bias conditions, while we observe lowering of V_{th} for positive (forward) bias V_b, in agreement with the predictions of image force lowering. A similar image force lowering effect for Au/Si Schottky diodes was observed using BEEM by Davies and Craighead.[20]

IV. DISCUSSION

Since we have made estimates of both the trapped charge density following injection and of the lateral extent of the trapped charge, we can now estimate the total number of electrons trapped in a certain area. It turns out that the number of trapped electrons is actually rather small. For example, assuming a trapped charge density $\sim1.3\times10^{13}$ cm^{-2} spread over the 50×50 nm^2 injected area shown in Fig. 4(d), we conclude that only about 300–350 electrons became trapped during the entire injection scan. This projects into only about 15 electrons for each out of the 25 I_c-V_T/I_c-t used for the injection. This suggests that it may ultimately be possible to use BEEM to directly observe the trapping or de-trapping of *individual* electrons in the buried SiO_2 film.[21] In fact, we have often observed abrupt fluctuations in I_c-t curves measured over particular sample areas (such as that shown in Fig. 3), which are reminiscent of discrete electron trap "switching" events reported by other techniques for the SiO_2 system.[21,22] Further work is clearly required to determine whether these fluctuations are in fact caused by discrete trapping events, and we will continue to study this in future experiments.

Past macroscopic studies of oxide charging indicate that some traps are already present in the "as-fabricated" oxide film while others can be *generated* as a consequence of the hot-electron current flowing through the oxide.[12,13] It is of interest to consider whether as-fabricated or generated traps are dominant in our measurements. If we assume that the traps we observe were already present in the oxide at the moment the injection started, we can roughly estimate their cross-section. The cross-section of such traps is $\sigma=eA/\tau I$, where A is the total area through which the "charging" current I is flowing and τ is the average time required to fill a trap.[23,24] We can estimate A from the approximate width of the suppressed region (Fig. 7), roughly 25^2 nm^2. A very conservative estimate of the current I is the final current in Fig. 3, i.e., about 2 pA. This is in fact the lowest possible limit for the following additional reasons. I_c is the current of electrons that make it all the way through the oxide. However, the current responsible for charging can be constituted of I_c *and* electrons which travel a distance into the oxide, and then are scattered back into the metal.[5–7] Also, if scattering of electrons is significant in the region of trapping, then the effective current is higher by the factor l/l_0, where l_0 is the total depth of the region of trapping and l is the average path of an electron through this region.[23] Finally, from Fig. 3 we estimate conservatively $\tau=10$ s as the characteristic time of the suppression decay. This yields $\sigma\leqslant5\times10^{-20}$ cm^2, which is much smaller than cross sections for typical as-fabricated electron traps in SiO_2 films.[24] We therefore conclude that filling of as-fabricated traps is not the main cause of the suppression of I_c and that electron traps *generated* by hot electron flux are the dominant effect and the rate-determining

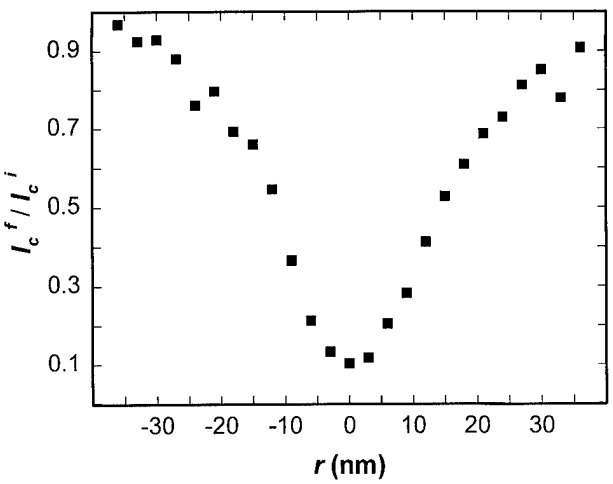


FIG. 7. An average profile of suppression after injection in the center (r=0 nm) calculated from several different suppressed regions generated on different sample locations (similar to Fig. 6). The FWHM of the average profile of about ~25 nm can be taken as an estimate of the extent of the area affected by injection at one point.

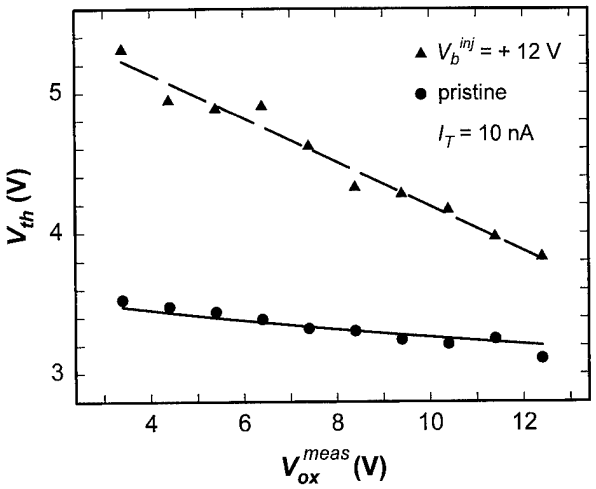


FIG. 9. Dependence of V_{th} on V_{ox}^{meas} applied during I_c-V_T curve measurements, for pristine (circles) and forward-bias injected (triangles) regions. Curves show model fits according to Eq. (1). The fit yields the approximate depth of trapped charge $\bar{x}$=3.9 nm, the trapped charge density n_{trap} = 1.3 × 10^{13} cm^{-2} (for the injected area), and V_0=3.8 V.

process in our experiments. In this regard, we note that past macroscopic experiments[12] reported that generated traps were dominant for injected charge densities $Q_c \gtrsim 1$ C cm^{-2} which is in agreement with our experiments where the typical charge densities are of the order of ~100 C cm^{-2}.

It should be noted that the lateral extent of the trapped charge as determined in Figs. 6–8 appears surprisingly large in comparison with the estimated depth of the trapped charge in the oxide $\bar{x} \cong 4$ nm, as determined from Eq. (1) and Fig. 9 for the injection at V_T=5.4 V. Since lateral scattering of the BEEM electrons in the 5 nm thick metal film is usually assumed to be relatively small,[7] then most of this "widening" must presumably occur in the oxide. Below, we consider several factors possibly contributing to this apparent large value of the width of the trapped charge compared with its depth. (1) A build-up of trapped electrons directly under the tip could scatter subsequent incoming BEEM electrons *sideways*, thus widening the distribution. Monte Carlo simulations similar to those performed by Ludeke *et al.*[5–7] could help determine the importance of this effect. (2) It is possible that the simple charge trapping model considered here somehow *underestimates* the depth of the centroid of the actual trapped charge distribution. In our simple model we have assumed the trapped charge in the form of a thin sheet, whereas the true distribution has some finite spread around the centroid depth in the oxide. We have, however, compared our experiments with a model which permits a finite *uniform* spread, and found that it does not significantly affect our estimate of the centroid. Also note that the charge centroid depth was determined for injection done with V_T = 5.4 V, whereas the lateral width of the suppressed region shown in Figs. 6 and 7 was determined for injection at a single point done at V_T = 10 V. It is possible that the charge centroid may be deeper for this higher-voltage injection. (3) It is possible that the large lateral extent of the trapped charge is a consequence of the actual mechanism for trap generation. In particular, we note that a leading model for trap creation[13] assumes that injected hot electrons must *completely traverse* the oxide film to liberate/activate a trap-creating species (thought to be hydrogen-related), which then diffuses back across the oxide to create traps near the injecting interface. Lateral diffusion of the injected electrons and/or the back-diffusing hydrogen could naturally produce a lateral spread in generated traps comparable to the oxide film thickness, (about 25 nm in our samples). While we do not yet know whether this particular mechanism is important in our experi-

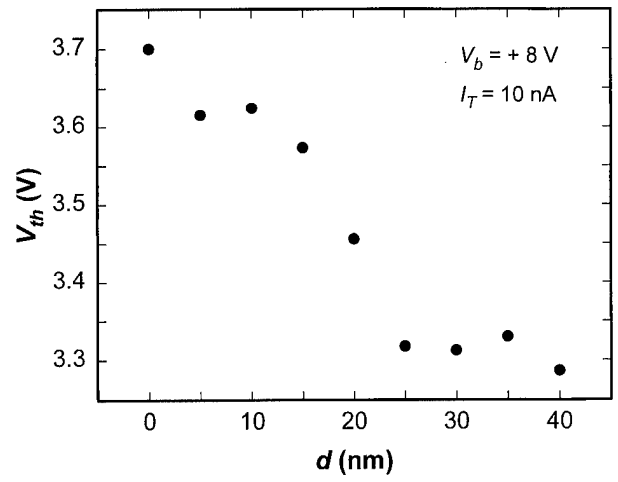


FIG. 8. Dependence of the apparent average threshold voltage V_{th} on the spacing d between adjacent of I_c-V_T curves. 25 I_c-V_T curves measured on a 5×5 grid of spacing d were averaged together and used to determine the threshold voltage V_{th}. If the I_c-V_T curves are spaced too closely, the residual charge from the previous I_c-V_T curve will influence the following I_c-V_T curves, resulting in an increase in V_{th}. For $d \gtrsim 25$ nm the averaged V_{th} is independent of spacing.

ments, we note here that BEEM should actually be a very useful tool to study this (and other) possible trap generating mechanisms.

V. SUMMARY

We used BEEM to locally inject electrons into a ~25 nm thick SiO_2 film built into a MOS structure. Some of these electrons are trapped in the bulk of the oxide and cause an increase in the local barrier height and suppression of the BEEM current. We devised a method using BEEM to determine locally the density of this trapped charge and a its depth in the oxide. We estimated the lateral extent of the charge trapped during injection at one point, and concluded that generation of traps created in the oxide by hot BEEM electrons is a likely dominant trapping mechanism in our experiments. We also observed lowering of the barrier due to the image force effect, and were able to extract the intrinsic barrier height at the Pt-SiO_2 interface.

ACKNOWLEDGMENTS

The authors wish to acknowledge the advice and help of Darrell Jones. This work is supported by the Office of Naval Research, Grant No. 00014-93-0607, and by NSF Grant No. DMR93-57535.

[1]R. W. Keyes, *The Physics of VLSI Systems* (Addison-Wesley, New York, 1987).
[2]E. H. Nicollian and J. R. Brews, *MOS Physics and Technology* (Wiley, New York, 1982).
[3]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).
[4]Cuberes *et al.*, J. Vac. Sci. Technol. B **12**, 2646 (1994).
[5]R. Ludeke, A. Bauer, and E. Cartier, Appl. Phys. Lett. **66**, 730 (1995).
[6]R. Ludeke and A. Bauer, J. Vac. Sci. Technol. A **13**, 614 (1995).
[7]R. Ludeke, A. Bauer, and E. Cartier, J. Vac. Sci. Technol. B **13**, 1830 (1995).
[8]B. Kaczer, Z. Meng, and J. P. Pelz, Phys. Rev. Lett. **77**, 91 (1996).
[9]W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988); L. D. Bell and W. J. Kaiser, Phys. Rev. Lett. **61**, 2368 (1988).
[10]E. L. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford University Press, New York, 1989).
[11]V_{ox} (the voltage drop across the oxide) differs slightly from V_b (the applied external voltage) due to band-bending in the Si substrate (Ref. 3) and workfunction difference between Pt and Si (Ref. 7).
[12]D. J. DiMaria and J. W. Stastiak, J. Appl. Phys. **65**, 2342 (1989).
[13]D. J. DiMaria, E. Cartier, and D. Arnold, J. Appl. Phys. **73**, 3367 (1993).
[14]It is generally known that x rays generated during e-gun deposition can corrupt the quality of oxide films. We have found no problem with oxide breakdown with our 25 nm thick oxide films for e-beam depositions done at an acceleration voltage of 4.1 kV.
[15]T. R. Albrecht *et al.*, Appl. Phys. Lett. **55**, 1727 (1989).
[16]C. J. Roberts *et al.*, J. Vac. Sci. Technol. B **9**, 841 (1991).
[17]H. D. Hallen *et al.*, J. Vac. Sci. Technol. B **9**, 585 (1991); A. Fernandez *et al.*, Appl. Phys. Lett. **57**, 2826 (1990).
[18]Here we make the implicit assumption that the wavefunction amplitude of a BEEM electron at the metal-oxide interface is independent of the potential deep in the oxide (Ref. 8). A similar separation of the interface and the bulk oxide is assumed in Refs. 5, 6, and 7. Therefore the probability of interface damage by BEEM electrons should also be *independent* of the potential deep in the oxide.
[19]D. J. DiMaria, J. Appl. Phys. **47**, 4073 (1976).
[20]A. Davies and H. G. Craighead, Appl. Phys. Lett. **64**, 2833 (1994).
[21]M. E. Welland and R. H. Koch, Appl. Phys. Lett. **48**, 724 (1986).
[22]K. S. Ralls *et al.*, Phys. Rev. Lett. **52**, 228 (1984).
[23]D. R. Wolters and A. T. A. Zegers-van Duynhoven, in *The Physics and Technology of Amorphous SiO_2*, edited by R. A. B. Devine (Plenum, New York, 1988), p. 391.
[24]D. J. DiMaria, in *The Physics of SiO_2 and its Interfaces*, edited by S. T. Pantelides (Pergamon, New York, 1978), p. 160.

Contactless capacitance–voltage and photoluminescence characterization of ultrathin oxide–silicon interfaces formed on hydrogen terminated (111) surfaces

Tamotsu Hashizume, Satoshi Koyanagi, and Hideki Hasegawa
Research Center for Interface Quantum Electronics, and Graduate School of Electronics and Information Engineering, Hokkaido University, Sapporo 060, Japan

(Received 21 January 1996; accepted 16 April 1996)

Electronic properties of the interfaces between Si and ultrathin ($\lesssim$10 Å) oxides formed by various low-temperature processes were characterized in contactless fashion, using contactless capacitance–voltage and photoluminescence surface state spectroscopy techniques together with x-ray photoelectron spectroscopy measurement. Hydrogen (H) terminated Si(111) surfaces were used as the initial surface. Ultrathin oxides were formed at low temperatures by chemical oxidation processes (hot HNO_3, H_2SO_4+H_2O_2), long-time air exposure, and low-temperature oxidation processes below 350 °C. The initial H-terminated surfaces showed presence of Fermi-level pinning at $E_0 = E_V + 0.65$ eV due to high density of amphoteric discrete state probably originating from Si dangling bonds. On the other hand, all the ultrathin oxide–Si interfaces exhibited very limited capacitance variations with voltage at low capacitance levels similar to GaAs metal–insulator–semiconductor systems, and indicated that the Fermi level is pinned near the hybrid orbital charge neutrality level E_{HO} due to presence of interface states with narrow U-shaped continuous distributions. Low-temperature oxidation at 350 °C slightly weakens such pinning. The present work indicates difficulty of realizing unpinned ultrathin oxide–silicon interface by low-temperature processes. © *1996 American Vacuum Society.*

I. INTRODUCTION

Since feature sizes used in silicon ultralarge-scale integrated circuits (ULSIs) have been scaled down into the submicron range, the importance of the quality of thin oxide layers as well as their formation process and preformation surface treatments related to surface roughness on the atomic scale has been recognized. Obviously, importance of atomic scale control of surfaces and interfaces will further increase in the future in the Si ULSI industry. In addition, the realization of next-generation advanced devices in the quantum regime such as single electron transistors,[1] will also depend strongly on successful formation of a highly controlled ultrathin oxide–Si interface, because availability of ultrathin pinning-free tunneling barriers is crucial in such devices.

High-temperature dry oxidation is the standard technique to produce gate oxide in current ULSI processes. However, it is getting increasingly difficult to form extremely thin (a few 10 Å) oxide films with precise control of the thickness and electrical bulk and interface properties using the high-temperature dry oxidation technique only. This is due to the atomic-scale roughness and presence of native oxides/suboxides on the initial Si surfaces as well as due to rapid initial oxidation rates of the Si surface at elevated temperatures, whose values depend sensitively on initial surface conditions and the temperature/gas conditions.

One approach to overcome this difficulty is to perform a suitable initial surface treatment to realize a well-defined atomically flat surface and subsequently to grow a well-controlled ultrathin ($\lesssim$10 Å) "pre-oxide" at low temperatures before final oxidation is made to a desired thickness at an elevated but optimal temperature with a well-controllable oxidation rate. In fact, such an approach has been recently attempted by Ohmi *et al.*[2] who have found it useful to realize an atomically flat interface and high-quality oxide with a precise control of thickness.

As a possible candidate for the preprocessing surface treatment, particular attention has been paid recently to the hydrogen termination of the silicon surface. According to the literature,[3–6] it is capable of producing atomically well-defined and flat Si surfaces with strong surface passivation properties and high reproducibility.

As for controlled formation of ultrathin preoxides, one could try various low-temperature processes to find out a simple and optimal process. In order to successfully develop a most suitable surface treatment and preoxide formation process, surface/interface characterization techniques play important roles.

For characterization of atomic scale structures of surfaces and interface, surface-science-type characterization techniques have recently made great progress which include high-resolution x-ray photoelectron spectroscopy (XPS),[7] scanning tunneling microscopy (STM)/atomic force microscopy (AFM),[8,9] grazing angle x-ray diffraction (XRD),[10] crystal truncation rod (CTR) scattering,[11] Fourier transform infrared-absorption spectroscopy (FTIR),[3,12] etc.

However, there has been no well-established suitable method to characterize more device-oriented electronic properties of the hydrogen terminated surfaces and thin oxidized surfaces. The standard metal–insulator–semiconductor (MIS) capacitance–voltage (C–V) technique fails due to

Schottky or tunneling MIS conduction in such an ultrathin regime. Recently use of the old surface photovoltage (SPV) technique[6,13] which detects surface potential change due to photoinduced band flattening has been proposed.[14] However, it requires a complicated parameter fitting analysis due to presence of the Dember effect and therefore it is difficult to judge the quantitative validity of the results.

In order to overcome this difficulty, we previously proposed the photoluminescence surface state spectroscopy (PLS^3) technique and applied it to silicon surfaces.[15,16] More recently, we applied a contactless $C-V$ technique developed by Sakai *et al.*[17] to the characterization of hydrogen terminated and air-exposed silicon (111) surfaces together with the PLS^3 technique, and showed that these two techniques are complementary to each other and provide a powerful means for characterization of free surfaces.[18]

In a continuation of efforts to establish well-defined preprocessing conditions for formation of well-controlled extremely thin thermal oxides on Si, the purpose of the present article is to characterize the electronic properties of the ultrathin oxide ($\lesssim$10 Å)–Si(111) interfaces prepared by various low-temperature oxidation processes. The contactless $C-V$ and the PLS^3 techniques were used together with x-ray photoelectron spectroscopy (XPS) measurement. As the initial surface treatment prior to formation of ultrathin oxides, hydrogen termination was applied in order to take its full average of well-defined atomic structure, flatness, and high reproducibility. Low-temperature processes to grow ultrathin oxide include chemical oxidation processes (hot HNO_3, $H_2SO_4+H_2O_2$), a long-time air exposure, and a low-temperature thermal oxidation below 350 °C.

XPS band-bending, $C-V$, and PLS^3 measurements gave results that are consistent with each other. The initial H-terminated surfaces showed the presence of Fermi-level pinning at $E_0=E_V+0.65$ eV due to high density of amphoteric discrete state probably due to Si dangling bonds. On the other hand, the results on all the ultrathin-oxide-covered surfaces indicated that the Fermi level is pinned near the charge neutrality level E_{HO} due to high-density narrow U-shaped interface state continuum being similar to GaAs MIS interfaces.[19] Low-temperature oxidation at 350 °C slightly weakens such pinning.

The present work indicates that realization of an unpinned ultrathin oxide–silicon interface by a low-temperature process only is unexpectedly difficult even with the thermal oxidation.

II. EXPERIMENT

A. Sample preparation and XPS analysis

Wafers used in the present study were n-type silicon (111) with the carrier concentration of $1–2\times10^{15}$ cm^{-3}. The following surfaces were prepared and characterized:

(1) a thick thermally oxidized surface;
(2) a hydrogen terminated surface;
(3) chemical oxidized surfaces;
(4) an air-exposed surface; and
(5) low-temperature thermally oxidized surfaces in a vacuum chamber.

The ultrathin oxidized surfaces of (3)–(5) were formed on the hydrogen-terminated surface (2).

The thermally oxidized surfaces were prepared in dry O_2 at 1000 °C for 3 h and annealed at for 30 min in dry N_2. This surface was used as a reference.

Hydrogen termination of the surface was carried out according to the method of Higashi *et al.*[3] Namely, thick thermal oxide (about 150 nm) films were formed in dry oxygen at 1000 °C. Then, after removing the oxide films in buffered hydrofluoric acid, native oxide films were formed by immersion of the surface in a $HCl–H_2O_2$ solution at 80 °C for 10 min. Finally, the surface was immersed in a 40% NH_4F solution. After each process, the sample was rinsed for 10 min in de-ionized water. According to the literature,[3,5] such a process gives an atomically flat hydrogen-terminated silicon (111) surface.

Chemical oxidation of the hydrogen terminated surfaces was done following the recent work by Ogawa and Hattori,[20] i.e., one immersed into a hot HNO_3 solution (60 °C) for 5 min, and the other into a solution of H_2SO_4:H_2O_2=4:1 for 10 min. Formation of ultrathin oxide by air exposure was done leaving the hydrogen-terminated surface in an air–nitrogen mixture for 2 weeks. For comparison, as-received wafers with saturated formation of native oxide by air exposure was also studied.

Ultrathin low-temperature thermal oxides[2] were formed by heating the hydrogen terminated surface at 300–350 °C for 30 min in dry oxygen with a pressure of 133 Pa in an ultrahigh-vacuum (UHV) chamber (base pressure $<10^{-7}$ Pa). Chemical status of the various surfaces were investigated by XPS. For this, a Perkin–Elmer PHI 5100C spherical capacitor energy analyzer was used under excitation by Mg $K\alpha$ ($h\nu=1253.6$ eV).

B. Contactless $C-V$ method

The principle of the contactless $C-V$ method used in this study is shown in Fig. 1(a). Instead of depositing a gate electrode directly on the sample surface, $C-V$ measurement is performed from the field electrode that is placed above the sample, being separated from the sample surface by a thin "air gap." Thus, a MIS assessment of "free" surfaces becomes possible by being assisted by an air-gap insulator.

In the present study, such contactless $C-V$ measurement was performed using a commercial measurement system (CV-8000, Dainippon Screen Mfg. Co., Ltd.). In this system, an electrode head shown in Fig. 1(b) is used in order to maintain parallelism and a constant distance between the field electrode and the sample surface. The head has three additional surrounding "parallelism electrodes" which maintain a constant air-gap distance of 300–350 nm via a piezo mechanism with capacitance feedback. The capacitance is then measured between the field electrode and the sample mounted on the sample stage. The area of the electrode for $C-V$ measurement was 7.5×10^{-3} cm^{-2} and Cr was used as

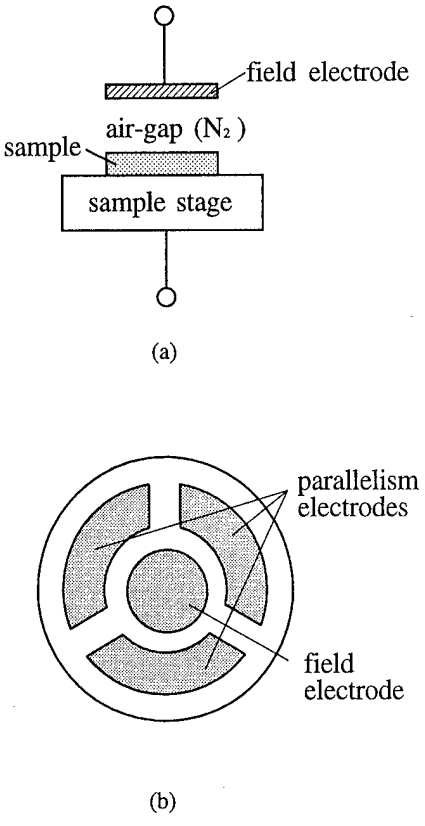


FIG. 1. (a) Basic principle of the contactless $C-V$ method and (b) electrodes used for the contactless $C-V$ system.

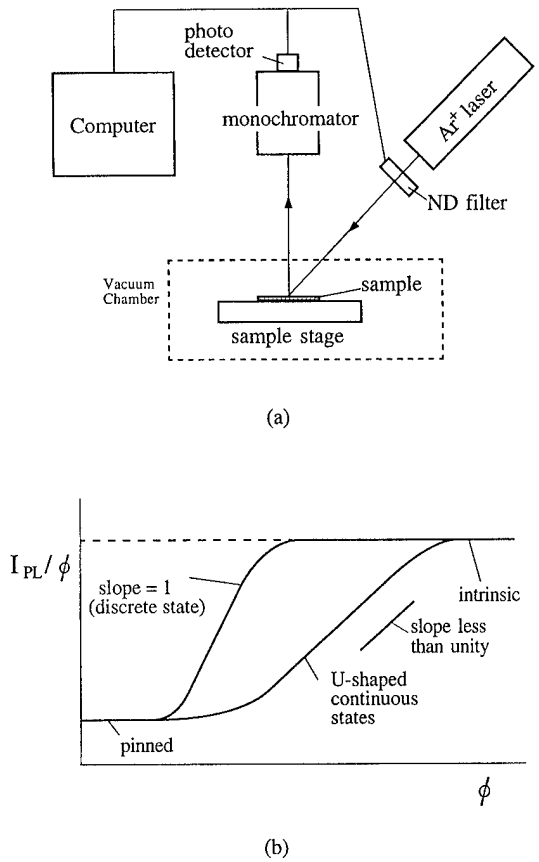


FIG. 2. (a) Experimental setup of the PLS^3 technique. (b) Schematic illustration of typical behavior of band-edge PL efficiency.

the field plate metal. The absolute value of the air-gap distance can be measured optically using the "Goos–Haenchen effect" of optics in this system.[17]

Formation of an ohmic contact at the back surface of the samples was not necessary, because the capacitance between the sample and the sample stage is much larger than the capacitance to be measured.

C. Photoluminescence surface state spectroscopy (PLS^3)

It is well known that the intensity of band-edge photoluminescence (PL) is very sensitive to surface properties; however, it is generally used for qualitative assessment of the surface quality. The photoluminescence surface state spectroscopy (PLS^3) technique recently developed by our group[15,16] allows, on the other hand, a quantitative assessment of the surface state properties of free surfaces in a contactless fashion.

The experimental setup and the principle of the PLS^3 technique are shown in Fig. 2. In this system, the band-edge PL efficiency I_{PL}/ϕ, where I_{PL} is the band-edge PL intensity and ϕ the excitation intensity, is measured as a function of ϕ, using an automated setup schematically shown in Fig. 2(a). A typical behavior of the PL efficiency after such a measurement is shown in Fig. 2(b). In high-quality Si single crystals, the bulk recombination is of little importance at low excitation intensities and the PL efficiency takes a low constant value due to a relatively large surface recombination velocity due to Fermi-level pinning. On the other hand, at high excitation intensities, the PL efficiency approaches an intrinsic value limited by bulk radiative and Auger recombination processes.

In the transition region between these two limits, the PL efficiency becomes very sensitive to surface state properties. Here, as the light intensity is increased, the Shockley–Read–Hall (SRH) recombination rate of each surface state becomes saturated. At the same time, the quasi-Fermi levels for electrons and holes scan the surface band gap as the photoinduced band flattening proceeds and this gradually increases the number of surface states causing recombination. Thus, the former tends to increase the PL efficiency whereas the latter tends to reduce it, making the behavior of the PL efficiency in this region strongly reflect the N_{SS} distribution. Roughly speaking, the slope of the PL efficiency gives a good indication concerning the shape of the distribution. It is equal to unity for discrete states due to complete saturation of the SRH process. On the other hand it becomes less steep for continuous distribution, reflecting the aforementioned increase of surface states taking part in recombination. The density (N_{SS}) distribution can be determined by fitting the measured I_{PL}/ϕ with the results of rigorous computer simulation taking account of all possible recombination processes. The procedure for the computer analysis was given in

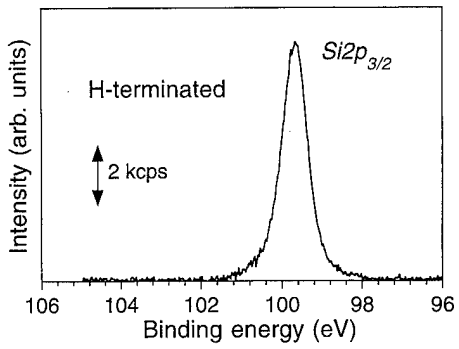


FIG. 3. XPS core-level spectrum of the Si $2p_{3/2}$ of H-terminated Si(111) surface.

detail in Ref. 15. In the present study, a PL measurement system which is fully automated by computer control was used. The excitation source was an Ar^+ laser light with a wavelength of 514.5 nm.

III. RESULTS AND DISCUSSION

A. XPS analysis

Figure 3 shows the spectrum of Si $2p_{3/2}$ core level obtained on the H-terminated surfaces. The Si $2p_{1/2}$ peak was subtracted from the original signal by employing a spin-orbit-splitting value of 0.6 eV and the Si $2p_{1/2}$ to Si $2p_{3/2}$ intensity ratio of 1/2 based on the previous results.[21] As can be seen, there is no oxide component around 103 eV on the H-terminated surface. In addition, the peak position of the Si—Si bond was found to be 99.65 eV consistent with those reported by other workers.[22,23]

Figure 4 compares the Si $2p_{3/2}$ spectra obtained after oxidation of the H-terminated surfaces by various low-temperature processes. In Fig. 4(a), the peak position and height of the main Si peak corresponding to Si—Si bond were shifted and adjusted for each oxidation process so that they agree with those of the H-terminated surface. It is clearly seen in Fig. 4(a) that an ultrathin oxide was formed after each of the oxidation processes tried. The oxide film thickness can be calculated by analyzing these Si $2p$ photoelectron spectra, if the electron escape depth is known. By assuming that the electron escape depth in Si and Si oxide are 2.6 and 3.4 nm, respectively,[20] the thickness estimates shown in Table I were obtained for the various low-temperature oxides. These values are in excellent agreement with those in the literature.[2,20]

In order to see the intermediate oxidation status more carefully, oxide peaks are shown in a magnified scale in Figs. 4(b) and 4(c). The positions corresponding to the oxidation status of Si of Si^{1+}, Si^{2+}, Si^{3+}, and Si^{4+} are shown by arrows using the data by Himpsel *et al.*[21] The relative amounts of the suboxide components were separated by the curve fitting and the results are summarized in Table I in terms of the integrated intensity ratio with respect to the main Si^0 peak. As seen in Table I, not only the amount of the SiO_2 component (Si^{4+}) but also the amount of each suboxide component is different for different oxidation processes and conditions.

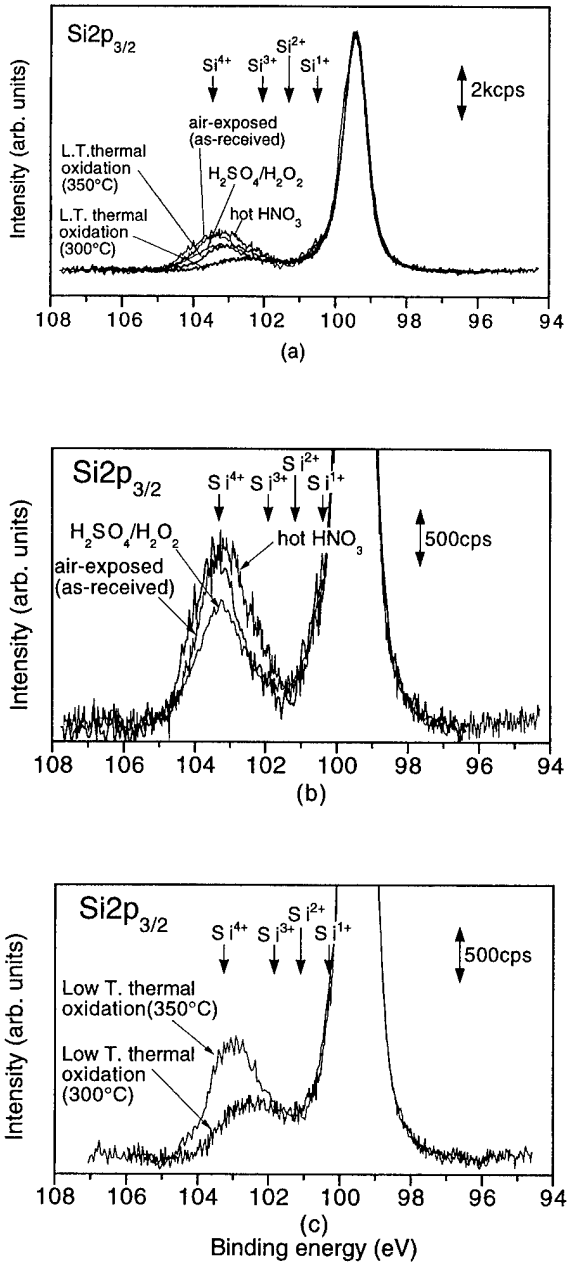


FIG. 4. (a) XPS spectra of the Si $2p_{3/2}$ of H-terminated Si(111) surfaces after formation of ultrathin oxide by various processes. The peak position and height of the main Si peak corresponding to Si—Si bond were shifted and adjusted for each oxidation process so that they agree with that of the initial H-terminated surface. Spectra are shown in magnified scale for (b) air-exposed and chemically oxidized surfaces, and (c) low-temperature thermally oxidized surfaces.

For example, the Si^{1+} phase is emphasized at the surface chemically oxidized in H_2SO_4 solution, whereas the surface chemically oxidized in HNO_3 solution exhibited a larger Si^{3+} component in agreement with the spectra reported by Ogawa and Hattori.[20] For the low-temperature thermal oxidation, a remarkable change in the interface suboxide phase was observed by a slight change in the oxidation temperature:

TABLE I. Summary of suboxide thickness and components for various oxidation methods at low temperatures.

Method	Oxide thickness (nm)	Si^{1+}/Si^0	Si^{2+}/Si^0	Si^{3+}/Si^0	Si^{4+}/Si^0
Air exposure	1.0±0.1	0.04±0.01	0.02±0.01	0.03±0.01	0.24±0.005
Chemical oxidation: H_2SO_4	0.8±0.1	0.07±0.01	0.02±0.01	0.02±0.01	0.20±0.005
HNO_3	1.1±0.1	0.06±0.01	0.01±0.01	0.08±0.01	0.26±0.005
Thermal oxidation: 300 °C	0.4±0.1	0.04±0.01	0.02±0.01	0.07±0.01	0.02±0.01
350 °C	0.8±0.1	0.05±0.01	0.02±0.01	0.02±0.01	0.22±0.005

Namely, the Si^{3+} component was dominant in the oxidized interface prepared at 300 °C, whereas a sharp Si^{4+} peak appeared at 350 °C, keeping, however, the amount of other suboxide components more or less the same.

Finally, an attempt was made to determine the zero-bias Fermi-level positions of the sample surfaces from the XPS peak shifts of Si $2p$ core levels. To do this, one needs to know the energy separation between the Si $2p$ core level and the valence-band edge. Since the resolution of our XPS equipment was not high enough to resolve the positions of Si $2p_{3/2}$ and Si $2p_{1/2}$ peaks very accurately by curve fitting, we adopted the procedure by Schwaltz *et al.*[24] who used a weighted average peak energy of Si $2p$ level of $E_V-E(\mathrm{Si}\,2p)=99.01$ eV. This value was derived from a high-resolution work on nearly intrinsic Si wafers by Himpsel *et al.*[25] which gave $E_V-E(\mathrm{Si}\,2p_{3/2})=98.86$ eV and $E_V-E(\mathrm{Si}\,2p_{1/2})=99.46$ eV. The estimated Fermi-level positions on various sample surfaces are summarized in Fig. 5.

As seen in Fig. 5 the Fermi level at the H-terminated surface lies at 0.60–0.70 eV from the valence-band edge. On the other hand, a clear energy shift of the core level was seen on the surfaces after oxidation. The resultant positions depended rather weakly on the oxidation process and they lie in the range of 0.30–0.40 eV from the valence-band edge as shown in Fig. 5. Obviously, the present determination of the E_F position depends on the accuracy of the value of $E_V-E(\mathrm{Si}\,2p)$. For example, a more recent work by Himpsel, Hollinger, and Pollak[26] gave $E_V-E(\mathrm{Si}\,2p_{3/2})=98.74$ eV and $E_V-E(\mathrm{Si}\,2p_{1/2})=99.34$ eV, and this would shift our result upward by 0.12 eV, making the E_F positions on the H-terminated and oxidized surfaces lie at 0.72–0.82 and 0.42–0.52 eV, respectively. In order to check the validity of the reference value of $E_V-E(\mathrm{Si}\,2p)$, we further attempted to directly measure the valence-band spectra of H-terminated Si surfaces. This gave $E_V-E_F=0.62\pm0.05$ eV and this seems to favor the result in Fig. 5 rather than the more recent result by Himpsel and co-workers. This latter work by Himpsel and co-workers[26] assumed that deposition of a Ga overlayer or a (Cs+O) overlayer on the clean Si surface brings the Fermi level exactly to the valence- or conduction-band edge, respectively. Such an assumption might be wrong, judging from the result on Schottky barriers formation.[27] On the other hand, the result in Fig. 5 agrees very well also with $C-V$ results shown in the following subsection.

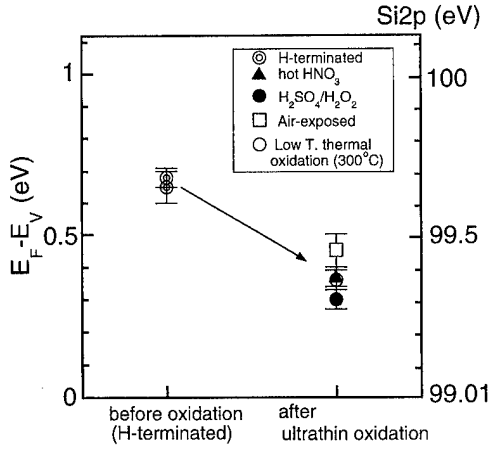


FIG. 5. Zero-bias Fermi-level positions of the sample surfaces determined from the XPS peak shifts of Si $2p$ core levels before and after ultrathin oxidation of the H-terminated surfaces.

B. Contactless $C-V$ measurements

$C-V$ measurement was performed at the small signal frequency of 500 kHz. The insulator capacitance C_I was estimated from the optically determined air-gap distance. All measurements were carried out in pure N_2 (99.9999%) atmosphere, utilizing a "nitrogen gap" rather than an air gap.

The results of $C-V$ measurement for the samples having high-temperature thermal oxide film (100 nm) are shown in Fig. 6. The calculated ideal $C-V$ curve is shown by a broken line. It was calculated from the doping density and the insulator capacitance including the air-gap contribution, assuming flatband at zero bias, and ignoring the effect of work-function difference between the metal field plate and the semiconductor surface. A well-behaved $C-V$ curve was obtained for the high-temperature thermally oxidized surface with the presence of some amount of positive charges. The hysteresis which is seen only in the deep depletion region corresponds to the nonequilibrium supply of the minority carriers. Pulsed $C-V$ measurement was also performed in which gate voltages were supplied by pulse wave forms. As expected, the response of minority carriers disappeared, and the $1/C^2-V$ plot gave the correct carrier concentration.

The contactless $C-V$ curves taken on a H-terminated surface are shown in Fig. 7(a). Flattening of capacitance took place under positive and negative bias sweeps with a very large hysteresis resulting in a very limited range of $C-V$ variation. Complete flattening under positive bias indicates

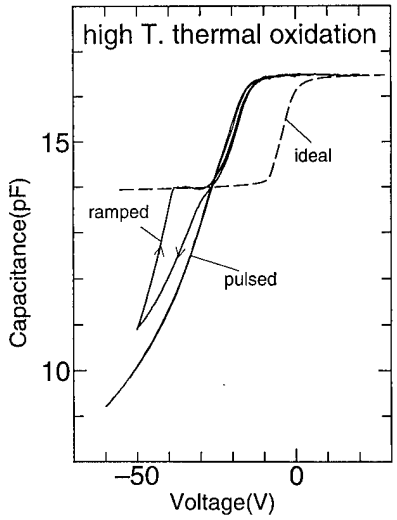


FIG. 6. Contactless $C-V$ curves for the samples having high-temperature thermal oxide film (100 nm). The calculated ideal $C-V$ curve is shown by the broken line.

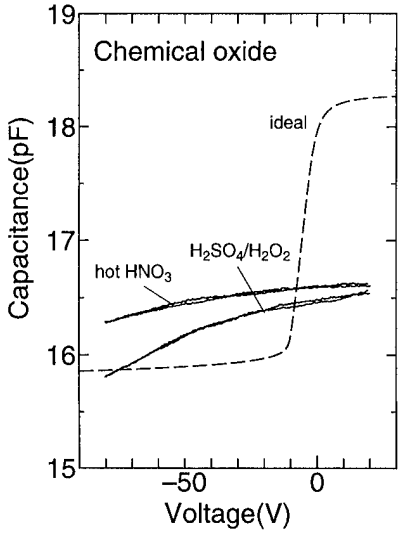


FIG. 8. Contactless $C-V$ curves after chemical oxidation of H-terminated surfaces.

existence of a strong Fermi-level pinning at slightly above the midgap. By comparing the measured capacitance and the ideal $C-V$ curve, the pinning position of E_F corresponding to the flattening position of capacitance was estimated to be 0.69 eV above the valence-band edge. This estimation includes a possible error of ± 0.1 eV mainly caused by the measurement error of the air-gap distance of $\Delta d_{\text{air}} \approx 5$ nm. The pinning position thus obtained is in good agreement with the XPS band bending data shown in Fig. 5. This flattening persisted in pulsed $C-V$ curves and takes place at the same position.

On the other hand, the width of hysteresis under a negative bias increased as the sweep rate of bias decreased. This large hysteresis was also observed for the pulsed $C-V$ curve. The flat portion in the negative going sweep corresponds to the high-frequency inversion capacitance, indicating that an inversion layer is produced rapidly through a rapid surface generation of minority carriers. However, this inversion level is not stable under positive-going sweep and the capacitance quickly comes up to an upper plateau. Furthermore, our previous analysis of the capacitance transients[18] also showed anomalous two-step transients under a negative bias, indicating that the inversion is not stable on this surface. As discussed in detail previously,[18] the whole behavior can only be explained by presence of amphoteric discrete states which have donor and acceptor characters. Such a state may be due to Si dangling bonds surrounded by Si—H bonds.

Another remarkable feature of the H-terminated surface was reproducibility of its electronic properties. As an example, Fig. 7(b) shows the $C-V$ curve taken after the second H termination done on chemical oxidation (hot HNO_3) of the initial H-terminated surface. It is very interesting that the characteristic $C-V$ behavior with a large hysteresis is completely reproduced after removal of the chemical oxide by a NH_4F solution, i.e., reappearance of the H-terminated Si(111) surface, although the $C-V$ behavior of the chemically oxidized surface is totally different as shown in the following. This reproducibility makes the H termination particularly attractive as the initial surface treatment for atomically controlled processing.

Figure 8 shows the $C-V$ curves after two-types of chemi-

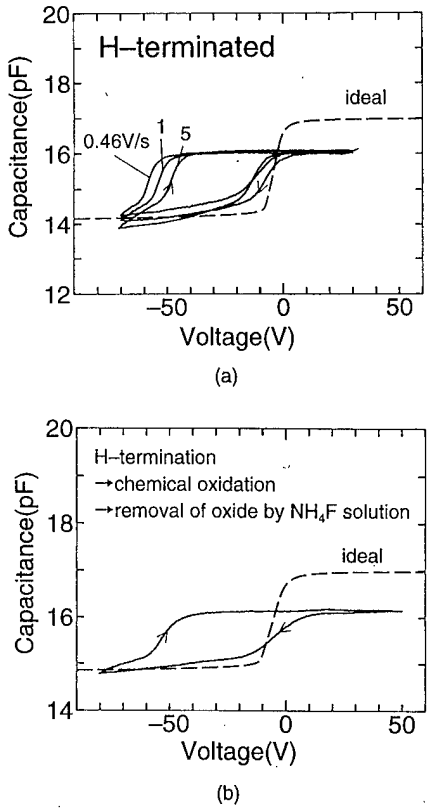


FIG. 7. Contactless $C-V$ curves taken (a) on the initial H-terminated surface and (b) after second H termination done on chemically oxidized surface.

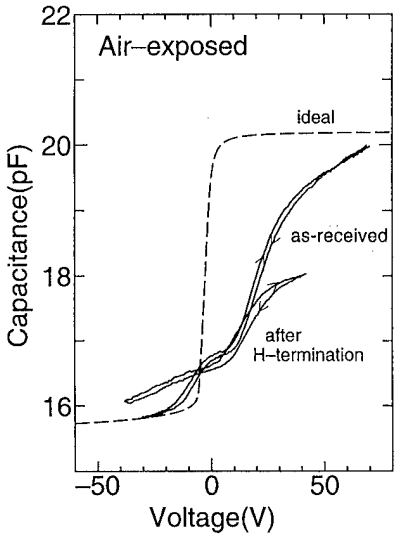


FIG. 9. Contactless $C-V$ curves obtained at the air-exposed surfaces.

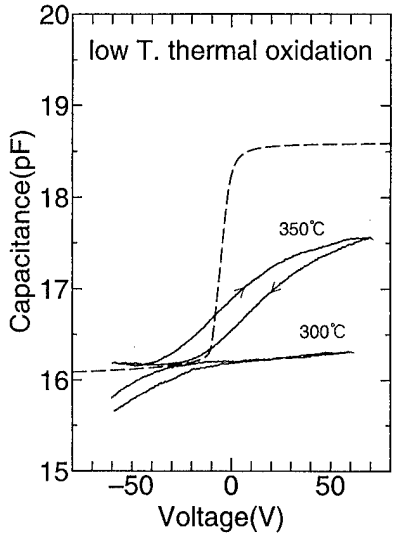


FIG. 10. Contactless $C-V$ curves after low-temperature thermal oxidation of H-terminated surfaces.

cal oxidation processes of H-terminated surfaces. As seen in Fig. 8, the behavior is completely different from that of the H-terminated surface. An extremely limited capacitance change is seen at a lower capacitance level. The shape of the curves is similar to that encountered in GaAs MIS systems[19] where the $C-V$ curve under accumulation bias shows a strong frequency dispersion and eventually flattens out to a small value which is far below the insulator capacitance. This behavior indicates presence of a high density of interface states which cause a strong Fermi-level pinning at a much deeper position than the H-terminated surface. The capacitance analysis indicates that the pinning position is at about 0.35–0.40 eV above the valence band edge. This agrees again with the XPS band-bending result. It should be noted that the observed energy position is close to the hybrid orbital charge neutrality level, E_{HO},[28] of 0.32 eV, or to the Tersoff midgap energy level[29] of 0.36 eV for Si.

Figure 9 shows the results of $C-V$ measurements on surfaces after air exposure of the H-terminated surface. The result obtained on the as-received air-exposed surface is also included. Again, a very limited range of capacitance variation was observed, indicating Fermi-level pinning near E_{HO}. In both samples, the capacitance flattens under a positive bias, and then further increases as shown in Fig. 9. This shape of the curve is again similar to that of an n-type GaAs MIS capacitor[19] with strong frequency dispersion of capacitance. Thus, it is most likely that the second increase of capacitance will flatten out at high measurement frequencies.

Figure 10 shows the $C-V$ curves after low-temperature thermal oxidation of the H-terminated surfaces. When the surface was oxidized at 300 °C, a very flat $C-V$ curve was obtained. For the surfaces after oxidation at 350 °C, on the other hand, the capacitance change was enhanced as seen in Fig. 10. This appears to be correlated with the result of XPS analysis. Namely, the surfaces having a large amount of suboxide components exhibited a small capacitance variation, whereas increase of Si^{4+} component at a higher oxidation temperature led to a larger capacitance variation.

On the basis of the measured $C-V$ results, one should be able to determine the N_{SS} distribution by applying the Terman method. However, this was not attempted here because of the following ambiguities. First, the Terman method assumes that a true high-frequency limit is reached in the $C-V$ measurement. In the present contactless $C-V$ measurement system with a complex capacitance feedback mechanism, the measurement frequency was limited to a single frequency of 500 kHz. This frequency is high enough for the standard thermally oxidized Si MOS capacitors to reach the high-frequency limit. However, since rather nonstandard interfaces are dealt here, it may not be high enough judging from our previous experience on GaAs MIS systems[19] where the true high-frequency limit was reached at microwave frequencies. Thus, a nonflattened portion of $C-V$ curves under high-accumulation bias requires a more careful examination in the future to determine the N_{SS} distribution quantitatively. Another ambiguity lies in the determination of the insulator capacitance. Although the present system can determine the air-gap distance within ± 5 nm on thermally oxidized surfaces by optical method, further investigation is necessary as to its applicability to the surfaces covered with other types of oxides.

On the other hand, relative comparison of the density of interface states between different surfaces can still be made by comparing the range of capacitance variation with bias: namely, as seen in Figs. 6, 8, 9, and 10, the capacitance corresponding to the capacitance flattening decreases in the order of the high-temperature thermally oxidized surface, the low-temperature thermally oxidized surface at 350 °C, the air-exposed surface, the chemically oxidized surface, and the low-temperature thermally oxidized surface at 300 °C. In $C-V$ curves, the lower the flattening capacitance level under positive bias is, the narrower the surface potential sweep

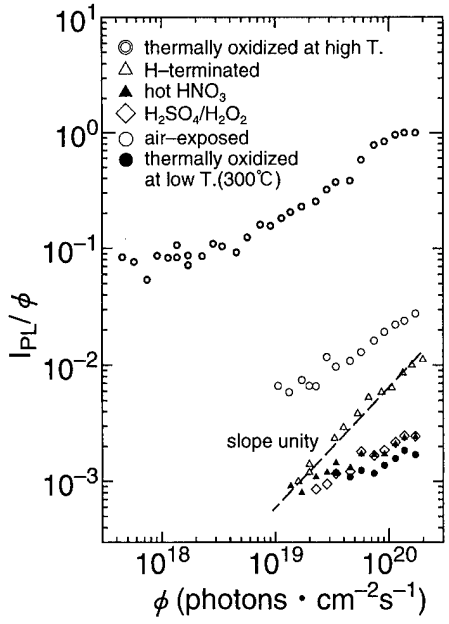


FIG. 11. The result of PLS[3] measurements for various Si(111) surfaces. The PL efficiencies are plotted as a function of excitation intensity.

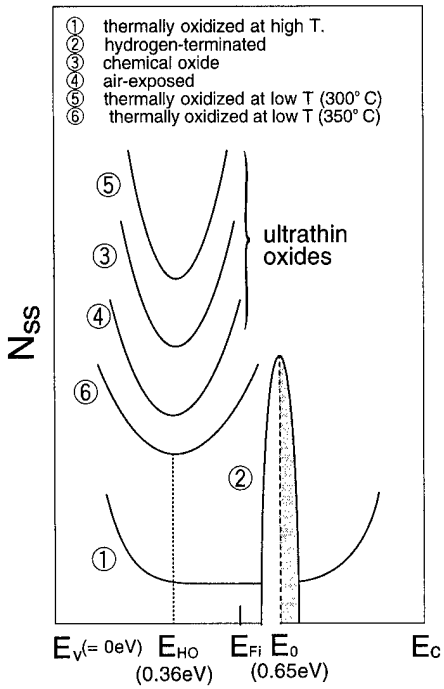


FIG. 12. Schematic illustration of N_{SS} distributions estimated from the results of XPS, $C-V$, and PLS[3] measurements for the various Si surfaces.

with bias is, and, therefore, the higher the interface state density is. Thus, it can be concluded that the interface state density increases in the aforementioned order.

C. PLS[3] measurements

Figure 11 summarizes the result of PLS[3] measurements where the measured values of the PL efficiency are plotted against excitation intensity for various surfaces investigated here. It should be noted that efficiency values are normalized by the PL intensity from the reference metal–oxide–semiconductor (MOS) sample and given in arbitrary units, since absolute determination of the efficiency was extremely difficult. It is seen in Fig. 11 that the PL efficiency takes very different values for different samples. More precisely, the PL intensity decreases in the order of the high-temperature thermally oxidized surface, the air-exposed surface, the chemically oxidized surfaces, and the low-temperature thermally oxidized surface at 300 °C. One would naturally expected that, the higher the PL efficiency, the lower the interface state density. In fact, this order in the PL efficiency values corresponds qualitatively well to the $C-V$ results mentioned in the previous section concerning the flattening levels of the $C-V$ curves under positive bias. Thus, PLS[3] and $C-V$ results are consistent with each other and indicate that interface state density increases in the aforementioned order.

As mentioned previously, the shape of the PL efficiency versus ϕ curve provides a very reliable information concerning the shape of the distribution of the surface states causing surface recombination. On the hydrogen-terminated surface, the slope of the PL efficiency was exactly unity, indicating that discrete surface state exists within the forbidden gap. This is again consistent with the interpretation of the $C-V$ curves shown in Fig. 7. For all the oxide-covered surfaces, on the other hand, the slope of the curve was less than unity and it continuously decreases with the increase of the photon flux density. This indicates that the N_{SS} distribution is U shaped for these surfaces.

The PLS[3] method can determine not only the shape of N_{SS} distribution but also magnitude by fitting the data with the computer simulation; but, this requires knowledge of electron and hole capture cross sections of interface states as well as the absolute values of the PL efficiency. Due to lack of knowledge of these, together with the availability of a limited amount of data taken at the high excitation range due to weak PL intensities, quantitative determination of the N_{SS} distributions by this method was not attempted in this article.

D. Models of N_{SS} distribution for hydrogen terminated and ultrathin-oxide-covered Si(111) surfaces

The models of N_{SS} distribution deduced combining all the results of XPS, $C-V$, and PLS[3] measurements are shown in Fig. 12 for various Si surfaces studied here.

Starting from the H-terminated surface, the unity slope of PL efficiency indicates that there exists a high density of discrete states which cause Fermi-level pinning. $C-V$ results indicate the amphoteric nature of these states. The pinning position of the H-terminated surface determined by the XPS analysis shown in Fig. 5 is in fairly good agreement with that estimated from the flattening value of the capacitance in $C-V$ curves. It may be said that it lies at a position of 0.65 $\pm$0.05 eV.

As for the concentration of the amphoteric discrete states, it was impossible to determine its value due to danger of

dielectric breakdown in $C-V$ measurement. From the applied voltage sweep corresponding to the flattening of capacitance, it should be larger than 1×10^{12} cm^{-2}.

This value is much larger than the number of recombination centers of 10^7–10^9 cm^{-2} obtained by Yablonovitch *et al.*[30] for Si surfaces immersed in HF acid using the photodecay method. This large difference seems to be due to the fact that the H-terminated surface in the present study is exposed directly to dry N_2, whereas the sample studied by Yablonovitch *et al.* was held in equilibrium with hydrogen-containing liquid. In fact, a recent SPV study[14] of a H-terminated surface in liquid has shown that the surface state density is very low in such a liquid, unlike the present result.

As for the structural origin of the discrete states, it is highly unlikely that they are due to H termination itself, because a theoretical calculation[31,32] and recent STM/scanning tunneling spectroscopy (STS) observations[3,33] indicate that no gap states are produced on perfectly H-terminated surfaces. We have recently proposed[18] that these states are due to Si dangling bonds on the H-terminated surface. Although STM study has demonstrated perfectly H-terminated surfaces in nanometer field sizes, it is highly likely that hydrogen atoms will leave from the surface on the real surface of wide extension, leaving dangling bonds, which possess amphoteric properties.

Since no theoretical calculation has so far made to our knowledge on the validity of such a structural model, a future theoretical work is needed.

As for the surfaces covered with various low-temperature ultrathin oxides, their $C-V$ and PLS[3] behaviors were rather similar to each other and were completely different from that of the initial H-terminated surface. As shown in the models in Fig. 12, N_{SS} distributions of these surfaces are all U shaped, having a minimum at 0.35 eV above the valence-band maximum, although the magnitude of N_{SS} itself showed considerable process dependence.

According to the disorder-induced gap state (DIGS) model,[28] disorder of bonds near the surface produce a donor- and acceptorlike state continuum within the energy gap whose boundary is given by the hybrid orbital charge neutrality level E_{HO}. Thus, when the state density is high, the Fermi level pinned at E_{HO} which lies at about $E_V+0.32$ eV for Si. This seems to be more or less the situation which took place on all the ultrathin-oxide-covered surfaces studied here, and may be due to the presence of nonstoichiometric and nonuniform oxides. The present work indicates rather unexpected difficulty in obtaining pinning-free ultrathin-oxide–Si interfaces by low-temperature processing. A further work is necessary to establish a process for growth of pinning-free ultrathin oxides. In this connection, the result on the low-temperature thermal oxidation at 350 °C gives hope for future successful optimization.

IV. CONCLUSION

Electronic properties of ultrathin-oxide-formed Si(111) surfaces were characterized in contactless fashion, using contactless capacitance–voltage ($C-V$) and photoluminescence surface state spectroscopy (PLS[3]) techniques together with x-ray photoelectron spectroscopy (XPS). Hydrogen (H) terminated Si surfaces were used as the initial surface prior to ultrathin oxidation. Ultrathin oxides were formed by chemical oxidation (hot HNO_3, $H_2SO_4+H_2O_2$), long-time air exposure, and low-temperature oxidation (below 350 °C). The main conclusions are listed below.

(1) The initial H-terminated surfaces showed the presence of Fermi-level pinning at $E_0=E_V+0.65$ eV due to a high density of amphoteric discrete state, probably originating from Si dangling bonds.

(2) All the ultrathin-oxide-covered surfaces exhibited very limited capacitance variation with bias similarity to GaAs capacitors. The results indicate that the Fermi level is pinned near the hybrid orbital charge neutrality level E_{HO} due to presence of a high-density interface states with narrow U-shaped continuous distributions.

(3) Low-temperature thermal oxidation at 350 °C slightly weakens such pinning.

The present work indicates difficulty of realizing unpinned ultrathin-oxide–silicon interfaces by the low-temperature processes studied here.

ACKNOWLEDGMENTS

The authors express their sincere thanks to T. Sakai, Dainippon Screen Mfg. Co., Ltd., for his useful discussion, and to T. Yoshida for his technical assistance.

[1]A. Fujiwara, Y. Takahashi, K. Murase, and M. Tabe, Appl. Phys. Lett. **67**, 2957 (1995).
[2]T. Ohmi, M. Morita, A. Teramoto, K. Makihara, and K. S. Tseng, Appl. Phys. Lett. **60**, 2126 (1992).
[3]G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari, Appl. Phys. Lett. **56**, 656 (1990).
[4]P. Jakob and Y. J. Chabal, J. Chem. Phys. **95**, 2897 (1991).
[5]T. Hattori, J. Vac. Sci. Technol. B **11**, 1528 (1993).
[6]Th. Dittrich, H. Angermann, H. Flietner, Th. Bitzer, and H. J. Lewerentz, J. Electrochem. Soc. **141**, 3595 (1994).
[7]K. Ohnishi and T. Hattori, Jpn. J. Appl. Phys. **33**, L675 (1994).
[8]R. S. Becker, G. S. Higashi, Y. J. Chabal, and A. J. Becker, Phys. Rev. Lett. **65**, 1917 (1990).
[9]Y. Morita, K. Miki, and H. Tokumoto, Appl. Phys. Lett. **59**, 2187 (1991).
[10]A. Ourmazd and P. H. Fuoss, Appl. Surf. Sci. **41/42**, 365 (1989).
[11]N. Kashiwagura, Y. Kashihara, M. Sakata, J. Harada, S. W. Wilkins, and A. W. Stevenson, Jpn. J. Appl. Phys. **26**, L2026 (1987).
[12]T. Takahagi, I. Nagai, A. Ishitani, H. Kuroda, and Y. Nagasawa, J. Appl. Phys. **64**, 3516 (1988).
[13]K. Heiling, H. Flietner, and J. Reineke, J. Phys. D **12**, 927 (1979).
[14]S. Rauscher, Th. Dittrich, M. Aggour, J. Rappich, H. Flietner, and H. J. Lewerenz, Appl. Phys. Lett. **66**, 3018 (1995).
[15]T. Saitoh, H. Iwadate, and H. Hasegawa, Jpn. J. Appl. Phys. **30**, 3750 (1991).
[16]T. Saitoh and H. Hasegawa, Appl. Surf. Sci. **56–58**, 94 (1992).
[17]T. Sakai, M. Kohno, S. Hirae, I. Nakatani, and T. Kusuda, Jpn. J. Appl. Phys. **32**, 4005 (1993).
[18]S. Koyanagi, T. Hashizume, and H. Hasegawa, Jpn. J. Appl. Phys. **35**, 630 (1996).
[19]H. Hasegawa and T. Sawada, Thin Solid Films **103**, 119 (1983).
[20]H. Ogawa and T. Hattori, Appl. Phys. Lett. **61**, 577 (1992).
[21]F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B **38**, 6084 (1988).
[22]S. Watanabe, M. Shigeno, N. Nakayama, and T. Ito, Jpn. J. Appl. Phys. **30**, 3575 (1991).

[23]C. H. Bjorkman, J. L. Alay, H. Nishimura, M. Fukuda, T. Yamazaki, and M. Hirose, Appl. Phys. Lett. **67**, 2049 (1995).
[24]G. P. Schwartz, M. S. Hybertsen, J. Bevk, R. G. Nuzzo, J. P. Mannaert, and G. J. Gualtieri, Phys. Rev. B **39**, 1235 (1989).
[25]F. J. Himpsel, P. Heimann, T.-C. Chiang, and D. E. Eastman, Phys. Rev. Lett. **45**, 1112 (1980).
[26]F. J. Himpsel, G. Hollinger, and R. A. Pollak, Phys. Rev. B **28**, 7014 (1983).
[27]J. D. van Otterloo, Surf. Sci. **104**, L205 (1981).
[28]H. Hasegawa and H. Ohno, J. Vac. Sci. Technol. B **4**, 1130 (1986).
[29]J. Tersoff, Phys. Rev. B **30**, 4874 (1984).
[30]E. Yablonovitch, D. L. Allara, C. C. Chang, T. Gmitter, and T. B. Bright, Phys. Rev. Lett. **57**, 249 (1986).
[31]M. Schlütter and M. Cohen, Phys. Rev. B **17**, 716 (1978).
[32]K. C. Pandy, IBM J. Res. Dev. **22**, 250 (1978).
[33]Y. Morita, K. Miki, and H. Tokumoto, Jpn. J. Appl. Phys. **30**, 3570 (1991).

Rapid thermal N_2O oxynitride on Si(100)

Z. H. Lu,[a] R. J. Hussey, and M. J. Graham
Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

R. Cao
Stanford Synchrotron Radiation Center, Stanford, California 94309-0210

S. P. Tay[b]
Telecom Microelectronics Centre, Northern Telecom Limited, Nepean, Ontario K2H 8V4, Canada

(Received 16 February 1996; accepted 13 April 1996)

High-resolution x-ray photoelectron spectroscopy (XPS) in conjunction with secondary-ion-mass spectrometry was used to study the chemical nature and distribution of N in oxynitride films formed by rapid thermal N_2O processes (RTPs) or conventional furnace methods. The kinetics of furnace oxide growth in N_2O are slower than that in O_2. During reoxidation the oxidation rate increased to that in pure O_2 and the N in the SiO_2–Si interface region is displaced into the bulk of the oxide. High-resolution synchrotron Si $2p$ core-level photoemission spectroscopy (PES) was used to study the oxide–Si(100) interface suboxide structures produced by RTP with and without the presence of N. XPS N $1s$ studies indicated that there are two types of N in the RTP oxynitride films. The chemical bond configuration of the first type of N is similar to that of N in Si_3N_4 and is mainly distributed within the first 1 nm from the interface. The second type of N is distributed mainly outside of the first 1 nm region, and the N is likely bonded to two Si and one oxygen atom. PES studies showed that Si formed suboxides with oxygen at the interface for all oxynitride films. It is found that there is no change in the Si^{+1} structure while there is a dramatic decrease in the Si^{+2} and Si^{+3} states with the inclusion of N in the oxide. Both the XPS and PES results are explained in terms of a strain reduction as N is incorporated in the film near the interface region, where Si_3N_4 functions as a buffer layer which reduces the stress caused by the large Si lattice mismatch between the bulk Si and the oxide overlayer. About 1/5 of the Si^{+2} and 1/3 of Si^{+3} atoms at the SiO_2–Si interface have been replaced by the Si_3N_4 buffer layer at the oxynitride–Si interface. © *1996 American Vacuum Society.*

I. INTRODUCTION

The gate-dielectric–Si(100) interface has been one of the prime issues in the development of deep submicron Si-based ultralarge-scale integrated circuit (ULSI) technology. Because of their superior electrical properties, nitrided oxides or oxynitrides are currently under intensive study for use as gate dielectrics.[1–9] Most reported work has focused on the growth method and the electrical properties of these thin films. There have been some articles dealing with the physical and chemical structure of the oxynitride films. For example, it has been reported[1,5–8] that the total amount of N is very small (a few monolayers). The N accumulates at the SiO_2–Si interface and leads to improved electrical properties;[10] however, questions remain regarding the effect of N on the oxide–Si interface structure. The introduction of N into the films using NH_3 was hitherto the usual method, but subsequent reoxidation was required to reduce electron trapping caused by hydrogen incorporation.[11] The alternative use of N_2O to introduce N into the SiO_2 film avoids this problem and constitutes a single-step process by direct oxidation in N_2O without further reoxidation.[12–14] There has been a great deal of interest in the oxidation of Si in N_2O over the last few years because of these factors and studies include examination both of the N distribution in and electrical properties of the SiO_2 oxide films.[10,12,15–17] In these studies very thin oxide films have been grown in N_2O ambients both in conventional furnaces and by rapid thermal oxidation (RTO) methods.[10,12,15–19] Oxide growth rates were determined for temperatures in the range 950–1200 °C under a variety of conditions. N accumulation at or near the SiO_2–Si interface has been determined by Auger electron spectroscopy (AES)[10,12,15] and secondary-ion-mass spectrometry (SIMS).[10,13,14,16,18–20] The nature of the N-containing species at or near the SiO_2–Si interface formed during oxidation in N_2O has been examined by x-ray photoelectron spectroscopy (XPS).[15,21,22] The aim of this article is to study the chemical nature and depth distribution of N in oxynitride films and the effects of N on the interface structure using XPS, SIMS, and photoemission spectroscopy (PES). The surface-analytical data are complemented by kinetic studies in O_2 and N_2O.

II. EXPERIMENT

Two types of oxide and oxynitride films were grown. The first is grown in low-pressure O_2 and N_2O in a quartz tube heated in a conventional furnace. The experimental proce-

[a]Electronic mail: lu@sg1.chem.nrc.ca
[b]Present address: AG Associates, 4425 Fortran Drive, San Jose, CA 95134.

dures used as well as details of the determination of oxide thicknesses and N content in the films, and the SIMS analysis are described more fully elsewhere.[17] This type of oxide was used to study oxide growth kinetics. The second type of oxide and oxynitride is grown at Northern Telecom's standard 0.5 μm complementary metal–oxide–semiconductor (CMOS) process facilities. The high quality of these films was confirmed by electrical measurements on both n- and p-channel MOS field-effect transistors (MOSFETs) fabricated on these oxynitride films. In order to focus on the effect of N on the interface structure, we concentrate on oxynitride films processed at 1050 °C. This avoids the complications of other effects such as growth temperature on the interface suboxide distributions. In a previous study[23] we have established, in the case of furnace-grown oxide, that high-temperature treatment leads to an increase of Si suboxides. Here we focus our study on several representative growth conditions used in fabrication of CMOS circuits:

(a) rapid thermal oxidation using pure oxygen (RTO) on a bare Si surface;
(b) rapid thermal oxidation using N_2O gas (RTN) on a bare Si surface;
(c) RTN treatment of the RTO oxide (RTN/RTO); and
(d) rapid thermal reoxidation of RTN/RTO (ReOx RTN/RTO) films.

III. RESULTS AND DISCUSSION

A. Growth kinetics

The growth of the film formed during the oxidation of Si(100) in N_2O at 1050 °C and 10 Torr pressure can be best expressed in terms of a direct logarithmic equation of the form

$$x = k_{\log} \log(t + t_0) + A, \tag{1}$$

where x is the oxide thickness, $k_{\log}$ is the logarithmic rate constant, t is the time, t_0 is the time at $x=0$, and A is a constant. This is in contrast to film growth in pure O_2 at 1050 °C, where a parabolic rate relation was found to best fit the kinetic data.[24] This is shown in Fig. 1(a), where kinetic data are compared for the oxidation of Si(100) in O_2 and N_2O at 10 Torr pressure and 1050 °C. The extent of oxidation is much less in N_2O than in O_2 at any given time. These data support the view that the formation of a N-rich interfacial layer retards oxidant diffusion to the substrate surface; the more N at the interface, the less is the extent of oxidation.[17] Lange *et al.*[16] and Ting *et al.*[15,25] also found that the oxide thicknesses formed during RTO are much less in N_2O than in O_2. They also found that the oxide growth rate became very slow after the film thicknesses reached 2–6 nm dependent on the temperature. They also support the view of a N-rich interfacial layer retarding oxidant diffusion to the substrate. Several reoxidation experiments were also carried out, in which samples were first oxidized in N_2O for 1 h and then oxidized in $^{18}O_2$ (nominally 97.0%–99.0% ^{18}O and containing a few ppm H_2O) for periods from 1 to 16 h at 1050 °C. The kinetics of reoxidation are compared with

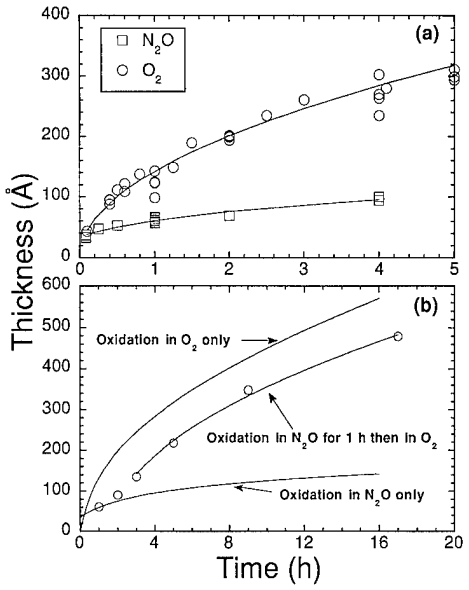


FIG. 1. (a) A comparison of the oxidation kinetics for Si(100) at 1050 °C in N_2O and in pure O_2 at 10 Torr pressure. The points on the solid lines are for individual experiments, the lines being least-squares fits to a direct logarithmic equation for N_2O and to a parabolic equation for pure O_2, respectively. (b) Kinetics for Si(100) previously oxidized in N_2O for 1 h followed by reoxidation in $^{18}O_2$ for various times, both oxidations being at 1050 °C and 10 Torr pressure. The points on the solid line (a parabolic least-squares fit to the data) are for individual experiments. Kinetic curves for oxidation in O_2 and N_2O are shown for comparison.

those for oxidation in N_2O and in O_2 only. The data for N_2O have been extrapolated to longer times. As shown in Fig. 1(b) during subsequent oxidation in $^{18}O_2$ the oxidation rate increased gradually becoming similar to that in pure O_2 after about 2 h. The longer term data points from separate runs are fitted to a parabolic rate equation of the form

$$x^2 = k_p t + B, \tag{2}$$

where k_p is the parabolic rate constant and B is a constant. SIMS analysis showed that during reoxidation in $^{18}O_2$ the amount of N in the region near the SiO_2–Si interface becomes reduced, the N being displaced into the bulk of the oxide. The displaced N-enriched zone moves progressively outward with increasing time; the amount originally at the SiO_2–Si interface is reduced significantly within 2 h of reoxidation. After 16 h the amount displaced into the bulk of the film relative to the total oxide becomes very small and the N is now concentrated near the gas–SiO_2 interface. The total percentage of N in the reoxidized film relative to that in the oxide formed in N_2O was found to be reduced to about half its original value after 2 h of reoxidation and this value remains roughly constant out to 16 h of reoxidation. This indicates that some N is lost to the ambient initially, but the remainder is stable within the oxide thereafter. The replacement of N by O in the SiO_2 network occurs because of the diffusion of O to the SiO_2–Si interface, where new oxide forms. Some of the N in the very thin Si_3N_4 and oxynitride layers adjacent to the SiO_2–Si interface (see below) is replaced by O and these oxynitride layers are displaced out-

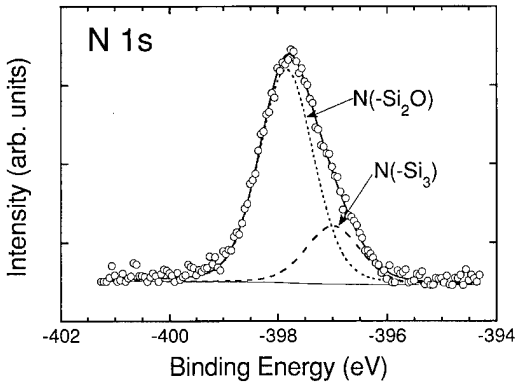


FIG. 2. N 1*s* XPS spectrum recorded on a RTN etchback sample. The open circles are as-recorded data. The dotted and dashed lines are curve-fitted peaks with the chemical state as labeled. The solid line is the sum of these two peaks.

ward by the new oxide formation at this interface. This displacement process has also been observed by Saks *et al.*,[11] Okada *et al.*,[20] and Green *et al.*[7] These SIMS results may be correlated with the kinetic data, which show that after about 2 h the reoxidation kinetics are similar to that for oxidation with no prior oxynitride present [see Fig. 1(b)]. This is the point at which the amount of total N in the oxide film becomes constant. Evidently this level of N in the oxide film is now too low to cause any significant effect on the inward diffusion of oxygen and oxidation essentially proceeds at the same rate as in pure O_2.

B. Chemical structure and distribution of N

X-ray photoelectron spectroscopy (XPS) was used to determine the nature and distribution of N in the film. The measurements were carried out on a PHI 5500 system which is equipped with a monochromated Al $K\alpha$ source and a hemispherical electron analyzer. Thin oxide films of various thicknesses were obtained by etchback in a dilute (1%) HF solution. No thickness variation was found on the etchback samples. Si 2*p* core levels were recorded and were used to calculate the film thickness using the method described in Ref. 26 for an electron free path of 3.5 nm. In order to increase the depth resolution, the measurements were made at a 30° take-off angle. Figure 2 shows a high-resolution N 1*s* spectrum taken on an etchback RTN sample. The spectrum was taken with a 5.8 eV pass energy. It normally takes several hours to obtain such a spectrum on an oxynitride sample which has a very dilute N concentration (as low as a few at. %). As can be seen in the figure, there are two components in the spectrum. Curve-fitting analysis indicated one peak located at 396.97 eV and another with a chemical shift of 0.85 eV at 397.82 eV. Measurements on Si_3N_4 films yield a N 1*s* peak at 397.0 eV. This suggests that the chemical environment of the N at 396.97 eV in the oxynitride is similar to that in Si_3N_4, i.e., one N bonded to three Si atoms and is referred to as N(-Si_3) in the figure and the following text. Based on its binding energy, the nature of N at 397.82 eV is tentatively assigned to N bonded to two Si and one O, which

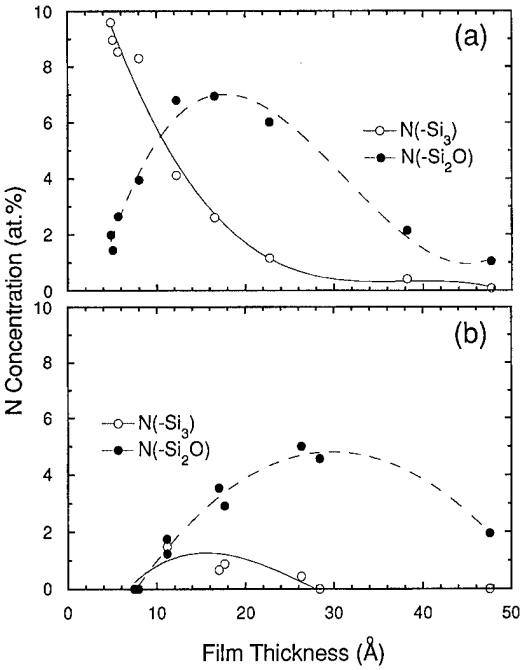


FIG. 3. N concentration profiles as a function of oxynitride film thickness for (a) the RTN sample and (b) the reoxidized RTN/RTO sample.

is more electronegative than N and thus causes the N 1*s* shift to the higher binding energy with the formation of a N—O bond. This N species is referred to as N(-Si_2O) in the figure and following text. From synchrotron Si 2*p* core-level measurements to be presented, it was found that the oxynitride–Si interface is formed by suboxides. This means that the N(-Si_3) second-nearest-neighbor atoms are either N or O. Theoretical calculations[27] on the SiO_2–Si system found that the core-level shift is determined by first-nearest-neighbor atoms and contributions from the second-nearest-neighbor atoms are negligible.

The concentration of N in the oxynitride can be calculated from the integrated peak intensities of N 1*s*, Si^{+4} (i.e., SiO_2) 2*p*, and O 1*s*. Their relative sensitivity factors were determined from thermally produced SiO_2 and Si_3N_4 samples. We have found that there are some residual contaminants such as hydrocarbon and water present on the etchback samples. In order to eliminate this artificial fluctuation in the N profile, we calculate the N concentration using only the N 1*s* and the Si^{+4} 2*p* peaks. Figure 3(a) shows N distributions obtained from the RTN. The data show that the N(-Si_3) is distributed mainly within the first 1 nm region of the interface while the N(-Si_2O) is distributed mainly about 2 nm away from the interface on as-grown oxynitride films. There is no apparent difference in the N distribution between the RTN and the RTN/RTO films. A possible explanation for the observed N distribution is related to the Si atomic density mismatch. It is known that there is a large mismatch in the Si

atomic density across the oxide–Si interface, from 2.2×10^{22} atoms cm^{-3} in the SiO_2 to 5×10^{22} atoms cm^{-3} in bulk Si. The transition of this interface is realized through the formation of an extra sacrificial buffer layer of suboxide.[23] Details of this suboxide structure are presented in the following text. It is also known that the Si density in Si_3N_4 is 4×10^{22} atoms cm^{-3}, between that of bulk Si and SiO_2. The inclusion of Si_3N_4 near the oxide–Si interface therefore creates a buffer layer which reduces the mismatch-induced strain at the oxide–Si(100) interface.

Figure 3(b) shows the N profile on the reoxidized oxynitride film (ReOx RTN/RTO film). The reoxidation was done at 1050 °C for 1 min which produces a nominal 8 nm oxide on a bare surface. Here the high diffusion barrier of the oxynitride film restricts the oxide growth to about 1 nm. The new oxide film, as shown in the figure, is formed between the Si substrate and the "old" oxynitride. The reoxidation process is also shown to reduce the amount of Si_3N_4, which is likely converted into N(-Si_2O). These observations appear to contradict our above conclusion based on a strain energy minimization principle that a dense Si nitride layer between the Si substrate and oxide film is favored. This paradox may be explained by the growth kinetics. New oxide is formed on the substrate surface by continuous flow of O atoms; there is no mobile N associated species readily available. The N in the old film is chemically bonded to the oxynitride network and a large activation energy is required to break the network into mobile species. As the new oxide grows and the Si_3N_4 species move away from the interface region, the strain on the N(-Si_3) species is increasing due to its moving into a less dense SiO_2 network. The strained N(-Si_3) is then gradually converted into a less strained N(-Si_2O) species, as observed.

C. The effect of N inclusion on interface suboxide structure

Considering the interface structure, it is well established that the SiO_2–Si interface is formed through a transition region made up of Si suboxides denoted as Si^{+1}, Si^{+2}, and Si^{+3}, where the Si atoms have one, two, and three first-nearest-neighbor oxygen atoms, respectively. Synchrotron photoelectron spectroscopy (PES) is probably the only technique that can identify unambiguously these suboxides on thin etchback oxides.[28,29] We have shown recently[23,30] that PES can be used to effectively monitor SiO_2–Si(100) interfacial changes under various industrial ULSI process conditions. Previous PES measurements on Si $2p$ core levels have found that the chemical shifts of these Si^{+x} peaks are 0.97, 1.8, and 2.6 eV for $x=1$, 2, and 3, respectively.[28,29] A recent theoretical first-principle calculation by Pasquarello and co-workers[27] found excellent agreement with these chemical shifts.

The synchrotron PES measurements were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL) using a 6 m toroidal-grading monochromator (TGM) beam line (experimental station 8-1). The photoelectrons were collected by an angle-integrated PHI double-pass cylindrical-mirror analyzer in an UHV chamber with a load-lock chamber. The photoelectron take-off angle was 45°. Si $2p$ core shell electrons were excited by 130 eV photons. Figure 4 shows a Si $2p_{3/2}$ spectrum recorded from the RTO sample with an etchback SiO_2 thickness of 1 nm. The data were analyzed using the same curve-fitting method as previously used for furnace and UHV-grown SiO_2–Si interfaces.[23,29] In Fig. 4 the dashed lines indicate various Si peaks with the oxidation states as labeled. As can be seen in the figure, the SiO_2 peak is very strong while the substrate and suboxide Si peaks are very weak, which requires several hours of data acquisition time in order to obtain usable data in the suboxide region.

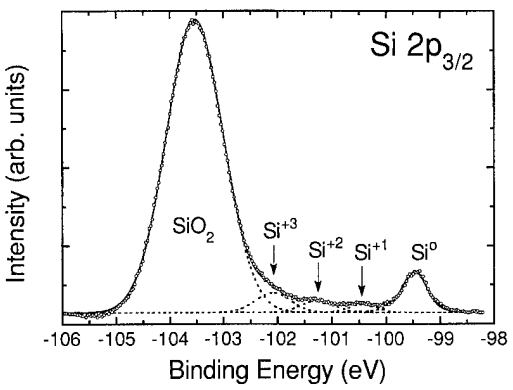


FIG. 4. Si $2p_{3/2}$ PES spectrum recorded on the RTO sample using 130 eV photons. The open circles are raw data with the $2p_{1/2}$ and the Shirley background subtracted. The dotted lines represent various Si species as labeled. The solid line the sum of all of these Si peaks.

Since the relative interfacial suboxide distribution is the main focus of this study, we have subtracted the SiO_2 contribution from Fig. 4 and the results are shown as Fig. 5(a) where the suboxides distribution is now much clearer. Figure 5(b) shows curve-fitting analysis on the RTN sample. The results show that the Si^{+1}, Si^{+2}, and Si^{+3} peaks have the same chemical shift as those found for the RTO sample. This indicates that the interface suboxide chemical bonds are Si—O bonds. The relative intensities (to that of substrate Si) of these suboxides, however, is quite different. The suboxide intensities from the RTN oxynitride–Si(100) is significantly reduced compared with those of the RTO SiO_2–Si(100) interface. Based on their relative intensities and using the photoionization cross section and mean free path given by Himpsel *et al.*,[28] the amount of various suboxides can be calculated and the results from all three samples are listed in Table I. The table shows that the amount of Si^{+1} is the same for all samples. This may not be surprising as Si^{+1} is located right at the substrate surface[28] and therefore is less sensitive to structural changes in the oxide film. The amounts of both Si^{+2} and Si^{+3}, however, decrease significantly by about 1/5 for Si^{+2} and 1/3 for Si^{+3} at the oxynitride–Si(100) interface. This can again be explained by a reduction in the mismatch-induced strain with the inclusion of N near the interface. In previous studies, we established[23,30] that a high number of Si^{+2} and Si^{+3} is produced at high temperature. This was attributed to a structural relaxation to minimize SiO_2–Si in-

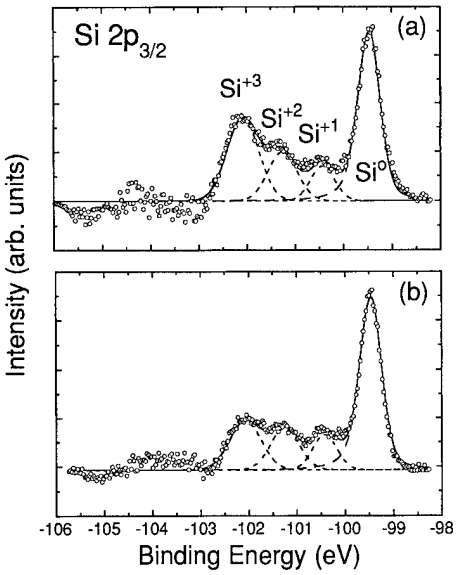


FIG. 5. Si $2p_{3/2}$ PES spectra recorded from (a) the RTO sample and (b) the RTN sample. The $2p_{1/2}$, the Shirley background, and the SiO_2 contributions have been subtracted. The open circles are the raw data. The dotted lines are contributions from the substrate silicon Si^0 and the various interface suboxides Si^{+x}, as labeled. The solid line is the sum of all of these Si peaks.

terfacial strain through the production of extra sacrificial layers of Si^{+2} and Si^{+3} states as a Si density buffer. Here, in the case of the oxynitride–Si interface, the above XPS data showed that N(-Si_3) is distributed mainly near the interface. The nitrided region with high Si density also functions as a buffer layer to help reduce the strains caused by the density mismatch at the interface. This means that a relatively small amount of sacrificial Si^{+2} and Si^{+3} is required at the interface to accommodate the interface lattice mismatch. Upon reoxidation, the Si^{+3} is shown to increase while Si^{+1} and Si^{+2} remain unchanged. This can be explained by the above observation of N moving away from the interface region. An atomic density buffer layer on the reoxidized film therefore has to be assumed again by suboxides. However, the amount of suboxide is still less than that of pure oxide. This may be explained by the fact the oxynitride–Si interface is less strained even without Si_3N_4 as a buffer layer compared with that of the SiO_2–Si interface. It is also possible that there may be a small amount of N (below the XPS sensitivity ~0.1 at. %) existing in the suboxide region. The dramatic effect of small amount of impurity on the Si lattice structure is well known.[31]

TABLE I. The amount of suboxide N_{SiO_x}, at the dielectric–Si(100) interface formed by various rapid thermal processes (RTPs) at 1050 °C.

	N_{SiO_x} (10^{14} atoms cm^{-2})		
RTP process	Si^{+1}	Si^{+2}	Si^{+3}
RTO	2.22	3.49	4.16
RTN	2.25	2.98	2.43
RTN/RTO	2.24	2.90	2.91
ReOx RTN/RTO	2.31	2.70	3.61

IV. CONCLUSIONS

A. Growth kinetics

The oxidation of *p*-type Si(100) in N_2O at 10 Torr pressure at 1050 °C produced oxide films significantly thinner than those which would have been produced during oxidation in O_2 under the same conditions. During oxidation in N_2O a N-rich layer is produced near to the SiO_2–Si interface, which inhibited the influx of oxidant probably because of slower oxygen diffusion through this layer. The conformity of the oxidation kinetics to a direct logarithmic rate relation lends support to the notion of a reduction in rate caused by a second phase at the reaction zone. Reoxidation caused the N in the SiO_2–Si interface region to be displaced into the bulk of the oxide and a reduction in the total N content of the film to a constant 50% of its original value. The oxidation rate on reoxidation increased to that for oxidation in pure O_2 and the time at which this occurred corresponded to that where the total N content of the film became constant.

B. N structure and distribution

XPS studies found two types of N present in the N_2O oxynitride films: one with a N 1*s* binding energy at 396.97 eV attributed to N in the form of Si_3N_4, and another with a chemical shift of 0.85 eV at 397.82 eV attributed to N situated in a more electronegative environment such as N bonded to two Si and one O atom. XPS etch profiling found that Si_3N_4 is located within the first 1 nm from the interface and N(-Si_2O) is mainly distributed adjacent to the Si_3N_4 layer. Both types of N are pushed away from the interface upon reoxidation. The amount of Si_3N_4 species is also found to decrease significantly after reoxidation. The observed data are explained by the strain-energy minimization principle.

C. Interface structure

Synchrotron Si 2*p* core-level studies found that the topological link between oxynitride film and the Si surface is made by various Si suboxides, similar to that for the SiO_2–Si interface. The amounts of the various suboxides, however, are found to differ dramatically. The data showed that the amount of Si^{+2} and Si^{+3} states at the oxynitride–Si interface is about 30% less than that at the SiO_2–Si interface, while the amount of Si^{+1} states remains the same. The current results, together with previous data[23,30] on the effect of temperature where suboxides are found to vary significantly with growth and anneal temperature, are explained by a new interface model where the functionalities of suboxides are twofold. The first function of the suboxide is to provide a topological link between the crystalline Si substrate and the amorphous SiO_2 network. The current theoretical model is focused only on this type of suboxide structure.[32] The second function of the suboxides, mainly Si^{+2} and Si^{+3}, is to act as "buffer" layers to minimize strains caused by the large (about 50%) Si density gradient across the interface. In the

case of the oxynitride–Si interface, the role of the "buffer" suboxides has been replaced by Si_3N_4. Since the Si atomic density in the nitride is between that in bulk Si and that in SiO_2, the interface is less strained. The less strained oxynitride–Si interface may be correlated with the observations[1] of a lower amount of interface states generated under current stress and radiation, compared with conventional thermal SiO_2–Si interfaces. The results also indicate that oxynitride films are very attractive in terms of a reduced interface transition region for future 0.1 μm or less ULSI manufacturing which uses 3.4 nm or less gate dielectrics.[33]

ACKNOWLEDGMENT

The Synchrotron measurements were carried out at SSRL, which is funded and operated by the U.S. DOE.

[1]H. Hwang, W. Ting, B. Maiti, D. L. Kwong, and J. Lee, Appl. Phys. Lett. **57**, 1010 (1990).
[2]Y. Okada, P. J. Tobin, R. I. Hegde, J. Liao, and P. Rushbrook, Appl. Phys. Lett. **61**, 3163 (1992).
[3]Y. Okada, P. J. Tobin, V. Lakhotia, W. A. Feil, S. A. Ajuria, and R. I. Hegde, Appl. Phys. Lett. **63**, 194 (1993).
[4]E. C. Carr and R. A. Buhrman, Appl. Phys. Lett. **63**, 54 (1993).
[5]M. Bhat, J. Ahn, D. L. Kwong, M. Arendt, and J. M. White, Appl. Phys. Lett. **64**, 1168 (1994).
[6]H. T. Tang, W. N. Lennard, M. Zink-Allmang, I. V. Mitchell, L. C. Feldman, M. L. Green, and D. Brasen, Appl. Phys. Lett. **64**, 3473 (1994).
[7]M. L. Green, D. Brasen, K. W. Evans-Lutterodt, L. C. Feldman, K. Krisch, W. Lennard, H. T. Tang, L. Manchanda, and M. T. Tang, Appl. Phys. Lett. **65**, 848 (1994).
[8]E. C. Carr, K. A. Ellis, and R. A. Buhrman, Appl. Phys. Lett. **66**, 1492 (1995).
[9]R. I. Hegde, P. J. Tobin, K. G. Reid, B. Maiti, and S. A. Ajuria, Appl. Phys. Lett. **66**, 2882 (1995).
[10]T. Hori, H. Iwasaki and K. Tsuji, IEEE Trans. Electron Devices **ED-36**, 340 (1989).
[11]N. S. Saks, D. I. Ma, D. M. Fleetwood, and M. E. Twigg, in *Symposium on Silicon Nitride and Silicon Oxide Thin Insulating Films*, edited by V. J. Kapoor and W. D. Brown (The Electrochemical Society, New York, 1994), Vol. 94-16, p. 395.
[12]J. Ahn, W. Ting, T. Chu, S. N. Lin, and D. L. Kwong, J. Electrochem. Soc. **138**, L39 (1991).
[13]P. J. Tobin, Y. Okada, S. A. Ajuria, V. Lakhotia, W. A. Feil, and R. I. Hegde, J. Appl. Phys. **75**, 1811 (1994).
[14]Y. Okada, P. J. Tobin, V. Likhotia, W. A. Feil, and R. I. Hegde, Appl. Phys. Lett. **63**, 194 (1993).
[15]W. Ting, H. Hwang, J. Lee, and D. L. Kwong, Appl. Phys. Lett. **57**, 2808 (1990).
[16]P. Lange, H. Bernt, E. Hartmannsgruber, and F. Naumann, J. Electrochem. Soc. **141**, 259 (1994).
[17]R. J. Hussey, T. L. Hoffman, Y. Tao, and M. J. Graham, J. Electrochem. Soc. **143**, 221 (1996).
[18]D. Krüger, R. Kurps, and G. Weidner, Semiconductor Sci. Technol. **8**, 1706 (1993).
[19]G. Weidner and D. Krüger, Appl. Phys. Lett. **62**, 294 (1993).
[20]Y. Okada, P. J. Tobin, V. Likhotia, W. A. Feil, R. I. Hegde, J. L. Liao, P. P. Rushbrook, and L. J. Arias, Jr., J. Electrochem. Soc. **140**, L87 (1993).
[21]M. Bhat, J. Ahn, D. L. Kwong, M. Arendt, and J. M. White, Appl. Phys. Lett. **64**, 1168 (1994).
[22]Z. H. Lu, S. P. Tay, R. Cao, and P. Pianetta, Appl. Phys. Lett. **67**, 2836 (1995).
[23]Z. H. Lu, S. P. Tay, T. Miller, and T. C. Chiang, J. Appl. Phys. **77**, 4110 (1995).
[24]R. J. Hussey, G. I. Sproule, D. F. Mitchell and M. J. Graham, in 12th International Corrosion Congress, Houston, TX, September 1993 (unpublished), Vol. 5B, p. 3831.
[25]W. Ting, H. Hwang, J. Lee, and D. L. Kwong, J. Appl. Phys. **70**, 1072 (1991).
[26]Z. H. Lu, J.-M. Baribeau, and D. J. Lockwood, J. Appl. Phys. **76**, 3911 (1994).
[27]A. Pasquarello, M. S. Hybertsen, and R. Car, Phys. Rev. Lett. **74**, 1024 (1995).
[28]F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B **38**, 6084 (1988).
[29]Z. H. Lu, M. J. Graham, D. T. Jiang, and K. H. Tan, Appl. Phys. Lett. **63**, 2941 (1993).
[30]Z. H. Lu, M. J. Graham, S. P. Tay, D. T. Jiang, and K. H. Tan, J. Vac. Sci. Technol. B **13**, 1626 (1995).
[31]Z. H. Lu and A. Yelon, Phys. Rev. B **41**, 3284 (1990).
[32]A. Pasquarello, M. S. Hybertsen, and R. Car, Appl. Phys. Lett. **68**, 625 (1996).
[33]See, for example, *The National Technology Roadmap for Semiconductors* (Semiconductor Industry Association, San Jose, 1994).

Surface passivation of $In_{0.53}Ga_{0.47}As$ ridge quantum wires using silicon interface control layers

Hajime Fujikura, Satoshi Kodama, Tamotsu Hashizume, and Hideki Hasegawa
Research Center for Interface Quantum Electronics and Graduate School of Electronics and Information Engineering, Hokkaido University, Sapporo 060, Japan

(Received 21 January 1996; accepted 17 February 1996)

Surface state effects in $In_{0.53}Ga_{0.47}As$ one-dimensional quantum wires and the effectiveness of the Si interface control layer (Si ICL)-based passivation technique are investigated using photoluminescence (PL) as the probe. Scanning electron microscope and x-ray photoelectron spectroscopy measurements were made to characterize the structure and the interface properties. The $In_{0.53}Ga_{0.47}As$ quantum wires embedded in $In_{0.52}Al_{0.48}As$ barrier material were fabricated by selective molecular beam epitaxy growth on patterned InP substrates. Unpassivated near-surface quantum wires showed an exponential decrease of PL intensity with reduction of surface-to-well distance, t_{ws}, similarly to the near-surface quantum wells. By applying the Si ICL-based passivation, a nearly complete recovery of PL intensity was achieved with an observed maximum recovery factor of 250 for the InGaAs quantum wire directly passivated with SiO_2/Si ICL ($t_{ws}=0$). The mechanism for the PL recovery is explained in terms of suppression of surface states by passivation. © *1996 American Vacuum Society.*

I. INTRODUCTION

Recent intensive research efforts in the study of quantum effects indicate that compound semiconductor quantum nanostructures[1–8] such as quantum wires and quantum dots may become fundamental building blocks of the next generation ultralarge-scale integrated circuits (ULSIs) based on quantum effects. For successful large-scale planar integration, such structures should be placed near surface.

However, the surface of compound semiconductors is generally known to posses a high density of surface states and an interaction between the surface states and the quantum structure may spoil the proper operation of the devices. In fact, the presence of strong surface state effects has been reported recently by several groups[9–12] on photoluminescence (PL) properties of near-surface quantum wells. According to these reports, PL intensity from a quantum well lying near surface is reduced exponentially as the distance between the surface and the well is reduced. This reduction can be explained in terms of the interaction between the quantum well states and the surface states, leading to leak of confined carriers to the surface and subsequent nonradiative recombination. Therefore, a suitable passivation technology applicable to compound semiconductor quantum structures must be established for the successful planar integration of quantum devices.

As a powerful approach for successful passivation of compound semiconductor surfaces, we proposed and analyzed a new passivation structure including a silicon interface control layer (Si ICL).[13–20] In this structure, an ultrathin Si layer grown by molecular beam epitaxy (MBE) is inserted between the compound semiconductor and the outer thick SiO_2 passivation dielectric. The basic idea is to achieve coherent termination of surface bonds of the compound semiconductor by the Si ICL and subsequently to passivate Si ICL by Si-based dielectric. This technique has been successfully applied to fabrication of InGaAs metal-insulator-semiconductor field effect transistors (MISFETs).[16,19] More recently, this passivation technique has been applied to GaAs/AlGaAs near-surface quantum wells and has realized more than 10^3 times recovery of PL intensity at the surface-to-well distance of 5 nm.[20]

As an alternative approach for passivation, use of low-energy ion-beam hydrogenation has also been reported to be powerful for recovery of PL intensity of near-surface quantum wells.[21,22] However, this technique seems to be applicable only to native oxide covered bare surfaces of AlGaAs, and does not seem to compatible with a standard dielectric layer passivation.

Being encouraged by the success of the Si ICL-based passivation of the two-dimensional quantum wells, the purpose of the present article is to investigate the surface state effects in the $In_{0.53}Ga_{0.47}As$ one-dimensional quantum wires and to demonstrate the effectiveness of the Si ICL passivation technique on such structures. The sample structure is schematically shown in Fig. 1(a). The surface state effects and the effectiveness of passivation were investigated by measuring the PL intensity from the wires as a function of the wire-to-surface distance, t_{ws}. Scanning electron microscope (SEM) and x-ray photoelectron spectroscopy (XPS) measurements were also made to characterize the structure and the interface properties of the samples.

The cross-sectional views of the specific unpassivated and passivated wire structures used in this study are shown in Figs. 1(b) and 1(c), respectively. They are $In_{0.53}Ga_{0.47}As$ quantum wires embedded in $In_{0.52}Al_{0.48}As$ barrier material. As compared with the standard AlGaAs/GaAs system, this InP-based $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ material system posses a larger conduction band discontinuity and superior electron transport properties, both of which are favorable for constructing high performance quantum devices operating at high temperatures.

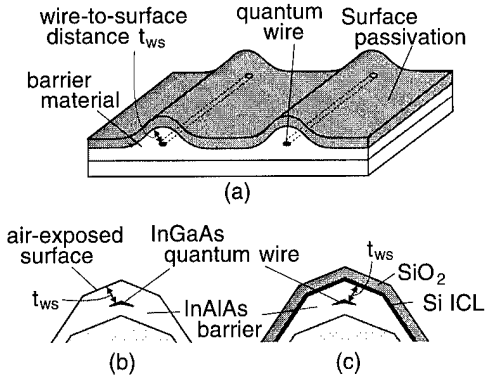


FIG. 1. (a) Schematically illustration of the passivated quantum wire. The surface state effects and the effectiveness of passivation were investigated by measuring the PL intensity from the wires as a function of the wire-to-surface distance t_{ws}. (b) and (c) are the cross-sectional views of the specific unpassivated and passivated $In_{0.53}Ga_{0.47}As$ wire structures used in this study, respectively.

The wires were fabricated by selective MBE growth on patterned InP substrates. Details of wire growth as well as PL and cathodoluminescence (CL) properties of deeply embedded wires with large wire-to-surface distances have been published elsewhere.[7,8]

In this article, it is shown that unpassivated near-surface quantum wires showed an exponential decrease of PL intensity with reduction of surface-to-well distance and that a nearly complete recovery of PL intensity can be achieved by the Si ICL-based passivation technique with an observed maximum recovery factor of 250 for the quantum wire directly passivated with SiO_2/Si ICL (t_{ws}=0). The mechanism for the PL recovery is explained in terms of suppression of surface states by passivation.

II. EXPERIMENT

All the samples were fabricated using an ultrahigh vacuum (UHV)-based integrated fabrication/characterization system where a standard solid-source MBE chamber, a photoenhanced chemical vapor deposition (photo-CVD) chamber, an XPS chamber, and other chambers are connected by an UHV transfer chamber.

The sequence for sample preparation is shown in Figs. 2(a) and 2(b). The preparation starts with formation of mesa patterns on InP substrates with (001) terrace width, W, of 1 μm as shown in Fig. 2(a), using photolithography and wet chemical etching. The orientation of the mesa stripe was $\langle\bar{1}10\rangle$ direction and the pitch, P, was 4 μm. Then, wires were formed by following the UHV-based sequence shown in Fig. 2(b). As the first step, a thick InGaAs buffer layer was grown on the patterned InP substrates. During the growth of this layer, sharp ridge structures consisting of (311)A sidewall facets were formed on the mesa. Subsequent growth of an InAlAs lower barrier layer resulted in formation of a narrow (411)A facet region on the top of the ridge. Then, InGaAs was grown preferentially on the (411)A facet region of the InAlAs ridge, leading to formation of an InGaAs quantum wire surrounded by (311)A and (411)A high-index facets.

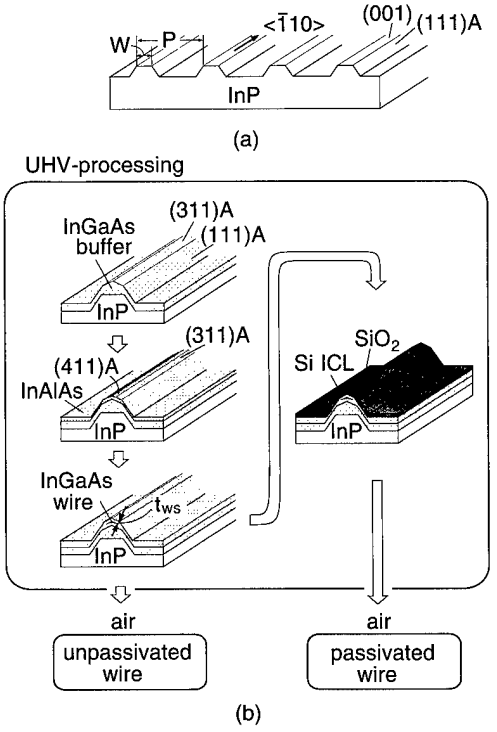


FIG. 2. (a) Mesa-patterned InP substrates used in this study. The (001) terrace width, W, and the pitch, P, were 1 and 4 μm, respectively. (b) UHV-based selective MBE growth and passivation sequence for fabrication of passivated and unpassivated InGaAs quantum wires.

Details of the conditions and the mechanism of such a selective growth have been discussed elsewhere.[7] In the present study, the growth conditions were adjusted to obtain wires whose base width was 100 nm. Then, the top InAlAs barrier layer with a thickness of t_{ws} was grown. Here, t_{ws} is measured perpendicular to the (311)A facet and is defined as the wire-to-surface distance in this study. Sample with t_{ws}=0–300 nm were grown, where t_{ws}=0 corresponds to direct deposition of Si ICL passivation structure on the $In_{0.53}Ga_{0.47}As$ wire as the top barrier. All the MBE layers were undoped. For the preparation of the unpassivated wires, the samples were taken out from the UHV system and were exposed to air.

For the preparation of passivated quantum wires, a Si ICL and an outer thick SiO_2 layer were grown successively on the top of the wire by *in situ* MBE and photo-CVD processes, respectively, without breaking vacuum. For the MBE growth of the Si ICL the Si Knudsen cell (K cell) was used as the Si source and the substrate temperature was 250 °C. The thickness of the Si ICL should be thin enough to maintain pseudomorphic lattice matching, but thick enough to prevent penetration of oxygen radicals during the photo-CVD process. Based on the previous studies by our group,[16] thickness of the Si ICL was chosen to be 1 nm. For the photo-CVD SiO_2 formation, a gas mixture of SiH_4, N_2O, and Ar was used with an ArF excimer laser (193 nm) irradiation. The deposition temperature was 250 °C.

The cross-sectional structures of the fabricated InGaAs quantum wires were characterized by scanning electron mi-

croscope (SEM). To investigate interface structure, XPS measurements were made on separately prepared planar samples, a Perkin–Elmer spectrophotometer was used with an excitation source of Al $K\alpha$ line (1486.6 eV). To characterize the optical properties of the unpassivated and passivated InGaAs ridge quantum wires, low temperature PL measurements were performed using Ar^+ laser (514.5 nm) as an excitation source.

III. RESULTS AND DISCUSSION

A. SEM observation of the InGaAs quantum wire

Figure 3(a) shows schematically the cross-sectional structure of the fabricated wire. SEM photographs taken at the main wire portion and at the bottom groove portion are shown in Figs. 3(b) and 3(c), respectively. As seen clearly in Fig. 3(b), the well-defined InGaAs quantum wire with a lateral width of 100 nm was formed at the top of the InAlAs ridge structure. It is also seen that a very thin quantum wells is formed on the (311)A sidewall facets. Additionally, wells having lager thickness were also formed at the bottom of the grooves as seen in Fig. 3(c). In the groove region, considerable waving of the growth boundary and fluctuation of the well width took place due to irregular growth involving large amounts of mass transfer.

B. XPS analysis of surface and interface

To clarify the effect of the insertion of the Si ICL on the chemical structure of the interface, air-exposed surfaces and a passivated interfaces of InAlAs and InGaAs were analyzed by the XPS method using separately grown planar samples. In the passivated samples, the thickness of the outer SiO_2 layer was kept to be 1.5 nm in order to allow photoelectrons to escape from the interface region.

The results of XPS analysis of InAlAs samples are summarized in Fig. 4, where the In $3d$, Al $2p$, and As $2p$ core level spectra from air exposed and Si ICL-passivated samples are compared. It is clearly seen that large amounts of In, Al, and As oxide components exist on the air exposed surface and that they are drastically reduced by the SiO_2/Si ICL-passivation process. Although certain amounts of Al and As oxide components as well as the elementary As component still exist at the Si ICL-passivated interface, the interface structure becomes much more coherent and stoichiometric.

As for the InGaAs, the XPS results have already been reported by our group elsewhere.[19] It has been found that the surface oxides were drastically reduced again by the Si ICL passivation process as compared with the air-exposed InGaAs surface.

C. PL properties of unpassivated wires

For a systematic investigation of the surface state effects on the optical properties of the InGaAs ridge quantum wires, the unpassivated InGaAs ridge quantum wires having various values of the wire-to-surface distance, t_{ws}, were fabricated and their PL properties were studied.

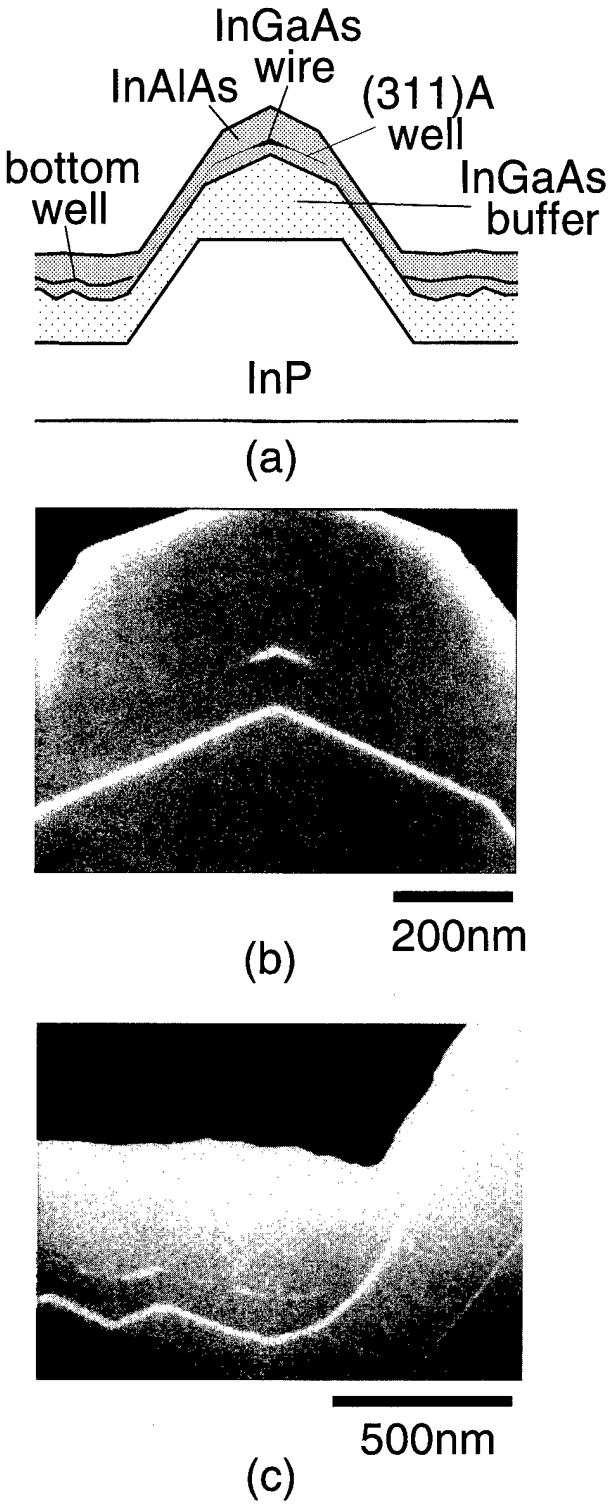


FIG. 3. Results of the SEM observation of the InGaAs quantum wire. (a) The schematic cross-sectional structure of the fabricated wire. (b) and (c) The SEM photographs taken at the main wire portion and at the bottom groove portion, respectively.

Three examples of the PL spectra of the unpassivated wires taken at 20 K are shown in Fig. 5. The wires had the values of t_{ws} of 300, 30, and 5 nm. Gaussian deconvolution of the PL spectra are given by dashed curves.

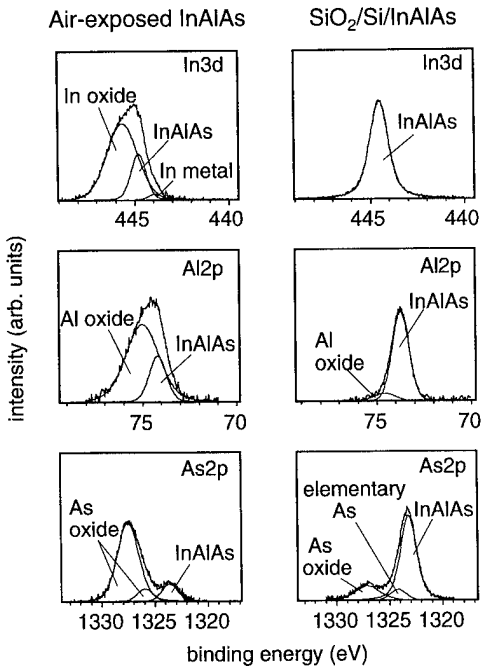


FIG. 4. The results of XPS analysis of planar InAlAs samples. The In $3d$, Al $2p$, and As $2p$ core level spectra from air exposed and Si ICL-passivated samples are compared.

In the PL spectrum of the wire having t_{ws} of 300 nm, an intense peak was observed at 1.03 eV. The peak has two shoulders on the higher and lower energy sides. The positions of the main peak and two shoulders after Gaussian deconvolution are indicated by arrows in Fig. 5.

As for the identification of the spatial origin of these peaks, we performed previously a detailed low temperature

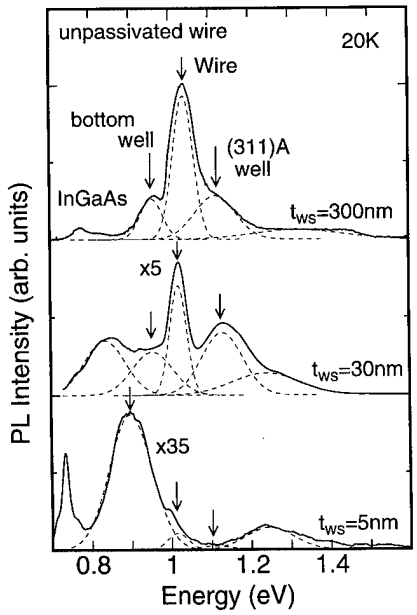


FIG. 5. Three examples of the PL spectra of the unpassivated wires taken at 20 K. The wires had the values of t_{ws} of 300, 30, and 5 nm. Gaussian deconvolution of the PL spectra are given by dashed curves.

spatially resolved CL study.[7,8] According to this study, the intense peak seen at 1.03 eV originates from the wires themselves.[7,8] Two shoulders seen at the higher and lower energy sides of the main peak also come from the parasitic quantum wells formed on the (311)A sidewall facets and those formed at the bottom of the grooves, respectively. Additionally, a peak and a weak tail were observed at around 0.8 eV and above 1.2 eV, respectively. The peak at around 0.8 eV, which corresponds to the energy gap of InGaAs, is due to the emission from the InGaAs buffer layer. The peak observed below 0.8 eV may correspond to the impurity related transition in the InGaAs buffer layer. On the other hand, the tails seen above 1.2 eV were identified to be emissions from the InAlAs barrier, including deep levels, since they were also found in the PL spectrum of InAlAs layers grown on the mesa patterned substrate without InGaAs wire and buffer layers.[6]

With the reduction of the wire-to-surface distance, t_{ws}, the essential features of the PL spectra kept unchanged, as can also be seen in Fig. 5, although the PL intensity decreased as discussed fully later.

The position of the main wire peak was held remarkably constant with the reduction of t_{ws}, indicating stability of the selective growth process employed here. It also indicates that peak energy shifts caused by interaction of quantum confined states and the surface states as reported by Moisson *et al.*[9] for *in situ* PL studies of AlGaAs/GaAs near surface quantum wells are absent here. Our previous study on *ex situ* PL of AlGaAs/GaAs near-surface quantum wells did not detect such shifts. On the other hand, considerable differences were observed in the energy positions of the subpeaks from the two types of parasitic quantum wells. The origin of these difference may be attributed to the fluctuations of the quantum well thickness caused by fluctuation in the long-range material transfer during growth.

The observed variation of the PL intensity of the main wire peak from the unpassivated quantum wires is summarized in Fig. 6 as a function of the wire-to-surface distance t_{ws}. Here, the PL intensities are normalized by its value obtained for the wire with t_{ws}=300 nm.

The result of Gaussian deconvolution was used to accurately determine the main peak intensity in each sample. As seen in Fig. 6, the PL intensity decreases gradually as t_{ws} is reduced from 300 down to 10 nm. Then, it starts to drop exponentially as t_{ws} falls below 10 nm. Thus, the PL intensity of the unpassivated wire with t_{ws}=5 nm was 60 times smaller than that of the wire having t_{ws} of 20 nm and 400 times smaller than that of the wire having t_{ws}=300 nm.

As in the case of the previous reports on quantum wells,[9,10,20] the mechanism of the strong exponential reduction of PL intensity from the near-surface quantum wire with the reduction of the surface-to-wire distance t_{ws} below 10 nm can be explained in terms of the interaction between confined quantum wire states and surface states. Namely, surface states provide a t_{ws}-dependent channel for nonradiative surface recombination in competition with radiative transition between electron and hole confined quantum states.

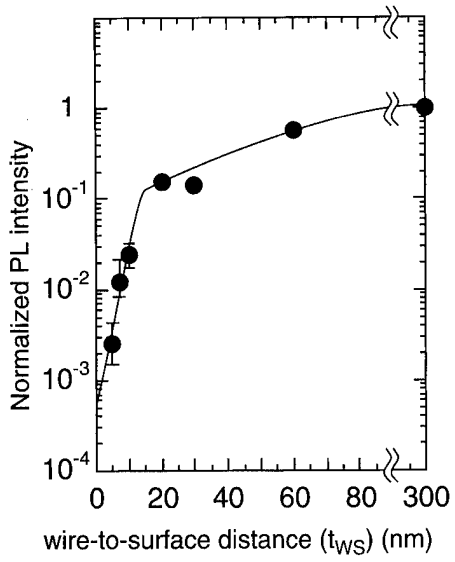


FIG. 6. The observed variation of the PL intensity of the main wire peak from the unpassivated quantum wires as a function of the wire-to-surface distance, t_{ws}. Here, the PL intensities are normalized by its value obtained for the wire with t_{ws}=300 nm.

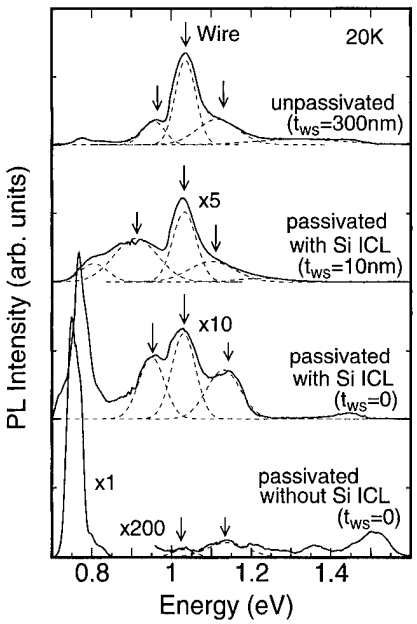


FIG. 7. PL spectra of the wires passivated with Si ICL having t_{ws} of 10 nm and 0 nm are compared with that of an unpassivated wire having t_{ws}=300 nm. The PL spectrum of the wire directly passivated with SiO_2 without insertion of Si ICL is also shown at the bottom.

In addition, absence of the appreciable shifts of PL peak energy from the wire with the change of surface-wire distance suggests that the observed reduction of PL intensity in this study is not due to the surface states with spatially extended wave function but due to the highly localized one, as has been discussed previously.[20]

On the other hand, the more gradual PL intensity reduction that took place from t_{ws}=300 nm down to t_{ws}=10 nm in Fig. 6 is a new feature that was not observed in the near-surface quantum wells. This is not due to the surface states. We believe that this was caused by the particular sample structure used here. Namely, the thickness of the bottom InAlAs barrier was limited to about 100 nm in order to realize a narrow wire with the width of 100 nm using the present selective MBE growth. Thus, when t_{ws} is much larger than the penetration depth of the excitation light about 100 nm, almost all the excitation light is absorbed in the top InAlAs layer and generated carriers diffuse and recombine mainly in the InGaAs wire. On the other hand, as the thickness of the top InAlAs barrier layer t_{ws} become smaller than the penetration depth, the carriers are generated mainly in the thin bottom InAlAs barrier layer and InGaAs buffer layers. Only a certain portion of carriers generated in the bottom InAlAs barrier can supply carriers to the wire, because some part of them escape into the InGaAs buffer layer. Thus, the decrease of t_{ws} reduces the net carrier supply into the wire, resulting in the PL intensity reduction. This model is consistent with the fact that the relative peak intensity of the peak due to the InGaAs buffer becomes larger with the reduction of t_{ws}, as can be seen in Fig. 5.

As for the variation of PL intensities of subpeaks, the reduction of the PL intensity of the (311)A well showed nearly the same trend as that of the quantum wire, as can be seen in Fig. 5. This is exactly what is expected from the explanations given above for the two ranges of t_{ws}. On the other hand, the bottom well showed weaker dependence on t_{ws}, also seen in Fig. 5. This seems to be due to the larger thickness of the top InAlAs barrier layer caused by material transfer, which can be seen in Fig. 3(c).

D. Recovery of PL intensity in wires with Si ICL-based passivation

Figure 7 compares the PL spectra of the wires passivated with Si ICL having t_{ws} of 10 and 0 nm with that of an unpassivated wire having t_{ws}=300 nm. The PL spectrum of the wire directly passivated with SiO_2 without insertion of Si ICL is also shown at the bottom. By comparing Fig. 7 with Fig. 5, it is clearly seen that the wires passivated with the Si ICL technique give rise to much increased PL emission from the wire, although the shape of the PL spectrum remained the same. Additionally, the emission intensities from the parasitic quantum wells on the (311)A facets and at the bottom grooves are also increased. On the other hand, the wire with t_{ws}=0 passivated by SiO_2 without Si ICL showed a much weaker PL intensity. It is also noted in Fig. 7 that the intensities of the emission peak from the InGaAs buffer layers are very much increased in the two types of the passivated wires, with t_{ws}=0. This is obviously caused by increased carrier flow into the InGaAs buffer layer due to the absence of the top InAlAs barrier as discussed above.

Figure 8 shows the PL intensity of the quantum wires passivated by the Si ICL-based process by open circles as a function of the wire-to-surface distance t_{ws}. For comparison, the PL intensity data on the unpassivated wires and that on the wire passivated by SiO_2 without Si ICL are also included in Fig. 8 by filled circles and open square, respectively. As

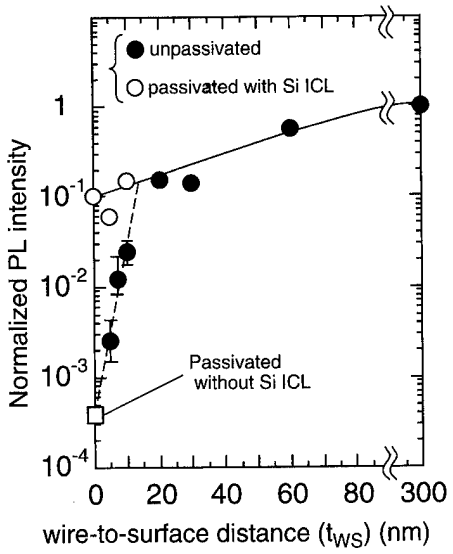


FIG. 8. PL intensity of the quantum wires passivated by the Si ICL-based process by open circles as a function of the wire-to-surface distance t_{ws}. For comparison, the PL intensity data on the unpassivated wires and that on the wire passivated by SiO_2 without Si ICL are also included by closed circles and open square, respectively.

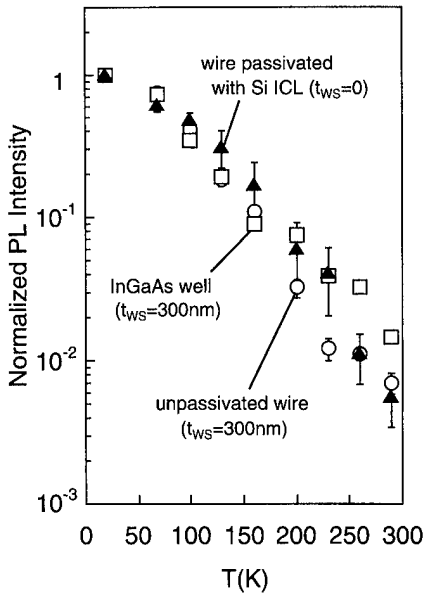


FIG. 9. PL intensity of the wire with t_{ws}=0 nm passivated with Si ICL as a function of measurement temperature. Data on the PL intensities of the unpassivated wire with t_{ws}=300 nm and that of a planar InGaAs quantum well simultaneously grown with the unpassivated wire having t_{ws} of 300 nm are also included by the open circles and open squares, respectively. Each of the PL intensities are normalized by their values at 20 K.

seen in Fig. 8, the PL intensity of the unpassivated wire is reduced by the factor of 400 by decreasing the thickness of the top InAlAs barrier layer from 300 to 5 nm, as shown in Fig. 6. In contrast to this, the wire passivated with Si ICL showed PL emission only 10 times smaller than that of the wire with t_{ws}=300 nm even in the severest case of t_{ws}=0. On the other hand, the emission intensity of the wire directly passivated by SiO_2 without Si ICL is 2500 times smaller than that of the wire with t_{ws}=300 nm. Thus, the Si ICL-based passivation technique is very powerful in enhancing PL emission from near-surface quantum wires.

It is also noted in Fig. 8 that the Si ICL-passivated wires no longer show as strong an exponential decrease of PL intensity as the unpassivated wire. In fact, the data points of the wires passivated with Si ICL lie on a line that is obtained by naturally extrapolating the data points of the gradually decreasing portion of the unpassivated wires having t_{ws} above 10 nm. This indicates that the nonradiative recombination process through the surface states become almost negligible in the wires passivated with Si ICL having t_{ws} values smaller than 10 nm. Thus, it can be said that the Si ICL-based passivation process can almost entirely remove surface state induced PL deterioration from the InGaAs ridge quantum wires formed near surface.

The observed PL recovery can be attributed to reduction of surface states achieved by the Si ICL-based passivation process. Referring to the XPS results in Fig. 4 and following the previous discussion made on a similar improvement obtained in the case of Si ICL-based passivation of near-surface quantum wells,[20] reduction of the surface states can be regarded as a consequence of the formation of a more ordered, coherent, and stoichiometric interface that minimizes bond disorder at the interface.

Finally, in view of the application of such passivated wires to the devices operating at a higher temperature, the temperature dependence of the PL intensity of the wire passivated with Si ICL was investigated. Figure 9 compares the PL intensity of a wire with t_{ws}=0 nm passivated with Si ICL with that of an unpassivated wire with t_{ws}=300 nm as a function of measurement temperature. For the purpose of comparison, the data on the PL intensity of a planar InGaAs quantum well simultaneously grown with the unpassivated wire having t_{ws}=300 nm is also included. In Fig. 9, the PL intensities are normalized by their values at 20 K. As seen in Fig. 9, the PL intensities of these three samples showed the same trend. Namely, the PL intensities were almost constant up to 100 K. Then, they started to decrease rapidly above 100 K over two orders of magnitude reduction up to 290 K, although the peaks maintained still reasonably strong intensity even at 290 K. Thus, this reduction in PL intensity with temperature is not due to surface passivation but is related either to the InGaAs crystal or to the InGaAs–InAlAs interface. Similar decreases of the PL intensities have been reported in the GaAs-based wires by several other groups.[3–5] However, quantitatively speaking, the present InGaAs quantum wires show much smaller decrease of the PL intensity with temperature than these GaAs-based wires.

IV. CONCLUSIONS

In view of the urgent necessity to establish a suitable surface passivation technique applicable to near-surface compound semiconductor quantum structures, the Si ICL-based passivation technique was applied to passivation of

$In_{0.53}Ga_{0.47}As$ ridge quantum wires having different wire-to-surface distance t_{ws} that were formed by selective MBE. The main results are listed below.

(1) The PL intensity of the unpassivated wires decrease exponentially with reduction of wire-to-surface distance t_{ws} below about 10 nm. The behavior can be explained in terms of tunneling capture and subsequent nonradiative recombination of carriers from the quantum wire by surface states present on air exposed surfaces.

(2) A nearly complete recovery of PL intensity can be achieved by use of the Si ICL-based passivation process. An maximum recovery factor of the PL intensity of 250 was obtained in the severest case of $t_{ws}=0$. The wire with $t_{ws}=0$ passivated by the Si ICL technique showed reasonably strong PL emission even at room temperature.

(3) The recovery of PL intensity is due to reduction of interface state densities by the Si ICL process, which realizes reasonably ordered and coherent interface structures, as indicated by the XPS analysis.

[1]R. Bhat, E. Kapon, D. M. Hwang, M. A. Koza, and C. P. Yun, J. Cryst. Growth **93**, 850 (1988).
[2]R. Nötzel, J. Temmyo, H. Kamada, and T. Fukura, Appl. Phys. Lett. **65**, 457 (1994).
[3]M. Walther, E. Kapon, J. Christen, D. M. Hwang, and R. Bhat, Appl. Phys. Lett. **60**, 521 (1992).
[4]M.-S. Lee, Y. Kim, M.-S. Kim, S.-I. Kim, S.-K. Min, Y. D. Kim, and S. Nahm, Appl. Phys. Lett. **63**, 3052 (1993).
[5]H. Akiyama, S. Koshiba, T. Someya, K. Wada, H. Noge, Y. Nakamura, T. Inoshita, A. Shimizu, and H. Sakaki, Phys. Rev. Lett. **72**, 924 (1994).
[6]H. Fujikura, H. Tomozawa, M. Akazawa, and H. Hasegawa, Appl. Surf. Sci. **60/61**, 702 (1992).
[7]H. Fujikura and H. Hasegawa, J. Electron. Mater. **25**, 619 (1996).
[8]H. Fujikura and H. Hasegawa, Jpn. J. Appl. Phys. **35**, 1333 (1996).
[9]J. Moisson, K. Elcess, F. Houzay, J. Y. Marzin, J. M. Gérard, F. Barthe, and M. Bensoussan, Phys. Rev. B **41**, 12 945 (1990).
[10]Y.-L. Chang, I.-H. Tan, Y.-H. Zhang, D. Bimberg, J. Merz, and E. Hu, J. Appl. Phys. **74**, 5144 (1993).
[11]Z. Sobiesierski and D. I. Westwood, Superlatt. Microstruct. **12**, 267 (1992).
[12]Z. Sobiesierski, D. I. Westwood, D. A. Woolf, T. Fukui, and H. Hasegawa, J. Vac. Sci. Technol. B **11**, 1723 (1993).
[13]H. Hasegawa, M. Akazawa, K. Matuzaki, H. Ishii, and H. Ohno, Jpn. J. Appl. Phys. **27**, L2265 (1988).
[14]H. Hasegawa, M. Akazawa, H. Ishii, A. Uraie, H. Iwadate, and E. Ohue, J. Vac. Sci. Technol. B **8**, 867 (1990).
[15]M. Akazawa, H. Hasegawa, and E. Ohue, Jpn. J. Appl. Phys. **28**, L2095 (1989).
[16]M. Akazawa, H. Ishii, and H. Hasegawa, Jpn. J. Appl. Phys. **30**, 3744 (1991).
[17]H. Hasegawa, S. Kodama, S. Koyanagi, and T. Hasizume, Jpn. J. Appl. Phys. **34**, L495 (1995).
[18]S. Kodama, S. Koyanagi, T. Hasizume, and H. Hasegawa, Jpn. J. Appl. Phys. **34**, 1143 (1995).
[19]S. Suzuki, S. Kodama, and H. Hasegawa, Solid-State Electron. **38**, 1679 (1995).
[20]S. Kodama, S. Koyanagi, T. Hashizume, and H. Hasegawa, J. Vac. Sci. Technol. B **13**, 1794 (1995).
[21]V. Emiliani, B. Bonanni, C. Presilla, M. Capizzi, A. Frova, Y.-L. Chang, I.-H. Tan, J. L. Merz, M. Colocci, and M. Gurioli, J. Appl. Phys. **75**, 5114 (1994).
[22]Y.-L. Chang, W. Widdra, S. I. Yi, J. Merz, W. H. Weinberg, and E. Hu, J. Vac. Sci. Technol. B **12**, 2605 (1994).

Structure of the SiN_x/GaAs (110) interface modified with ultrathin Si and sulfur passivation

L. J. Huang[a] and W. M. Lau
Department of Materials Engineering and Department of Physics, University of Western Ontario, London, Ontario N6A 5B9, Canada

H. T. Tang, W. N. Lennard, and I. V. Mitchell
Department of Physics, University of Western Ontario, London, Ontario N6A 3K7, Canada

D. Landheer and J.-M. Baribeau
National Research Council of Canada, Ottawa, Ontario O1R 1C7, Canada

S. Ingrey
Bell Northern Research, Ottawa, Ontario K1Y 4H7, Canada

(Received 22 January 1996; accepted 20 April 1996)

The SiN_x/GaAs (110) interface structure of the metal-insulator-semiconductor (MIS) capacitors modified with ultrathin Si or Si/Ge (5 to 15 Å) heterolayers and sulfur passivation were studied by capacitance-voltage (*CV*) measurements, high resolution medium energy ion scattering and extended x-ray absorption fine structure with synchrotron radiation. The results showed that Si/S multilayer passivation can improve the interface properties of the MIS capacitors made on *n*-type GaAs (110). The element depth profiling in a subnanometer scale and the atomic structure measurements showed that a Si or Si/Ge heterolayer on sulfur passivated GaAs (110) can substantially limit Ga(As) precipitation at the interface but cannot completely suppress the incorporation of Ga(As) into the Si heterolayer. These results can be used to interpret the observed *CV* characteristics of the capacitors and to understand the limitation of the passivation approach.
© *1996 American Vacuum Society.*

I. INTRODUCTION

Understanding the mechanism for reducing the density of state at the insulator/GaAs interface is crucial for the potential for metal-insulator-semiconductor (MIS) device technology of GaAs. Such a problematic interface has been studied for years.[1–10] The predominant issues have been attributed to interfacial disordering, including that from local stress and defects, oxygen induced deep levels, and arsenic segregation. All of these defects can lead to GaAs surface Fermi level pinning and make modulation of the GaAs surface potential very difficult. To eliminate the interface defects, an interface with atomic abruptness is naturally a model to pursue. The approach of using a thin interfacial control layer (ICL) of pseudomorphic silicon by Hasegawa and co-workers[3] has been considered a success. Limited improvement of the capacitance-voltage (*CV*) characteristics with the Si ICL approach for the MIS capacitors made on GaAs (100) has been encouraging many researchers to improve the quality of the silicon ICL with the expectation that a more ordered interface will further lower the interface state density and hence substantiate device performance. This includes the use of a Si/Ge multilayer to replace a silicon single layer with the hope for a better lattice match (atomic abruptness) since the lattice of germanium is closer to that of GaAs than that of silicon. In addition, reasonably good pseudomorphic growth of silicon on germanium has proved feasible.[8–10] On the other hand, to bring down the cost for a complete *in situ* fabrication, sulfur passivation has been identified as an effective approach to inhibit GaAs surface oxidation and form an ordered surface for *ex situ* device fabrication. Nevertheless, the interface atomic structure of the MIS capacitors fabricated with these approaches remains unclear and the present results are still far from that for practical applications of MIS capacitors.

In this article, MIS capacitors were fabricated on a GaAs (110) substrate with its interface modified by Si or Si/Ge and sulfur passivation. To understand the *CV* characteristics of the MIS capacitors, high resolution medium energy scattering (MEIS) and extended x-ray absorption fine structure (EXAFS) with synchrotron radiation were used to measure the interface structure. The results are nontrivial for understanding the MIS capacitor performance using the present interface passivation approach.

II. EXPERIMENT

Both *n*-type (Si doped) and *p*-type (Zn doped) GaAs (110) wafers supplied by American Xtal Technology were used for the fabrication of the MIS capacitors. The substrate etch pit densities were less than $10^3/cm^2$ and both the *n*- and *p*-dopant concentration was 1×10^{17} at./cm^3. The as-received wafers were treated using two cycles of UV/ozone oxidation for 10 minutes followed by etching in an 1:60 hydrofluoric (HF) acid solution for 2 minutes. The samples were then rinsed with deionized water for 30 seconds and blown dry with pure nitrogen gas. The oxide-free samples were subject

[a]Present address: Conner Peripherals, 311 Turquoise Street, Milpitas, CA 95035; Electronic mail: LHuang@hydra.uwo.ca

to gas phase polysulphide passivation for 20 minutes and the as-passivated samples were immediately transferred into a VG V80 molecular beam epitaxy system for e-beam deposition of Si or Si/Ge ICLs. The layer thickness varied from 3 Å to 20 Å and the substrate temperature during deposition was 250 °C. For MIS capacitor fabrication, the samples after ICL deposition were immediately deposited with 350 Å silicon nitride in a remote plasma chemical vapor deposition system and then 400 Å of aluminum top contacts were evaporated onto the nitride surface through a shadow mask. The diameter of the contacts varied from 0.2 to 1.0 mm. The capacitors were annealed at 450 °C for 15 minutes in a forming gas mixture of 4.2% hydrogen in nitrogen. Back contacts to the capacitors were made by In/Ga eutectic. *CV* data for the fabricated MIS capacitors were collected with a Hewlett-Packard 4061A semiconductor component test system. For the quasi-static (QS) measurements the potential of the gate was scanned from inversion to accumulation, while for the high frequency measurements it was scanned from negative to positive bias and then immediately scanned back to negative bias. A detailed discussion of the *CV* analysis can be found in our previous publications.[11]

To understand the *CV* characteristics of the MIS capacitors, MEIS and EXAFS were employed to study the interface structure. MEIS was carried out with the 1.7 MV Tandetron accelerator at the University of Western Ontario by using a $^4He^+$ collimated ion beam at 280 keV and a high resolution hemispherical electrostatic energy analyzer. The spectra were collected in a double crystal alignment configuration in which the scattering angle between the incoming channelling string and outgoing ion direction was 90°. The energy resolution for all the MEIS data collected was about 0.24 keV, which is equivalent to about 4 Å for a GaAs single crystal. EXAFS spectra were measured at the Canadian Synchrotron Radiation Facility (Aladdin) on the 1 GeV electron storage ring, University of Wisconsin-Madison. The x-ray was monochromatized and the photon resolution at 1800 eV (silicon *k*-edge at 1839 eV) was about 0.8 eV. The total electron yield (TEY) mode was used for the EXAFS measurements. In order to eliminate the unrelated structures in the measurements, all the TEY spectra were normalized to the simultaneously recorded TEY signal from a gold mesh reference monitor. The detailed description of the above experimental systems and the relevant experimentation can be found in the literature or our previous publications.[12]

III. RESULTS AND DISCUSSION

A. *CV* characteristics

Figure 1 shows the *CV* characteristics for the MIS capacitors fabricated on *n*-type GaAs (110) with the interface modified by (a) HF etch, (b) HF etch and 15 Å Si ICL, and (c) sulfur passivation and 15 Å Si ICL. Some measurable parameters from the curves are listed in Table I. One sees that the capacitor treated with HF etch had only a very poor high frequency *CV* characteristic [Fig. 1(a)]. The large frequency dispersion indicated a strong Fermi level pinning away from

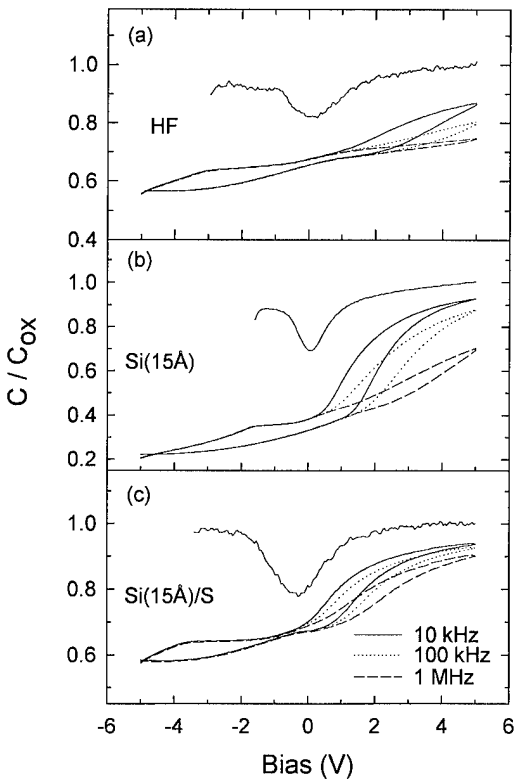


FIG. 1. The *CV* characteristics for the MIS capacitors made on *n*-type GaAs (110) with the interface modified by (a) HF etch, (b) HF etch and 15 Å Si ICL, and (c) sulfur passivation and 15 Å Si ICL.

the conduction band minimum (CBM) and the shift of the high frequency characteristics to the right with increasing frequency shows a high density of shallow states at the insulator/semiconductor interface. In addition, the difference between the maximum and minimum values of the high frequency curves was very small and the amount of hysteresis was large (about 1.1 V), which came from a large amount of negative charge trapping in slow states. The very large offset between the QS and the high frequency *CV* curve in accumulation also suggested that there was a high density of interface states near the CBM edge. All of these characteristics indicated that such interface could not be used for any practical applications. Applying a Si ICL to the interface of the capacitor [Fig. 1(b)] could, however, only improve the *CV* characteristics to a limited amount. The difference between the maximum and minimum values of the high fre-

TABLE I. *CV* characteristic parameters measured for the *n*-type samples discussed in the text.

Passivation approach	Frequency dispersion (V)	Hysteresis (V)	QS dip	Offset between quasi-static and 1 MHz *CV* curve	D_{it} (10^{12} eV^{-1} cm^{-2})
HF etch	···	1.1	0.11	0.27	···
Si	1.8	0.8	0.19	0.29	5.1
Si/S	1.0	0.7	0.21	0.10	3.4

quency curves increased by a factor of 2, the amount of hysteresis (0.7 V) decreased by a factor of 1.6, and the offset between the QS and high frequency curves in accumulation decreased by a factor of 1.5. But the overall *CV* characteristics were still poor. The observed large frequency dispersion showed that the interface Fermi level was still pinned away from the CBM. The QS curve at the negative bias was dominated by the slow movement of charges and the hysteresis was still very large (about 1 V). These characteristics are typical for the MIS capacitors made on GaAs (100), and the results observed had no substantial difference as compared to those in the literature.[2–5] Additional improvements were observed for the capacitor whose interface was modified by sulfur passivation and a 15 Å Si ICL. Improvement of the capacitor interfacial properties is evident from the measured *CV* characteristics [Fig. 1(c)] such as an even deeper QS integral dip and a smaller offset between the QS curve and the high frequency data in accumulation compared with that shown in Fig. 1(b). The frequency dispersion was also substantially reduced and the measured hysteresis was decreased to about 0.5 V. Nonetheless, the frequency dispersion was still present and the deep depletion was seen from the *CV* characteristics. By applying a Si/Ge/S multilayer interface modification approach, the *CV* characteristics can be further improved[10]; however, there was no substantial improvement in the calculated interface state densities ($\sim 10^{12}$ eV^{-1} cm^{-2}) by comparing QS and high frequency (1 MHz) capacitance (*high-low* method).

B. Interface structure

It was previously proposed[13] that using a Si ICL could prevent the As at the interface from precipitation by stabilizing the free As and hence eliminating one of the critical factors leading to interface Fermi level pining. It was also suggested[5] that the dangling bonds in Si or the high conduction band density of Si (assuming a fixed 0.2 eV relative conduction band offset between Si and GaAs) would limit further improvement of the MIS interface characteristics. To clarify the interface structure of the MIS capacitors with interface modification by the Si ICL and sulfur passivation, high resolution MEIS and EXAFS were used to measure the interface element profiling and atomic structure with Si or Si/S interface passivation. This enabled us to know the element distribution and atomic structure at the interface. Figure 2 shows the MEIS spectrum for an *n*-type GaAs (110) cleaned by UV/ozone oxidation and with a 1:60 HF etch. The spectrum was collected in a double alignment configuration in which ⟨101⟩ was chosen as the incoming ion channelling direction and the outgoing ions were collected at a scattering angle of 90° along ⟨121⟩. The surface peak shown in the spectrum was fitted for ^{75}As, ^{71}Ga, and ^{69}Ga. In the present peak fitting approach, the ^{75}As, ^{71}Ga, and ^{69}Ga peak positions were determined by the corresponding ion scattering energy $E=kE_0$, where E_0 was the analytical beam energy, and k was the kinematic factor for ^{75}As, ^{71}Ga, and ^{69}Ga, respectively. The ratio of the ^{71}Ga and ^{69}Ga peak area intensity was also fixed at 0.576 according to the natural

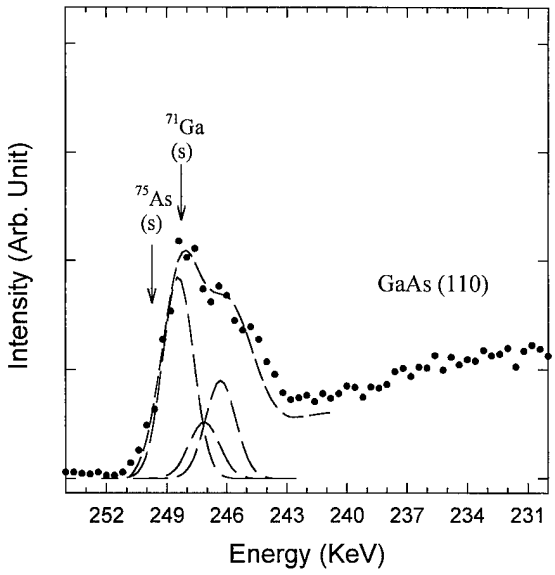


FIG. 2. MEIS spectrum for a HF cleaned *n*-type GaAs (110).

isotope abundance. A Gaussian peak was chosen for the fitting and the data were fitted to the χ^2 minimum. It was found that the surface As to Ga ratio was about 1, i.e., the surface was close to stoichiometry. The number of surface GaAs atoms, $(Nt)_{GaAs}$, can be calculated from the measured area, A, of the surface peaks in the spectrum by[14]

$$(Nt)_{GaAs}=A/(\Omega Q\sigma)\cos\,\theta,$$

where σ is the cross section, Ω is the solid angle of the detector, Q is the total number of $^4He^+$ ions of the analytical beam, and θ is the angle between the direction of the $^4He^+$ beam and the sample normal. From the spectrum, we found that the calculated areal density for the surface peak was about $3\times10^{15}/cm^2$. Figure 3 shows the MEIS spectra for a 15 Å Si heterolayer on sulfur passivated [Fig. 3(a)] and UV/ozone/HF cleaned [Fig. 3(b)] GaAs (110). It can be measured from the spectra that the edge of the As surface peak shifted about 0.9 keV towards the lower energy end, which indicated that there was no As or Ga on the surface of the Si ICL. Such an energy shift was equivalent to a depth of 9 Å. Since the samples were exposed to air during the transfer to the MEIS analytical chamber, surface oxidation would occur. If we assume that the oxide thickness was the same as that for native silicon oxide ($\sim$ 10 Å), or that the top 5 Å of the silicon was oxidized, the front surface of the As or Ga was distributed below 3–4 Å of the Si ICL. The striking feature shown in Fig. 3 is that for the sample with sulfur passivation, the surface As intensity was substantially lower. The measured atomic areal density for the surface As was about $3\times10^{14}/cm^2$, or more than 5 times lower than that of the virgin GaAs (110) surface shown in Fig. 2, while for the sample without sulfur passivation the measured As areal density was about $1.2\times10^{15}/cm^2$, which is about 4 times higher than that with sulfur passivation and about 0.8 times lower than that of virgin GaAs (110). This could be one of the reasons for the improvement of the *CV* characteristics of the

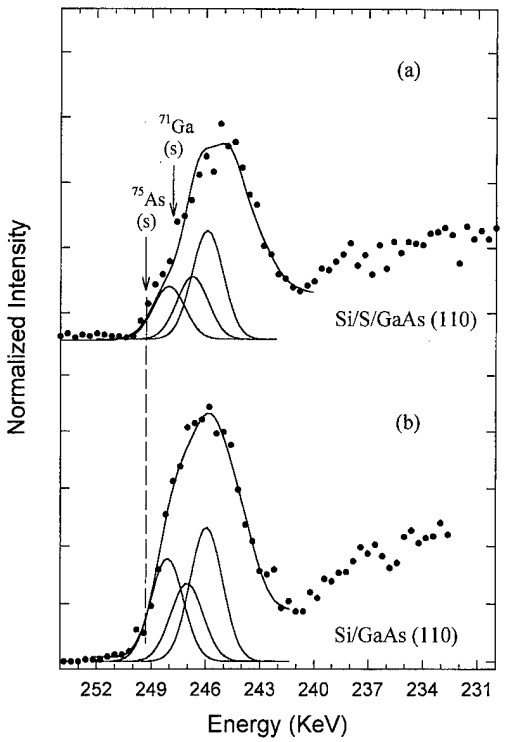


FIG. 3. MEIS spectra for a 15 Å Si heterolayer on (a) sulfur passivated and (b) UV/ozone/HF cleaned *n*-type GaAs (110).

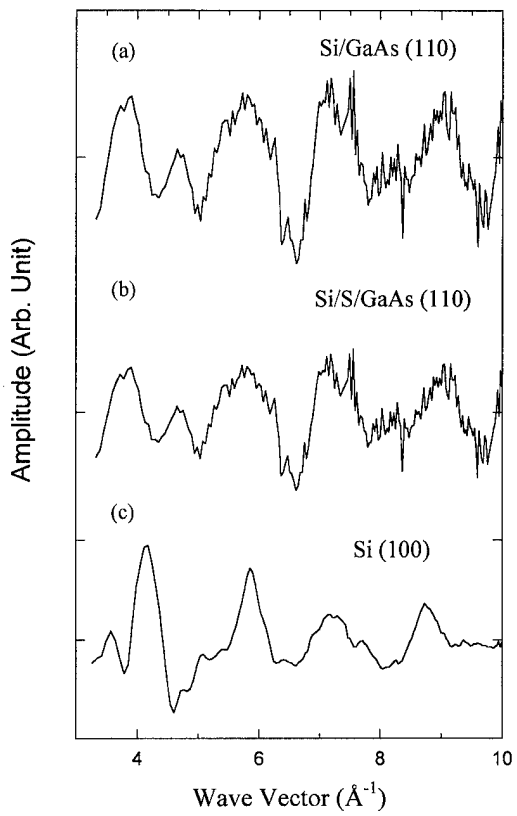


FIG. 4. The k^2 weighted EXAFS spectra for a 15 Å Si heterolayer on (a) sulfur passivated and (b) UV/ozone/HF cleaned *n*-type GaAs (110). The k^2 weighted EXAFS spectrum of Si (100) is shown in (c).

corresponding MIS capacitors with sulfur passivation (Fig. 1); i.e., sulfur passivation removed the surface As by forming S-Ga bonds which further suppressed As precipitation, since the As segregation would induce high interface state density.

The results that the As(Ga) peak moved into the Si ICL as shown in Fig. 3 also indicated that both Si or Si/S passivation could not inhibit the site exchange between Ga and Si, and completely suppress the As precipitation at the interface as a considerable amount of As(Ga) had incorporated into the Si ICL. This implied that the proposed S-Ga terminated GaAs surface by sulfur passivation would mostly be destroyed after the Si ICL deposition. The actual ICL was in fact a layer mixed with Si-Ga(As). The amounts of sulfur diffused into the Si ICL could be measured using particle induced x-ray emission.[15] The total amount of sulfur at the interface was measured to be about 4×10^{14} at./cm^2, or about half a monolayer. Therefore, even we assumed that all the sulfur diffused into the Si ICL, the concentration of sulfur in the "mixed" layer would be less than 5 at.%. Our results[15] showed that sulfur diffused towards both the Si ICL and the substrate. Therefore, the amount of sulfur in the ICL was much less than that of Ga(As), which was about 20 at.% in this mixed layer. The atomic structure of this mixed layer was measured by EXAFS. In Fig. 4 the k^2 weighted EXAFS data are shown for the Si ICL on GaAs (110) with and without the sulfur passivation. The data measured from a Si (100) single crystal are also plotted for comparison. Compared to the Si EXAFS data, two extra peaks at wave vector 4.8 and 8.2 Å^{-1} could be seen clearly. Figure 5 shows the Fourier transformation (FT) of the corresponding EXAFS spectra shown in Fig. 4. Since the local structure of thin amorphous or microcrystalline silicon film is close to that of a silicon single crystal,[12] the data showed clearly that there were foreign atoms in this Si ICL whose distance to silicon host atoms contributed the additional peak between the Si-Si first neighbor and the second neighbor in both the Si and Si/S heterolayers on GaAs (110). Since such a distance is longer than that for Si-Si and there were large amounts of Ga(As) atoms in the heterolayer, it is reasonable to assume that this peak was contributed by the Si-Ga(As) neighbors. Such a distance and the Si-Si neighbor distances in these heterolayers can be calculated from the FT data based on the experimental phase function, $\Phi(k) = \sin(2kr + \rho_{ab})$, where k is the photoelectron wave vector and $\rho_{a,b}$ is a theoretical atomic pair phase function for a Si-Si absorber-scatter pair.[16] The measured absorption coefficient, $\chi(k)$, was weighted with k^n (n=1,2,3) for different regions of the data in k space. By filtering and back-transforming the FT data, the calculations showed that the first shell Si-Si nearest interatomic distances in the Si ICL are 2.43 Å with sulfur passivation and 2.37 Å without sulfur passivation (compared to 2.35 Å for the first neighbor distance in a Si (100) single crystal). The distances calculated from the foreign atom contributed peak using a k^2 weight is 3.34 Å for the sample with sulfur passi-

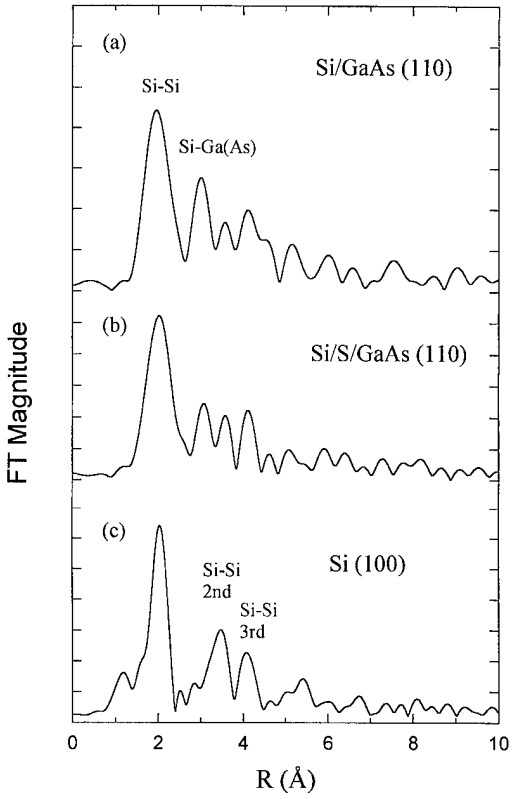


FIG. 5. Fourier transformed data for the corresponding spectra showed in Fig. 4.

vation and 3.32 Å for the sample without sulphur passivation. Slightly longer distances for the Si-Si and Si-Ga(As) in the heterolayer with sulfur passivation could be due to either the trace amount of sulfur incorporation during the break-up of the Ga-S bonds or the defects in these layers. The data shown in Fig. 5 also indicated that for the heterolayers with and without sulfur passivation the Si-Si bonds were still dominating but the relative amplitude for the Si-Ga(As) in the heterolayer with sulfur passivation was lower compared to that without sulfur passivation. This observation supported the results measured by MEIS that sulfur passivation can further suppress As precipitation.

If the improvement of the *CV* characteristics of the MIS capacitors whose interface was modified by the Si ICL and the sulfur passivation was only interpreted as being caused by less As precipitation at the insulator/GaAs interface, and the lack of long-range atomic abruptness or the Si-Ga mixing in the Si ICL is to be blamed for the limitation of further interface state density reduction, one could expect that Si/Ge/S multilayer passivation could be a better choice since Ge is better lattice-matched to GaAs and the pseudomorphic growth of Si on Ge is readily achievable. This had been confirmed by the results by Falta *et al.* where Ge growth on an epitaxial GaAs layer led to a better interface than that by Si growth.[17] The growth on the GaAs also led to surface segregation of both Ga and As; however, the mixing of the heterolayer was not an issue. This difference could come from the dopant modulated growth. We have applied the Si/Ge/S approach to both *n*- and *p*-type GaAs (110) substrates.[8] The *CV* characteristics of the MIS capacitors made on the *n*-type substrate indeed improved impressively but the results on the *p*-type substrate was just the opposite to that of Si/S passivation. Our subsequent interface atomic abruptness measurements by x-ray scattering showed surprisingly that the Si/Ge/S passivation led to a smoother interface for the MIS capacitors on the *p*-type substrate than that on the *n*-type substrate but yet had worse *CV* characteristics.[18] The atomic abruptness, therefore, does not seem to be the very factor which determines the *CV* characteristics of the MIS capacitors discussed here. Our preliminary results[18] using x-ray absorption near-edge structure measurement and x-ray photoemission spectroscopy showed that the band gap of the Si-Ga mixed layer was much wider than that for the bare Si ICL or Si/Ge multilayer. We speculate that band gap engineering might be a key for further improvement of the interface properties of the MIS capacitors made on GaAs. Additional work towards this direction is currently under way.

IV. SUMMARY

In summary, we found that using Si/S or Si/Ge/S multilayer interface passivation for the MIS capacitors made on GaAs (110) could improve the *CV* characteristics of these capacitors. Both Si ICL and sulfur passivation can suppress As precipitation at the interface, and such an effect is substantial by applying Si and S or Si/Ge and S at the same time. However, the sulfur passivation could not prevent the Ga(As)-Si site exchange or Ga(As) from incorporation into the Si ICL and the mixture was dominated by the Ga incorporation.

ACKNOWLEDGMENTS

The authors thank Dr. K. H. Tan of the Canadian Synchrotron Radiation Center for his assistance in some synchrotron radiation measurements. This work was supported by the Natural Sciences and Engineering Council of Canada. Support from Bell Northern Research, NATO collaborative research Grant No. CRG.950125, and the Ontario Center for Materials Research are also greatly appreciated.

[1]Z. Wang, M. E. Lin, D. Biswas, B. Mazhari, N. Teraguchi, Z. Fan, X. Gui, and H. Morkoc, Appl. Phys. Lett. **62**, 2977 (1993); D. S. L. Mui, D. Biswas, J. Reed, A. L. Demirel, S. Strite, and H. Morkoc, *ibid.* **60**, 2511 (1992).
[2]J. L. Freeouf, D. A. Buchanan, S. L. Wright, T. N. Jackson, and B. Robinson, Appl. Phys. Lett. **57**, 1919 (1990).
[3]H. Hasegawa, M. Akazawa, H. Ishii, and K. Matsuzaki, J. Vac. Sci. Technol. B **7**, 870 (1989).
[4]A. Callegari, P. D. Hoh, D. A. Buchanan, and D. Lacey, Appl. Phys. Lett. **54**, 332 (1989); A. Callegari, D. K. Sadana, D. A. Buchanan, A. Paccagnella, E. D. Marshall, M. A. Tischler, and M. Norcott, *ibid.* **58**, 2540 (1991).
[5]A. Paccagnella, A. Callegari, J. Batey, and D. Lacey, Appl. Phys. Lett. **57**, 258 (1990).
[6]R. W. Grant and J. R. Waldrop, J. Vac. Sci. Technol. B **5**, 1015 (1987).
[7]C. J. Sandroff, M. S. Hedge, L. A. Farrow, C. C. Chang, and J. P. Harbison, Appl. Phys. Lett. **54**, 362 (1989); C. J. Sandroff, R. N. Nottenburg, J. C. Bischoff, and R. Bhat, *ibid.* **51**, 33 (1987).

[8]L. J. Huang, K. Rajesh, W. M. Lau, S. Ingrey, D. Landheer, J. P. Noel, and Z. H. Lu, J. Vac. Sci. Technol. A **13**, 792 (1995).
[9]Z. H. Lu, D. Landheer, J. M. Baribeau, L. J. Huang, and W. M. Lau, Appl. Phys. Lett. **64**, 1702 (1994).
[10]J. Reed, Z. Fan, G. B. Gao, A. Botchkarev, and H. Morkoc, Appl. Phys. Lett. **64**, 2706 (1994).
[11]D. Landheer, Z. H. Lu, J. -M. Baribeau, L. J. Huang, and W. M. Lau, J. Electron. Mater. **23**, 943 (1994).
[12]L. J. Huang, W. M. Lau, H. T. Tang, I. V. Mitchell, W. N. Lennard, P. J. Schultz, and M. Kraisa, Phys. Rev. B **50**, 18 453 (1994).
[13]J. L. Freeouf and J. M. Woodall, Appl. Phys. Lett. **39**, 727 (1981).
[14]W. K. Chu, J. W. Meyer, and M.-A. Nicolet, *Backscattering Spectroscopy* (Academic, New York, 1978).
[15]H. Xia, W. N. Lennard, L. J. Huang, W. M. Lau, D. Landheer, J. M. Baribeau, and S. Ingrey (unpublished).
[16]P. A. Lee, P. H. Citrin, P. Eisenberger, and B. M. Kincaid, Rev. Mod. Phys. **53**, 769 (1981).
[17]J. Falta, M. Copel, F. K. LeGoues, and R. M. Tromp, Phys. Rev. B **47**, 9610 (1993).
[18]L. J. Huang *et al.* (unpublished).

Structural and electronic properties of Sb islands on GaAs (110)

R. Magri,[a] F. Manghi, and C. Calandra
Instituto Nazionale per la Fisica della Materia and Dipartimento di Fisica, Università di Modena, Via Campi 213/A, I-41100 Modena, Italy

(Received 22 January 1996; accepted 22 April 1996)

We present the results of theoretical calculations of the electronic structure of the Sb/GaAs(110) interface in the submonolayer coverage regime performed in the full *ab initio* self-consistent pseudopotential scheme. Different structural models for the edges of the extended Sb islands have been considered and their equilibrium geometry has been determined by total energy minimization. The single particle band structure shows interface states arising mainly from the incomplete bonding of the Sb adatoms at the terrace edge to the substrate, which fall within the optical gap. Distinct features are associated to different island terminations. The interface turns out to be metallic in all the considered cases with a partially occupied peak at the Fermi level. We have also studied the effect of including explicitly the on-site Hubbard electron-electron correlation in the calculation of the quasiparticle spectrum, obtaining the observed semiconducting interface when the Coulomb interaction parameter U is larger than 3 eV. The interface states within the optical band gap can be present also at higher coverages when some disorder exists, evidentiating a general mechanism for the Fermi level pinning at this interface. © *1996 American Vacuum Society.*

I. INTRODUCTION

The interface between Sb and the GaAs(110) surface can be considered the prototypical example of a metal-semiconductor interface as the deposition of one Sb monolayer on GaAs is known to produce a well characterized 1×1 ordered structure.[1] Theoretical calculations have been performed in the first-principles pseudopotential approach[2–4] and in the semi-empirical tight-binding scheme[5] assuming in both cases a full monolayer deposition with Sb atoms periodically arranged along zigzag chains following the same geometry of the Ga and As atoms at the topmost layer of the GaAs(110) surface. This structure has been denoted as *epitaxial continued layer structure* (ECLS) and has provided a description of the electronic structure of the full monolayer deposition consistent with the experimental evidence deduced from direct[6] and inverse[7] photoemission spectroscopy and by scanning tunneling microscopy (STM) and spectroscopy measurements.[8] In the ECLS the Sb-derived states have been found in the valence band and close to the bottom of the conduction band; the absence of interface states in the band gap implies that the ordered Sb monolayer cannot alter the flat band condition at the GaAs surface.

Photoemission measurements[9,10] have shown that the Fermi level pinning in this system is strongly correlated to surface disorder: thermal annealing has been shown to alter the Fermi level position and a greater band-bending has been found for unannealed (disordered) samples. Moreover in the submonolayer deposition regime STM measurements[8] have found electronic states within the optical gap which are localized near the edges of the Sb terraces. These states are assumed to be responsible for the Fermi level pinning at the Sb/GaAs(110) interface. The band-gap states were found on samples whose Sb coverage was about 0.7 ML. They form two separate bands, one (seen only in some spectra) filled and the other (observed in most spectra) empty. The surface Fermi level turns out to be pinned between the two bands. The states within the optical gap were observed only when the spectrum was acquired at the edges of the Sb terraces and probably arise from dangling bonds localized near the edge of the islands but their precise structure is unknown.

The shape of these features may change at different locations near the edge of the terraces but the presence of a large peak near the top of the band gap and of a minimum of the measured density of states at the Fermi level seem to be a general result. These features have been observed also for other metal adsorbates at the submonolayer coverage regime such as Au/GaAs(110), Fe/GaAs(110), and even in the case of a simple two-atom cluster of Sb on GaAs(110).[11] The tails of states of these features are degrading toward the minimum at the Fermi level, in some cases giving rise to a small gap, in other cases there is almost no gap.

Little is known about the submonolayer case. Recent theoretical calculations[2] seem to confirm the experimental observation[12] that in the submonolayer coverage regime antimony tends to form flat islands, approximately one monolayer thick with a local 1×1 symmetry. However, no calculations have been performed so far for extended Sb islands. Preliminary tight-binding calculations of the band structure of the half monolayer 2×1 missing chain model (every second $\bar{1}10$ Sb chain is missing in a ECLS) did not show any electronic state within the optical band gap.[13]

In this article we present a theoretical investigation of the electronic structure of the Sb/GaAs(110) interface at submonolayer coverages in order to identify the interface states associated with the incomplete bonding of Sb atoms to the substrate as possible candidates for pinning states. We show that electronic states within the optical gap are indeed obtained when the Sb zigzag chains are truncated. They are

[a]Electronic mail: magri@c220.unimo.it

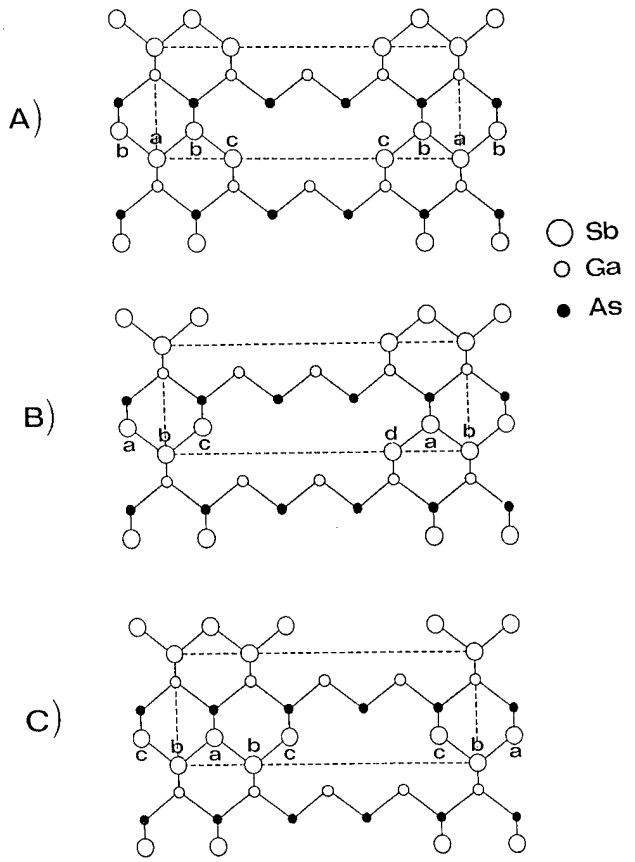


FIG. 1. Top view of the Sb/GaAs(110) interface for the submonolayer coverage: (A) θ=0.625 ML deposition with Sb terraces terminated with Sb on Ga atoms, structure A; (B) half monolayer deposition, structure B; and (C) θ=0.625 ML deposition with Sb terraces terminated with Sb on As atoms, structure C. a, b, c, and d indicate the inequivalent Sb adatoms for each structure.

associated mainly to Sb dangling bonds, however, the dangling bonds of substrate atoms contribute somewhat to the band-gap bands. The states are strongly localized near the Sb adatoms at the terrace edges. Our first-principles calculations predict metallic interfaces, with a partially occupied peak at the Fermi level for all the considered geometries.

II. COMPUTATIONAL DETAILS

As shown in Fig. 1 we have considered different cases corresponding to different coverages: a deposition of 0.625 ML [cases (A) and (C)], realized by truncating the Sb chains and eliminating 3 over 8 Sb adatoms; a deposition of half monolayer [case (B)] where 4 over 8 adatoms have been eliminated. Since Sb atoms bonded to Ga and As substrate atoms are not equivalent, the configuration corresponding to 0.625 ML can be obtained in two different ways, with terraces terminated with Sb atoms bonded to Ga substrate atoms [case (A)] or with Sb atoms bonded to As substrate atoms [case (C)]. These geometries correspond to 4×1 surface unit cells.

In the ECLS there are two inequivalent Sb adatoms in the surface unit cell: Sb on Ga atoms of the substrate, Sb(Ga), and Sb on As atoms of the substrate, Sb(As). In the structures A, B, and C of Fig. 1 there are, respectively, 3, 4, and 3 inequivalent Sb adatoms, since we have to add to Sb(Ga) and Sb(As) also the Sb adatoms at the terrace edges, which we indicate in the following as Sb_e(Ga) and Sb_e(As). In Fig. 1 these inequivalent Sb adatoms are indicated with the letters a, b, c, and d.

We calculated the equilibrium geometry for each structure by performing full *ab initio* total energy calculations in the pseudopotential approach,[14,15] using the density functional theory[16] in the local density approximation (LDA). The Ceperley-Alder form of exchange and correlation energy has been used.[17,18] The Sb covered GaAs(110) surface has been modelled by a repeated slab. Each slab contained five atomic (110) layers of GaAs plus two Sb overlayers on both sides of the slab. The vacuum region had a thickness of three atomic layers. The equilibrium geometry has been determined by a total energy converged to less than 1 meV per atom and forces on the atoms less than 0.01 eV/Å. The search of the minimum energy configuration has been performed using a plane wave cutoff 8 Ry, while the electronic structure has been calculated using an higher cutoff, 11 Ry.

This computational recipe has provided well converged geometries for the fully relaxed clean GaAs(110) surface and the ECLS, which are in good agreement with the experimental results[19–22] and other previous calculations[2,5,23,24] and for their band structures.

In the case of the three structural models of Fig. 1, we have found that the Sb atoms lower toward the substrate. The shortening of the distance of the adatoms from the substrate is very pronounced for the adatoms at the terrace edge, that move away toward the free surface area, trying to saturate their partially occupied dangling bonds by forming new bonds with other substrate atoms, restoring in this way the threefold coordination they had originally in the zigzag chain before truncation. However, probably because the model islands are only few atoms wide in the $\bar{1}10$ direction, these new bonds with the substrate are not completely formed and the saturation of the dangling bonds is only partial. The buckling between the inequivalent Sb atoms turns out to be enormously enhanced with respect to the ECLS case.

III. RESULTS AND DISCUSSION

A. Electronic states at the overlayer terrace edges

In Fig. 2 the total density of states (DOS) of structures A, B, and C are compared with the total DOS of the ECLS. The DOS have been calculated for the optimized minimum energy atomic configurations using the plane wave cutoff 11 Ry and sets of 27, 9, and 10 equivalent **k** points for the ECLS, the A(C) structures and the B structure, respectively. To compare the positions of the peaks we have aligned the onsets of the fundamental GaAs related structures. As a consequence the entire DOS of the ECLS has been rigidly shifted 1.2 eV towards a lower energy, while the DOS of the B structure has been shifted 0.4 eV towards a higher energy with respect to

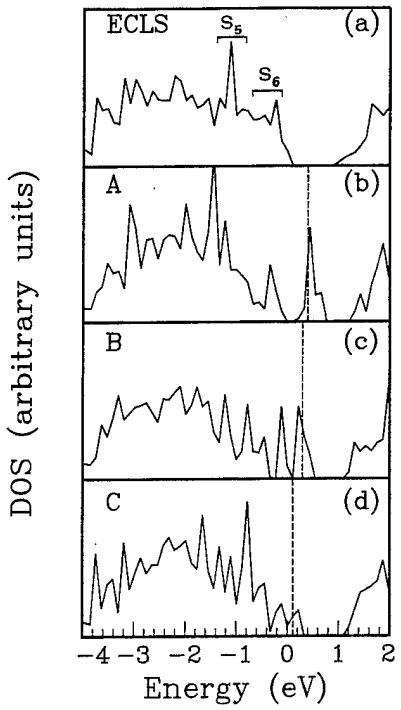


FIG. 2. Total DOS of the: (a) ECLS, (b) A, (c) B, and (d) C structures in an energy region around the optical bandgap. The calculated Fermi levels have been indicated by dashed lines. The zero of the energy axis is at the VBM of the ECLS.

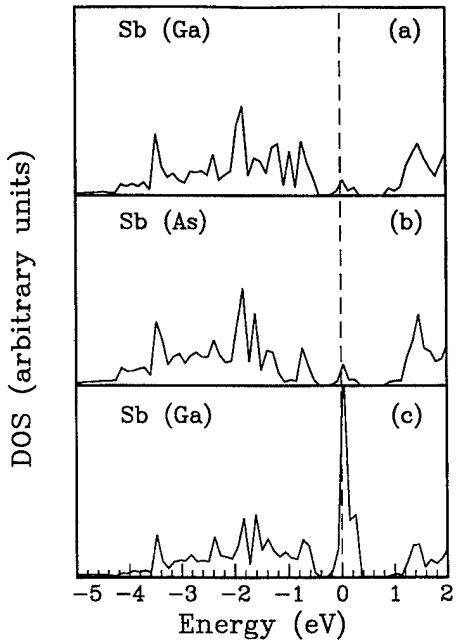


FIG. 3. LDOS of the inequivalent Sb surface sites in structure A. a, b, and c refer to the Sb adatoms (see Fig. 1).

the DOS of the A and C structures. The DOS are depicted in an energy region including the band-gap region of the ECLS.

It is possible to recognize in the DOS of the ECLS the surface S_5 and S_6bands, which are attributed to the bonding of the Sb chains to the substrate. The S_5 band (related to the Sb—As bonds) contributes mostly to the peak at about −1.1 eV below the VBM, while the features lying between −1.0 eV and the VBM can be attributed to the contribution of the S_6 band (related to the Sb—Ga bonds). Obviously, since there are more inequivalent Sb adatoms in the A, B, and C structures, and they are in different relaxed configurations with respect to the ECLS, more peaks appear in the DOS and the contribution of the Sb-substrate bonds is much more widespread over the energy region ranging from −1.5 eV to the band-gap peaks.

We can see that the effect of truncating the Sb chains is mainly to introduce prominent features in the region of the optical band gap. The peaks have different shapes and positions passing from structure A, to structures B and C.

Structure A [case (b)] presents a well defined peak about 0.2 eV above the top of the broad valence-band feature where another sharp feature appears, much more distinct than in the case of the ECLS. In the case of structure B [case (c)] the gap separating the band-gap states from the valence-band states is smaller, about 0.1 eV. The band-gap states in this case are organized in two sharp peaks separated by a very small gap. Finally, in case of structure C there are no distinct gaps between the band-gap states and the valence band maximum. They form an extension of the broad valence-band feature and the peaks are much smaller than in the previously considered cases.

For all the considered geometries the electronic states appearing in the band-gap region are only partially occupied, predicting metallic interfaces.

To gain insight about the nature of these states we analyzed the atom and angular momentum resolved local density of states (LDOSs) and the electronic charge density distribution associated to them. In Figs. 3–5 the calculated LDOS of the inequivalent Sb adatoms in structures A, B, and C, respectively, are depicted. From the analysis of these results we can deduce that a big contribution to the partially occupied peaks at the Fermi level in the band-gap region comes from the Sb adatoms at the terrace edge. For example in the case of structure A, the p states of planar x and z character (see below), which give rise to the p^2 planar bonding between Sb adatoms in the chain become dangling bonds after the chain truncation. These dangling bonds contribute up to the 50% of the prominent peak at the Fermi level and are by far the predominant contribution. The p states related to the Ga and As dangling bonds of the free substrate also contribute to the peak, but their contribution is much smaller, only few percent. We can also interpret these states in terms of one-dimensional Tamm states, showing weight predominantly on the last atom of the chain and decaying into the body of the chain.

We can understand the outcome of a metallic interface for all the considered geometries through a simple electron counting rule. We consider for simplicity an unrelaxed *ideal* configuration. Each Sb adatom at the terrace edge has five valence electrons. In an unrelaxed position it forms only two bonds, one with the substrate and the other with only one other Sb adatom. Two dangling bonds (if we still assume a

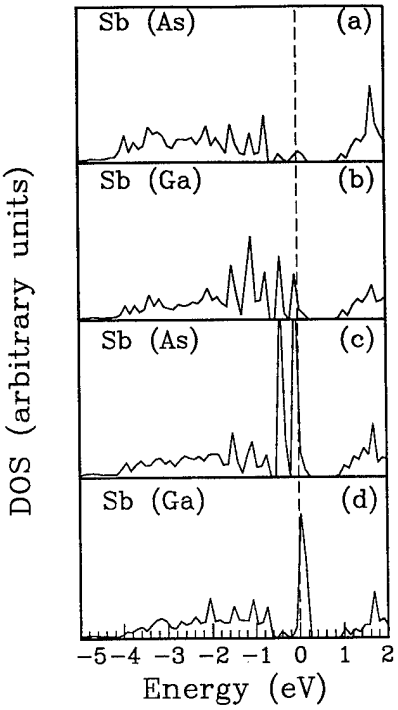


FIG. 4. Same as Fig. 3 for structure B.

tetrahedral configuration around the atom, which is approximately true) have to be occupied by three electrons, which results in one dangling bond only half occupied.

The situation is only slightly modified if we include in the description the atomic relaxation. In the case of structures A and C, the Sb adatoms at the terrace edges are both on Ga (structure A) or on As (structure C) atoms, and for symmetry

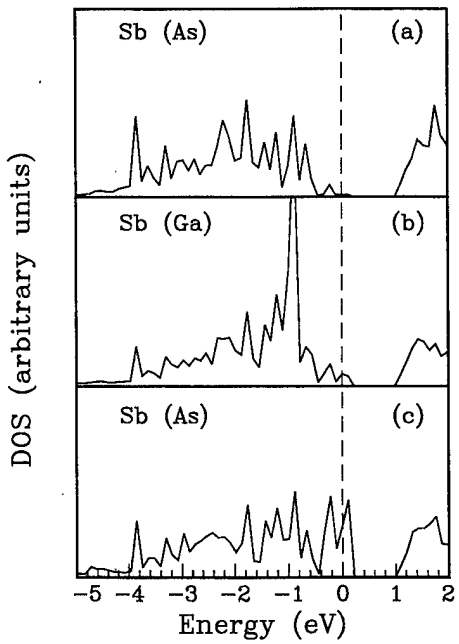


FIG. 5. Same as Figs. 3 and 4 for structure C.

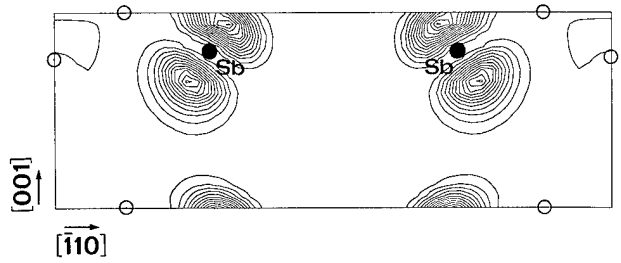


FIG. 6. Contour plots of the valence charge density relative to the electronic states within the intra-gap peak of structure A plotted along the surface plane. The contour step is 15 in units of electrons per unit cell volume.

their energies have to be the same. Therefore even in this case we should have partially occupied dangling bonds. Moreover the atoms of the GaAs substrate, in its free portion, are only partially relaxed. We can suppose that since these substrate atoms are not fully relaxed, their dangling bonds overlap the band region and contribute mostly to the states at the band edges and somewhat to the band-gap peaks[25] (indeed, in the A and C structures, the dangling bonds of the As atoms in the substrate are not anymore in the same number of the dangling bonds of the Ga atoms).

The situation is different for the configuration B, where the Sb adatoms at the terrace edges are one on a Ga substrate atom, the other on an As atom. The energies of the two dangling bonds can be now different and, in theory, one could be fully occupied and the other empty. In practice the interplay with the dangling bonds of the free substrate and the only partial saturation of the Sb dangling bonds with the substrate gives rise again to a metallic interface.

The second important observation about the LDOS of the Sb adatoms, Figs. 3–5, is that the terrace edges on As atoms produce two distinct peaks, while those on Ga atoms produce only one peak at higher energy. This can be concluded by comparing Fig. 3(c) with Fig. 4(d) and Fig. 5(c) with Fig. 4(c). The identification of the band-gap peaks has been obtained by inspection of the electronic charge density distribution associated to the states falling in this energy region. The assignment of these DOS features to band-gap states has been easily performed because of the very characteristic nonbonding and spatially localized nature of these states. As an example we show in Fig. 6 and Fig. 7 the contour plots of the electronic charge density associated to the states within the prominent band-gap peak in the case of structure A [see also Fig. 2(b)], plotted along the surface plane and along a ($\bar{1}$10) plane perpendicular to the surface and passing through the Sb_e(Ga) atoms, respectively. These plots reveal the nature of these states as nonbonding states, strongly localized at the terrace edges with the typical shape of dangling bonds of p character with maxima of charge pointing out of the surface. Identical results have been obtained in the case of the band-gap states of structures B and C. The investigation of the electronic charge distributions associated to the double peak of the Sb(As) termination and the single peak of the Sb(Ga) termination reveals that the nature of these states is essentially the same. The fact that, in the case of the Sb (As)

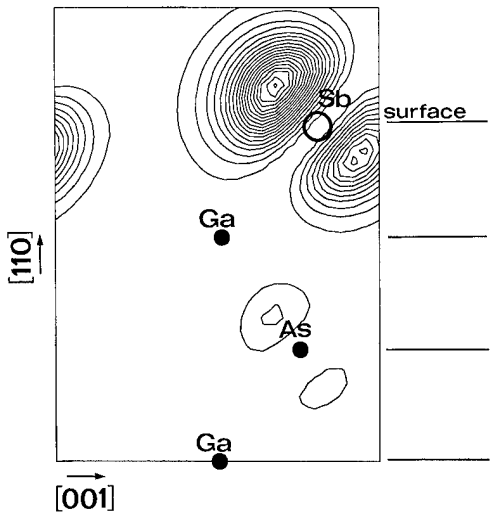


FIG. 7. Same as Fig. 6 but plotted along the $(\bar{1}10)$ plane perpendicular to the surface plane passing through the Sb_e(Ga) atoms. The contour step is 15 in units of electrons per unit cell volume.

termination, the states give rise to a doublet is probably due to a different dispersion of the four bands derived from the four terrace edges in the unit cell (two on each side of the slab). In the limit of an infinite distance between the terrace edges, these four bands would be degenerate with a single peak well within the optical band gap. The single peak would be closer to the VBM in the case of the Sb (As) termination. However, because of the finite (and relatively small) dimension of the slab used to model the system, these states interact between them and split in their bonding and antibonding combinations. In the case of the Sb (As) termination these four bands are very close to (even slightly overlapping) the states at the top of the valence band. The interaction with these other states influences further their dispersion, giving rise to the double peak.

Concluding this subsection we have found electronic states within the band gap when the Sb surface chains of the ECLS are truncated. These electronic states are organized in distinct peaks which can be assigned mainly to dangling bonds of Sb adatoms bonded to Ga and As atoms of the substrate, respectively, and are strongly localized on the terrace edges, in agreement with the experimental result. However our first-principles calculations predict metallic interfaces for all the considered configurations, while the experiment shows a semiconducting interface with a very small gap between the filled and the empty bands. In particular we have been unable to individuate the prominent empty peak shown in Fig. 4 of Ref. 8. One possible reason could be that the atomic relaxation of the Sb adatoms has been limited by the small dimension along the $[\bar{1}10]$ direction of our model islands. In a more realistic (and large) island the position of the Sb adatoms at the edge could give rise to a more effective saturation of their dangling bonds with the dangling bonds of the substrate, changing their energy. Again, it is possible to think of some geometry in which the interplay between edges on Ga atoms and edges on As atoms could create a new phase at the island boundary leading to electronic states similar to those seen in the experiment.

Another explanation could be related to the following observations:

(a) The band-gap states are highly localized in the space, nonbonding, and because of the lower atomic coordination of the adatoms at the terrace edges they are more similar to atomic p states.
(b) The LDA method, like all the theories based on a mean field approach does not take into a proper account the dynamical fluctuations of the orbital occupation and often provides an inaccurate description of the dynamical correlation between electrons localized on the same atom. Because of the high spatial localization of the terrace-edge states this effect might be not negligible.
(c) The experiment probes excited states with $N-1$ and $N+1$ particles. The experimental result should therefore be compared more correctly with a quasiparticle spectrum instead of the spectrum of the eigenstates of the ground state of the system provided by the LDA theory.

And, most importantly:

(d) The experimental results seem to be independent of the actual geometry of the terrace edge.

The hypothesis that the Coulomb repulsion U between electrons in the partially occupied band-gap states could play a substantial role has been suggested also in Ref. 11. Theoretical work on Au/GaAs(110)[26] and on alkali-metal adsorbates on GaAs(110)[27] has given an explanation of the experimental results in terms of *negative-U* systems, in which doubly occupied levels are energetically favored.

In this article we try, instead, to interpret the experiment on the Sb/GaAs(110) system in the submonolayer coverage regime in terms of a Mott–Hubbard system. We have calculated a quasiparticle spectrum where the electron-electron correlation beyond the mean-field approach of the LDA has been treated explicitly through an Hubbard term. We will illustrate this calculation in the next subsection, applying the method to the case of structure A.

B. Effects of the inclusion of the on-site electron-electron correlation

The Hubbard hamiltonian has often been used to modify single particle eigenvalues, but most of the studies are based either on a crude description of the single particle states or on simplified approaches to treat the electron-electron correlation. Starting from the first-principles single particle results we obtained the quasiparticle spectrum for electron and hole states adopting a multiband extension of a recent approach which treats the on-site Hubbard repulsion beyond the mean field theory.[28] The method allows to include both the full details of the single particle band structure and to treat correlation effects through a non-perturbative solution of the Hubbard hamiltonian. In this approach the many-body states of the interacting system are expanded in terms of uncorrelated states with a different number of electron-hole pairs.

The expansion is here truncated to include three-body interactions. The interactions between configurations with the same number of hole-electron pairs are represented by a set of two-body scattering t-matrices, which can be exactly calculated. The Fadeev theory[29] is used to find the total scattering matrix and, ultimately, the resolvent giving the energy of the many-body system and the quasiparticle spectrum.

In this theory the quasiparticle states are given in terms of the spectral functions for hole and electron states as

$$D^{\pm}_{\mathbf{k},n}(\omega) = -\frac{1}{\pi}\mathrm{Im}\frac{1}{\omega - \varepsilon_{\mathbf{k},n} - \Sigma^{(\pm)}_{\mathbf{k},n}(\omega)}, \tag{1}$$

with $\omega \leqslant \varepsilon_f$ for $(-)$ and $\omega \geqslant \varepsilon_f$ for $(+)$ and

$$\Sigma^{(\pm)}_{\mathbf{k},n}(\omega) = \sum_{\beta} |c^n_{\beta}(\mathbf{k})|^2 \left[\mp \sum_{\alpha} U_{\alpha,\beta} N^{h,e}_{\alpha} \pm \Sigma^{\pm}_{\beta}(\omega) \right], \tag{2}$$

for the hole $(-)$ and electron $(+)$ self-energy. Here α and β label different atomic orbitals, $\{c^n_{\beta}(\mathbf{k})\}$ the orbital coefficients of the single particle state $(\mathbf{k},n)$, $U_{\alpha,\beta}$ the Coulomb repulsion between electrons on orbitals α and β on the same site, $N^{h,e}_{\alpha}$ is the percentage of empty (filled) states of orbital character α over the total. Ultimately we calculate the total density of quasiparticle states as

$$D^{\pm}(\omega) = \sum_{\mathbf{k},n} D^{\pm}_{\mathbf{k},n}(\omega). \tag{3}$$

A detailed description of the method is given in a previous paper.[28] The inputs of the calculation of the quasiparticle spectra are simply an orbital resolved single particle density of states, the single particle eigenstates, and the values of the intra-site correlation parameter $U_{\alpha,\beta}$. In order to obtain the orbital composition of the crystal wavefunctions $\psi_{\mathbf{k},n}$ from the first-principles pseudopotential calculations, (which are expanded in plane waves), we calculated their projections onto s, p, and d symmetry atomic orbitals. The projection coefficients are given by[30]

$$C(\tau_i, lm, n\mathbf{k}) = \int_0^{R_{ws}} \langle \psi_{n\mathbf{k}} | Y_{lm} \rangle \langle Y_{lm} \psi_{n\mathbf{k}} \rangle r^2 \, dr, \tag{4}$$

where the integration extends over the Wigner–Seitz radius R_{ws} centered at the lattice site τ_i and Y_{lm} are the spherical harmonics. Brackets imply integration over angular coordinates. The projection coefficients are normalized such that

$$\sum_{\tau_i} \sum_{lm} C(\tau_i, lm, n\mathbf{k}) = 1, \tag{5}$$

and symmetrized over the star of the k-point. Through the projection coefficients we are able to obtain the orbital resolved single particle density of states $n_{\tau_i, lm}(\varepsilon)$,

$$n_{\tau_i, lm}(\varepsilon) = 2 \sum_{n\mathbf{k}} \delta(\varepsilon - \varepsilon_{n\mathbf{k}}) C(\tau_i, lm, n\mathbf{k}). \tag{6}$$

We have found by inspection of the electronic charge distribution that only the electronic states within the optical band gap are highly localized in the space (apart the nonbonding s states of the same Sb adatoms whose energies are however, much lower and are not the objects of our investigation). We assumed that the on-site electron-electron correlation has not been adequately described in the framework of the LDA only for these states. These states have a planar p_x (p_y), and p_z character. Figure 8 shows the angular momentum decomposed LDOS of a Sb_e(Ga) atom relative to the p symmetry states. Only the planar p_x (p_y for the other terrace edge atom in the surface unit cell) and p_z states contribute substantially to the band-gap peak, while their weight on other energy regions is quite small. The other p_y (p_x) orbital contributes mainly to the energy region corresponding to the top of the broad valence feature. The electronic charge distribution associated to this energy region (Fig. 9) reveals indeed bonding states with maxima of charge between the Sb_e adatoms and the substrate Ga atoms and between the As and Ga atoms of the substrate. The electronic charge is much less localized (the contour step in figure is 6) than in the case of the electronic states lying within the optical gap shown in Fig. 7 (where the contour step is 15).

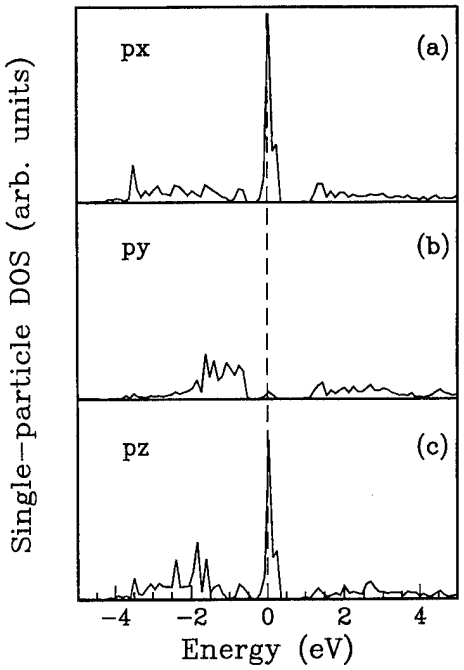


FIG. 8. Decomposition of a Sb_e(Ga) adatom p states in the p_x, p_y and p_z components. The dashed line indicates the position of the Fermi level.

Since the key quantity measuring the "strength" of the electron-electron correlation is the ratio U/W (W being the orbital band width), we have included electron-electron correlation only on those states with a narrow bandwidth, that is between the p_x (p_y) and p_z states. We assume a model planar Hamiltonian,

$$\hat{H} = \hat{H}_{mf} + \sum_{i,j=1}^{2} U_{i,j} \hat{n}_{i,\sigma} \hat{n}_{j,-\sigma}, \tag{7}$$

where i,j run over the planar p states of Sb_e(Ga) and $\hat{H}_{mf}$ indicates the density functional, LDA, hamiltonian which has been self-consistently solved and treats all the correlation effects in a sort of mean field approach. Because the value of

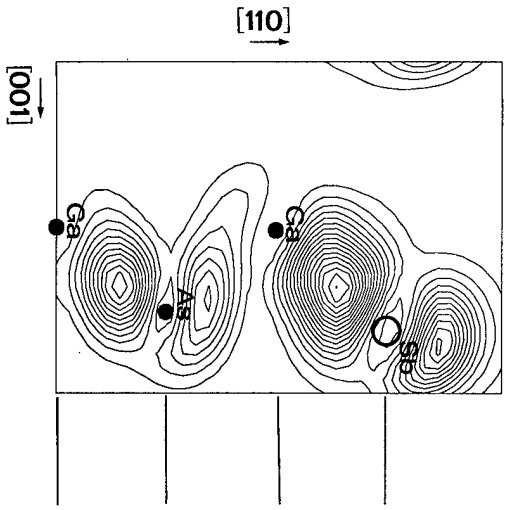


FIG. 9. Contour plots of the valence charge density of the electronic states falling in the peak at the top of the broad valence feature in structure A drawn along a $(\bar{1}10)$ plane perpendicular to the surface plane passing through Sb_e(Ga). The contour step is 6 (in unit of electrons per unit cell volume).

the Hubbard Coulomb integral U is reasonably on the same order for all the p orbitals, we take the Coulomb repulsion parameters $U_{i,j}=U$ for all i,j.

Figure 10 shows our results for $D^{\pm}(\omega)$, the quasiparticle density of states, when the Coulomb correlation parameter ranges from 0 to 6 eV. Figure 10(a) is the total density of states of Figure 2(b), translated in the formalism of the spectral function [Eq. (1)] for hole and electron states. Increasing

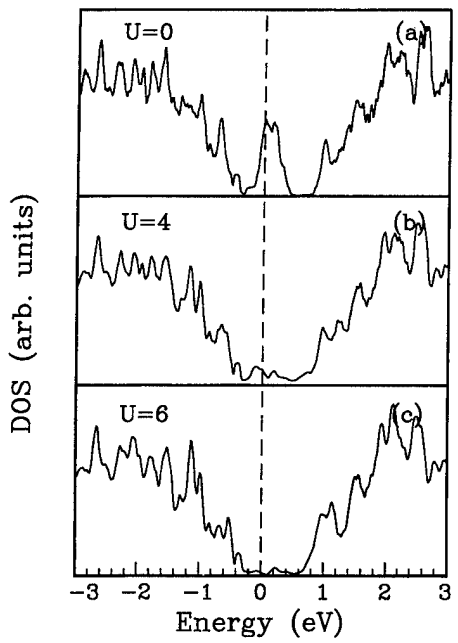


FIG. 10. Density of quasiparticle states in the energy region around the optical gap in the cases: (a) without the Hubbard on-site correlation; (b) with an Hubbard parameter $U=4$; and (c) with an Hubbard parameter $U=6$. The Fermi level position has been indicated by a dashed line.

the Hubbard correlation in the Sb_e(Ga) p symmetry states we can see that the peak at the Fermi level tends to disappear, while the features at the valence and conduction band edges broaden and move toward the Fermi level. The separation between the *valence* peak and the *conduction* peak is about 1 eV, which is in reasonable agreement with the experimental results.

We have analyzed the DOS projected on the atomic sites in order to recognize and assign the band-gap features to those observed in the experimental spectra. As U increases a feature in the calculated DOS broadens at about $E_f+1.0$ eV. This peak can be associated to the substrate dangling bonds, mainly to states on the substrate Ga atom. Moreover the feature in the occupied part of the spectrum which moves closer to the Fermi level is related to terrace edge states pushed down by the Hubbard interaction plus some minor contribution from substrate states. We can explain the result of the calculation in this way: once the localized states are removed from the gap through the Hubbard mechanism, the band gap states are dominated by features associated with substrate states. This could explain the experimental outcome that the peak at about 1.0 eV from the Fermi level is substantially independent from the actual geometry of the adsorbate while the tail of states in the valence band seems to be more affected by it.

Considering that the model applied is very simple and the correlation effects have been taken into account only in the case of two orbitals, which constitute only half of the wave functions of the band-gap states (probably this fact has led to the need to double the value of the parameter U necessary to obtain these results), this result is quite interesting, as we can infer that the on-site electron-electron correlation could play a substantial role in determining the energies of the terrace edge states.

IV. SUMMARY

In conclusion we have studied the Sb/GaAs(110) interface in the submonolayer coverage regime using first-principles band structure methods. Different structural models for the Sb island edges have been examined and we have shown that the pinning of the Fermi level found for Sb deposited on the GaAs(110) surface could be associated to unsaturated Sb and substrate dangling bonds which may as well be present even at a full monolayer coverage when some residual disorder still exist. Terrace edge states have been found within the optical bandgap with distinctive features associated to the possible different terminations of the Sb islands. A better agreement with the experiment about the Fermi level position has been achieved when the effects of the on-site electron-electron correlation have been explicitly included in the calculation.

ACKNOWLEDGMENTS

We acknowledge financial support from MURST and CNR. This work was also partially supported by the Centro Di Calcolo Interuniversitario Dell'Italia Nord-Orientale (CINECA) under the grant "SuperCalcolo."

[1]J. Carelli and A. Kahn, Surf. Sci. **116**, 380 (1982).
[2]F. Bechstedt, W. G. Schmidt, and B. Wenzien, Europhys. Lett. **25**, 357 (1994); W. G. Schmidt, B. Wenzien, and F. Bechstedt, Phys. Rev. B **49**, 4731 (1994).
[3]G. P. Srivastava, Phys. Rev. B **46**, 7300 (1992).
[4]C. M. Bertoni, C. Calandra, F. Manghi, and E. Molinari, Phys. Rev. B **27**, 1251 (1983).
[5]C. Mailhiot, C. B. Duke, and D. J. Chadi, Phys. Rev. B **31**, 2213 (1985).
[6]P. Martensson, G. V. Hansson, M. Lähdeniemi, K. O. Magnusson, S. Wiklund, and J. M. Nichols, Phys. Rev. B **33**, 7399 (1986).
[7]W. Drube and F. J. Himpsel, Phys. Rev. B **37**, 855 (1988).
[8]R. M. Feenstra and P. Martensson, Phys. Rev. Lett. **61**, 447 (1988).
[9]F. Schäffler, R. Ludeke, A. Taleb-Ibrahimi, G. Huges, and D. Rieger, J. Vac. Sci. Technol. B **5**, 998 (1987).
[10]R. Cao, K. Miyano, T. Kendelewicz, I. Lindau, and W. E. Spicer, Surf. Sci. **206**, 413 (1988).
[11]M. Feenstra, Appl. Surf. Sci. **56–58**, 104 (1992).
[12]Martensson and R. M. Feenstra, Phys. Rev. B **39**, 7744 (1989).
[13]Manghi and C. Calandra, Vacuum **41**, 693 (1990).
[14]Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).
[15]R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
[16]Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964).
[17]D. M. Ceperley and B. I. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[18]Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[19]B. Duke, A. Paton, W. K. Ford, A. Kalin, and J. Carelli, Phys. Rev. B **26**, 803 (1982).
[20]K. Ford, T. Guo, D. L. Lessor, and C. B. Duke, Phys. Rev. B **42**, 8952 (1990).
[21]Y. Tong, W. N. Mei, and G. Xu, J. Vac. Sci. Technol. **2**, 393 (1984).
[22]W. Puga, G. Xu, and S. Y. Tong, Surf. Sci. **164**, L789 (1985).
[23]J. Chadi, Phys. Rev. B **19**, 2074 (1979).
[24]X. Qian, R. M. Martin, and D. J. Chadi, Phys. Rev. B **37**, 1303 (1988).
[25]Calandra, F. Manghi, and C. M. Bertoni, J. Phys. C **10**, 1911 (1977).
[26]G. Allan and M. Lannoo, Phys. Rev. Lett. **66**, 1209 (1991).
[27]O. Pankratov and M. Scheffler, Phys. Rev. Lett. **71**, 2797 (1993).
[28]C. Calandra and F. Manghi, Phys. Rev. B **50**, 2061 (1994).
[29]J. Igarashi, J. Phys. Soc. Jpn. **52**, 2827 (1983); **54**, 260 (1985).
[30]S. L. Richardson, M. L. Cohen, S. G. Louie, and J. R. Chelikowsky, Phys. Rev. B **33**, 1177 (1986).

Room-temperature chlorination of As-rich GaAs (110)

William C. Simpson[a)]
Department of Physics, University of California, Riverside, California 92521 and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

David K. Shuh
Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Jory A. Yarmoff[b)]
Department of Physics, University of California, Riverside, California 92521 and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

(Received 22 January 1996; accepted 29 March 1996)

As-rich GaAs (110) is prepared by ion bombardment and annealing, followed by chlorination and reannealing. The surface is then reacted at room temperature with Cl_2 gas and examined with soft x-ray photoelectron spectroscopy of the Ga and As $3d$ core levels. After low exposures ($<5\times10^4$ L), the surface appears to passivate with half a monolayer of Cl adsorbed, primarily as AsCl. Following sufficiently large ($>5\times10^4$ L) exposures, however, the surface begins to etch, as indicated by the continuous uptake of chlorine and the formation of As and Ga chlorides. After the largest exposures, the distribution of As chlorides still favors the monochloride, whereas the Ga chlorides favor $GaCl_2$. It is proposed that the heavily reacted surface is covered with $-AsCl-GaCl_2$ treelike structures. The addition of Cl to form $GaCl_3$ from $GaCl_2$ is identified as the rate-limiting step in the overall etching reaction. © *1996 American Vacuum Society.*

I. INTRODUCTION

It is becoming increasingly important to understand the basic mechanisms involved in the chemical etching of semiconductor wafers as processing technology is applied to III–V substrates. The interactions of Cl_2 with GaAs are particularly significant, since chlorine is the primary chemical etchant involved in GaAs device fabrication. The study of Cl_2/GaAs reactions is complicated somewhat by the fact that many GaAs surface stoichiometries are stable and that the composition and crystallinity of the starting surface affect the path of the reaction.[1,2]

For example, in the room-temperature reaction of Cl_2 with GaAs(110), there is a competition between passivation and etching, as well as between Ga–Cl and As–Cl bonding. Under certain conditions, GaAs(110) is observed to passivate against further reaction with Cl_2 at room temperature, presumably via the formation of an AsCl overlayer.[2] To explain this result, surface order was suggested to be the controlling factor in determining whether or not passivation occurs. On the other hand, soft x-ray photoelectron spectroscopy (SXPS) studies of the room-temperature reaction of Cl_2 with cleaved GaAs (110), which is highly crystalline, observed both Ga and As chloride formation accompanied by continuous chlorine uptake, which is suggestive of etching.[3–5] Moreover, scanning tunneling microscopy (STM)[6] and high-resolution electron energy loss studies (HREELS)[7] of the etching reaction showed that there is a preference, at least initially, for Ga–Cl bonding over As–Cl bonding.

A possible means of reconciling these apparently contradictory results lies in the sample preparation technique used in each study. In Refs. 3 and 5–7, clean starting surfaces were prepared by cleavage in vacuum, whereas in Ref. 4 they were prepared by ion bombardment and annealing (IBA). In contrast, GaAs(110) wafers were prepared in Ref. 2 by IBA followed by exposure to Cl_2 and reannealing to ~535 °C, which likely produced a starting surface different than those in Refs. 3–7.

In this investigation, it is shown that cleaning GaAs(110) with Cl_2 in this manner produces an As-rich surface. The subsequent room-temperature reaction of Cl_2 with such an As-rich surface is then followed with core-level SXPS. At low exposures, the reaction appears to passivate with a half-monolayer (ML) of AsCl formed, as reported in Ref. 2. After sufficiently large Cl_2 exposures, however, further reaction does occur, generating AsCl and $GaCl_2$ as the primary surface products.

II. EXPERIMENTAL PROCEDURE

The SXPS measurements were performed in a multichamber ultrahigh vacuum (UHV) system on beamline UV-8a at the National Synchrotron Light Source. Spectra were collected using a 3 m toroidal grating monochromator and an angle-integrating ellipsoidal mirror analyzer (EMA). Ga $3d$ and As $3d$ spectra were collected with photon energies of 57 and 79 eV, respectively, to generate core-level photoelectrons with kinetic energies of ~35 eV, which maximizes the surface sensitivity of the measurements. The combined resolution of the analyzer and monochromator at these photon energies is better than 150 meV.

Prior to being placed in the UHV system, GaAs(110) wafers (n type, Si doped, 10^{17} cm^{-3}) were chemically etched in a dilute solution of $HNO_3/H_2O_2/H_2O$, rinsed in isopropyl al-

a)Present address: Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352.
b)Corresponding author; Electronic mail: yarmoff@ucrac1.ucr.edu

cohol, and then dried with N_2 gas. The samples were degassed in UHV, then ion bombarded with 500 eV Ar^+ ions and annealed to ~550 °C in order to remove any remaining oxides. As a final step, the wafers were exposed to a small (<100 L) dose of Cl_2 and reannealed to ~550 °C for ~10 min (*note*: 1 L=1 Langmuir=10^{-6} Torr s). Exposure to Cl_2 gas was carried out in a separate UHV dosing chamber (base pressure=8×10^{-10} Torr). Following chlorination, the sample was transferred to the spectrometer chamber (base pressure=2×10^{-10} Torr) for measurement. The sample remained in UHV throughout the cleaning, dosing, transferring, and measuring procedures.

III. RESULTS

The Ga and As 3*d* core-level spectra were numerically fit, via the method outlined in Ref. 4, to a sum of spin-orbit split, Gaussian-broadened Lorentzian line shapes. A Lorentzian width of 0.15 eV was used, as were spin-orbit splittings and branching ratios of 0.45 eV and 0.65 for Ga 3*d* and 0.68 eV and 0.65 for As 3*d*, respectively. In addition to the signal arising from atoms in the bulk, the spectra collected from the clean surface have contributions from surface Ga and As atoms, which are in a tricoordinate geometry, and therefore have different core-level binding energies (BEs) than the bulk atoms. The magnitude and direction of shifts obtained for these surface-shifted core levels (SSCLs) are identical to those previously reported for cleaved GaAs.[3]

Examples of spectra from the reacted surfaces are shown in Figs. 1 and 2, which contain As 3*d* and Ga 3*d* spectra collected after exposures of 1.5×10^3 L and 1.2×10^6 L of Cl_2. In both figures, the filled circles are the background-subtracted raw data, dashed lines are the individual contributions to the spectra from each chemical species, and the solid line is the sum of all these contributions. Except for the low-BE Ga peak, each spectral component is labeled by the chemical species with which it is associated.

In addition to the signal originating from bulk atoms, the spectra collected from the reacted surfaces have several other core-level components, which have been identified previously. The As 3*d* spectra contain features shifted ~0.45 and ~1.2 eV higher in BE than the bulk component, which are identified as arising from AsCl and $AsCl_2$, respectively.[1,3–5,8] Additional Ga 3*d* components are also observed, shifted ~0.9 and ~1.7 eV higher in BE than the bulk component, and are identified as due to GaCl and $GaCl_2$, respectively.[1,3–5]

There are also contributions in the spectra from tricoordinate Ga and As atoms, which have been observed previously in the reactions of Cl_2 and XeF_2 with GaAs.[3,9–11] The production of subsurface tricoordinate Ga and As atoms has been shown to be a consequence of the etching reaction. Because of the valences of Ga and As, tricoordinate configurations are reasonably stable. Ga, with three valence electrons, prefers a tricoordinate bonding geometry, while As, with five valence electrons, has a stable lone pair of electrons remaining when the other three are involved in bonding.

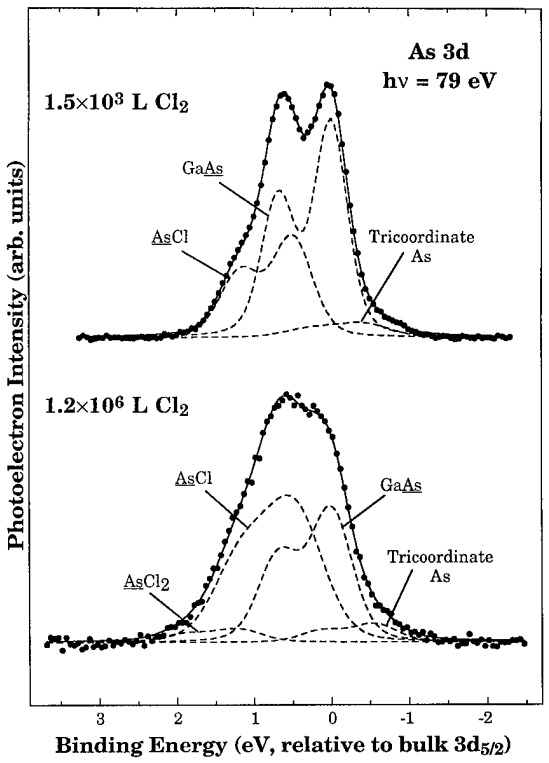


FIG. 1. High-resolution SXPS spectra of the As 3*d* core level collected from GaAs(110) exposed at room temperature to (a) 1.5×10^3 L of Cl_2 and (b) 1.2×10^6 L of Cl_2. Dots are the background-subtracted raw data. The dashed lines are the individual fit components. The solid line is the sum of the fit components.

Finally, for Cl_2 exposures above ~5×10^4 L, an additional component shifted 0.5–0.6 eV lower in BE than the bulk component is present in the Ga 3d spectra (see Fig. 2). A similar shift, 0.5–0.6 eV to lower BE, is observed for Ga surface atoms on clean GaAs(001)-4×6, which persists with chlorination.[1] It is therefore proposed that this low-BE shift is due to Ga–Ga bonds and that Ga dimers, or possibly some other type of Ga cluster, may be responsible for this shift. The persistence of the low-BE shifted component with increasing chlorination may indicate that Ga atoms in such a configuration are reasonably inert to further reaction.

A simple calculation based on the areas of the core-level components provides a quantitative estimate of the coverage of each chemical species present on the surface following reaction. The coverage of each species is simply proportional to its relative contribution to the total core-level intensity, if attenuation within the overlayer and chemically induced variations in the photoionization cross sections are neglected. The proportionality constant is determined straightforwardly by comparing intensity ratios of the reacted surfaces to those of a surface with known composition. This is done by assuming that the ratio of the SSCL area to the total core-level area, in a spectrum collected with the same photon energy from a cleaved GaAs(110) surface, represents the signal from 0.5 ML of surface atoms. Coverages estimated in this manner, using the core-level intensities reported in Ref. 3 for cleaved GaAs(110), are given in Fig. 3.[12] The amount of Cl adsorbed

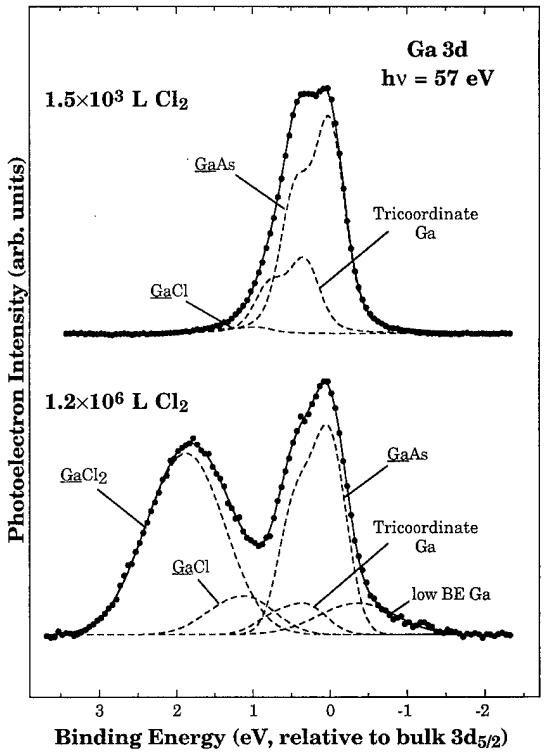


FIG. 2. High-resolution SXPS spectra of the Ga 3*d* core level collected from GaAs(110) exposed at room temperature to (a) 1.5×10^3 L of Cl_2 and (b) 1.2×10^6 L of Cl_2. Dots are the background-subtracted raw data. The dashed lines are the individual fit components. The solid line is the sum of the fit components.

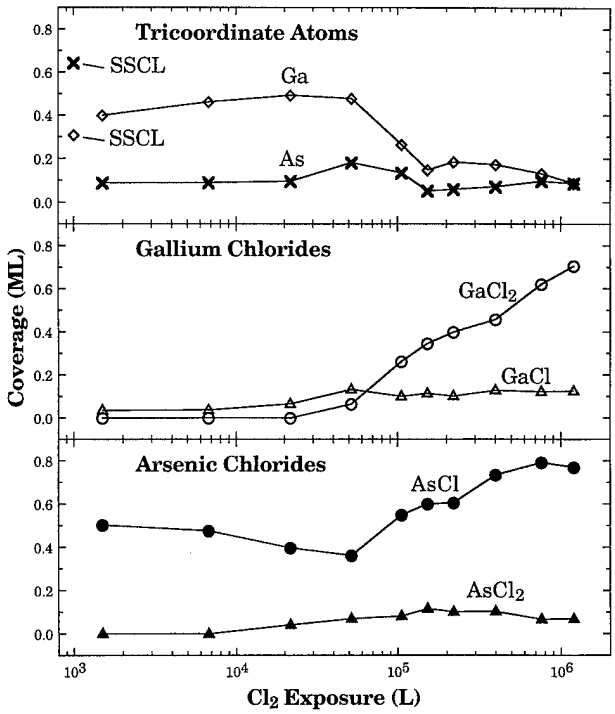


FIG. 3. Coverage of each surface species, in monolayers (ML), as a function of Cl_2 exposure.

on the surface is found by summing the coverages of each of the Ga and As chlorides, appropriately weighted by the number of Cl atoms, and is given in Fig. 4.

IV. DISCUSSION

The starting surface, which was cleaned by Cl_2 adsorption and annealing, is slightly As-rich. This is apparent from the coverage estimates, described above, which indicate that there is ~0.65 ML of surface-shifted As and only ~0.35 ML of surface-shifted Ga on the clean surface. These data are shown at the left side of the top panel of Fig. 3 by the two points labeled "SSCL." The atoms that contribute to the SSCL components are all exposed at the surface, as their total coverage is not more than ~1 ML. Experimentally, it has been shown that up to 1 ML of As can be grown on GaAs(110), and that the excess As is stable on the surface for temperatures below ~575 K.[13] The structure of As-rich GaAs(110) is unknown, but a theoretical investigation suggests that the excess As atoms form chains, bonding to both Ga and As atoms below.[14] Presumably, the Cl_2 adsorption and annealing procedure generates an As-rich surface via the preferential formation and removal of Ga chlorides. Note that, in other systems, halogen adsorption followed by annealing also produces a group-V terminated surface via the preferential removal of group-III halides.[15–20]

From Figs. 3 and 4, it is seen that this As-rich surface is remarkably resistant to chlorine attack. That is, once ~0.5 ML of Cl is adsorbed (primarily as AsCl), there is minimal Cl uptake over an order-of-magnitude increase in Cl_2 exposure. As previously noted, this apparent passivation to chlorine attack is not observed for cleaved GaAs(110),[3,5] nor for sputtered and annealed GaAs(110),[4] but it is seen for sputtered and annealed GaAs(110) wafers that are further cleaned by exposure to Cl_2 and reannealing.[2] Since the cleaved surfaces, which are highly crystalline, show no sign of passivation, surface order does not appear to be the only factor

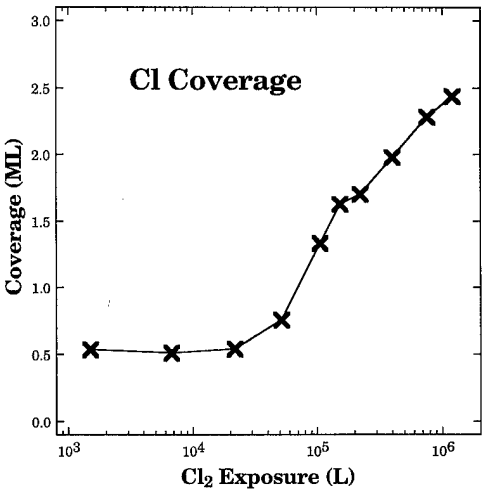


FIG. 4. Cl coverage, in monolayers (ML), as a function of Cl_2 exposure.

governing whether or not the surface passivates. Instead, it is suggested that excess surface As is also involved in impeding the reaction. How this occurs is unclear, however, especially since the surface is not entirely covered with As, but has only ~0.15 ML of excess As. A possibility is that As migrates to defects, thereby impeding the dissociation of Cl_2 by tying up the active sites.

It is unexpected for the principal chloride on the passivated surface to be AsCl, since photoemission[3,5] and STM[6] studies indicate that there is a clear preference for Ga–Cl bonding in the initial stages of the reaction on GaAs(110). It is possible that differences in the atomic structure or the morphology of As-rich GaAs(110), as compared to cleaved GaAs(110), are responsible for altering the distribution of surface reaction products. STM measurements of As-rich GaAs(110), both before and after reaction with Cl_2, would be of great use for understanding this unexpected result.

This apparently passive surface is not absolutely unreactive, however, since the reaction does continue providing that the Cl_2 exposure is sufficiently large. That is, for exposures greater than $\sim 5\times10^4$ L, the amount of AsCl increases noticeably and a measurable amount of $GaCl_2$ forms. The chlorine uptake also begins to increase at this point (see Fig. 4), signaling the onset of etching. Once etching has begun, the Cl coverage increases linearly with the logarithm of the Cl_2 exposure, as is observed for the chlorination of GaAs(001).[1] This linear dependence of the coverage on the logarithm of the exposure implies that the sticking coefficient decays exponentially with coverage, i.e., $S(\Theta)=a\exp(-b\Theta)$. Thus, although excess surface As does slow the reaction significantly, once this As is removed the etching reaction continues normally.

By carrying out the reaction at a sufficiently low temperature, in this case room temperature, the lifetimes of the reaction intermediates are sufficiently long that a determination of the rate-limiting step can be made. It is seen from Fig. 3 that, following the largest Cl_2 exposures, the principal As surface reaction product continues to be AsCl, whereas $GaCl_2$ is the primary Ga product. The Cl_2/GaAs reaction occurs in a stepwise fashion,[1,3] with the primary volatile etch products at or near room temperature being $AsCl_3$ and $GaCl_3$.[21,22] Because the crystal structure of GaAs constrains the availability of reaction sites, the etching reaction cannot be simply divided into discrete decoupled pathways. Hence, a build-up of AsCl and $GaCl_2$ on the surface suggests that the rate-limiting step in the overall etching reaction is either

$$GaCl_2+Cl\rightarrow GaCl_3 \quad (1)$$

or

$$AsCl+Cl\rightarrow AsCl_2. \quad (2)$$

The rate of one of these steps probably limits the other, which then results in the simultaneous build-up of both $GaCl_2$ and AsCl.

After the highest Cl_2 exposures, the coverages of AsCl and $GaCl_2$ are nearly identical (see Fig. 3). A likely reason for there being equal amounts of AsCl and $GaCl_2$ on the surface, and for the stability of one of them to limit the removal of the other, is that they form a common moiety. If such an assumption is made, then the simplest and most reasonable arrangement having AsCl and $GaCl_2$ bound together, and to the surface, is $-AsCl-GaCl_2$. In this moiety, the Ga and As atoms are bound to each other, while the Ga atom is also bonded to two Cl atoms and the As atom is bonded to a single Cl atom and to the surface. This structure has three requisite properties: (1) it consists of an equal number of AsCl and a $GaCl_2$ moieties; (2) it can bind to the surface, in this case through the As atom; and (3) both Ga and As are tricoordinately bonded, which they prefer in the absence of a crystal field. Any other possible arrangement does not satisfy these criteria. In this configuration, it must be the terminal–$GaCl_2$ group that is the more stable of the two species, i.e., the one that limits the reaction rate. Otherwise, it would quickly form $GaCl_3$ and desorb. Furthermore, due to its stability, the terminal–$GaCl_2$ group prevents further chlorination of the AsCl moiety by tying up an otherwise available reaction site. Thus, it appears that the rate-limiting step in the reaction is that given by Eq. (1).

Note that the formation of an $-AsCl-GaCl_2$ structure in the Cl_2 etching of GaAs is analogous to the generation of fluorosilyl trees (e.g., $-SiF-SiF_3$ and $-SiF_2-SiF_3$) in the XeF_2/Si etching reaction.[23] An interesting experiment, which could confirm the existence of such a structure, would be to study heavily chlorinated GaAs(110) with molecular beam scattering or secondary ion mass spectrometry, and to look for desorbing species that contain both Ga and As atoms.

The Cl_2 exposure at which significant chlorine uptake begins on As-rich GaAs(110) ($\sim 5\times10^4$ L) is uncommonly large compared to the exposures that cause appreciable reaction on other GaAs surfaces.[1,3–5] In addition, prior to the onset of reaction, the AsCl coverage dips slightly and the $AsCl_2$ and GaCl coverages begin to rise (see Fig. 3). The gradual transformation of the surface product distribution away from AsCl, combined with the very large exposures needed to produce appreciable reaction, suggests that the reaction nucleates at defects on an otherwise unreactive AsCl-stabilized surface. Thus, barring such defects, an AsCl-covered surface would be inert to further reaction.

V. SUMMARY

In summary, As-rich GaAs(110) was exposed at room temperature to large doses of Cl_2 and examined with SXPS. Following the initial Cl_2 exposures, the surface appears to passivate with half a monolayer of AsCl formed, in a manner similar to that observed previously.[2] The reaction continues, however, following much larger Cl_2 exposures, with both As and Ga chlorides formed on the surface. For the most heavily reacted surfaces, the distribution of chlorides favors AsCl and $GaCl_2$. It is concluded that the chlorination of $GaCl_2$ to form $GaCl_3$ is the rate-limiting step in the reaction, and that $GaCl_2$ stabilizes AsCl on the surface, possibly through the formation of a treelike $-AsCl-GaCl_2$ structure.

ACKNOWLEDGMENTS

This work is based upon work supported by the U.S. Army Research Office under Grant No. DAAH04-94-G-0410. This work was performed at the NSLS, which is supported by the Department of Energy (Division of Materials Sciences and Division of Chemical Sciences of Basic Energy Sciences) under Contract No. DE-AC0276CH0016. The authors recognize the contributions of J. Yurkas and A. Marx for the construction of the EMA. J. Yurkas and C. Costas are also thanked for their efforts in upgrading the 3 m beamline.

[1]W. C. Simpson, D. K. Shuh, W. H. Hung, M. C. Håkansson, J. Kanski, U. O. Karlsson, and J. A. Yarmoff, J. Vac. Sci. Technol. A **14**, 1815 (1996).
[2]D. J. D. Sullivan, H. C. Flaum, and A. C. Kummel, J. Chem. Phys. **101**, 1582 (1994).
[3]D. K. Shuh, C. W. Lo, J. A. Yarmoff, A. Santoni, L. J. Terminello, and F. R. McFeely, Surf. Sci. **303**, 89 (1994).
[4]W. C. Simpson, W. M. Tong, C. B. Weare, D. K. Shuh, and J. A. Yarmoff, J. Chem. Phys. **104**, 320 (1995).
[5]F. Stepniak, D. Rioux, and J. H. Weaver, Phys. Rev. B **50**, 1929 (1994).
[6]J. C. Patrin and J. H. Weaver, Phys. Rev. B **48**, 17913 (1993).
[7]J. Pankratz, H. Niehaus, and W. Mönch, Surf. Sci. **307–309**, 211 (1994).
[8]R. D. Schnell, D. Rieger, A. Bogen, K. Wandelt, and W. Steinmann, Solid State Commun. **53**, 205 (1985).
[9]W. C. Simpson, T. D. Durbin, P. R. Varekamp, and J. A. Yarmoff, J. Appl. Phys. **77**, 2751 (1995).
[10]W. C. Simpson, P. R. Varekamp, D. K. Shuh, and J. A. Yarmoff, J. Vac. Sci. Technol. A **13**, 1709 (1995).
[11]P. R. Varekamp, W. C. Simpson, D. K. Shuh, T. D. Durbin, V. Chakarian, and J. A. Yarmoff, Phys. Rev. B **50**, 14 267 (1994).
[12]Note that the photon energies used here are ~5 eV higher than those used in Ref. 3. This introduces only a negligible variation in the electron escape length, however, and therefore does not significantly affect the coverage calibration.
[13]T. T. Chiang and W. E. Spicer, J. Vac. Sci. Technol. A **7**, 724 (1988).
[14]J. E. Northrup, Phys. Rev. B **44**, 1349 (1991).
[15]K. Jacobi, G. Steinert, and W. Ranke, Surf. Sci. **57**, 571 (1976).
[16]R. G. Jones, N. K. Singh, and C. F. McConville, Surf. Sci. **208**, L34 (1989).
[17]A. Ludviksson, M. Xu, and R. M. Martin, Surf. Sci. **277**, 282 (1992).
[18]A. P. Mowbray and R. G. Jones, Vacuum **41**, 672 (1990).
[19]P. R. Varekamp, M. C. Håkansson, D. K. Shuh, J. Kanski, L. Ilver, Z. Q. He, J. A. Yarmoff, and U. O. Karlsson, Vacuum **46**, 1231 (1995).
[20]P. R. Varekamp, M. C. Håkansson, J. Kanski, D. K. Shuh, M. Björkqvist, M. Göthelid, W. C. Simpson, U. O. Karlsson, and J. A. Yarmoff, Phys. Rev. B (in press).
[21]H. Hou, Z. Zhang, S. Chen, C. Su, W. Yan, and M. Vernon, Appl. Phys. Lett. **55**, 801 (1989).
[22]L. A. DeLouise, J. Chem. Phys. **94**, 1528 (1991).
[23]C. W. Lo, D. K. Shuh, V. Chakarian, T. D. Durbin, P. R. Varekamp, and J. A. Yarmoff, Phys. Rev. B **47**, 15648 (1993).

Interaction of hydrogen ions with oxidized GaAs(100) and AlAs(100) surfaces

Ying-Lan Chang[a)]
Center for Quantized Electronic Structures (QUEST), University of California, Santa Barbara, California 93106

R. Cao, W. E. Spicer, and P. Pianetta
Stanford Synchrotron Radiation Lab (SSRL), Stanford University, Stanford, California 94305

Song Shi and Evelyn Hu
Center for Quantized Electronic Structures (QUEST), University of California, Santa Barbara, California 93106

James Merz
Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 22 January 1996; accepted 20 April 1996)

We have performed photoemission experiments, using a tunable soft x-ray synchrotron radiation source to study the chemical changes of oxidized GaAs and AlAs surfaces subject to exposure from hydrogen ions. Results indicate that the net effects for hydrogen ion irradiation are (i) the reduction of arsenic and (ii) the growth of the cation oxide components. The reduction of arsenic can result from the formation/desorption of arsine. The oxide overlayer after hydrogen ion treatments is dominated by cation oxides which are the more stable chemical species as described in the phase diagram. This oxide layer should then remain stable in atmosphere. These results can provide insight into the chemical reaction between hydrogen ions and oxidized AlGaAs surfaces. © *1996 American Vacuum Society.*

I. INTRODUCTION

Surface passivation has been an important issue for the fabrication of semiconductor devices. The ability to reduce or control surface states can result in enhanced optical and electronic device properties, particularly with the increasing surface-to-volume of the devices involved.

The common approach to obtain surface passivation has been to remove the oxide layers and terminate the surface dangling bonds. However, those treatments lack long-term stability.[1–4] The study of the stability of the GaAs surface passivated by different methods has shown that the reason for the subsequent degradation of electronic properties is reoxidation of the surface.[5] To improve the chemical stability, one needs to understand the chemical reactions that may take place between the passivated surface and the atmosphere, and try to prevent the passivated surface from reoxidation.

We have previously demonstrated a new approach to surface passivation, utilizing hydrogen ion irradiation. The assessment of passivation is related to the luminescence efficiency of a near-surface/surface quantum well (QW).[6–8] The passivation process takes place in the presence of the native oxide. The improvement in luminescence has been found to persist for periods as long as three years.[6,9] *In situ* Auger electron spectroscopy (AES) and temperature programmed desorption measurements suggest that the passivation is linked to the removal of arsenic in the form of arsine[9,10] with the consequent reduction in formation of arsenic antisite defects which may be responsible for high surface recombination velocity.[11]

A more direct assessment of the hydrogen-ion surface chemistry requires the use of techniques such as photoemission spectroscopy (PES). This article describes *in situ* PES measurements carried out with a tunable soft x-ray synchrotron radiation source. Such experiments provide good surface sensitivity and valuable information about the evolution of surface oxide components due to hydrogen ion treatments; therefore, they can provide further insight into the mechanism of hydrogen passivation and long-term stability.

II. EXPERIMENT

In our earlier studies, the substrates used had surface layers of $Al_{0.3}Ga_{0.7}As$ cladding GaAs QWs.[6] To simplify the analysis of the data, these studies were carried out on GaAs and AlAs surfaces separately, and the behavior of $Al_xGa_{1-x}As$ surfaces will be extrapolated from the behavior of the binary substrates.

The GaAs epilayer was grown by molecular beam epitaxy (MBE) on GaAs(100) substrates. The sample was treated by buffered HF solution to remove the surface oxide layer due to long-term air exposure. Then the sample was immediately transferred to the ultrahigh vacuum (UHV) chamber for hydrogen ion treatments and photoemission measurements. The total exposure time during this sample transfer is ~15 min.

The AlAs epilayer was also grown by MBE on GaAs(100) substrates and was capped by As before removal from the MBE chamber. To prevent any degradation of the protecting As cap, the samples were kept in a nitrogen box at

[a)]Present address: Hewlett–Packard Co., Hewlett–Packard Laboratory, Palo Alto, CA 94304; Electronic mail: ying-len chang@hpl.hp.com

all times between removal from the MBE chamber and insertion to the UHV analysis chamber. The As capping layer was thermally desorbed in the photoemission chamber by heating the sample at 450 °C for ~15 min. Subsequently, an oxide layer was grown by exposing this clean AlAs surface either to air for ~15 min or to oxygen in the UHV chamber.

Hydrogen ions were generated by using an ion sputter gun with a differential gas inlet, operated at an ion energy of 100 eV. The flux was measured using the ion current collected on the sample. A typical ion current density of 1.5 $\mu A/cm^2$ was used. The hydrogenation process was done at room temperature.

The photoemission measurements were performed on beam line 8-1 at the Stanford Synchrotron Radiation Laboratory with tunable synchrotron radiation as the source for photoelectron spectroscopy. Photon energy could be tuned in the range between 10 and 200 eV. Typically, the photon energy used for Ga-3*d* and As-3*d* core level spectra, and valence band spectra is 80 eV, and the photon energy used for Al-2*p* core level spectra is 120 eV. Under these measurement conditions, surface sensitive (~5–10 Å escape depth) valence band and Ga, As, Al core level spectra could be obtained.[12]

The investigation of the surface chemical reaction requires appropriate deconvolution of the core level spectra. Although the resolution of our PES system is about 0.3 eV, several factors, e.g., the inhomogeneity of the native oxide, may contribute to the broadening of the core level peaks. The large linewidth complicates the interpretation of the core level spectra. For clarity, we studied the interaction between hydrogen ions and oxidized binary compounds, including GaAs and AlAs surfaces. The results on these surfaces can provide some insight into the chemical reaction on ternary compounds, such as AlGaAs.

III. RESULTS

A. Chemical reaction of hydrogen ions with oxidized GaAs surface

The As-3*d* core level spectrum for a native-oxide-covered GaAs surface is shown in Fig. 1(a). Deconvolution of the As-3*d* core level spectrum reveals the substrate component (at ~33.2 eV) and As_2O_3 (at ~30.3 eV). Elemental As was not observed on this native-oxide-covered surface. Previous results on a clean GaAs surface indicate that the surface Fermi level can be pinned at submonolayer oxygen coverage,[11] which is much lower than the coverage required for the formation of elemental As.[12,13] As a matter of fact, elemental As has only been observed on surfaces subjected to heavy oxidation, e.g., ionized oxygen treatments, etc.[12]

The As-3*d* core level spectrum for the oxidized GaAs surface after *in situ* room temperature hydrogen ion treatments with exposure of 10^{16} cm^{-2} is shown in Fig. 1(b). The As(oxide)/As(substrate) ratio was found to decrease by about 11% after hydrogen ion treatments.

The Ga core level spectra similarly show a substrate component (at ~55.1 eV) and Ga_2O_3 (at ~54.0 eV), as shown in Figs. 2(a) and 2(b). Contrary to As, the Ga(oxide)/Ga(substrate) ratio increases by about 27% after hydrogen ion treatments.

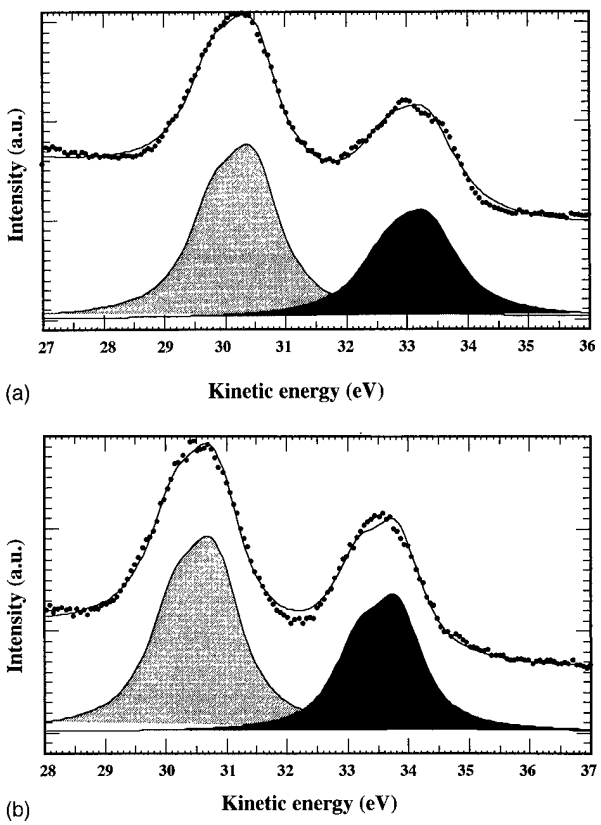


FIG. 1. The As-3*d* core level spectra for an oxidized GaAs surface (a) before and (b) after room temperature hydrogen ion treatments. The hydrogen ion exposure was 10^{16} cm^{-2}. The dots represent the raw data, the solid line through the dots is the curve fit to the data, and the various components are labeled according to the shading: The dark area represents the substrate components, and the light area represents the oxide component.

The As/Ga integrated intensity ratio decreases by about 33% after room temperature hydrogen ion treatment with exposure of 10^{16} cm^{-2}. The reduction of the As/Ga ratio is consistent with our previous observation by AES measurements.[10]

It was also noticed that after room temperature hydrogen passivation at low exposure, a small amount of As_2O_3 still exists on the surface, as shown in Fig. 1(b).

B. Chemical reaction of hydrogen ions with oxidized AlAs surface

Studies of the AlAs surface produced similar results; that is, the As(oxide)/As(substrate) ratio decreases after hydrogen ion treatment.

Since Al oxides rapidly in air, only the oxide component can be observed in the Al-2*p* core level spectrum for the native-oxide-covered AlAs surface. It is therefore difficult to observe the modification of the Al core level components after hydrogen ion treatments. The studies of interaction between hydrogen ions and oxidized AlAs surfaces had to be

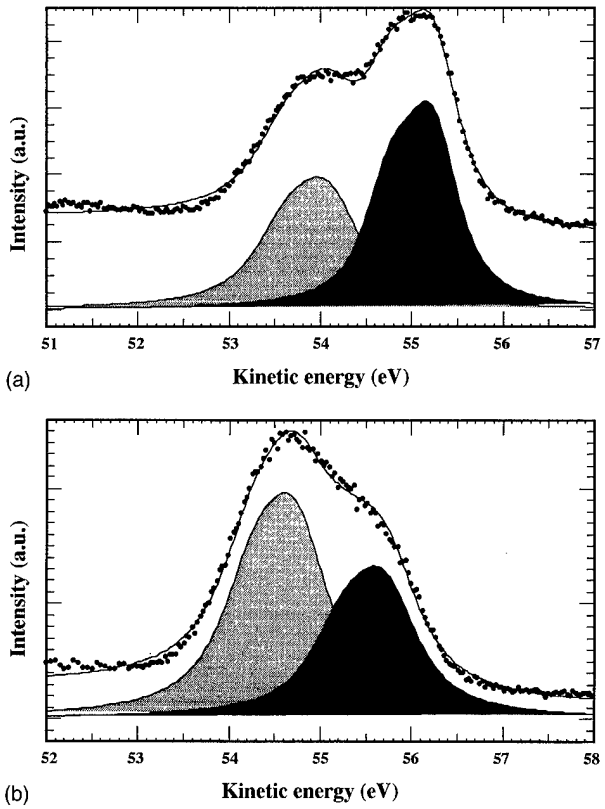


FIG. 2. The Ga-3*d* core level spectra for an oxidized GaAs surface (a) before and (b) after room temperature hydrogen ion treatments. The hydrogen ion exposure was 10^{16} cm^{-2}. The dots represent the raw data, the solid line through the dots is the curve fit to the data, and the various components are labeled according to the shading. The dark area represents the substrate components, and the light area represents the oxide component.

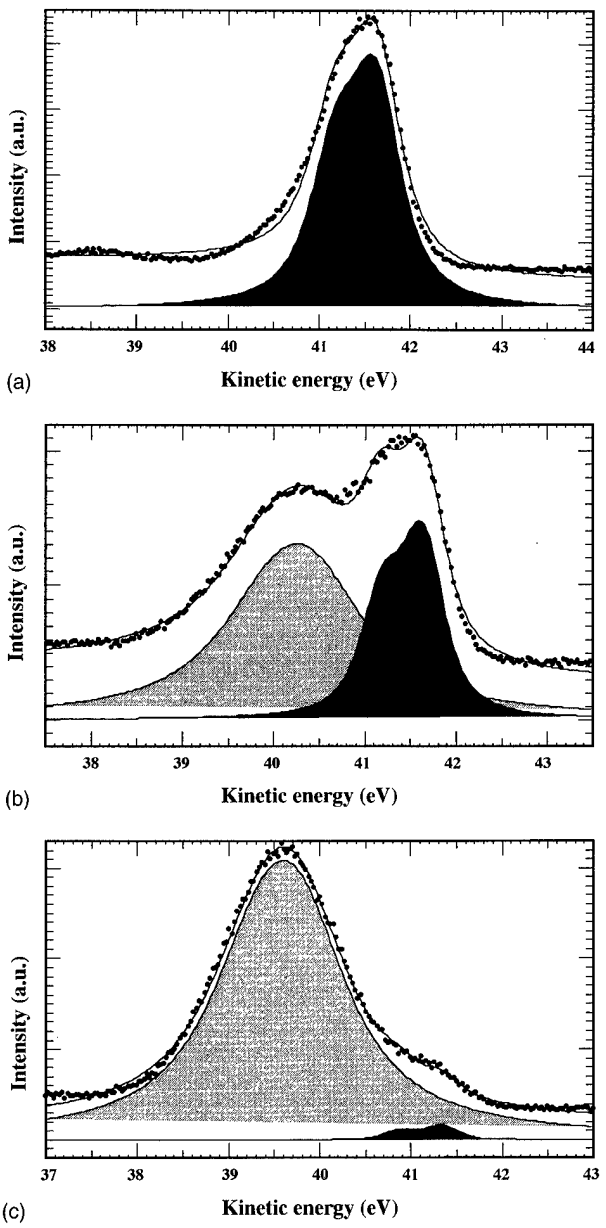


FIG. 3. The Al-2*p* core level spectra for a clean AlAs surface before and after different surface treatments. (a) Clean (prepared by As-decapping), (b) after 10^6 L oxygen exposure, and (c) after room temperature hydrogen ion treatment with exposure of 10^{16} cm^{-2}. The dots represent the raw data, the solid line through the dots is the curve fit to the data, and the various components are labeled according to the shading: The dark area represents the substrate components, and the light area represents the oxide component.

carried out by capping the grown AlAs layer with arsenic, and then carefully desorbing the arsenic and doing a controlled, *in situ* oxidation of the surface.

The Al-2*p* core level spectra for the clean AlAs surface and the surface treated with 10^6 L oxygen exposure are shown in Figs. 3(a) and 3(b), respectively. Both the substrate and oxide components can be clearly observed in Fig. 3(b). The spectrum obtained from the oxide covered surface after hydrogen ion treatment is presented in Fig. 3(c). After room temperature hydrogen ion treatments, the increase of Al(oxide)/Al(substrate) ratio due to hydrogen ion irradiation is clearly evident.

IV. DISCUSSION

As mentioned above, the reduction of As from the oxidized surfaces was observed after hydrogen ion treatments. Our earlier experiments validated the evolution of AsH_3 from the hydrogenated substrate and the attendant reduction in surface arsenic concentration.[9,10] Two possibilities for the mechanism of reduction involve either (1) reaction of hydrogen and arsenic at the oxide-semiconductor interface or (2) the reaction of the hydrogen ions with the arsenic oxide.

Reaction (1) requires the formation of AsH_3 below the surface of the native oxide, and the subsequent desorption of the AsH_3. This process can be limited by the transport of AsH_3 molecule through the oxide to the uppermost surface.

However, this may not be an issue if the formation of AsH_3 takes place through the oxide. Such a reaction would account for the preferential reduction of the arsenic bound as As_2O_3. What could then explain the increase of Ga_2O_3?

The interaction between hydrogen and As_2O_3 can produce H_2O.[14] If the generated H_2O remained physisorbed on the surface of the substrate, further oxidation of GaAs could take place, with gallium oxidizing more readily than arsenic and some portion of the As_2O_3 being formed reacting with the

hydrogen and being desorbed as arsine. That is, during the treatments, the reduction of As_2O_3 and As is accompanied by formation of Ga_2O_3 and the net reaction can be written as

$$2GaAs+As_2O_3+12H \rightarrow Ga_2O_3+4AsH_3.$$

As a result, the As/Ga ratio is reduced. However, the oxygen concentration remains unchanged, as observed by our AES measurements.[10]

Evidence that this reaction can be taking place has been provided by previous *in situ* real-time Fourier transform infrared spectroscopy measurements which were carried on for a closely related process, that of hydrogen plasma treatment of an oxidized GaAs surface.[14] In this study, it was found that the concentration of physisorbed H_2O on the GaAs surface increases during hydrogen treatments. The formation of Ga_2O_3 was observed as a result of the oxidation of the GaAs through reaction with the physisorbed water. Although the oxidized GaAs surface is treated by hydrogen ions, rather than hydrogen plasma in the present case, it is likely that similar reactions are taking place.

We have previously reported that the enhancement of luminescence efficiency from a near-surface GaAs/AlGaAs QW due to hydrogen passivation remains stable in atmosphere.[9] Regarding the long-term stability of passivation effects, two things need to be considered. First, the surface termination layer needs to remain intact after the exposure to atmosphere. Second, the overlayer needs to be tightly packed so that the diffusion of oxygen through the overlayer can be minimized.

Based on our PES results shown before, the chemical reaction taking place on oxidized GaAs and AlAs surfaces appears to be similar. That is, the net effects for the hydrogen ion treatments on oxidized surfaces are (i) the reduction of As and (ii) the growth of cation oxide components. One can expect that similar results can be obtained on an oxidized AlGaAs surface after hydrogen ion irradiation.

That is, hydrogen ion treatments not only reduce the amount of As and the density of midgap states at the oxide/semiconductor interfaces,[11] but also improve the chemical stability of the oxides.[5] As a result, the chemical reaction between the surface layer and the atmosphere can be eliminated. Contrary to other passivation approaches, we *use* this oxide as an overpassivation layer, rather than *remove* it to achieve a clean surface.

It has been suggested that the *complete* removal of As and As_2O_3 is important in surface passivation.[5] However, the peak corresponding to As_2O_3 is still observable in the surface sensitive PES spectra taken from samples after hydrogen ion treatments. On the basis of the phase diagram,[5] if As_2O_3 does not directly contact the substrate, the reaction resulting in the formation of free As will not take place. Although we do not have direct evidence, we believe that this is likely for an As deficient surface. Moreover, since the thickness of the oxide overlayer is small, the presence of a small amount of As_2O_3 may be important in forming a closely packed oxide.

V. CONCLUSIONS

In summary, we report *in situ* monitoring of the chemical changes of oxide-covered GaAs and AlAs surfaces caused by hydrogen ion treatments. Experimental evidence shows that the hydrogen ion irradiation decomposes the arsenic oxide and reduces the amount of arsenic at the oxide/semiconductor interface. The oxide overlayer after hydrogen ion treatments is dominated by the cation oxides, which are the more stable chemical species as described in the phase diagram.[5] This oxide layer remains stable in atmosphere, and can serve as an effective capping layer for the hydrogen ion-treated semiconductor surface, preventing surface reoxidation.

It was also noted that a small amount of As_2O_3 was still present on the surface after hydrogen passivation. It is likely that the residual As_2O_3 does not directly come in contact with the semiconductor substrate, therefore does not react with the substrate and form free arsenic at the interface. The oxide-semiconductor interface remains arsenic deficient. As a result, the improved interface quality can be preserved, as indicated by the long-term stability of the passivation effects.[6,9]

[1]N. A. Ives, G. W. Stupian, and M. S. Leung, Appl. Phys. Lett. **50**, 256 (1987).
[2]E. Yablonovitch, B. J. Scromme, R. Bhat, J. P. Harbison, and T. J. Gmitter, Appl. Phys. Lett. **54**, 555 (1989).
[3]M. Oshima, T. Scimeca, Y. Watanabe, H. Oigawa, and Y. Nannichi, Jpn. J. Appl. Phys. **32**, 518 (1993).
[4]W. S. Hobson, U. Mohideen, S. J. Pearton, R. E. Slusher, and F. Ren, Electron. Lett. **29**, 2199 (1993).
[5]N. N. Berchenko and Yu. V. Medvedev, Russ. Chem. Rev. **63**, 623 (1994).
[6]Y.-L. Chang, I-H. Tan, Y. H. Zhang, J. Merz, E. Hu, A. Frova, and V. Emiliani, Appl. Phys. Lett. **62**, 2697 (1993).
[7]Y.-L. Chang, I-H. Tan, Casper Reaves, J. Merz, E. Hu, S. DenBaars, A. Frova, and V. Emiliani, Appl. Phys. Lett. **64**, 2658 (1994).
[8]J. M. Moison, K. Eless, F. Houzay, J. Y. Marzin, J. M. Gerard, F. Barthe, and M. Bensoussan, Phys. Rev. B **41**, 12945 (1990).
[9]Y.-L Chang, S. I. Yi, S. Shi, E. Hu, W. H. Weinberg, and J. Merz, J. Vac. Sci. Technol. B **13**, 1801 (1995).
[10]Y.-L. Chang, W. Widdra, S. I. Yi, J. Merz, W. H. Weinberg, and E. Hu, J. Vac. Sci. Technol. B **12**, 2605 (1994).
[11]W. E. Spicer, T. Kendelewicz, N. Newman, R. Cao, and E. R. Weber, J. Vac. Sci. Technol. B **6**, 1245 (1988).
[12]P. Pianetta, Ph. D. thesis, Stanford University (1978).
[13]A. M. Green and W. E. Spicer, J. Vac. Sci. Technol. A **11**, 1061 (1993).
[14]E. S. Aydil, Z. Zhou, K. P. Giapis, Y. Chabal, J. A. Gregus, and R. A. Gottscho, Appl. Phys. Lett. **62**, 3156 (1993).

Study on interface abruptness of $In_xGa_{1-x}As/In_yGa_{1-y}As_zP_{1-z}$ heterostructures grown by gas-source molecular beam epitaxy

W. G. Bi[a] and C. W. Tu
Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, California 92093-0407

(Received 23 January 1996; accepted 20 April 1996)

We report a study of the effect of source switching sequences on the interface abruptness of InGaAs/InGaAsP quantum wells (QWs) grown by gas-source molecular beam epitaxy. Four methods were investigated: i.e., at each interface during growth interruption, (1) introducing a residual group-V source evacuation time (RSE) t where all sources are shut off; (2) no RSE; (3) no RSE but adding a group-III In+Ga prelayer of m monolayers; or (4) using RSE plus the In+Ga prelayer. InGaAs/InGaAsP QWs grown by the different methods were characterized by low-temperature photoluminescence (PL). The results show that by optimizing t or m, both RSE and prelayer methods can improve the heterostructure interface abruptness, which is evidenced by the stronger PL intensity and narrower PL linewidth, and that combining the two with optimized t and m gives the best result. This is due to the minimization of the memory effect by using RSE and of the As/P exchange by using the group-III prelayer. © *1996 American Vacuum Society.*

I. INTRODUCTION

InGaAs/InGaAsP quantum well (QW) structures are used extensively for optoelectronic device applications, such as long-wavelength lasers and modulators.[1,2] It is important to obtain QW structures with abrupt heterointerfaces, as interface roughness will randomize quantum-confined levels, increase the linewidth of optical emission, and greatly influence the properties of heterostructure devices. There are two main issues that are needed to be solved when growing high-quality arsenide/phosphide heterostructures. One is the so-called memory effect or carryover,[3–7] i.e., the residual group-V source will be incorporated into the succeeding layer, forming a compositionally graded transitional layer at the heterointerface. This usually happens when phosphide-containing compounds are grown on arsenide-containing compounds as a result of the higher incorporation coefficient of As compared with P.[8–13] The reverse case can also happen when the P_2 required is much larger than As_2, where the residual P_2 will also be incorporated into the subsequent As-containing layers.[14] Another problem encountered is the exchange of As and P when switching the group-V sources.[15–19] The new group-V source will replace the old one in the underlying layer, resulting in a strained interfacial layer. It has been reported that relatively low-temperature growth can reduce the thickness of the intermixing layer[15] or can obtain abrupt interfaces.[20] The optical quality of the grown structure, however, is usually of poor quality, and post-growth annealing is required to improve the material quality. On the other hand, using one monolayer of In on an InGaAs layer before P_2 exposure has been proposed to improve the InGaAs/InP interface.[8] Due to the reaction of the deposited In metal with the As_2 in the underlying InGaAs layer, a transition region still exists by using this method, and ways to further improve the interface quality are needed.

In this article, we will investigate different ways to improve the interface quality of InGaAs/InGaAsP heterostructures by minimizing the memory effect and the As/P exchange. Besides the conventional method, three approaches are explored, and InGaAs/InGaAsP QWs with improved interfaces are obtained, as evidenced by stronger photoluminescence (PL) intensity and narrower PL linewidth.

II. EXPERIMENT

The experiment was carried out in a Varian Gen II MBE system, which had been modified to handle group-V gas sources. The hydride gas sources for arsenic and phosphorous dimers are pure arsine and phosphine, which were introduced into the growth chamber via two hydride injectors operated at 970 °C. The use of separate crackers is thought to be effective in minimizing the cross contamination of As/P. The samples were grown on a (100)-oriented semi-insulating InP substrate. The growth rate of $In_{0.84}Ga_{0.16}As_{0.32}P_{0.68}$ was 1 monolayer per second (ML/s), and that of $In_{0.53}Ga_{0.47}As$ was 0.3–0.4 ML/s. The flow rate of phosphine used was 2.8 sccm and that of arsine was 0.5 sccm. The typical chamber pressure was between 2×10^{-6} and 1×10^{-5} Torr. Reflection high-energy electron diffraction (RHEED) was used to examine the surface roughness of the As/P interface. The intensity of the RHEED specular beam was detected by an optical-fiber-coupled photodiode, and the signal was enhanced with a dual-channel differential amplifier before it was recorded by a computer. Low-temperature photoluminescence measurement was carried out by mounting the samples in a liquid He cryostat and using the 514.5 nm line of an argon ion laser as the excitation source. A thermoelectrically cooled Ge photodiode was used to detect the signal at the exit of a 50-cm monochromator through an amplifier. The pumping power density was ~2 W/cm^2 and the measurement temperature was 15 K.

[a]Electronic mail: wbi@sdcc3.ucsd.edu

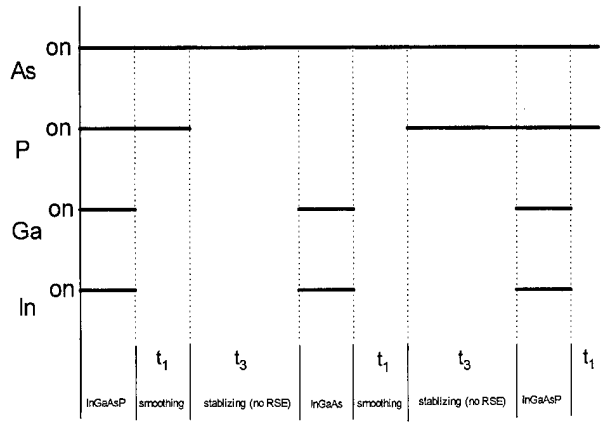


FIG. 1. Shutter switching sequences used for the growth of InGaAs/InGaAsP QWs (not drawn to scale in time), where the time period t_1 is used to smooth the grown surface and t_3 to stabilize the gas flow after switching the group-V sources (conventional method).

III. RESULTS AND DISCUSSION

Our control sample was grown by the source switching sequences shown in Fig. 1, where the solid lines indicate that the corresponding sources are turned on. This is a conventional way to grow arsenide/phosphide heterostructures; i.e., at each interface during growth interruption, after a continuing supply of a group-V source for t_1 sec to smooth the grown surface, a new group-V source is switched on for t_3 sec to stabilize the gas flow before growing the subsequent layer (the conventional method). As a result, a transient interfacial layer caused by both the memory effect and the As/P exchange might be formed by using this method. The memory effect is more pronounced at the InGaAs-on-InGaAsP interface because of the incorporation of the residual P_2 into the succeeding InGaAs layer, while the As/P substitution is the main source responsible for the transient layer at the InGaAsP-on-InGaAs interface. The latter can be evidenced by examining the change of the RHEED pattern of InGaAs under the exposure of As_2+P_2, as surface roughening is a direct evidence of As/P substitution. Figs. 2(a) and 2(b) show the RHEED pattern of an $In_{0.53}Ga_{0.47}As$ layer exposed to As_2 and that to As_2+P_2 for 12 sec at 500 °C, respectively. The freshly grown $In_{0.53}Ga_{0.47}As$ surface at 500 °C shows an initial (2×1) streaky RHEED pattern, which is quite stable under As_2 exposure. Upon being exposed to As_2+P_2, however, the RHEED pattern gradually becomes spotty within 12 sec [Fig. 2(b)], indicating that the $In_{0.53}Ga_{0.47}As$ surface becomes rough as a result of the As/P exchange. To solve these problems, we tried three other source switching sequences as illustrated in Figs. 3(a), 3(b),

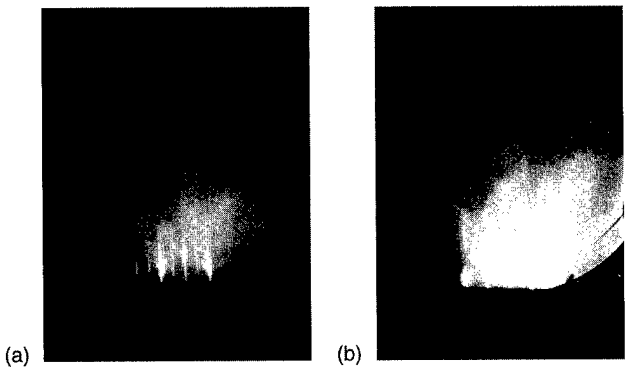


FIG. 2. RHEED patterns of InGaAs under As_2 (a) and As_2+P_2 (b) exposure for 12 sec at 500 °C measured from [110] azimuth.

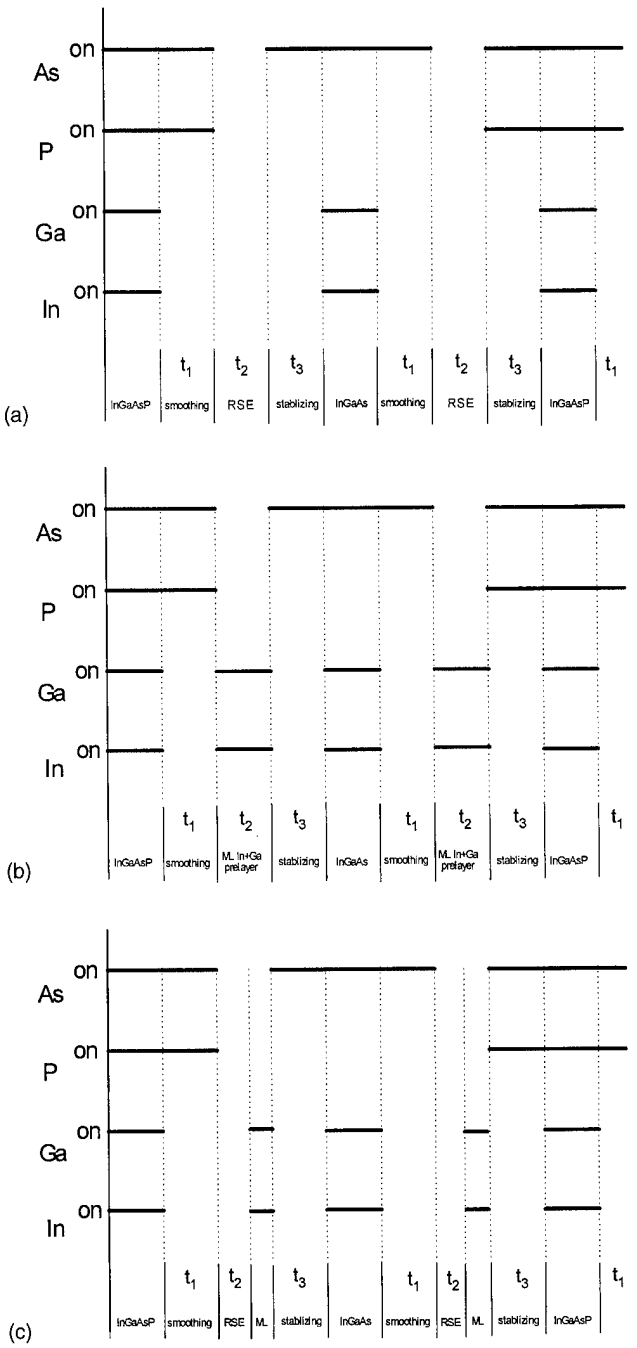


FIG. 3. Shutter switching sequences used for the growth of InGaAs/InGaAsP QWs (not drawn to scale in time): (a) RSE method, where a time period t_2 is introduced between t_1 and t_3 to evacuate the excess group-V species; (b) prelayer method, where one ML In+Ga is deposited between t_1 and t_3 to avoid As/P exchange; and (c) combined method, where RSE plus prelayer are used.

and 3(c). The first method used [Fig. 3 (a)] introduces a residual group-V source evacuation (RSE) time period t_2 between t_1 and t_3, where all sources are shut off (the RSE method). The second method [Fig. 3(b)] does not use RSE but inserts one monolayer (ML) of In and Ga metals (In +Ga) before the switching of the group-V sources (the prelayer method). In method three [Fig. 3 (c)], RSE plus one ML of In+Ga prelayer is used at each interface (the combined method). To compare the effectiveness of the three methods in improving interface abruptness, we grew four samples with the same sample structure, i.e., three $In_{0.53}Ga_{0.47}As/In_{0.84}Ga_{0.16}As_{0.32}P_{0.68}$ single quantum wells (SQWs) with well width of 0.5, 1, and 2 nm and InGaAsP barrier width of 20 nm, except that at each interface a different source switching method was used. The smoothing time t_1 was determined by examining the RHEED intensity recovery time after growth interruption,[14] as the magnitude of the RHEED intensity reflects the degree of surface roughness of the grown surface. The RSE time was optimized by growing different samples with the same sample structure but different RSE time previously.[14] A RSE time of 6 sec was used here.

Figure 4(a) and 4(b) show a comparison of the PL linewidth and PL intensity of the four samples, respectively. The open squares are the results of the sample grown with the combined method; the solid triangles, with the RSE method; the solid squares, with the prelayer method; and the open circles, with the conventional method. Clearly seen here is that as compared with the conventional method, both the RSE and the prelayer methods can improve the $In_{0.53}Ga_{0.47}As/In_{0.84}Ga_{0.16}As_{0.32}P_{0.68}$ interfaces, while combining the two gives best result. This is evidenced by the stronger PL intensity and narrower PL linewidth of the InGaAs/InGaAsP SQWs grown by these methods as compared with that grown with the conventional method. The improvement of using RSE is thought to be due to minimization of the memory effect. When no RSE is used, excess As and/or P accumulated at the grown surface during time t_1 might be incorporated into the succeeding layer, resulting in a transitional layer. By introducing a RSE time, the residual group-V sources will be evacuated, thus minimizing the memory effect. The reason the prelayer method also improves the interface abruptness is due to the elimination of As/P exchange at the interface. Usually As/P exchange will occur when an arsenide compound (phosphide compound) is exposed to phosphorous (arsenic) for a period of time, resulting in a rough surface of the grown layer, as has already been shown by observing the RHEED pattern when a freshly grown $In_{0.53}Ga_{0.47}As$ surface is exposed to a As_2+P_2 flux (Fig. 2). This As/P exchange can be eliminated by inserting an In+Ga prelayer before switching the group-V sources; e.g., after the $In_{0.53}Ga_{0.47}As$ growth and depositing one ML of In+Ga before supplying As_2 and P_2, the direct exposure of the InGaAs surface to P_2 is prevented, thus avoiding the exchange of As/P and leading to an improved interface. Studies on the InGaAs/InP interfaces[8] also show that by adding one ML of In on the InGaAs surface before P_2 beam exposure can improve the interface quality. The reason we use one ML In+Ga instead of In here is to avoid a strained transitional layer that might be formed by using this method. For InGaAs/InGaAsP growth, the deposited one ML of In+Ga or In will react with the group-V species in the underlying InGaAs layer and also with the As_2 and P_2 that are supplied subsequently in the time period of t_3 before growing the next InGaAsP layer. Quite obviously, if In instead was used, a strained layer of InAs and/or InAsP might be formed at the InGaAsP-on-InGaAs heterointerface, thereby deforming the shape of the well and influencing the optical properties of the heterostructures. By using a mixture of In +Ga, the effect can be minimized. When combining the RSE and prelayer methods, both memory and As/P substitution

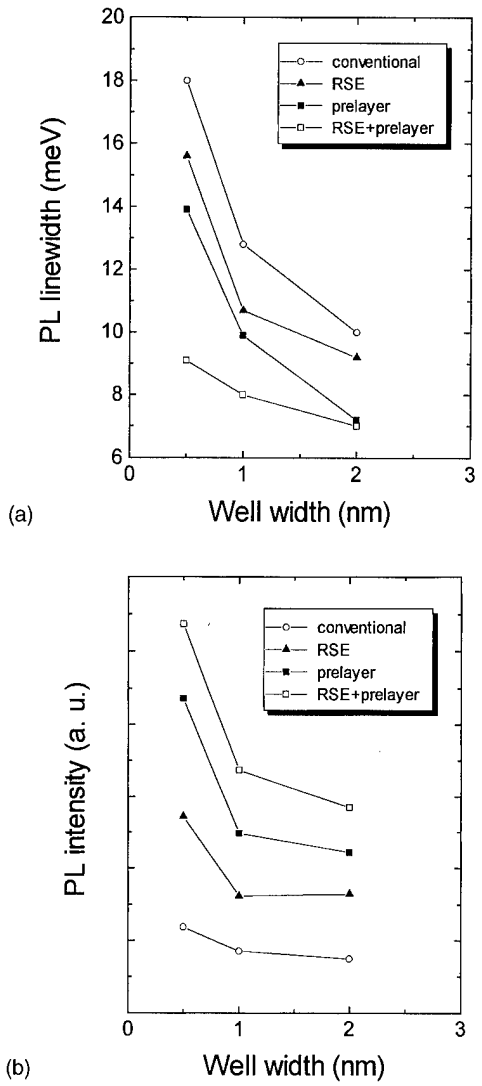


FIG. 4. A comparison of the PL linewidth (a) and intensity (b) of the InGaAs/InGaAsP QWs grown by the four different shutter switching methods. The open squares are the results of the sample grown with the combined method, the solid triangles are the results of the sample grown with the RSE method, the solid squares are the results of the sample grown with the prelayer method, and the open circles are results of the sample grown with the conventional method.

effects can be minimized, and the best interface is expected, as has been evidenced by our results.

IV. SUMMARY

In summary, InGaAs/InGaAsP QWs have been grown by GSMBE. Interface abruptness was investigated by changing the growth conditions and characterized by low-temperature PL. By introducing a short period of growth interruption where all sources are evacuated plus one monolayer of group-III elements before growing the next layer, both the memory effect and the As/P exchange can be minimized, resulting in improved interfaces of InGaAs/InGaAsP QWs.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation.

[1]P. J. A. Thijs, J. J. M. Binsma, L. F. Tiemeijer, P. I. Kuindersma, and T. V. Dongen, Microelectron. Eng. **18**, 57 (1992).
[2]M. Nido, K. Naniwae, T. Terakado, and A. Suzuki, Appl. Phys. Lett. **62**, 2716 (1993).
[3]A. Antoline *et al.*, J. Electron. Mater. **21**, 233 (1992).
[4]T. Y. Wang, E. H. Reihlen, H. R. Jen, and G. B. Stringfellow, J. Appl. Phys. **66**, 5376 (1989).
[5]M. E. Sherwin, F. L. Terry, Jr., G. O. Munns, J. S. Herman, E. G. Woelk, and G. I. Haddad, J. Electron. Mater. **21**, 269 (1992).
[6]P. J. Skevington, M. A. G. Halliwell, M. H. Lyons, S. J. Amin, M. A. Z. Rejman-Greene, and G. J. Davies, J. Cryst. Growth **120**, 328 (1992).
[7]H. Y. Lee, M. J. Hafich, and G. Y. Robinson, J. Cryst. Growth **105**, 244 (1990).
[8]T. Anan, S. Sugou, K. Nishi, and T. Ichihashi, Appl. Phys. Lett. **63**, 1047 (1993).
[9]M. R. Taylor, M. Hockly, A. Petford-Long, M. H. Lyons, and P. C. Spurdens, Inst. Phys. Conf. Ser. **100**, 305 (1989).
[10]A. G. Norman, B. R. Butler, G. R. Butler, G. R. Booker, and E. J. Thrush, Inst. Phys. Conf. Ser. **100**, 311 (1989).
[11]N. J. Long, A. G. Norman, A. K. Petford-LOng, B. R. Butler, C. G. Cureton, G. R. Booker, and E. J. Thrush, Inst. Phys. Conf. Ser. **117**, 69 (1991).
[12]R. Schwedler *et al.*, Appl. Surf. Sci. **63**, 187 (1993).
[13]H. Q. Hou and C. W. Tu, Appl. Phys. Lett. **60**, 1872 (1992).
[14]W. G. Bi and C. W. Tu, J. Electron. Mater. **25**, 1049 (1996).
[15]S. Nagao, M. Takashima, Y. Inoue, M. Kato, and H. Goto, J. Cryst. Growth **111**, 521 (1991).
[16]J. M. Vandenberg, S. N. G. Chu, R. A. Hamm, M. B. Panish, and H. Temkin, Appl. Phys. Lett. **49**, 1302 (1986).
[17]R. Averbeck, H. Riechert, H. Schlotterer, and G. Wiemann, Appl. Phys. Lett. **59**, 1732 (1991).
[18]K. Mahalingam, Y. Nakamura, N. Otsuka, H. Y. Lee, M. J. Hafich, and G. Y. Robinson, J. Electron. Mater. **21**, 129 (1992).
[19]M. Yano, H. Yokose, Y. Iwai, and M. Inoue, J. Cryst. Growth **111**, 609 (1991).
[20]Y. Nakamura, K. Mahalingam, N. Otsuka, H. Y. Lee, M. J. Hafich, and G. Y. Robinson, J. Vac. Sci. Technol. B **9**, 2445 (1991).

Effect of interface defect formation on carrier diffusion and luminescence in $In_{0.2}Ga_{0.8}As/Al_xGa_{1-x}As$ quantum wells

D. H. Rich,[a] K. Rammohan, H. T. Lin, and Y. Tang
Department of Materials Science and Engineering, University of Southern California, Los Angeles, California 90089-0241

M. Meshkinpour and M. S. Goorsky
Department of Materials Science and Engineering, University of California, Los Angeles, Los Angeles, California 90024

(Received 2 February 1996; accepted 17 April 1996)

We have examined the influence of strain relaxation on the excitonic recombination and diffusion in $In_{0.2}Ga_{0.8}As/Al_xGa_{1-x}As$ quantum-well (QW) samples designed for high-electron-mobility transistors, using spectrally and spatially resolved polarized cathodoluminescence (CL). Six molecular-beam epitaxial grown samples, with varying channel thicknesses ranging from 75 to 300 Å, were examined at various temperatures between 87 and 300 K. An increase in misfit dislocation density occurred with increasing channel thicknesses and resulted in changes in the dark line defect (DLD) density, polarization anisotropy, QW excitonic luminescence energy, and luminescence activation energy, as observed in CL. The influence of misfit dislocations on the ambipolar diffusion of excess carriers in a direction parallel to the dislocation line, in varying proximity to the DLDs, was examined with a CL-based diffusion experiment. The temperature dependence of the CL imaging was examined, enabling a study of the spatial variation of the activation energies associated with thermal quenching of the $GaAs/Al_{0.25}Ga_{0.75}As$ multiple QW and $In_{0.2}Ga_{0.8}As$ QW luminescence. The CL intensity exhibits an Arrenhius-type dependence on temperature and is controlled by thermally activated nonradiative recombination. The activation energies for both the $In_{0.2}Ga_{0.8}As$ QW and $Al_{0.25}Ga_{0.75}As$ MQW luminescence are found to vary spatially in close proximity to the misfit dislocations. We have utilized a new approach to obtain 2D images of the activation energies. The influence of the strain relaxation on the polarization and energy of the $In_{0.2}Ga_{0.8}As$ QW excitonic luminescence was examined with linearly polarized CL and CL wavelength imaging. A strain-induced modification of the luminescence energy and an increase in the polarization anisotropy was measured near DLDs. Thus, we find that certain DLDs exhibit significant polarization and energy variations in their optical properties, in addition to their more familiar nonradiative behavior. © *1996 American Vacuum Society.*

I. INTRODUCTION

The growth of high-quality strained films of $In_xGa_{1-x}As$ on GaAs substrates has several diverse applications in millimeter-wave electronic devices, such as high-electron-mobility transistors (HEMTs),[1–3] and near-infrared photonic devices, such as lasers[4–6] and light modulators.[7–10] The primary deleterious effect of strain relaxation is to introduce misfit dislocations at the $In_xGa_{1-x}As/GaAs$ interfaces which can serve as nonradiative recombination centers and impede the transport of carriers.[10–13] The generation of an asymmetrical density of 60° dislocations along the orthogonal ⟨110⟩ directions will also modify the biaxial symmetry of the strain tensor, and induce an asymmetry in the transport of carriers[10,12] and a polarization anisotropy in the excitonic luminescence.[14–16] Electron-beam probes which yield greater than a ~1-μm-scale resolution, such as in spatially and spectrally resolved cathodoluminescence (CL), can help advance our understanding of the influence of defects on the optical and transport properties of quantum heterostructures. A dislocation-induced asymmetry in the ambipolar diffusion of excess carriers in nipi-doped $In_{0.2}Ga_{0.8}As/GaAs$ multiple quantum wells (MQWs) has been previously measured using an electron-beam-induced absorption modulation technique.[10,12] In other studies, changes in the polarization anisotropy and energy of excitonic luminescence in thick $In_xGa_{1-x}As/GaAs$ films were found to correlate with the position of dark line defects (DLDs) from studies of linearly polarized cathodoluminescence (LPCL) and CL wavelength imaging (CLWI).[14–16]

In this study we aim to further explore the interplay between the strain relaxation, optical properties, and transport properties in $In_{0.2}Ga_{0.8}As/Al_xGa_{1-x}As$ QW samples, designed for HEMTs, using LPCL and CLWI. A four-band $\mathbf{k}\cdot\mathbf{p}$ calculation is used to examine the energy and polarization variations caused by the strain relaxation. A self-consistent-field calculation using the transfer-matrix method is used to determine the electron and hole eigenstates in the channel and the excitonic transition energies.

The temperature dependence of luminescence in QWs and superlattices has been investigated previously;[17,18] however, only recently have we begun to examine the effects of ther-

[a]Electronic mail: danrich@alnitak.usc.edu

mal quenching of luminescence by misfit dislocations. Thermal quenching of the luminescence has been interpreted in several ways by different authors, and has been attributed to either thermal dissociation of excitons and thermally activated nonradiative recombination,[19] or due to thermal emission of carriers out of the QWs, resulting in a reduction of luminescence intensity at higher temperatures.[20] The temperature dependence of the CL imaging was examined here in the 87–250 K range, enabling a study of the spatial variation of the activation energies associated with thermal quenching of luminescence near individual dislocations. The CL intensity exhibits an Arrenhius-type dependence on the temperature and is controlled by thermally activated nonradiative recombination. We have utilized a new approach to obtain 2D images of the activation energy.

In particular, we have also examined the change in the ambipolar diffusion length of carriers parallel to and in close proximity to misfit dislocations using a CL-based diffusion experiment. The changes in diffusion length are measured in varying proximity to DLDs, and correlations with the CL intensity, activation energy, polarization anisotropy, and luminescence energy are established. A detailed study of the CLWI, LPCL, activation energy imaging, and ambipolar diffusion length variations is performed for two samples whose $In_{0.2}Ga_{0.8}As$ channel thicknesses are just beyond the critical thickness in order that fundamental optical and transport properties can be examined near individual dislocations with a ~ 1 μm spatial resolution. An examination of relaxation- and defect-induced changes in carrier diffusion and luminescence, on a μm scale, is further essential in enabling an evaluation of fundamental design and growth parameters necessary for optimized HEMT performance.

II. EXPERIMENT

Six HEMT devices were grown by molecular-beam epitaxy.[21–24] Each sample consisted of the following layers (in order from substrate to the surface) grown on semi-insulating GaAs(001) substrates: a 1750-Å-thick GaAs undoped buffer layer; a $GaAs/Al_{0.25}Ga_{0.75}As$ MQW ($\sim$42-Å-thick QWs); a 3000-Å-thick undoped GaAs barrier; an $In_{0.2}Ga_{0.8}As$ QW (channel) varying in thickness from 75 to 300 Å; a 530 Å layer of $Al_{0.25}Ga_{0.75}As$, containing a δ-doped Si layer ($\sim 5\times10^{12}$ cm^{-2}) within about 30 Å of the channel; and a 50-Å-thick undoped GaAs cap. These samples were previously analyzed with transmission electron microscopy (TEM), triple-axis x-ray diffraction (XRD), and Hall measurements to establish a relationship between degradation of the device performance and the formation of $\langle 110\rangle$-oriented misfit dislocations.[21,22] That study established that the onset of a substantial device degradation occurred when misfit dislocations formed along both $\langle 110\rangle$ directions. High-resolution XRD further showed that the average In composition x in the samples varied from $0.203 \leqslant x \leqslant 0.214$.

Scanning monochromatic CL, panchromatic CL, CLWI, and LPCL were performed with a modified JEOL 840-A scanning electron microscope.[14–16] A rotatable linear polarizer was mounted *in vacuo* to perform polarization measurements.[25] The light collected was dispersed by a 0.25 m monochromator and detected with a liquid-nitrogen-cooled Ge p-i-n diode detector. CL spectra were obtained with a spectral resolution of 1 nm. An electron-beam energy of 12 keV with varying beam currents from 0.1 to 10 nA was used to probe the samples. The temperature of the samples was varied between 87 and 300 K, for the various CL measurements.

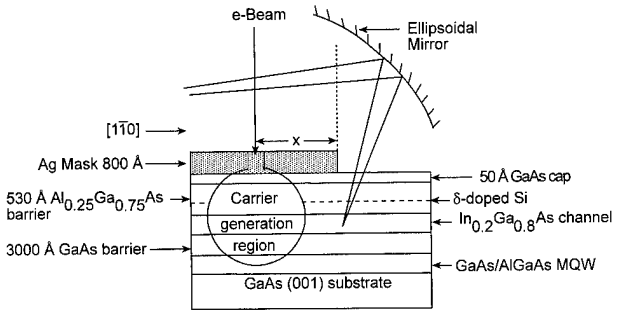


FIG. 1. Schematic diagram of the HEMT sample structure showing the ambipolar diffusion length experiment.

In CLWI, the wavelength λ_m at which the intensity of luminescence is a maximum is mapped as a function of the spatial (x,y) position, and a gray scale or false color image representing these wavelengths is generated.[14–16,26] A scanning area of 128×94 μm^2 is discretized into 640×480 pixels. In order to determine $\lambda_m(x,y)$, a spectrum consisting of 20 wavelength points (obtained from 20 discrete monochromatic CL images) was obtained at each (x,y) position, thereby enabling a mapping of the $In_{0.2}Ga_{0.8}As$ channel interband transition energy. The wavelength ranges for the 150 and 185 Å $In_{0.2}Ga_{0.8}As$ samples was from 960 to 965 nm and 967 to 977 nm, respectively. Spectrally integrated CL images (panchromatic in the specified wavelength range) were likewise obtained by summing the 20 discrete monochromatic images at each (x,y) pixel position.

LPCL imaging and spectroscopy measurements were taken with the polarizer rotated to detect the excitonic luminescence with $\mathbf{E}\perp[110]$ or $\mathbf{E}\parallel[110]$, where $\mathbf{E}$ is the electric field of the detected light. In order to emphasize the polarization variations, the ratio of the LPCL images at each (x,y) position is represented as $\log[I_\perp(x,y)/I_\parallel(x,y)]$, where $I_\perp$ and $I_\parallel$ are the pixel intensities under $\mathbf{E}\perp[110]$ and $\mathbf{E}\parallel[110]$ detection orientations, normalized to a 256-level gray scale. The $\perp$ and $\parallel$ subscripts are defined with respect to the [110] direction in this study.

The ambipolar diffusion lengths were measured using an approach illustrated in Fig. 1. The HEMT samples were coated with a 800-Å-thick Ag mask over part of the sample. The Ag film had lateral dimensions of 100×100 μm^2, and the edges of these squares were oriented along the $\langle 110\rangle$ directions. The e-beam energy of 12 keV was sufficiently large so that $\sim$80% of the beam penetrated the Ag film and

generated electrons and holes in the region just below the mask.[27] The Ag mask prevented light from radiative recombination in the generation region just below the mask from being detected by the CL collection system; however, luminescence from carriers which diffused along the $[1\bar{1}0]$ direction (as shown in Fig. 1) and recombined just beyond the edge of the mask was detected. By simultaneously scanning the e beam toward the edge of the mask and recording the integrated intensity of excitonic luminescence coming from the $In_{0.2}Ga_{0.8}As$ channel as a function of x, the distance from the mask edge, we have measured the ambipolar diffusion length L_D, of carriers in the channel. From a simple diffusion model, the CL intensity is proportional to $\exp(-x/L_D)$, as first demonstrated by Zarem *et al.*[28] for transport in $GaAs/Al_xGa_{1-x}As$ heterostructures using a similar CL experiment. Our diffusion length experiment was performed for various line scans parallel to and in varying proximity to the DLDs in the samples with 150 and 185 Å $In_{0.2}Ga_{0.8}As$ channel thickness. Owing to the formation of an orthogonal network of dislocations for channel thicknesses greater than 185 Å, we did not attempt this diffusion experiment in samples with thicker $In_{0.2}Ga_{0.8}As$ channels. A nonexponential dependence of the CL intensity would be expected for an e-beam crossing misfit dislocations in this situation, thus requiring more elaborate means to extract the diffusion lengths.

III. RESULTS AND DISCUSSION

A. Evaluation of the average strain relaxation in the HEMT samples

An anisotropy was observed in the density of ⟨110⟩-oriented dislocations, where dislocations first form along $[1\bar{1}0]$ and continue to have a greater density along this direction as the channel thickness increases. The two types of 60° dislocations are chemically inequivalent, owing to the difference in termination of the extra half-plane which, e.g., in the type-I (shuffle) set has a Ga and As termination, respectively, for the unreconstructed α ($[1\bar{1}0]$ line direction) and β ([110] line direction) dislocation cores. For a nonvicinal GaAs(001) substrate (i.e., nominally no misorientation) it is well established that for single thin $In_xGa_{1-x}As$ ($x \lesssim 0.2$) films grown on GaAs(001), α dislocations are the first to form in relaxing the strain.[13,29] This has previously been attributed to the different levels of stress required to nucleate α and β dislocations and the differences in α and β dislocation propagation velocities on nonvicinal GaAs(001) substrates.[30,31] Recently, studies have shown that a misorientation in the GaAs(001) substrate can also affect the degree of the asymmetry in the α and β dislocation density.[14,16,32] Assuming a predominance of 60°-type misfit dislocations in these samples, the average strain relaxation along a ⟨110⟩ direction is 0.02% for a linear dislocation density (LDD) of 1×10^4 cm^{-1}.[29] The maximum in-plane strain of 1.41% in the $In_{0.2}Ga_{0.8}As$ channel is therefore reduced in proportion to the orthogonal LDD. The average [110]- and $[1\bar{1}0]$-oriented LDD and resulting in-plane strains, ϵ_{110} and $\epsilon_{1\bar{1}0}$, are shown in Table I for the HEMT samples with various $In_{0.2}Ga_{0.8}As$ channel thicknesses. The linear dislocation densities were obtained from a combination of CL imaging and plan-view TEM. For CL imaging of DLDs, a maximum area of ~0.4 ×0.4 mm^2 limited by the field of view of the ellipsoidal mirror was used to determine the LDD.

TABLE I. Linear dislocation densities (LDD) along the [110] and $[1\bar{1}0]$ directions for the various $In_{0.2}Ga_{0.8}As$ channel thicknesses. The calculated ϵ_{110} and $\epsilon_{1\bar{1}0}$ strains for each sample are shown.

Channel thickness (Å)	[110] LDD (cm^{-1})	$[1\bar{1}0]$ LDD (cm^{-1})	ϵ_{110}	$\epsilon_{1\bar{1}0}$
75	$<1\times10^1$	$<1\times10^1$	0.014 129	0.014 129
150	3×10^3	$<1\times10^2$	0.014 068	0.014 127
185	5×10^3	4.4×10^2	0.014 027	0.014 120
205	1.7×10^4	5.5×10^2	0.013 785	0.014 117
250	4×10^4	1.8×10^3	0.013 313	0.014 092
300	1.16×10^5	2.5×10^3	0.011 769	0.014 078

In order to evaluate the average strain relaxation in the $In_{0.2}Ga_{0.8}As$ HEMT samples, we have examined spatially integrated CL spectra of these samples at room temperature, as shown in Fig. 2. An area of 128×94 μm^2 was scanned during the acquisition of the these spectra. The energy of the peak position is found to decrease from 1.217 to 1.179 eV as the channel thickness is increased from 150 to 300 Å. The first confined electron to heavy-hole (e1–hh1) transition en-

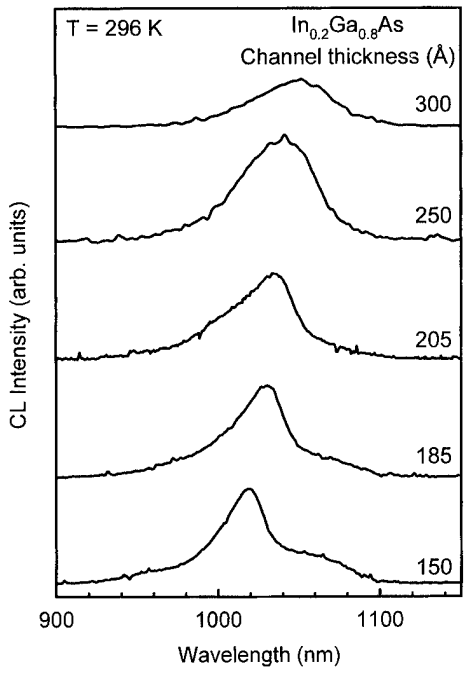


FIG. 2. Stack plot of spatially integrated CL spectra at room temperature for all HEMT samples.

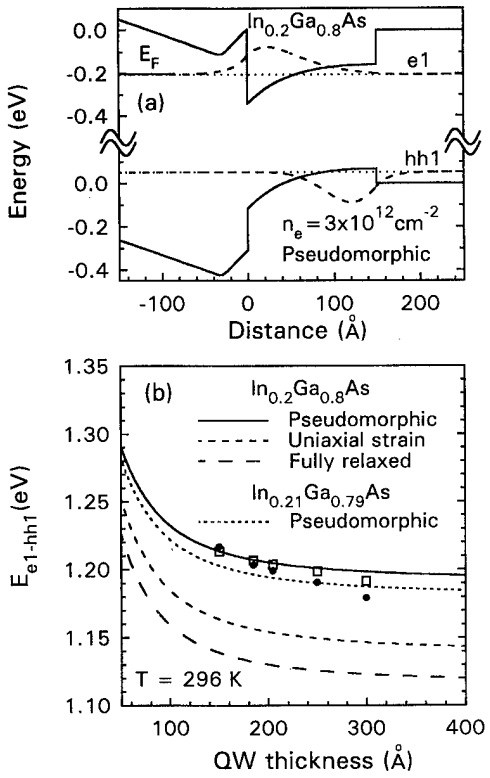


FIG. 3. Self-consistent-field calculation of the band profile for the $In_{0.2}Ga_{0.8}As$ HEMT sample (channel thickness of 150 Å) showing the (a) e1 and hh1 wave functions and Fermi-level position and (b) plots of the theoretical e1–hh1 luminescence energy position vs the $In_{0.2}Ga_{0.8}As$ QW width for the (i) fully strained, (ii) fully relaxed, (iii) uniaxially strained, and (iv) partially relaxed strain (squares) conditions using the data in Table I. The experimental CL peak positions from Fig. 2 are shown with dots in (b). Calculations of e1–hh1 vs the QW width are also shown for a fully strained $In_{0.21}Ga_{0.79}As$ QW, to illustrate the affect of an In composition variation.

ergy was calculated as a function of the channel thickness for the cases of

(i) fully biaxially strained (pseudomorphic),
(ii) uniaxially strained,
(iii) fully relaxed, and
(iv) partially relaxed $In_{0.2}Ga_{0.8}As$ channels.

The purpose of the top $Al_{0.25}Ga_{0.75}As$ barrier is to increase the confinement energy for electrons, thereby effectively increasing the electron density in the channel. The resulting asymmetrical QW structure (with a 150 Å width) subject to an electric field created by the ionized δ-doped Si layer is shown in Fig. 3(a).

The electron and hole envelope wave-function calculations were performed with a transfer-matrix method (TMM) technique using a self-consistent-field approximation that includes the Hartree term.[10,33–35] All occupied confined electron states, as determined by the Fermi-level position, were

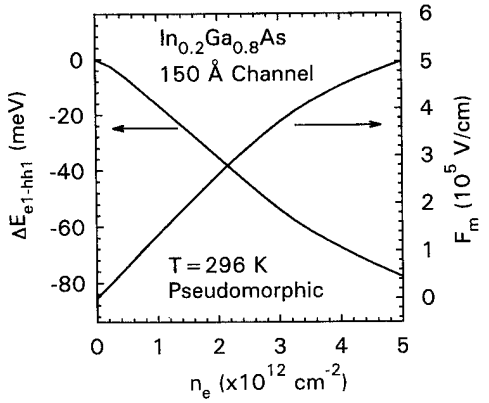


FIG. 4. Self-consistent-field calculation of the maximum electric field F_m in the $In_{0.2}Ga_{0.8}As$ channel and change in the e1–hh1 transition energy vs electron concentration n_e at room temperature.

used in calculating the field as a function of position in the channel. A similar method was employed to calculate the field self-consistently in a nipi-doped $In_{0.2}Ga_{0.8}As$ MQW structure, as discussed in Ref. 10. The wave functions and potential were calculated self-consistently using Airy functions in the TMM by discretizing the potential into discrete linear field regions with ~5 Å widths. The use of narrower widths only negligibly affected the calculated electron and hole eigenstates. The conduction to valence-band offset ratios ($\Delta E_c/\Delta E_v$) at both interfaces of the $In_xGa_{1-x}As$ QW were taken as linear interpolations between 70/30 and 60/40 for $GaAs/In_xGa_{1-x}As$ and $GaAs/Al_yGa_{1-y}As$ interfaces, respectively.[36] The results of a calculation showing the band diagram, ground-state electron and hole wave functions, and Fermi level for the pseudomorphic case is shown in Fig. 3(a). The ground-state electron and heavy-hole wave functions are shown superimposed on the calculated band profile in Fig. 3(a). The maximum electric field F_m in the $In_{0.2}Ga_{0.8}As$ channel (for a 150 Å width) and the e1–hh1 transition energy versus the electron concentration in the channel n_e at room temperature is shown in Fig. 4. The maximum field in the channel occurs at the $Al_{0.25}Ga_{0.75}As/In_{0.2}Ga_{0.8}As$ interface, resulting in Stark shifts of the confined electron and hole states. The field-induced Stark shift decreases for a reduced electron concentration in the channel and Si δ-doping concentration, resulting in a larger e1–hh1 transition energy as n_e decreases (as shown in Fig. 4).

The strain-induced changes in the $In_{0.2}Ga_{0.8}As$ band edges were calculated using the 4×4 Luttinger–Kohn and Pikus–Bir Hamiltonian for a general in-plane strain, $\epsilon_{110} \neq \epsilon_{1\bar{1}0}$, referred to the ⟨110⟩ dislocation directions.[37,38] The strain-induced energy change ΔE of the band gap involving the $j=3/2$ valence bands is given by the following solution of the orbital-strain Hamiltonian for $\mathbf{k}=0$:[15,37]

$$\Delta E = -a(\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}) \pm \tfrac{1}{2}\sqrt{4d^2\epsilon_{xy}^2 + b^2(2\epsilon_{zz}-\epsilon_{xx}-\epsilon_{yy})^2 + 3b^2(\epsilon_{xx}-\epsilon_{yy})^2}, \quad (1)$$

where $\epsilon_{xx} = \epsilon_{yy} = (\epsilon_{110} + \epsilon_{1\bar{1}0})/2$, $\epsilon_{xy} = (\epsilon_{1\bar{1}0} - \epsilon_{110})/2$, and $\epsilon_{zz} = -2\epsilon_{xx}C_{12}/C_{11}$; ϵ_{110} and $\epsilon_{1\bar{1}0}$ are the strains along [110] and [1$\bar{1}$0] directions, respectively. The constant a is the hydrostatic deformation potential; b and d are uniaxial deformation potentials associated with strains of tetragonal and rhombohedral symmetries, respectively, which remove the degeneracy of the bands as indicated by the $\pm$ sign; C_{11} and C_{12} are elastic constants; these constants for $In_xGa_{1-x}As$ are found by interpolating between values for GaAs and InAs.[38,39] The low symmetry of a $\epsilon_{110} \neq \epsilon_{1\bar{1}0}$ strain in the $In_{0.2}Ga_{0.8}As$ channel required the use of the Luttinger–Kohn Hamiltonian to determine the $In_{0.2}Ga_{0.8}As$ effective mass of holes along the [001] growth direction. Standard hole masses and band gaps are used for $Al_xGa_{1-x}As$. The Luttinger parameters are taken as a linearization, again, between the values for InAs and GaAs.[38] The effective masses and strain-modified barrier heights were then used in the TMM calculation to calculate the e1–hh1 transition energies for

(i) the pseudomorphic case ($\epsilon_{110} = \epsilon_{1\bar{1}0} = 0.0141$),
(ii) the uniaxial strain condition ($\epsilon_{110}=0$ and $\epsilon_{1\bar{1}0} = 0.0141$),
(iii) the fully relaxed case ($\epsilon_{110} = \epsilon_{1\bar{1}0} = 0$), and
(iv) the partially relaxed case with general strain values, ϵ_{110} and $\epsilon_{1\bar{1}0}$, as obtained from the measured dislocation densities shown in Table I for each channel thickness.

The results are shown in Fig. 3(b) as solid lines, medium-dashed lines, long-dashed lines, and squares for cases (i)–(iv), respectively. The experimental e1–hh1 peak energies (dots) are found to lie closest to the theoretical curve for the pseudomorphic case ($\epsilon_{110} = \epsilon_{1\bar{1}0} = 0.0141$).

The deviation between the experimental peak positions and pseudomorphic calculation increases as the channel thickness increases, consistent with an increased strain relaxation of the $In_{0.2}Ga_{0.8}As$ channel. The use of the measured strains of Table I resulted in a better agreement between the experimental and calculated e1–hh1 transition energies, as shown in Fig. 3(b). The CL imaging and plan-view TEM revealed an increase in the dislocation density with increasing channel thickness as shown in Table I. It is apparent that even for the case of the largest channel width of 300 Å, the largest dislocation density yields a strain of $\epsilon_{110}=0.0118$, and the film is still ~84% strained in the [110] direction. Therefore, the TMM calculations, when incorporating the observed strain relaxation, explain the salient features of the CL peak energies for varying channel thicknesses. Further, the deviation between the experiment and calculations for the partially relaxed case [dots and squares, respectively, in Fig. 3(b)] is likely due to a variation in the In composition x. A calculation of the e1–hh1 energy for a fully strained $In_{0.21}Ga_{0.79}As$ QW (short-dashed lines) is shown in Fig. 3(b), indicating a ~10 meV decrease in the calculated e1–hh1 energy will occur for cases (i) and (iv) above if $x=0.21$ is used instead of $x=0.20$ in these calculations.

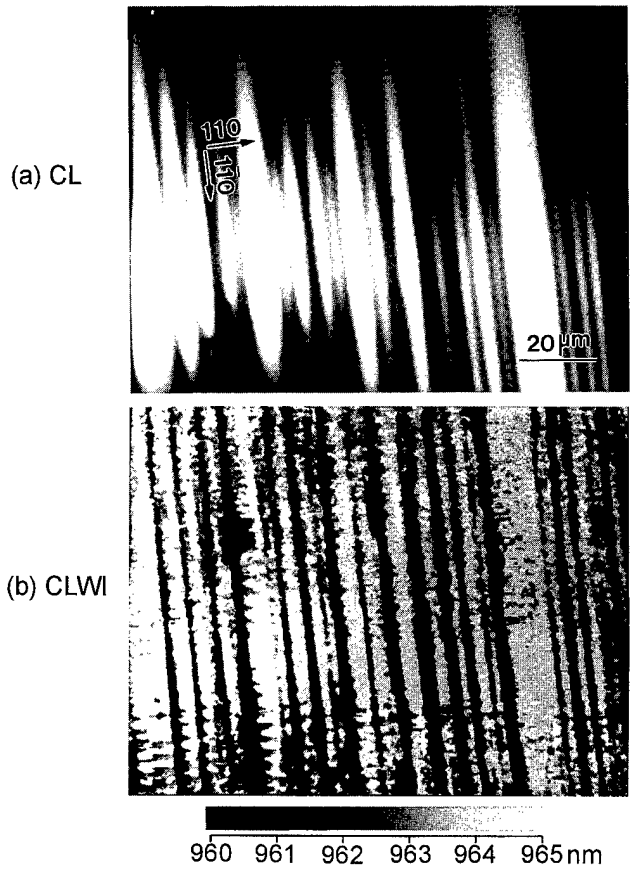


FIG. 5. (a) CL intensity and (b) CLWI images of the same regions in the $In_{0.2}Ga_{0.8}As$ HEMT sample with 150 Å channel thickness.

B. Local CL energy and polarization variations in close proximity to dislocations

Recently, we have established that there are significant energy and polarization variations in the optical transitions in close proximity to misfit dislocations, in addition to the non-radiative behavior of DLDs.[14–16] The previous systems examined were partially relaxed $In_xGa_{1-x}As/GaAs$ films which had linear dislocations densities greater than $\sim 1\times10^4$ cm^{-1} or one dislocation per μm. Since the carrier diffusion length is ~1 μm, defect densities greater than 1×10^4 cm^{-1} result in DLDs that are composed of bunches of dislocations, which cannot be resolved individually with conventional CL imaging. Bunches of dislocations that are formed with very narrow dislocation spacings can result in nearly complete strain relaxation along one ⟨110⟩ direction with a partial strain remaining in the orthogonal direction. This can further result in a quasi-uniaxial stress leading to a large $I_\parallel/I_\perp$ polarization anisotropy and a reduction in the excitonic transition energy, as previously reported.[14–16]

We have pursued a similar analysis here for the case of dislocations separated by a length on average greater than the carrier diffusion length, so as to examine the effects of isolated dislocations. Figures 5 and 6 show (a) CL intensity and (b) CLWI images of samples with 150 and 185 Å channel thicknesses, respectively. The false-color scale represents the

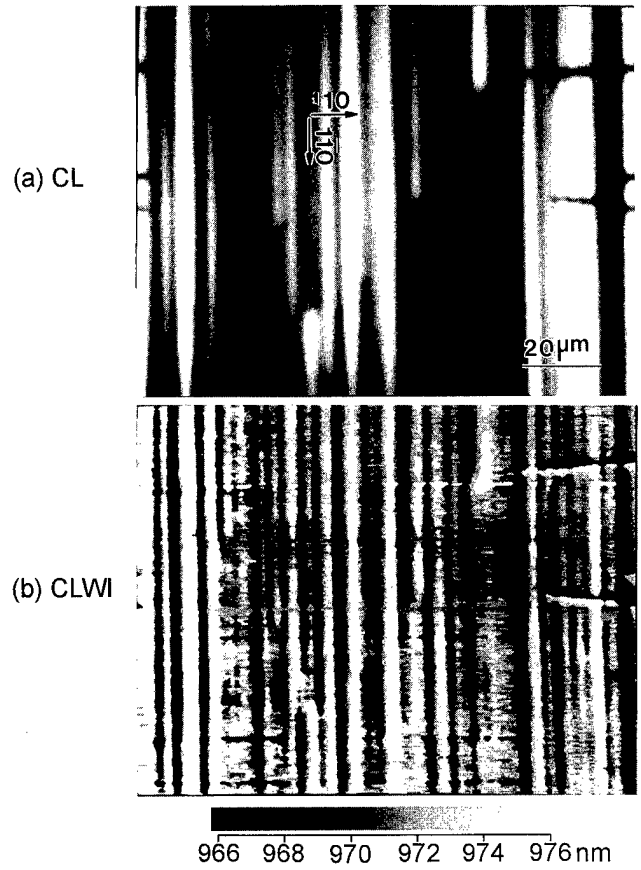


FIG. 6. (a) CL intensity and (b) CLWI images of the same regions in the $In_{0.2}Ga_{0.8}As$ HEMT sample with 185 Å channel thickness.

wavelength position λ_m of the peak CL intensity. A particularly striking feature is observed in the CLWI images. The wavelength of emission is found to decrease near the DLD position, showing a blue shift correlated with a defect-induced CL intensity reduction. Figures 7 and 8 each show a line scan analysis for an arbitrary line scan taken perpendicular to the [1$\bar{1}$0] dislocation direction and illustrate a one-to-one correlation between the blue shift and the DLD position. A maximum increase of ~5 and ~10 meV is seen at the DLD centers for the 150 and 185 Å samples, respectively, in Figs. 5–8. This behavior appears contradictory to the previously observed red shift measured near DLD positions in $In_xGa_{1-x}As$ samples exhibiting a greater strain relaxation and greater dislocation-induced reductions in the ϵ_{110} strain.[14–16] This apparent discrepancy is, however, explained by the influence of the dislocations and associated point defects on the electron density in the channel. It is our hypothesis that these defects create localized deep levels and traps that reduce the effective electron density in the channel near DLDs, thereby simultaneously reducing the field in the channel. The reduction in both the electron density and field will concomitantly reduce the Stark shift, thereby resulting in a local increase in the e1–hh1 emission energy. A similar behavior was observed for defect-induced reductions in the electric field and Stark shifts for nipi-doped $In_{0.2}Ga_{0.8}As/GaAs$ MQWs.[10] The larger blue shift in the e1–hh1 transition energies for the 185 Å sample (Figs. 6 and 8) is evidently due to the enhanced dislocation density. Dislocation bunching and an enhanced point defect density are more prevalent for larger $In_{0.2}Ga_{0.8}As$ channel thicknesses, which should result in a greater local depletion of the electron density in the channel.

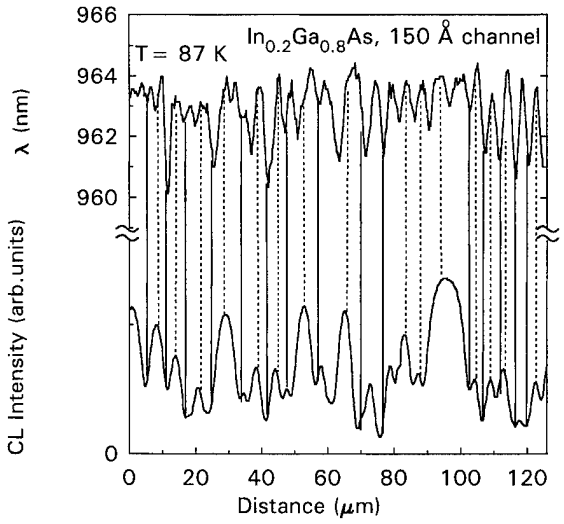


FIG. 7. Line scan analysis for an arbitrary [110]-oriented line for the images of Fig. 5 (150 Å $In_{0.2}Ga_{0.8}As$ channel width) showing the CL intensity and CLWI position correlations. Dashed and solid vertical lines are used to show correlations between a reduced CL intensity (DLDs) and a blue shift in the e1–hh1 transition energy.

The local strain relaxation appears to minimally effect the e1–hh1 transition energy. As shown in Fig. 3, the curve for a uniaxial strain condition is ~50 meV lower than that for the pseudomorphic case. A defect-induced reduction in the quan-

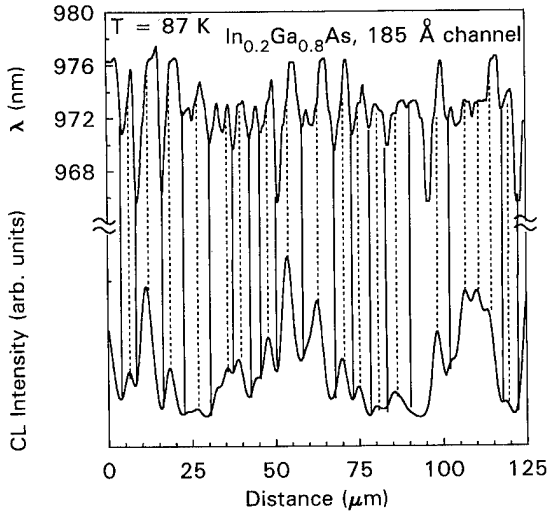


FIG. 8. Line scan analysis for an arbitrary [110]-oriented line for the images of Fig. 6 (185 Å $In_{0.2}Ga_{0.8}As$ channel width) showing the CL intensity and CLWI position correlations. Dashed and solid vertical lines are used to show correlations between a reduced CL intensity (DLDs) and a blue shift in the e1–hh1 transition energy.

tum confined Stark effect could also cause a blue shift of ~50 meV in the e1–hh1 transition energy, as seen from Fig. 4. That is, strain relaxation and defect-induced reductions in the field cause the e1–hh1 energy to shift in opposite directions, thereby possibly masking the effects caused by strain relaxation. However, from the CL polarization results discussed below, we show that the average strain tensor within the ~1 μm carrier diffusion length near DLDs is still well described by a biaxial strain with $\epsilon_{110} \approx \epsilon_{1\bar{1}0}$, for the 150 and 185 Å samples analyzed in Figs. 5–8.

Linearly polarized monochromatic CL images were taken with the polarizer rotated to detect emission of light with $\mathbf{E}\perp[110]$ and $\mathbf{E}\|[110]$ detection orientations at wavelengths of 962 and 972 nm for the 150 and 185 Å samples, respectively. In order to emphasize the polarization variations, the ratio of these images is displayed in Figs. 9 and 10. The pixels in the ratio image at a (x,y) position are presented as $\log[I_\perp(x,y)/I_\|(x,y)]$, where $I_\perp$ and $I_\|$ are the pixel intensities under $\mathbf{E}\perp[110]$ and $\mathbf{E}\|[110]$ detection orientations. The bright and dark bands present in the LPCL ratio images exhibit a local polarization anisotropy, which indicates the presence of μm-scale variations in ϵ_{110}. These bands correlate with the peaks and dips in the CL intensity image, as shown in the line scan analysis of Fig. 11 for the 185 Å sample. The maximum polarization anisotropies (minimum ratios) are $I_\perp/I_\| \approx 0.95$ and $I_\perp/I_\| \approx 0.85$ for the 150 and 185 Å samples, respectively. The spatially averaged $I_\perp/I_\|$ ratios for the LPCL images of Figs. 9(b) and 10(b) are 0.98 and 0.95, respectively. From the four-band $\mathbf{k}\cdot\mathbf{p}$ calculation discussed above, we have calculated $I_\perp/I_\|$ using the dipole approximation in Fermi's golden rule, i.e., $I_{\perp,\|} \propto |\langle u_e|\mathbf{E}_{\perp,\|}\cdot\mathbf{p}|u_h\rangle|^2$, where u_e is the electron wave function, u_h is the wave function of the uppermost hole state, and $\mathbf{p}$ is the linear momentum operator. Both u_e and u_h include the envelope wave functions and zone-center Bloch functions for the s- and p-type conduction- and valence-band states, respectively. The polarization ratio was calculated for a fixed $\epsilon_{1\bar{1}0} = 0.0141$ and variable ϵ_{110}, to simulate the effect of a transition from uniaxial to biaxial (pseudomorphic) strain. The results are shown in Fig. 12 for the 150 and 185 Å samples, where we also show the calculated e1–hh1 transition energy as a function of ϵ_{110}. From the measured average polarization anisotropy ratios above and the calculation in Fig. 12, the estimated values for ϵ_{110} are 0.0138 and 0.0132 (i.e., 98% and 94% strained) for the 150 and 185 Å samples, respectively. This is in reasonable agreement with the low strain relaxation and ϵ_{110} values observed in Table I from the linear dislocation densities for these samples. These results contrast with the situation previously studied for highly relaxed $In_xGa_{1-x}As$ samples with dislocation densities greater than $\sim 1\times10^5$ cm^{-1}, where dislocation bunching lead to a quasiuniaxial strain and a larger polarization anisotropy.[14–16] Thus, for individual dislocations studied here, within the minority carrier diffusion length of ~1 μm, the presence of dislocations with densities less than $\sim 1\times10^4$ leads to a measured luminescence behavior still well described by a biaxial strain, $\epsilon_{110} \approx \epsilon_{1\bar{1}0}$.

C. The CL temperature dependence and spatial variations in the activation energy

The integrated CL intensity I_T of the e1–hh1 transition in the $In_{0.2}Ga_{0.8}As$ channel and the excitonic luminescence of the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW was measured as a function of temperature for selected local regions in close proximity to and away from dislocations, labeled D and B, respectively to denote dark and bright regions in the $In_{0.2}Ga_{0.8}As$ CL imaging. The results are shown in Fig. 13 for both the 150 and 185 Å samples in a log I_T vs $1000/T$ plot. The CL intensities reduce as the temperature increases from 87 to 250 K. This reduction corresponds to the increase of thermally activated nonradiative recombination, which causes an Arrhenius dependence in the high-temperature range. In previous reports of the temperature-dependent luminescence efficiency,[17,18,40] the linear region of the Arrhenius behavior in the high-temperature range was characterized by one or two thermally activated nonradiative recombination processes. We use the following fitting equation for the temperature dependence of the $In_{0.2}Ga_{0.8}As$ QW CL intensity:

$$I_T = R\,\eta, \tag{2}$$

where $\eta = 1/(1+R_{nr}/R_r)$ is the quantum efficiency,[41,42] R is a coefficient which depends on the generation rate of electron–hole pairs and the relative weights of monomolecular and bimolecular recombination,[41] R_r is the radiative recombination rate which is assumed to be temperature independent, and R_{nr} represents the rate for nonradiative recombination and is assumed to have the following temperature dependence:

$$R_{nr} = R_{nr1} + R_{nr0}\exp(-E_a/kT), \tag{3}$$

where R_{nr0} and E_a are the temperature-independent prefactor and the thermal activation energy,[40,42] respectively, and R_{nr1} is the rate for nonthermally activated nonradiative recombination (i.e., independent of temperature). The temperature-dependent exponential term is due to the enhancement of the capture cross section of nonradiative recombination centers as seen by carriers as the temperature increases. The model of Eq. (3) is motivated by the existence of DLDs in the CL imaging at the lowest temperatures, indicating that there are nonradiative channels which are independent of any thermal activation. Therefore,

$$I_T = \frac{\theta}{1+\beta\exp(-E_a/kT)}, \tag{4}$$

where $\theta = R/(1+R_{nr1}/R_r)$ and $\beta = R_{nr0}/(R_r+R_{nr1})$ are independent of temperature and depend on the density of nonthermally and thermally activated nonradiative recombination centers. At the low-temperature limit, η saturates since thermally activated nonradiative recombination vanishes and $I_T = \theta$. The spatial variation in θ therefore accounts for the difference in I_T, the e1–hh1 emission intensity, when $T \lesssim 100$ K, as shown in Figs. 9, 10, and 13. The solid lines in Fig. 13 show a fit of Eq. (4) to the experimental CL data for the e1–hh1 transition at the corresponding B (bright) and D

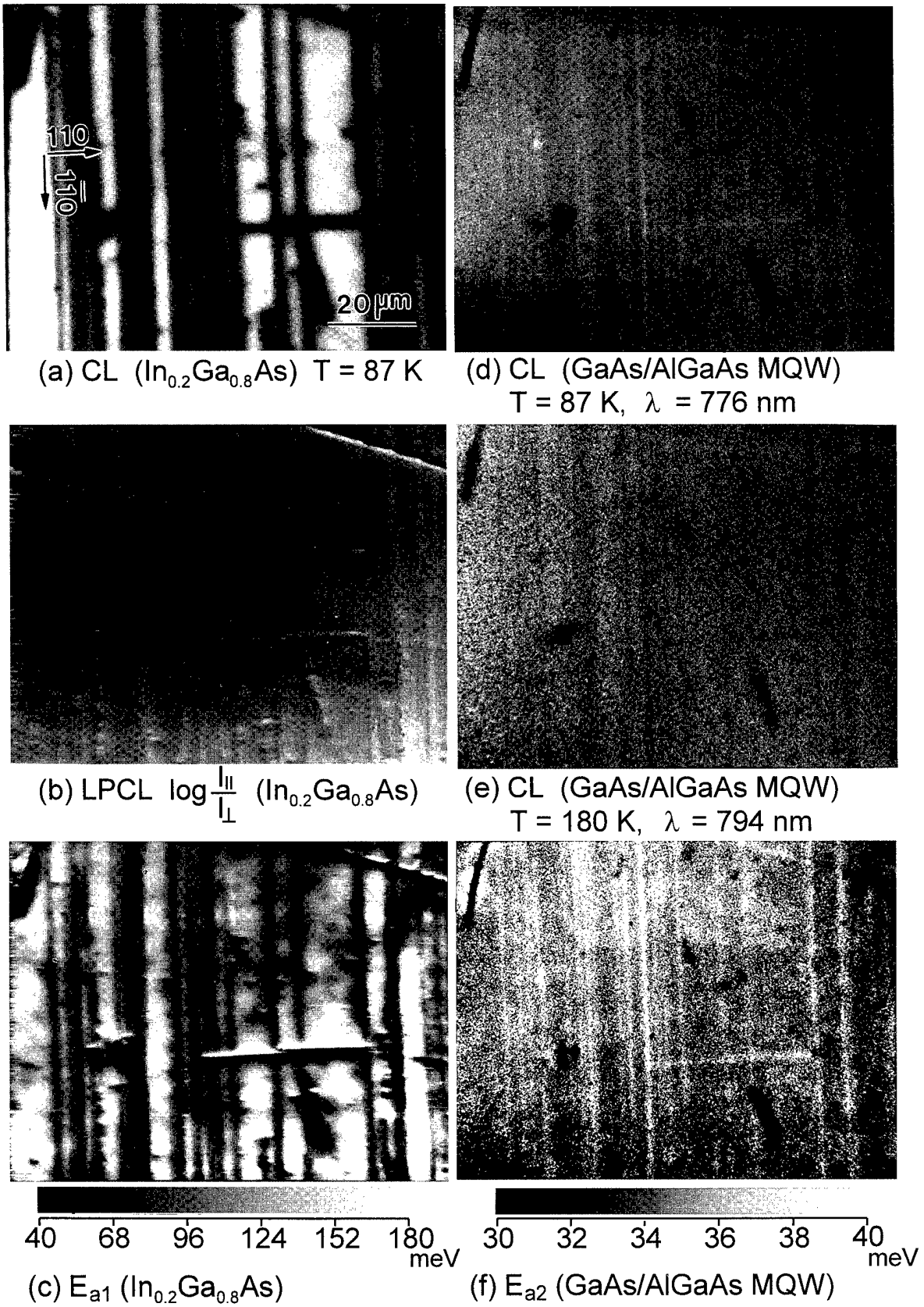


FIG. 9. CL imaging of the 150 Å $In_{0.2}Ga_{0.8}As$ HEMT sample showing (a) spectrally integrated CL intensity images for the e1–hh1 emission, (b) LPCL $I_{\perp}/I_{\parallel}$ ratio images, (c) activation energy E_{a1} images for the $In_{0.2}Ga_{0.8}As$ QW luminescence, monochromatic CL images for $GaAs/Al_{0.25}Ga_{0.75}As$ MQW at (d) $T=87$ K and (e) 180 K, and (f) activation energy E_{a2} images for the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW emission.

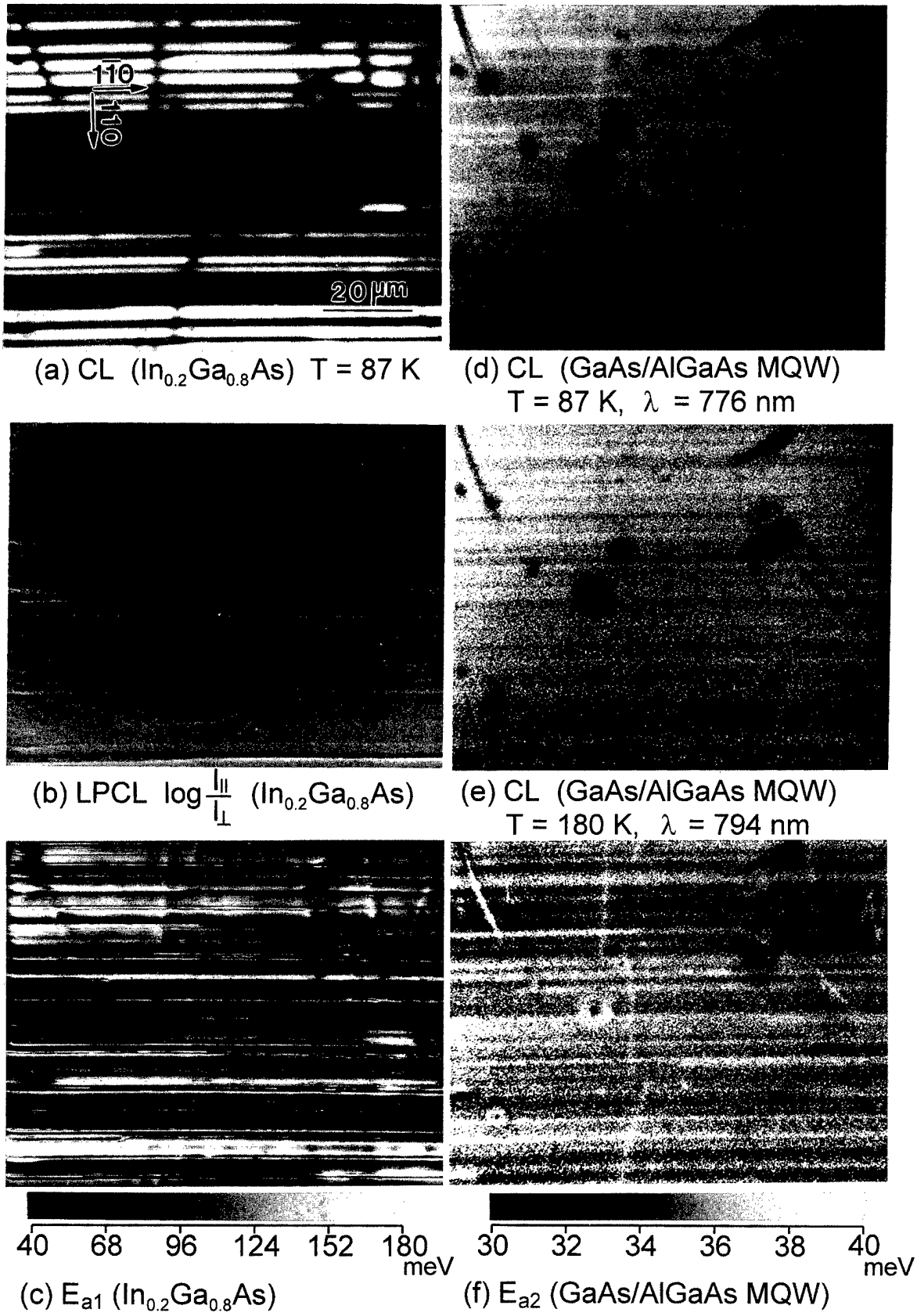


FIG. 10. CL imaging of the 185 Å $In_{0.2}Ga_{0.8}As$ HEMT sample showing (a) spectrally integrated CL intensity images for the e1–hh1 emission, (b) LPCL $I_{\perp}/I_{\parallel}$ ratio images, (c) activation energy E_{a1} images for the $In_{0.2}Ga_{0.8}As$ QW luminescence, monochromatic CL images for GaAs/$Al_{0.25}Ga_{0.75}As$ MQW at (d) $T=87$ K and (e) 180 K, and (f) activation energy E_{a2} images for the GaAs/$Al_{0.25}Ga_{0.75}As$ MQW emission.

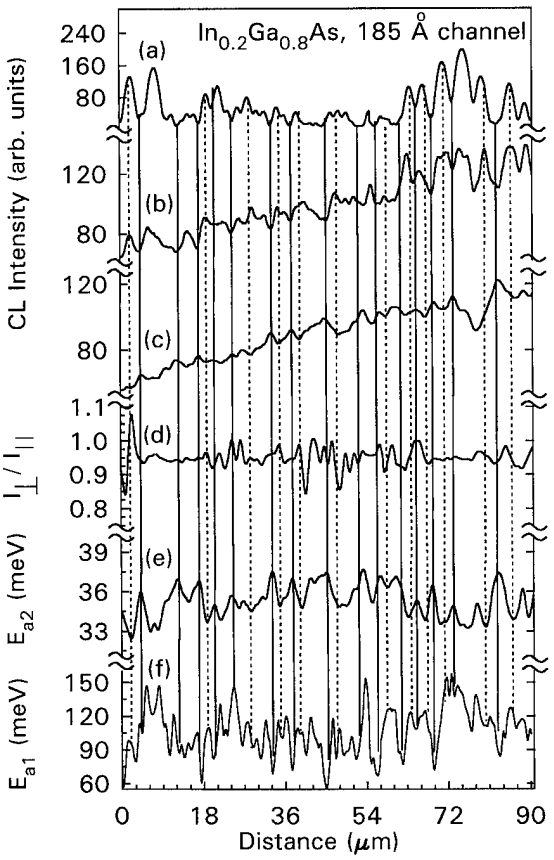


FIG. 11. Line scan analysis for an arbitrary [110]-oriented line for the images of Fig. 10 (185 Å $In_{0.2}Ga_{0.8}As$ channel width) showing (a) the CL intensity for the e1–hh1 emission, $GaAs/Al_{0.25}Ga_{0.75}As$ MQW emission intensity at (b) $T=180$ K and (c) 87 K, (d) LPCL $I_\perp/I_\parallel$ ratio, and activation energies (e) E_{a2} and (f) E_{a1}.

(dark) regions. The activation energies E_{a1}, were determined for each pixel position by fitting all 640×480 pixel intensity values for the monochromatic CL images of the e1–hh1 transition taken at 12 different temperatures. The results of E_{a1} for each (x,y) position are shown in Figs. 9(c) and 10(c), for

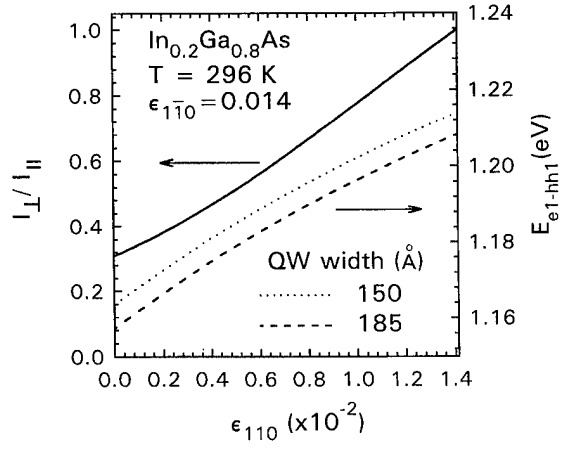


FIG. 12. Calculation of the e1–hh1 $I_\perp/I_\parallel$ emission ratio and $In_{0.2}Ga_{0.8}As$ QW transition energy vs ϵ_{110} for a fixed $\epsilon_{1\bar{1}0} = 0.0141$ at room temperature.

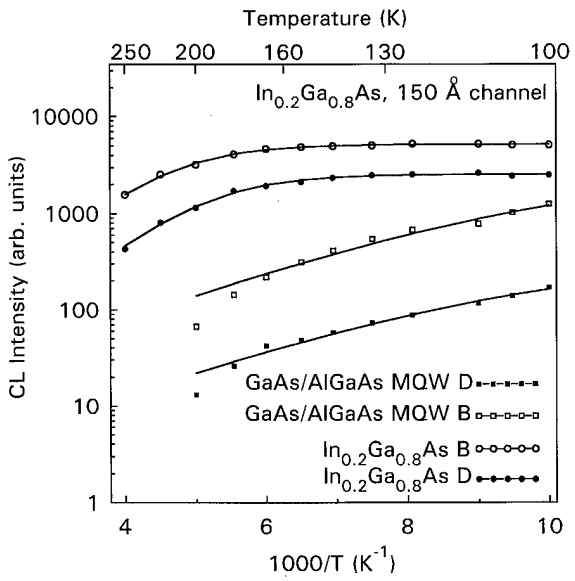


FIG. 13. CL intensity vs $1000/T$ for the same local bright (B) and dark regions (D) in the $In_{0.2}Ga_{0.8}As$ e1–hh1 and $GaAs/Al_{0.25}Ga_{0.75}As$ MQW monochromatic CL images. The solid lines running through the curves are a fit of Eq. (4) to the data to determine the activation energies E_{a1} and E_{a2} for the $In_{0.2}Ga_{0.8}As$ QW and $GaAs/Al_{0.25}Ga_{0.75}As$ MQW luminescence, respectively.

the 150 and 185 Å samples, respectively. A line scan analysis is also shown in Fig. 11(f) for the 185 Å sample.

For the luminescence originating from the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW, I_B, a low-temperature saturation of its intensity was not reached for the 87 K minimum in this study (as shown in Fig. 13). We have also fit the CL images of I_B, taken at the 12 different temperatures, with the model of Eqs. (2)–(4), obtaining the activation energy, E_{a2} for the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW excitonic luminescence. The results of two fits are shown by the solid lines in the log plot of Fig. 13 for the same regions, B and D, as indicated for the e1–hh1 $In_{0.2}Ga_{0.8}As$ luminescence. We further observe a change in the relative intensity of the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW emission near features which correspond to DLDs in the CL imaging of the e1–hh1 $In_{0.2}Ga_{0.8}As$ QW emission. At $T=180$ K, for both the 150 and 185 Å samples, the imaging of the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW emission [Figs. 9(e) and 10(e)] results in DLDs which correlate, one to one, with that of the e1–hh1 imaging, as also shown in the line scan data of Fig. 11. As the temperature is lowered, a reversal in the relative intensity of the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW emission occurs near defects, resulting in bright lines in the CL images, as shown in Figs. 9(d) and 10(d). This striking contrast reversal is also illustrated in the line scan analysis of Figs. 11(b) and 11(c). That is, high-temperature DLDs in the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW emission become bright line defects (BLDs) at lower temperatures. The activation energy for this emission E_{a2} increases near these defects, as indicated by the imaging and line scan analysis. This is in contrast to the decrease in E_{a1} near DLDs.

This behavior reflects salient differences in the thermal

activation of carriers which are in close proximity, but on opposite sides, of the $In_{0.2}Ga_{0.8}As/GaAs$ interface. The E_{a1} energy represents the activation energy for thermal reemission of carriers out of the $In_{0.2}Ga_{0.8}As$ QW, as has been observed for similar QW and MQW systems.[17–20] Once out of the $In_{0.2}Ga_{0.8}As$ QW, the carriers can be recaptured by the QW, recombine in the GaAs barrier, or diffuse to the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW where recombination can occur. An extremely weak GaAs near-band-edge luminescence was detected relative to emissions from the $In_{0.2}Ga_{0.8}As$ QW and $GaAs/Al_{0.25}Ga_{0.75}As$ MQW, indicating the very low impurity concentration and high quality of the GaAs barrier layer. Carriers which are generated in the GaAs barrier will then primarily diffuse to the underlying $GaAs/Al_{0.25}Ga_{0.75}As$ MQW or $In_{0.2}Ga_{0.8}As$ QW, where a higher carrier capture rate in these layers is expected, owing to quantum capture. The 3000 Å thickness of the GaAs barrier is less than the ~1 μm ambipolar diffusion length for high-quality and low-impurity GaAs, thereby enabling the GaAs barrier to serve as a conduit for efficient transport of electrons and holes to and from the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW and $In_{0.2}Ga_{0.8}As$ QW.

The defects created by the misfit dislocations further introduce other recombination channels for carriers in the $In_{0.2}Ga_{0.8}As$ QW. These defects enhance the probability for thermally assisted nonradiative recombination for carriers already residing in the $In_{0.2}Ga_{0.8}As$ QW, thereby resulting in a decrease in E_{a1} near DLDs. The analysis in Figs. 9(c), 10(c), and 11(f) shows that there is a maximum decrease in E_{a1} of ~60 and ~80 meV near DLDs for the 150 and 185 Å samples. The greater reduction in E_{a1} for the 185 Å sample reflects the enhanced defect density.

For carriers recombining in the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW, the misfit dislocations, likewise, introduce additional thermally assisted nonradiative channels which are accessible at higher temperatures. These channels may compete with carrier capture by the $In_{0.2}Ga_{0.8}As$ QW and subsequent radiative recombination. However, at lower temperatures, carrier capture by these defects is substantially reduced and simultaneously results in a reduced transfer of carriers from the $GaAs/Al_{0.25}Ga_{0.75}As$ MQW to the $In_{0.2}Ga_{0.8}As$ QW at defect sites, thereby resulting in a local increase of E_{a2}. These defect sites also appear to act as a barrier to transport into the $In_{0.2}Ga_{0.8}As$ QW at low temperatures, thereby enhancing the relative $GaAs/Al_{0.25}Ga_{0.75}As$ MQW emission near dislocations and resulting in BLDs correlated with the $In_{0.2}Ga_{0.8}As$ DLDs. The spatially averaged values of E_{a2} [from Figs. 9(f) and 10(f)] are 32.3 and 35.3 meV for the 150 and 185 Å samples. The larger value in the latter sample again likely reflects the larger relaxation-induced defect density in that sample.

D. Spatial variations in the carrier diffusion length

The diffusion length experiment was performed by scanning the e beam along the $[1\bar{1}0]$ direction (parallel to the DLDs) and recording the intensity of the e1–hh1 luminescence as a function of the distance x from the mask edge (as illustrated in Fig. 1). Typical scans are shown for the 150 and 185 Å $In_{0.2}Ga_{0.8}As$ samples at four different regions in each sample, labeled a–d, in Fig. 14. These regions are further identified in the CL line scan analysis of Figs. 15 and 16, which show the $In_{0.2}Ga_{0.8}As$ QW luminescence intensity versus distance (bottom scan) along [110] (i.e., perpendicular to the dislocation line direction). The dips in the CL intensity scan, again, represent the DLDs. For both samples, regions a and b are away from DLDs, while c and d are near the center of DLDs. From Fig. 14 it is apparent that a reduction in CL intensity at these regions is also accompanied by a greater negative slope, resulting in a smaller ambipolar diffusion length. This correlation is illustrated in Figs. 15 and 16 for several diffusion length measurements performed in varying proximity to DLDs at different temperatures. For both samples, regions near and far from the DLDs correspond to regions of smaller and larger diffusion lengths, respectively. A wide variation in L_D is observed from about 0.5 to 3 μm in Fig. 17, showing that defect regions can substantially influence the transport. No clear systematic variations in L_D with temperature are observed.

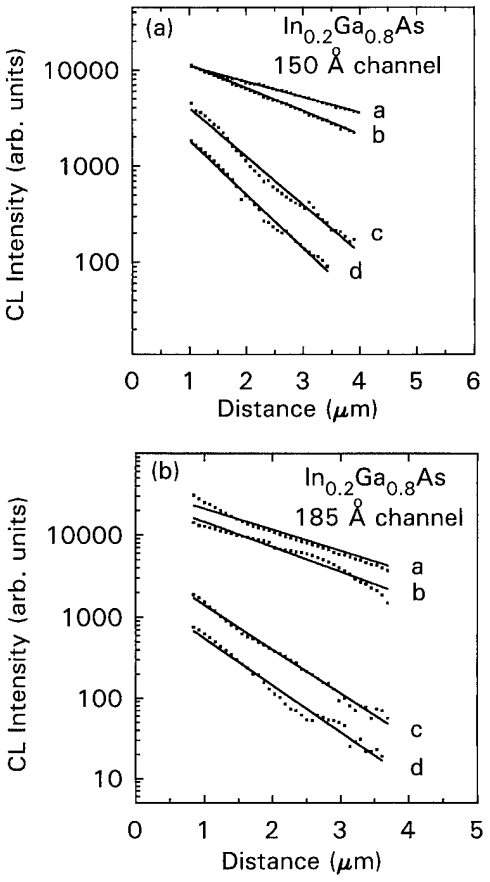


FIG. 14. CL intensity vs beam position x for different line scans (labeled a–d) parallel to the $[1\bar{1}0]$ dislocation line direction near and between DLDs for both the 150 and 185 Å $In_{0.2}Ga_{0.8}As$ HEMT samples.

From previous Hall measurements of these samples, the $In_{0.2}Ga_{0.8}As$ channel contains a large electron concentration of $\sim 3\times10^{12}$ cm^{-2} at 77 K.[21,22] The L_D measured here therefore reflects the diffusion of the minority carriers, i.e., holes.

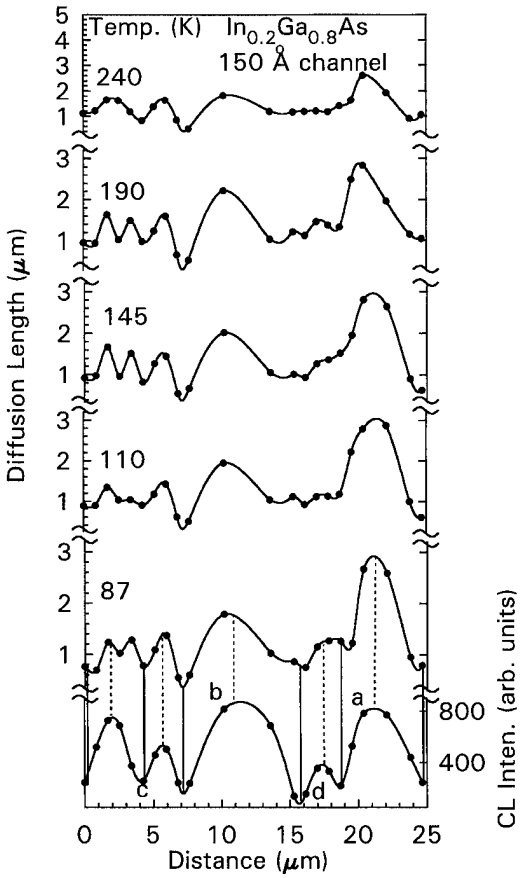


FIG. 15. CL intensity and diffusion length plot vs distance along [110] (perpendicular to the dislocation direction) for various temperatures for the 150 Å sample. The dashed and solid vertical lines show the spatial correlations between the DLD positions and a reduced diffusion length.

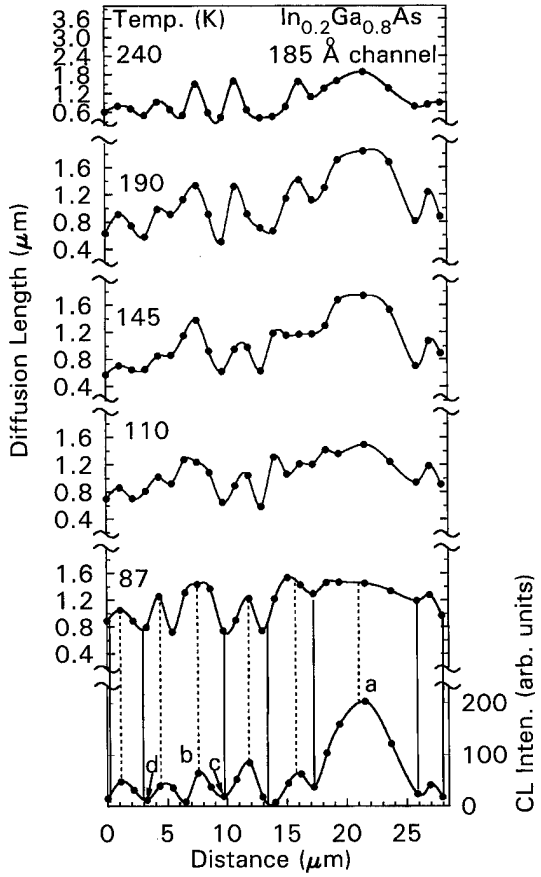


FIG. 16. CL intensity and diffusion length plot vs distance along [110] (perpendicular to the dislocation direction) for various temperatures for the 185 Å sample. The dashed and solid vertical lines show the spatial correlations between the DLD positions and a reduced diffusion length.

This hole diffusion length will involve an interplay between local changes in mobility μ_p, and minority-carrier recombination lifetime τ_p as $L_D=(D_p\tau_p)^{1/2}$, where μ_p and D_p are related by the Einstein relation $\mu_p=eD_p/kT$.[43] The presence of defects will evidently reduce both μ_p and τ_p as a result of enhanced scattering near defects and introduction of nonradiative recombination channels. As discussed in Sec. III A, the enhanced defect density near dislocations is expected to locally reduced the electric field. The reduction in the field is also expected to result in a decrease in τ_p since the electron and hole envelope wave-function overlap will also increase with a decrease in the field [see Fig. 3(a)]. However, without a quantitative measurement of the lifetime and its variations we refrain from attempting to deconvolve L_D into separate μ_p and τ_p terms here.

IV. SUMMARY AND CONCLUSIONS

A detailed study of fundamental optical and transport properties of two $In_{0.2}Ga_{0.8}As$ HEMT samples has been performed for channel thicknesses that just exceed the $In_{0.2}Ga_{0.8}As$ critical thickness. This facilitated a study of the excitonic luminescence polarization and wavelength, thermally activated nonradiative recombination, and carrier diffusion in a strain relaxation regime where the optical and transport properties near individual dislocations were examined. The dislocations and DLD network were found well separated on a μm scale and oriented primarily along one of the ⟨110⟩ directions. A plan-view TEM and CL imaging analysis was used to determine the linear dislocation densities for a greater range of samples with $In_{0.2}Ga_{0.8}As$ channel thicknesses below and well beyond the critical thickness. Theoretical calculations using a four-band $\mathbf{k}\cdot\mathbf{p}$ and TMM were used to calculate the band profile and eigenstates of the HEMT samples. The electric field in the channel and the electron and hole eigenstates were treated self-consistently. CL wavelength imaging showed a striking blue shift of ~5 and ~10 meV in the 150 and 185 Å samples, respectively, contrary to a previous red shift observed in undoped samples which exhibited a greater strain relaxation, as reported in Refs. 14–16. We hypothesize that a defect-induced reduction in the field occurs and leads to a reduced Stark shift of the electron and hole eigenstates.

A small polarization anisotropy is observed, consistent with the level of strain relaxation in each sample and theoretical calculations of the polarization ratio. The activation energies for the $In_{0.2}Ga_{0.8}As$ QW and $GaAs/Al_{0.25}Ga_{0.75}As$ MQW luminescence were obtained by temperature-

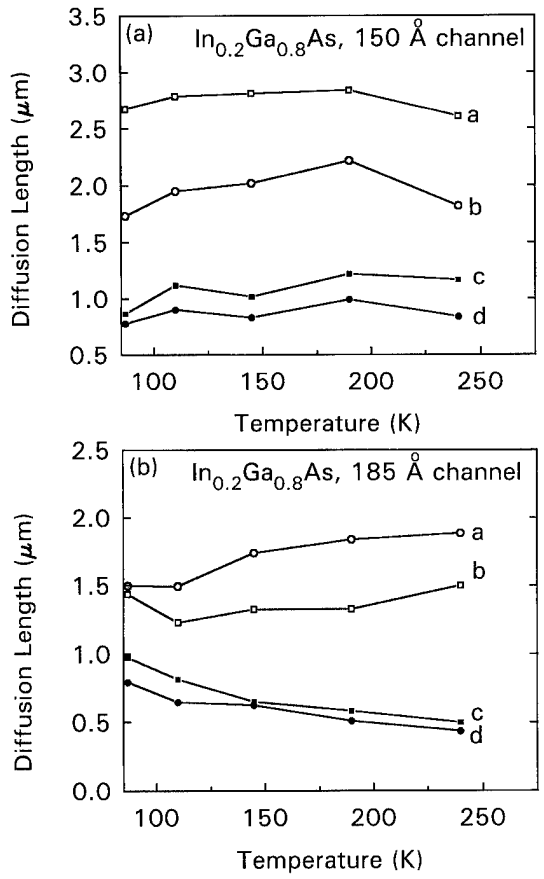


FIG. 17. Diffusion length vs temperature for various bright (a and b) and dark (c and d) regions in both 150 and 185 Å samples.

dependent CL measurements. Spatial variations in the activation energies are found to correlate with DLDs. These defects were found to lower the activation energy for thermal reemission of carriers from the $In_{0.2}Ga_{0.8}As$ QW and increase the activation energy for carrier transfer into the $In_{0.2}Ga_{0.8}As$ QW from the underlying GaAs barrier and $GaAs/Al_{0.25}Ga_{0.75}As$ MQW. The ambipolar diffusion length of carriers in the $In_{0.2}Ga_{0.8}As$ QW was quantified with a noncontact optical diffusion length measurement which utilized a one-dimensional line scan during monochromatic CL detection and a Ag mask which covered part of the sample. The diffusion lengths of holes parallel to the dislocation line direction were found to decrease by as much as a factor of ~5 in close proximity to DLDs, thereby revealing that an important transport parameter is deleteriously effected by the dislocation formation. Future time-resolved CL measurements involving lifetime determinations on a μm scale should help clarify the competing roles of mobility and the minority-carrier lifetime in determining the diffusion lengths. This study revealed that the optical and transport properties of $In_xGa_{1-x}As$ HEMTs are linked on a μm scale and are both intimately tied to the presence of misfit dislocations and dark line defects which form during strain relaxation.

ACKNOWLEDGMENTS

The work performed at USC was supported by the U.S. Army Research Office and the National Science Foundation (RIA-ECS). The research performed at UCLA was supported through funding from TRW, the UC MICRO Program, the National Science Foundation, and the National Defense Science and Engineering Graduate Fellowship Program. The authors thank D. C. Streit, T. R. Block and M. Wojtowicz of TRW, Electronic Technology Division, and Space and Electronics Group in Redondo Beach, CA for use of their HEMT samples.

[1]N. Moll, M. R. Hueschen, and A. Fischer-Colbrie, IEEE Trans. Electron Devices **ED-35**, 878 (1988).
[2]L. D. Nguyen, D. C. Radulescu, M. C. Foisy, P. J. Tasker, and L. F. Eastman, IEEE Trans. Electron Devices **ED-36**, 833 (1989).
[3]A. Fischer-Colbrie, J. N. Miller, S. S. Laderman, S. J. Rosner, and R. Hull, J. Vac. Sci. Technol. B **6**, 620 (1988).
[4]K. J. Beernink, P. K. York, and J. J. Coleman, Appl. Phys. Lett. **55**, 2585 (1989).
[5]G. M. Smith and J. J. Coleman, in *VLSI Electronics*, edited by N. G. Einspruch and W. R. Frensley (Academic, San Diego, 1994), Vol. 24, pp. 215–238.
[6]G. M. Yang, M. H. MacDougal, and P. D. Dapkus, Electron. Lett. **31**, 886 (1995).
[7]A. Larsson and J. Maserjian, Appl. Phys. Lett. **59**, 3099 (1991).
[8]D. S. McCallum, A. N. Cartwright, X. R. Huang, T. F. Boggess, A. L. Smirl, and T. C. Hasenberg, J. Appl. Phys. **73**, 3860 (1993).
[9]A. Madhukar, K. C. Rajkumar, L. Chen, S. Guha, K. Kaviani, and R. Kapre, Appl. Phys. Lett. **57**, 2007 (1990).
[10]D. H. Rich, H. T. Lin, and A. Larsson, J. Appl. Phys. **77**, 6557 (1995).
[11]D. H. Rich, T. George, W. T. Pike, J. Maserjian, F. J. Grunthaner, and A. Larsson, J. Appl. Phys. **72**, 5834 (1992).
[12]D. H. Rich, K. Rammohan, Y. Tang, H. T. Lin, J. Maserjian, F. J. Grunthaner, A. Larsson, and S. I. Borenstain, Appl. Phys. Lett. **64**, 730 (1994).
[13]E. A. Fitzgerald, G. P. Watson, R. E. Proano, D. G. Ast, P. D. Kirchner, G. D. Pettit, and J. M. Woodall, J. Appl. Phys. **65**, 2220 (1989).
[14]K. Rammohan, Y. Tang, D. H. Rich, R. S. Goldman, H. H. Wieder, and K. L. Kavanagh, Phys. Rev. B **51**, 5033 (1995).
[15]K. Rammohan, D. H. Rich, R. S. Goldman, and K. L. Kavanagh, Appl. Phys. Lett. **66**, 871 (1995).
[16]D. H. Rich, K. Rammohan, Y. Tang, H. T. Lin, R. S. Goldman, H. H. Wieder, and K. L. Kavanagh, J. Vac. Sci. Technol. B **13**, 1766 (1995).
[17]D. Bimberg, J. Christen, A. Steckenborn, G. Weimann, and W. Schlapp, J. Lumin. **30**, 562 (1985).
[18]K. Uno, K. Hirano, S. Noda, and A. Sakaki, in *Proceedings of the 19th International Symposium on GaAs and Related Compounds* (IOP, Bristol, 1993), p. 241.
[19]D. S. Jiang, H. Jung, and K. Ploog, J. Appl. Phys. **64**, 1371 (1988).
[20]U. Jahn, J. Menninger, R. Hey, and H. T. Grahn, Appl. Phys. Lett. **64**, 2382 (1994).
[21]M. Meshkinpour, M. S. Goorsky, G. Chu, D. C. Streit, T. R. Block, and M. Wojtowicz, Appl. Phys. Lett. **66**, 748 (1995).
[22]M. Meshkinpour, M. S. Goorsky, D. C. Streit, T. R. Block, M. Wojtowicz, K. Rammohan, and D. H. Rich, Mater. Res. Soc. Symp. Proc. **340**, 327 (1994).
[23]D. C. Streit, K. L. Tan, R. M. Dia, J. K. Liu, A. C. Han, and J. R. Velebir, IEEE Electron Device Lett. **EDL-9**, 621 (1988).
[24]M. Meshkinpour, M. S. Goorsky, D. C. Streit, T. R. Block, and M. Wojtowicz, Mater. Res. Soc. Symp. Proc. **378**, 783 (1995).
[25]D. H. Rich, A. Ksendzov, R. W. Terhune, F. J. Grunthaner, B. A. Wilson, H. Shen, M. Dutta, S. M. Vernon, and T. M. Dixon, Phys. Rev. B **43**, 6836 (1991).
[26]M. Grundmann, J. Christen, D. Bimberg, A. Hashimoto, T. Fukunaga, and N. Watanabe, Appl. Phys. Lett. **58**, 2090 (1991).

[27]T. E. Everhart and P. H. Hoff, J. Appl. Phys. **42**, 5837 (1971).
[28]H. A. Zarem, P. C. Sercel, J. A. Lebens, L. E. Eng, A. Yariv, and K. J. Vahala, Appl. Phys. Lett. **55**, 1647 (1989).
[29]K. L. Kavanagh, M. A. Capano, L. W. Hobbs, J. C. Barbour, P. M. J. Maree, W. Schaff, J. W. Mayer, D. Pettit, J. M. Woodall, J. A. Stroscio, and R. M. Feenstra, J. Appl. Phys. **64**, 4843 (1988).
[30]T. George, E. R. Weber, S. Nozaki, T. Yamada, M. Konagai, and K. Takahashi, Appl. Phys. Lett. **59**, 60 (1991).
[31]I. Yonenaga and K. Sumino, J. Appl. Phys. **65**, 85 (1989).
[32]R. S. Goldman, H. H. Wieder, and K. L. Kavanagh, Appl. Phys. Lett. **67**, 344 (1995).
[33]D. C. Hutchings, Appl. Phys. Lett. **55**, 1082 (1989).
[34]D. Campi and C. Alibert, Appl. Phys. Lett. **55**, 454 (1989).
[35]S. I. Borenstain, I. Grave, A. Larsson, D. H. Rich, B. Jonsson, I. Andersson, J. Westin, and T. Andersson, Phys. Rev. B **43**, 9320 (1991).
[36]G. Ji, D. Huang, U. K. Reddy, H. Unlu, T. S. Henderson, and H. Morkoç, J. Vac. Sci. Technol. B **5**, 1346 (1987).
[37]F. H. Pollak, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Wiley, New York, 1990), Vol. 32, pp. 17–53.
[38]S. L. Chuang, Phys. Rev. B **43**, 9649 (1991).
[39]S. Adachi, J. Appl. Phys. **53**, 8775 (1982).
[40]J. D. Lambkin, L. Considine, S. Walsh, G. M. O'Connor, C. J. McDonagh, and T. J. Glynn, Appl. Phys. Lett. **65**, 73 (1994).
[41]B. G. Yacobi and D. B. Holt, J. Appl. Phys. **59**, R1 (1986).
[42]J. I. Pankove, *Optical Processes in Semiconductors* (Dover, New York, 1971), p. 166.
[43]See, for example, N. W. Ashhcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), pp. 602–605.

Role of structural and chemical contributions to valence-band offsets at strained-layer heterojunctions: The GaAs/GaP (001) case

M. Di Ventra[a)]
Institut de Physique Appliquée, Ecole Polytechnique Fédérale de Lausanne, PHB-Ecublens, CH-1015 Lausanne, Switzerland

M. Peressi
INFM-Dipartimento di Fisica Teorica dell'Università di Trieste, Strada Costiera 11, I-34014 Trieste, Italy

A. Baldereschi
Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA), INR-Ecublens, CH-1015 Lausanne, Switzerland and INFM-Dipartimento di Fisica Teorica dell'Università di Trieste, Strada Costiera 11, I-34014 Trieste, Italy

(Received 21 January 1996; accepted 4 May 1996)

We calculate with the *ab initio* pseudopotential method the valence-band offset (VBO) of GaAs/GaP lattice-mismatched heterojunctions pseudomorphically grown on $GaAs_xP_{1-x}(001)$ substrates ($0 \leq x \leq 1$). The resulting value of 0.76 eV for the GaP substrate supports the highest among the experimental data available. We show that the VBO for any substrate can be easily obtained from that calculated for one substrate following standard band-structure theory since the interface-dependent contribution to the VBO, which is the electrostatic potential lineup, has no significant variation with the substrate. Such variation, however, can be taken into account accurately by a simple analytic scaling law, which is demonstrated here. Finally, we show that the VBO of GaAs/GaP (001) is due to *structural* and *chemical* effects, and we propose an operative method for identifying and separating the two contributions, whose dependence on strain is also discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

At variance with the case of lattice-matched (LM) interfaces for which a general theory has been definitely established,[1] the problem of valence-band offsets (VBOs) at lattice-mismatched (LMM) heterojunctions under arbitrary strain is far from being generally understood. With respect to LM heterojunctions where VBO tuning is possible only via local chemical modifications at the interface, LMM heterojunctions offer the additional degree of freedom of varying the strain—and hence the band-edge energies—of the two constituents, by changing the substrate. A theoretical comprehension of the mechanisms which are responsible for the VBO in these systems and its variation with strain is therefore necessary.

To this aim we study a prototypical case of LMM heterojunctions, and we choose the *isovalent, heteropolar, common-cation* GaP/GaAs (001) heterostructure, which is used in optoelectronic devices[2] and for which a few (not very accurate) experimental measurements[3,4] and only one *ab initio* theoretical investigation[5] have been reported.

We use the *ab initio* pseudopotential method, within the framework of density-functional theory in the local-density approximation (LDA). As in Ref. 1, we divide the VBO into two contributions: VBO $=\Delta E_v + \Delta V$. The band-structure term ΔE_v is the difference between the energies of the two valence-band edges measured with respect to the average electrostatic potential in each bulk material. It can thus be obtained from two separate calculations for each (appropriately strained) bulk material. The second term, ΔV, is the electrostatic potential lineup generated by the valence-electron charge distribution and by the ionic cores. It contains all interface-specific details and requires therefore the calculation of the self-consistent electronic density of the heterojunction. The latter is modeled with a periodically repeated supercell of 16 atoms, corresponding to a $(GaP)_4/(GaAs)_4$ (001) superlattice. Other technical details of the calculations are described in Ref. 6.

II. RESULTS

In order to study the variation of the VBO with strain we consider three different substrates: (i) GaP, (ii) $GaAs_{0.5}P_{0.5}$ corresponding to an in-plane lattice constant $a_\parallel$ equal to the average between the bulk lattice parameters a_i of the constituents (i denotes GaAs or GaP), and (iii) GaAs. The subscripts $\parallel$ and $\perp$ indicate quantities parallel or perpendicular to the interface plane. We consider only pseudomorphic growth conditions where $a_\parallel$ is constant throughout the structure and equals the substrate lattice parameter, and the overlayer material contracts or expands along the growth direction for accommodating the stress induced by the lattice mismatch. The theoretical calculations give for the latter a value of 3.7% in good agreement with the experimental value of 3.5%. For a given substrate, the equilibrium value of $a_{i\perp}$ for each material (see Table I) is determined according to the macroscopic theory of elasticity.[7] We found that for GaAs/GaP the latter describes very accurately the equilibrium position of the atoms also in the interface region.

[a)]Electronic mail: diventra@irrmasq11.epfl.ch

TABLE I. Perpendicular lattice constants $a_\perp$ (a.u.), their average value $\langle a_\perp \rangle$ (a.u.), parallel ($\epsilon_\parallel$), and perpendicular ($\epsilon_\perp$) components of the macroscopic strain tensor $\vec{\vec{\epsilon}}$ for strained GaAs/GaP(001) heterostructures grown on the three different substrates. The $a_\perp$'s are calculated using the macroscopic theory of elasticity. The change of the substrate has the effect of a *uniform* macroscopic strain throughout the structure.

$a_\parallel$ (a.u.)	10.20	10.40	10.60
$a_\perp$(GaAs)	10.94	10.77	10.60
$a_\perp$(GaP)	10.20	10.03	9.86
$\langle a_\perp \rangle$	10.57	10.40	10.23
$\epsilon_\parallel$	−0.019	0	+0.019
$\epsilon_\perp$	+0.016	0	−0.016

TABLE II. Valence-band offset at the GaP/GaAs(001) interface for different substrates. ΔV is the potential-lineup contribution. $\Delta E_{v,\mathrm{ave}}$ is the band-structure term relative to the average of the valence-band manifolds while ΔE_v is the same term but relative to the topmost valence states with spin-orbit effects included; VBO $=\Delta E_v+\Delta V$ and $\mathrm{VBO}_{\mathrm{ave}}=\Delta E_{v,\mathrm{ave}}+\Delta V$. All energies are in eV. Positive energy values indicate that the GaAs level is higher than that of GaP. An energy correction of about +0.15 eV must be added to the VBO values reported here in order to take into account self-energy corrections.

$a_\parallel$ (a.u.)	10.20	10.40	10.60
ΔV	0.16	0.16	0.16
ΔE_v	0.45	0.24	0.02
VBO	0.61	0.40	0.18
$\Delta E_{v,\mathrm{ave}}$	0.20	0.17	0.16
$\mathrm{VBO}_{\mathrm{ave}}$	0.36	0.33	0.32

The band-structure term is obtained from bulk calculations for GaP and GaAs in the proper strain configurations. Table II shows that $\Delta E_{v,\mathrm{ave}}$, which is the band-structure term relative to the averages of the valence band-edge manifolds, has a very small variation (less than 0.05 eV) when the substrate changes from GaAs to GaP. Conversely, variations of about 0.4 eV are found for ΔE_v, which is the band-structure term relative to the topmost valence states, whose shifts with respect to the average result from both strain and spin–orbit effects (see Refs. 6 and 7 for details).

The potential lineup ΔV, which must be added to ΔE_v to get the total VBO, is obtained from supercell calculations, using the technique of double macroscopic average[8] to get the profile of the charge density and of the corresponding electrostatic potential along the superlattice growth direction. Figure 1 shows the profile of the electrostatic potential originated by the valence-electron density and by the ionic cores, for the three strain configurations considered in this work. Electronic and ionic potentials are opposite in sign and give rise to a total electrostatic potential V which is almost two orders of magnitude smaller. Table II shows that, within our numerical accuracy, the electrostatic potential lineup ΔV has the same value of 0.16 eV for the three configurations. We conclude that the possibility of tuning the VBO with strain relies on the strain dependence of the band-structure term via the energy splittings produced by the strain at the valence-band top of the two materials. The above conclusion is consistent with the findings of Ref. 8, thus indicating a general trend for lattice-mismatched interfaces. The results for the total VBO, calculated both from the topmost valence states (VBO $= \Delta E_v+\Delta V$) and from the average of the valence-band manifolds ($\mathrm{VBO}_{\mathrm{ave}} = \Delta E_{v,\mathrm{ave}}+\Delta V$), are also summarized in Table II. Assuming that strain does not affect considerably the self-energy corrections to the LDA energies of GaAs and GaP, a correction of about +0.15 eV (Ref. 9) has to be added to our *ab initio* results for a meaningful comparison with the experimental data. The final theoretical estimate of the VBO is 0.76 eV for $a_\parallel = 10.20$ a.u., 0.55 eV for $a_\parallel = 10.40$ a.u., and 0.33 eV for $a_\parallel = 10.60$ a.u. Our result for the GaP substrate supports the highest value [0.80 eV (Ref. 4)] among the available experimental data.

In the present system the potential lineup is practically the same for the three configurations considered. A general study of the strain dependence of potential lineups has never been done as opposed to the case of the ΔE_v term.[7] Since ΔV is the contribution to the VBO which requires the strongest numerical effort, it would be of considerable importance to predict analytically its variation with the substrate. We will focus on this problem in what follows.

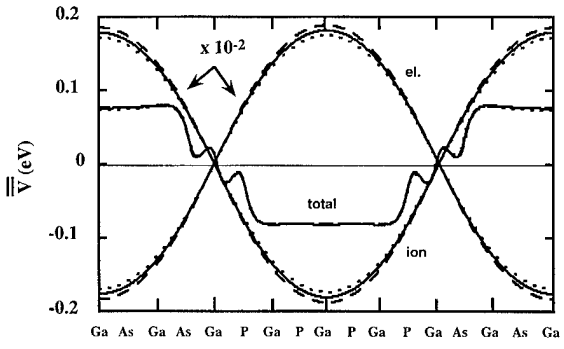


FIG. 1. Macroscopic average of the total electrostatic potential together with the two separate contributions due to electrons and ions, scaled by a factor 10^{-2}. Dashed, solid, and dotted lines give the results for GaP, $\mathrm{GaAs_{0.5}P_{0.5}}$, and GaAs substrates, respectively.

III. ANALYTIC SCALING LAW FOR ΔV

Let us consider the case (ii) as the reference configuration. When the substrate is $\mathrm{GaAs_{0.5}P_{0.5}}$, we have $a_\parallel = \langle a \rangle \equiv (a_1+a_2)/2$, and the average perpendicular lattice constant $\langle a_\perp \rangle \equiv (a_{1\perp}+a_{2\perp})/2$ is also equal to $\langle a \rangle$ (see Table I). Let ΔV be the lineup for this reference configuration. We will indicate with the superscript $'$ the quantities referring to a different configuration, corresponding to a different substrate, and, consequently, to a different *macroscopic* strain tensor $\vec{\vec{\epsilon}}$ of the supercell. The relevant $\vec{\vec{\epsilon}}$ components are those parallel and perpendicular to the interface plane, which are respectively $\epsilon_\parallel = (a'_\parallel/\langle a \rangle) - 1$ and $\epsilon_\perp = (\langle a'_\perp \rangle/\langle a \rangle) - 1$.

We show that, if the change of the substrate has the effect of a *uniform* strain throughout the structure, i.e., $a'_{i\perp} = (1+\epsilon_\perp)a_{i\perp}$, as it is indeed the present case (see Table I), then the lineup for the new configuration is

$$\Delta V' \simeq \frac{1}{(1+\epsilon_\parallel)^2}\Delta V. \tag{1}$$

In fact, if z is the direction perpendicular to the interface, we can write

$$\Delta V = \lim_{z\to+\infty} \bar{\bar{V}}(z) - \lim_{z\to-\infty} \bar{\bar{V}}(z) = \lim_{k_z\to 0^+} \tilde{V}(k_z) - \lim_{k_z\to 0^-} \tilde{V}(k_z),$$

where $\bar{\bar{V}}(z)$ is the macroscopic average of the electrostatic potential, $\tilde{V}(k_z)$ the corresponding Fourier transform, $\tilde{V}(k_z)=4\pi e\tilde{n}(k_z)/k_z^2$ and $\tilde{n}(k_z)$ the Fourier transform of the macroscopic average of the total charge density (i.e., electronic plus ionic) of the interface. Thus,

$$\begin{aligned}\Delta V &= 4\pi e\left(\lim_{k_z\to 0^+}\frac{\tilde{n}(k_z)}{k_z^2} - \lim_{k_z\to 0^-}\frac{\tilde{n}(k_z)}{k_z^2}\right),\\ &\simeq 4\pi e\left[\lim_{k_z\to 0^+}\frac{1}{k_z^2}\left(\tilde{n}(0)+k_z\frac{\partial\tilde{n}(k_z)}{\partial k_z}\right)\right.\\ &\quad\left. - \lim_{k_z\to 0^-}\frac{1}{k_z^2}\left(\tilde{n}(0)+k_z\frac{\partial\tilde{n}(k_z)}{\partial k_z}\right)\right]\\ &\simeq 4\pi e\left(\lim_{k_z\to 0^+}\frac{1}{k_z}\frac{\partial\tilde{n}(k_z)}{\partial k_z} - \lim_{k_z\to 0^-}\frac{1}{k_z}\frac{\partial\tilde{n}(k_z)}{\partial k_z}\right).\end{aligned}$$

The former step is valid up to linear order in k_z and the last step is justified since $\tilde{n}(0)=0$, i.e., the total charge of the system vanishes. A similar expression can be found for $\Delta V'$. If $\lim_{k_z'\to 0^\pm}\partial\tilde{n}'(k_z')/\partial k_z' = V_{\rm cell}/V'_{\rm cell}\lim_{k_z\to 0^\pm}\partial\tilde{n}(k_z)/\partial k_z$ then $\Delta V'/\Delta V = V_{\rm cell}k_z/V'_{\rm cell}k_z'$ where $V'_{\rm cell} = (1+\epsilon_\parallel)^2\times(1+\epsilon_\perp)V_{\rm cell}$, $V_{\rm cell}$ is the supercell volume of the reference configuration and $k_z'=(1+\epsilon_\perp)^{-1}k_z$. A uniform strain throughout the structure affects the position of the ionic cores, such that the above condition on $\partial\tilde{n}'/\partial k_z'$ is verified exactly for the ionic part. If the strain is small, we expect that the electronic charge distribution follows rigidly the ionic cores, and the condition is fulfilled also for the electronic charge density.

The predictions of Eq. (1) for GaAs/GaP (001) are in excellent agreement with the *ab initio* results, within the numerical precision of the latter. Equation (1) also holds separately for the electronic and ionic contributions, which are about two orders of magnitude larger than the total value. We stress that the scaling law, Eq. (1), can be extended to other systems, but we expect that it will not be valid for interfaces without common ions where local interface strains can be important, and contribute significantly to the potential lineup.

IV. STRUCTURAL AND CHEMICAL CONTRIBUTIONS TO THE POTENTIAL LINEUP

The ΔV term contains both *chemical* and *structural* contributions, which originate from the chemical difference between the two constituents and their different strain. More insight into the dependence of ΔV on strain will be possible if we succeed in separating these two contributions. Keeping in mind the case of LM heterojunctions, where the potential lineup is mainly due to the chemical difference between the elements constituting the junction (with a few exceptions for heterojunctions without common ions due to important local interface strains), it will be reasonable to define a *structural* term which vanishes in such a case.

For the sake of clarity let us focus on one strain configuration, say (iii), identified by a given $a_\parallel$ and the corresponding values $a_{1\perp}$ and $a_{2\perp}$. We consider the *virtual crystal* GaX which is a fictitious crystal whose anions X are described by a pseudopotential which is the arithmetic average between those of As and P. The perpendicular lattice constant of this virtual crystal equals $\langle a_\perp\rangle=(a_{1\perp}+a_{2\perp})/2$, while its parallel lattice constant is $a_\parallel$. This reference crystal has an electronic charge density n_0 whose macroscopic average is constant everywhere, and it can be seen as a homojunction with a uniform strain and zero potential lineup. Starting from this reference homojunction we can build up the real strained interface following two different two-step paths which differ in the initial step: (I) keeping the virtual anions, but introducing the non-uniform strain of the real heterojunction; (II) keeping the uniform strain of the virtual crystal, but transforming the homojunction into the heterojunction by replacing the virtual anions with the real ones.

In the non-uniformly strained homojunction obtained as intermediate system along the former path, the electronic charge density n_I has the same average asymptotic values—far from the interface—as the electronic charge density n of the real heterojunction. By construction n_I and n_0, and consequently their difference $\Delta n_{\rm struct}=n_I-n_0$, do not contain any term arising from chemical differences between the two constituents. Chemical effects instead are present in n, and in the difference $\Delta n_{\rm ch}=n-n_I$, which has a vanishing asymptotic value far from the interface on both sides. The electron density of the real heterojunction can thus be written as $n=n_0+\Delta n_{\rm struct}+\Delta n_{\rm ch}$. The upper panel of Fig. 2 shows the macroscopic average of the electronic charge densities $\Delta\bar{\bar{n}}_{\rm ch}(z)$ (dashed line) and $\Delta\bar{\bar{n}}_{\rm struct}(z)$ (dotted line). The normalization volume (average cell volume) is the average of the unit-cell volumes of GaAs and GaP under the proper strain.

The potential lineups generated by the electronic charge densities $\Delta n_{\rm ch}$ and $\Delta n_{\rm struct}$ can be identified as the chemical ($\Delta V_{\rm ch}$) and structural ($\Delta V^{\rm el.}_{\rm struct}$) electronic contributions to the total electronic potential lineup $\Delta V^{\rm el.}$. The total potential lineup ΔV contains also the ionic contribution, which is entirely of *structural* type. The inhomogeneous ionic core distributions corresponding to n and n_I are equal and therefore the electronic chemical term does not have an ionic counterpart, whereas the ionic contribution corresponding to $\Delta n_{\rm struct}$ gives a non vanishing dipole. The two terms $\bar{\bar{V}}_{\rm ch}(z)$ and $\bar{\bar{V}}_{\rm struct}(z)=\bar{\bar{V}}^{\rm el.}_{\rm struct}(z)+\bar{\bar{V}}^{\rm ion}_{\rm struct}(z)$ are displayed in the lower panel of Fig. 2 together with the total potential $\bar{\bar{V}}(z)$.

We can similarly define a chemical and a structural contribution to the total electronic charge density n, by following the second path. We find that the lineups are equal to those obtained following path I.[6] We can thus univocally

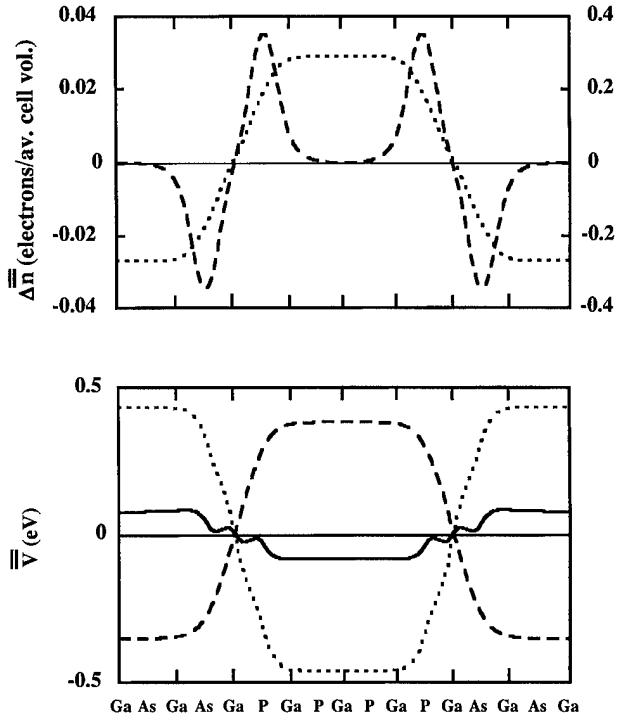


FIG. 2. Upper panel: macroscopic average of structural $[\Delta\bar{\bar{n}}_{\text{struct}}(z)$, dotted line, right scale] and chemical $[\Delta\bar{\bar{n}}_{\text{ch}}(z)$, dashed line, left scale] contributions to the electronic charge density of the $(GaP)_4/(GaAs)_4$ superlattice with GaAs substrate. The charge density is normalized with respect to the average of the unit-cell volumes of GaAs and GaP under the proper strain. Lower panel: macroscopic average of structural ($\bar{\bar{V}}_{\text{struct}}$, dotted line) and chemical ($\bar{\bar{V}}_{\text{ch}}$, dashed line) contributions to the total potential ($\bar{\bar{V}} = \bar{\bar{V}}_{\text{ch}} + \bar{\bar{V}}_{\text{struct}}$, solid line) of the $(GaP)_4/(GaAs)_4$ superlattice with GaAs substrate. The results of both panels have been obtained following path I.

define the chemical contribution ΔV_{ch} and the strain contribution ΔV_{struct} (including the ionic core contribution, which is the same in the two paths) whose sum gives the total potential lineup ΔV. We stress that ΔV_{ch} may also—in principle—depend on strain and hence be different for different configurations. Actually, ΔV_{ch} and ΔV_{struct} vary slightly with $a_{\parallel}$: e.g., ΔV_{struct} for configurations (i), (ii), and (iii) is 0.89, 0.88, and 0.87 eV, respectively. The variation is even smaller in the total lineup.

V. CONCLUSIONS

We have presented self-consistent pseudopotential calculations of the VBO at strained GaP/GaAs (100) heterojunctions. We have shown that the VBO can be tuned from 0.33 to 0.76 eV by varying the substrate from GaAs to GaP, and that this is a bulk effect rather than an interface feature. The value VBO_{ave}, calculated between the averages of the valence-band manifolds is, on the contrary, practically insensitive to the strain. These results confirm previous findings[8] indicating a general trend for lattice-mismatched heterojunctions.

We have demonstrated that once the potential lineup (which is the interface-dependent and computationally expensive contribution to the VBO) is known for one substrate, its value for any other substrate can be calculated analytically, with a very high accuracy, by a simple scaling law which is expected to be valid for the whole class of isovalent common-ion heterojunctions. For GaAs/GaP, this law predicts that the variation of the potential lineup with the in-plane lattice constant is within the numerical accuracy of the *ab initio* calculations (≃20 meV). Therefore the potential lineup can be considered as a constant and can be calculated for one strain configuration only. On the opposite, in order to reproduce the strain dependence of the VBO, it is important to take into account the large strain dependence of the bulk band-structure term, which however can be calculated easily with standard band-structure theory.

Finally, we have presented a method for separating *chemical* and *structural* contributions to the potential lineup ΔV and we have shown that, similarly to ΔV, both contributions vary negligibly for GaAs/GaP (001) with the in-plane lattice constant.

[1]See for instance: S. Baroni, M. Peressi, R. Resta, and A. Baldereschi, in *Proceedings of the Twenty-First International Conference on the Physics of Semiconductors,* edited by P. Jiang and H. Z. Zheng (World Scientific, Singapore, 1993), p. 689, and references therein.
[2]S. Ke, R. Wang, and M. Huang, Phys. Rev. B **49**, 10495 (1994), and references therein.
[3]P. L. Gouley and R. M. Biefeld, Appl. Phys. Lett. **45**, 749 (1984).
[4]M. E. Davis, G. Zeidenbergs, and R. L. Anderson, Phys. Status Solidi **34**, 385 (1969).
[5]R. G. Dandrea and A. Zunger, Phys. Rev. B **43**, 8962 (1991).
[6]M. Di Ventra, M. Peressi, and A. Baldereschi, Phys. Rev. B (to be published).
[7]C. G. Van de Walle, Phys. Rev. B **39**, 1871 (1989), and references therein.
[8]L. Colombo, R. Resta, and S. Baroni, Phys. Rev. B **44**, 5572 (1991); N. Tit, M. Peressi, and S. Baroni, *ibid.* **48**, 17 607 (1993); M. Peressi, L. Colombo, R. Resta, S. Baroni, and A. Baldereschi, *ibid.* **48**, 12 047 (1993).
[9]X. Zhu and S. G. Louie, Phys. Rev. B **43**, 14 142 (1991).

Atomic-scale structure of $InAs/InAs_{1-x}Sb_x$ superlattices grown by modulated molecular beam epitaxy

A. Y. Lew and E. T. Yu[a)]
Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, California 92093

Y. H. Zhang
Hughes Research Laboratories, Malibu, California 90265

(Received 22 January 1996; accepted 16 March 1996)

We present cross-sectional scanning tunneling microscopy (STM) studies of the atomic-scale structural and electronic properties of $InAs/InAs_{1-x}Sb_x$ superlattices grown by modulated molecular-beam epitaxy. In this technique the Group V composition is controlled by rapid modulation of the Group V beams rather than adjustment of the Group V flux ratios. A superlattice sample was grown at 475 °C consisting of 52 Å $InAs_{1-x}Sb_x$ alternating with 172 Å InAs for 30 periods on a GaSb (001) substrate. The $InAs_{1-x}Sb_x$ alloy layers consisted nominally of 7.8 Å InAs alternating with 5.2 Å InSb for four periods. X-ray diffraction was used to determine an average composition of $InAs_{0.76}Sb_{0.24}$ for the alloy layers. Constant-current STM images of the superlattice exhibit clear, electronically induced contrast between the InAs layers and $InAs_{0.76}Sb_{0.24}$ layers, and also reveal ordering within the $InAs_{0.76}Sb_{0.24}$ layers. Interfaces between the InAs layers and the $InAs_{0.76}Sb_{0.24}$ layers appear sharp, though there is evidence of more atomic intermixing at the InAs-on-$InAs_{0.76}Sb_{0.24}$ interface. Significant variation in composition within individual $InAs_{0.76}Sb_{0.24}$ alloy layers and apparent Sb incorporation from the $InAs_{0.76}Sb_{0.24}$ layers into the surrounding InAs layers are also observed. © *1996 American Vacuum Society.*

I. INTRODUCTION

III–V semiconductor material systems containing both arsenide and antimonide compounds offer a wide range of possibilities for designing heterostructure devices, including resonant-tunneling diodes,[1,2] infrared detectors,[3–5] and infrared lasers.[6–9] However, differences in heterointerface structure and composition arising from different growth techniques and parameters can have significant effects on heterostructure device performance. For example, $Ga_{1-x}In_xSb/InAs$ superlattices grown with InSb-like interfaces are found to have superior device characteristics than those grown with $Ga_{1-x}In_xAs$-like interfaces.[10,11] Spontaneous ordering in III–V ternary semiconductors has been shown to have substantial effects on the structural and optical characteristics of the materials.[12,13] Atomic ordering has also been observed in Group V alloys such as InAsSb (Refs. 14 and 15) and InAsP.[16] However, precise control over Group V composition in ternary compounds containing both As and Sb has proven to be extremely difficult to achieve in conventional solid-source molecular-beam epitaxy (MBE).[17–19]

Recently, dramatic improvements in compositional control and crystalline quality of $AlAs_xSb_{1-x}$, $GaAs_xSb_{1-x}$, and $InAs_xSb_{1-x}$ ordered alloys have been achieved using solid-source modulated molecular-beam epitaxy (MMBE).[20–22] In this technique, the Group V composition is controlled by rapid modulation of the As_2 and Sb_2 beams rather than adjustment of the Group V flux ratios. Modulation of the Group V beams is expected to lead to the growth of ordered rather than random alloys. The electronic and optical properties of heterostructures incorporating such layers will therefore depend significantly on the detailed structure of the alloys and of the heterojunction interfaces.

In this article we report cross-sectional scanning tunneling microscopy (STM) investigations of the structural and electronic properties of $InAs/InAs_{1-x}Sb_x$ superlattices grown by MMBE. $InAs/InAs_{1-x}Sb_x$ heterojunctions grown on InAs substrates exhibit a Type-II band alignment, and have shown potential for midwave infrared applications.[22,23] High-resolution constant-current images obtained by STM show that MMBE produces superlattices with sharp interfaces between the $InAs_{1-x}Sb_x$ layers and the InAs layers, and confirm that the MMBE technique leads to the growth of ordered rather than random alloys.

II. EXPERIMENT

The superlattice sample used in these experiments was grown in a solid-source VG V80 molecular beam epitaxy system using the MMBE growth technique. In this technique, the Group V composition during alloy growth is controlled by rapid modulation of the As_2 and Sb_2 beams rather than adjustment of the Group V flux ratios.[21] The superlattice sample consisted of 52 Å $InAs_{1-x}Sb_x$ alternating with 172 Å InAs for 30 periods, and was grown on top of a GaSb buffer on a *p*-type GaSb (001) substrate. Figure 1 shows a schematic diagram of the epitaxial layer structure and the Type II band-edge alignment expected in this system. The $InAs_{1-x}Sb_x$ layers consisted nominally of 7.8 Å InAs alternating with 5.2 Å InSb for four periods. However, because

[a)]Author to whom all correspondence should be addressed; Electronic mail: ety@ece.ucsd.edu

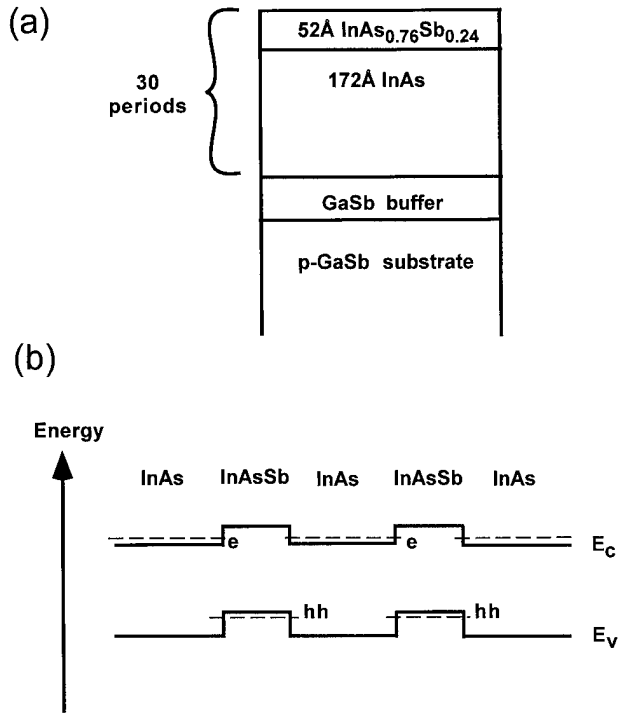


FIG. 1. (a) Schematic diagram of the epitaxial layer structure of the MMBE superlattice sample. (b) Schematic diagram of the Type II band-edge alignment expected for this structure. The conduction band, valence band, and the $InAs_{0.76}Sb_{0.24}$ heavy-hole and InAs electron energy levels are indicated.

the MMBE shutter times used were very short (4–6 s), excess As remaining in the growth chamber from the growth of the InAs layers was likely to have been incorporated into the growth of the InSb layers, resulting in InSb(As) layers; x-ray diffraction measurements yielded an average composition of $InAs_{0.76}Sb_{0.24}$ in the ternary alloy layers. A growth temperature of 475 °C and a growth rate of 1.3 Å/s were employed. Figure 2 shows a schematic diagram of the shutter sequence used to grow the $InAs_{0.76}Sb_{0.24}$ alloy layers and the surrounding InAs layers. Further details of the sample growth have been described elsewhere.[21]

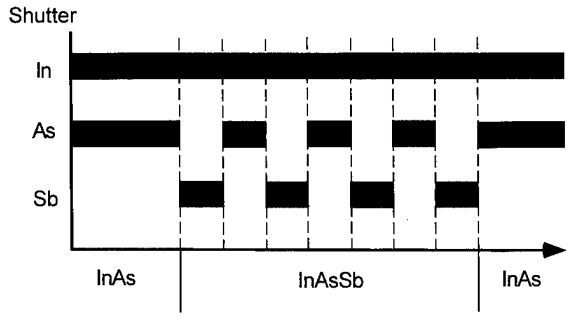


FIG. 2. Schematic diagram of the MMBE shutter sequence used to grow the $InAs_{0.76}Sb_{0.24}$ alloy layers and the surrounding InAs layers. The filled bars represent intervals when the corresponding shutters are open.

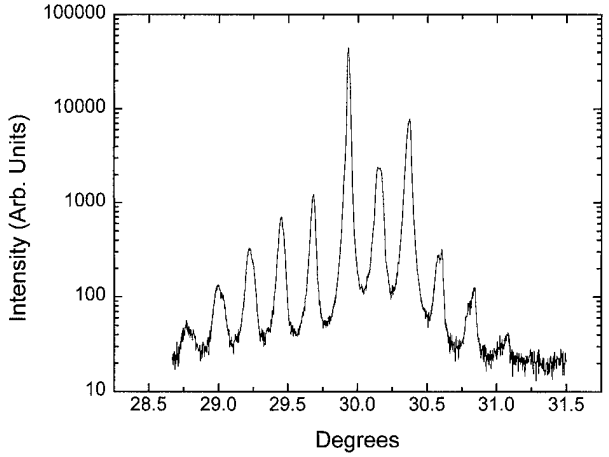


FIG. 3. High-resolution x-ray diffraction spectrum of the superlattice sample. The presence of many sharp satellite peaks confirms the overall high quality of the sample structure.

Sample growth and quality were monitored using high-resolution x-ray diffraction and *in situ* reflection high-energy electron diffraction (RHEED). The x-ray spectrum of the sample, shown in Fig. 3, exhibits many sharp satellite peaks, confirming the overall high quality of the sample structure. This x-ray spectrum was used to calculate the average period and composition of the epilayer region. RHEED data obtained during sample growth were similar to that observed in growth of $AlAs_xSb_{1-x}$ ordered alloys[21] using the MMBE technique. During growth of the $InAs_{0.76}Sb_{0.24}$ layers, a clear, streaky 1×3 reconstruction was seen when the Sb shutter was opened, but only a mixed 2×4 reconstruction was seen when the As shutter was opened. The RHEED data were inconclusive regarding interface abruptness in the ordered structures because the shutter times used (4–6 s) were comparable to the RHEED transient time.

Samples used for the STM studies were cleaved in ultrahigh vacuum at a base pressure of $<2\times10^{-10}$ Torr along the (110) cross-sectional plane, yielding an atomically flat, electronically unpinned surface on which tunneling experiments were performed. Commercially available Pt-Ir tips cleaned *in situ* by electron bombardment were used for these studies. Since all images presented here were obtained at negative sample bias, the features seen correspond to the Group V sublattice on the (110) cross-sectional surface.[24]

III. RESULTS

Figure 4 shows a high-resolution constant-current image of the MMBE sample, obtained at a sample bias voltage of −2.4 V and a tunneling current of 0.1 nA. Electronically induced contrast between the bright $InAs_{0.76}Sb_{0.24}$ layers and the darker InAs layers can clearly be seen in the image. The interfaces between the InAs layers and the outer InSb(As) layers of the $InAs_{0.76}Sb_{0.24}$ alloy layers appear sharp. Lateral variations in the $InAs_{0.76}Sb_{0.24}$ structure over length scales of approximately 100 Å can be seen. The image also shows contrast arising from the ordered structure along the growth

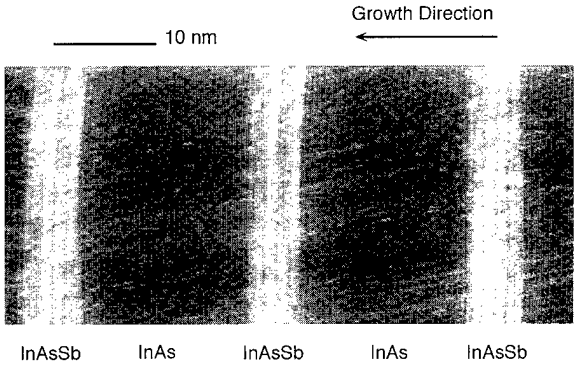


FIG. 4. High-resolution constant-current image of the MMBE sample. Electronically induced contrast between the $InAs_{0.76}Sb_{0.24}$ layers and the InAs layers can be seen. Also visible is the ordered structure within the $InAs_{0.76}Sb_{0.24}$ layers.

direction within the $InAs_{0.76}Sb_{0.24}$ layers. Features corresponding to the four-period InAs/InSb(As) superlattice structure of the $InAs_{0.76}Sb_{0.24}$ layers are visible. The electronically-induced contrast within the $InAs_{0.76}Sb_{0.24}$ layers is asymmetric along the growth direction, suggesting some nonuniformity in composition.

Figure 5(a) shows a high-resolution constant-current image of a single $InAs_{0.76}Sb_{0.24}$ layer surrounded by InAs, taken at a sample bias voltage of -2.4 V and a tunneling current of 0.1 nA. As in Fig. 4, contrast between the InAs layers and the $InAs_{0.76}Sb_{0.24}$ alloy layer can be seen clearly in the image. The growth direction is indicated in the figure. Figure 5(b) shows an averaged topographic line scan of the $InAs_{0.76}Sb_{0.24}$ layer and the surrounding InAs layers, taken over 200 individual scan lines covering an area approximately 200 Å × 170 Å shown in the image. The interfaces between the outer InSb(As) layers of the $InAs_{0.76}Sb_{0.24}$ alloy layer and the surrounding InAs appear sharp, though a more gradual change in contrast can be seen at the InAs-on-$InAs_{0.76}Sb_{0.24}$ interface, suggesting a larger degree of atomic intermixing at that interface than at the $InAs_{0.76}Sb_{0.24}$-on-InAs interface. Interface asymmetry has also been observed in other arsenide/antimonide systems, including InAs/GaSb,[25–29] $InAs/Ga_{1-x}In_xSb$,[30] and InAs/AlSb.[31] Also visible in the image is contrast caused by the ordered structure within the $InAs_{0.76}Sb_{0.24}$ layer produced by the MMBE growth technique. These features were observed consistently with different cleaved samples and tips. The averaged line scan makes more apparent the asymmetric appearance of the $InAs_{0.76}Sb_{0.24}$ layers in the constant-current images. Differences in topographic height, caused by differences in electronic structure, between the short-period superlattices and the darker InAs layers become larger as more $InAs_{0.76}Sb_{0.24}$ is grown. The maximum contrast in height (corresponding to maximum Sb content) is consistently observed at the next-to-last (i.e., the third) grown InSb-like layer in the $InAs_{0.76}Sb_{0.24}$ alloy region. This nonuniformity in the electronic structure in the $InAs_{0.76}Sb_{0.24}$ layers suggests the presence of atomic cross-incorporation within the $InAs_{0.76}Sb_{0.24}$ layers, possibly due to growth incorporation of residual material in the growth chamber during the MMBE shutter switching sequences. Finally, the apparent width of the $InAs_{0.76}Sb_{0.24}$ layer in the image is approximately 52 Å. If there was no atomic intermixing present in the $InAs_{0.76}Sb_{0.24}$ layer (i.e., in the four InAs/InSb short-period superlattices comprising the $InAs_{0.76}Sb_{0.24}$ layer), then only the four InSb layers and the three InAs layers between them should lead to visible contrast; the fourth InAs layer in the $InAs_{0.76}Sb_{0.24}$ structure would be indistinguishable from the surrounding InAs layers. This would lead to an apparent width of 44 Å for the $InAs_{0.76}Sb_{0.24}$ layer in the image. The fact that the full 52 Å width is seen suggests that some Sb is incorporated into the surrounding InAs layers, though this incorporation appears limited to areas within 1–2 monolayers of the InSb-like layers. From the averaged line scan, the Sb incorporation appears more prominent at the InAs-on-InSb(As) interface than at the InSb(As)-on-InAs interface. Taking into consideration the growth direction indicated at the top of Figure 5, this would suggest Sb riding up into grown InAs layers, a behavior which is consistent with that observed in growth of InAs/GaSb superlattices.[25]

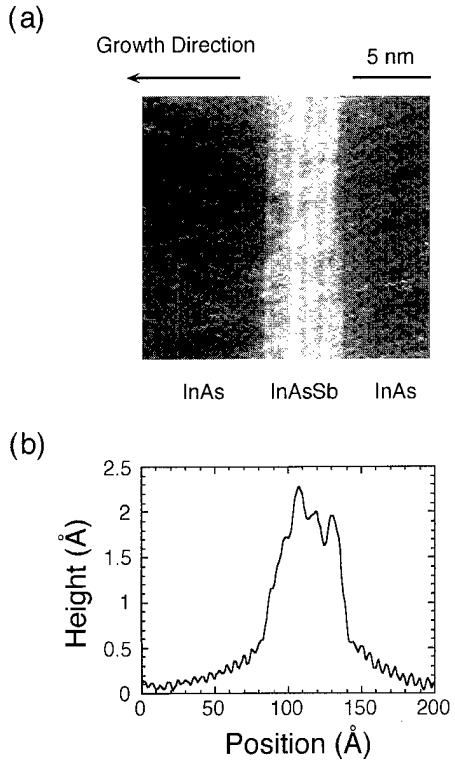


FIG. 5. (a) Constant-current image of a single $InAs_{0.76}Sb_{0.24}$ alloy layer and the surrounding InAs layers. (b) Topographic line scan of the $InAs_{0.76}Sb_{0.24}$ alloy layer and the surrounding InAs layers, averaged over an area approximately 200 Å × 170 Å shown in the image.

IV. CONCLUSIONS

We have used cross-sectional scanning tunneling microscopy (STM) to investigate the structural and electronic properties of $InAs_{0.76}Sb_{0.24}$ superlattices grown by modulated

molecular-beam epitaxy (MMBE). High-resolution constant current images of the superlattice structure reveal electronically-induced contrast between the InAs layers and the $InAs_{0.76}Sb_{0.24}$ layers. The interfaces between the InAs layers and the $InAs_{0.76}Sb_{0.24}$ layers appear sharp, though there is evidence of more atomic intermixing at the InAs-on-$InAs_{0.76}Sb_{0.24}$ interface. Contrast corresponding to the four-period InAs/InSb(As) superlattice structure can be seen in the $InAs_{0.76}Sb_{0.24}$ layers. The contrast within the $InAs_{0.76}Sb_{0.24}$ layers appears asymmetric along the growth direction; the maximum contrast in height is consistently observed at the third grown InSb-like layer in the $InAs_{0.76}Sb_{0.24}$ alloy regions. This non-uniformity in contrast suggests the presence of atomic cross-incorporation within the $InAs_{0.76}Sb_{0.24}$ layers, possibly caused by growth incorporation of residual material in the growth chamber during the MMBE shutter switching sequences. Further analysis of the data suggests Sb incorporation into areas within 1–2 monolayers of the InSb-like layers, possibly arising from Sb riding up into subsequently grown InAs layers. The STM data demonstrate that MMBE can be used to attain superior compositional control and crystalline quality in arsenide/antimonide ordered alloys, and also confirm that MMBE leads to the growth of ordered rather than random alloys.

ACKNOWLEDGMENTS

One author (Y. H. Z.) would like to acknowledge D. H. Chow and R. H. Miles for many stimulating discussions. Two authors (A. Y. L. and E. T. Y.) would also like to acknowledge the technical assistance of L. J. Ko. Part of the work done at UCSD was supported by Hughes Research Laboratories through the UC MICRO program, and by the National Science Foundation under Award No. ECS93-07986.

[1]G. A. Sai-Halasz, R. Tsu, and L. Esaki, Appl. Phys. Lett. **30**, 651 (1977).
[2]J. R. Söderström, D. H. Chow, and T. C. McGill, Appl. Phys. Lett. **55**, 1094 (1989).
[3]D. L. Smith and C. Mailhiot, J. Appl. Phys. **62**, 2545 (1987).
[4]D. H. Chow, R. H. Miles, J. R. Söderström, and T. C. McGill, Appl. Phys. Lett. **56**, 1418 (1990).
[5]R. H. Miles, D. H. Chow, J. N. Schulman, and T. C. McGill, Appl. Phys. Lett. **57**, 801 (1990).
[6]S. J. Eglash and H. K. Choi, Appl. Phys. Lett. **64**, 833 (1994).
[7]H. Lee, P. K. York, R. J. Menna, R. U. Martinelli, D. Z. Garbuzov, S. Y. Narayan, and J. C. Connolly, Appl. Phys. Lett. **66**, 1942 (1995).
[8]T. C. Hasenberg, D. H. Chow, A. R. Kost, R. H. Miles, and L. West, Electron. Lett. **31**, 275 (1995).
[9]R. H. Miles, D. H. Chow, Y. H. Zhang, P. D. Brewer, and R. G. Wilson, Appl. Phys. Lett. **66**, 1921 (1995).
[10]D. H. Chow, R. H. Miles, and A. T. Hunter, J. Vac. Sci. Technol. B **10**, 888 (1992).
[11]R. H. Miles, J. N. Schulman, D. H. Chow, and T. C. McGill, Semicond. Sci. Technol. **8**, S102 (1993).
[12]A. Gomyo, T. Suzuki, K. Kobayashi, S. Kawata, I. Hino, and T. Yuasa, Appl. Phys. Lett. **50**, 673 (1987).
[13]T. Suzuki, A. Gomyo, S. Iijima, K. Kobayashi, S. Kawata, I. Hino, and T. Yuasa, Jpn. J. Appl. Phys. **27**, 2098 (1988).
[14]H. R. Jen, K. Y. Ma, and G. B. Stringfellow, Appl. Phys. Lett. **54**, 1154 (1989).
[15]T. Y. Seong, G. R. Booker, A. G. Norman, and I. T. Ferguson, Appl. Phys. Lett. **64**, 3593 (1994).
[16]D. H. Jaw, G. S. Chen, and G. B. Stringfellow, Appl Phys. Lett. **59**, 114 (1991).
[17]C.-A. Chang, R. Ludeke, L. L. Chang, and L. Esaki, Appl. Phys. Lett. **31**, 759 (1977).
[18]T. H. Chiu, W. T. Tsang, J. A. Ditzenberger, S. N. G. Chu, and J. P. Van Der Ziel, J. Appl. Phys. **60**, 205 (1986).
[19]J. A. Lott, L. R. Dawson, E. D. Jones, and J. F. Klem, Appl. Phys. Lett. **56**, 1242 (1990).
[20]Y. H. Zhang and D. H. Chow, Appl. Phys. Lett. **65**, 3239 (1994).
[21]Y. H. Zhang, J. Cryst. Growth **150**, 838 (1995).
[22]Y. H. Zhang, Appl. Phys. Lett. **66**, 118 (1995).
[23]Y. H. Zhang, R. H. Miles, and D. H. Chow, IEEE Quantum. Electron. **1**, 749 (1995).
[24]R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. **58**, 1192 (1987).
[25]R. M. Feenstra, D. A. Collins, D. Z.-Y. Ting, M. W. Wang, and T. C. McGill, J. Vac. Sci. Technol. B **12**, 2592 (1994).
[26]R. M. Feenstra, D. A. Collins, D. Z.-Y. Ting, M. W. Wang, and T. C. McGill, Phys. Rev. Lett. **72**, 2749 (1994).
[27]B. R. Bennett, B. V. Shanabrook, R. J. Wagner, J. L. Davis, and J. R. Waterman, Appl. Phys. Lett. **63**, 949 (1993).
[28]B. V. Shanabrook, B. R. Bennett, and R. J. Wagner, Phys. Rev. B **48**, 17172 (1993).
[29]M. E. Twigg, B. R. Bennett, B. V. Shanabrook, J. R. Waterman, J. L. Davis, and R. J. Wagner, Appl. Phys. Lett. **64**, 3476 (1994).
[30]A. Y. Lew, E. T. Yu, D. H. Chow, and R. H. Miles, Appl. Phys. Lett. **65**, 201 (1994).
[31]B. R. Bennett, B. V. Shanabrook, and E. R. Glaser, Appl. Phys. Lett. **65**, 598 (1994).

Evidence for the occupation of DX centers in $In_{0.29}Al_{0.71}As$

A. P. Young[a] and H. H. Wieder[b]
University of California, San Diego, La Jolla, California 92093-0407

(Received 22 January 1996; accepted 17 April 1996)

We present the first results on $In_{0.29}Al_{0.71}As$ which conclusively show a large number of traps that are not in equilibrium with the conduction band at low temperatures, *the signature of a DX centerlike defect*. From low-temperature gated resistivity, Hall and Shubnikov–de Haas measurements of relaxed, $In_{0.29}Al_{0.71}As/In_{0.3}Ga_{0.7}As$ modulation doped heterostructures, we find a significant negative shift in the threshold voltage, $V_g(n_s \rightarrow 0)$, as the cooling bias is shifted to more negative values. Simultaneously, we observe a decrease in the mobility measured at the two-dimensional electron gas as the cooling bias is decreased, an indication of the "negative U" behavior of DX centers. Finally, we observe Fermi level pinning which we attribute to the occupation of a DX center level. From simulation of the heterostructure at large positive gate biases, and the change in the threshold voltage, we are able calculate the donor ionization energy of this trap to be 320 meV below the $In_xAl_{1-x}As$ conduction-band edge, over twice the depth of the DX center in $Al_xGa_{1-x}As$. © *1996 American Vacuum Society.*

I. INTRODUCTION

DX centers, a donor related point defect, have been found in many compound semiconductors, including the well-known AlAs/GaAs ternary alloy system.[1,2] These defects always move into the gap at some given alloy composition as if they were tracking a non-Γ conduction-band minimum. As suggested by Tachikawa,[3] this alloy dependence can be used to predict the appearance of DX centerlike behavior in materials with direct to indirect crossovers. $In_xAl_{1-x}As$, a high band-gap ternary alloy, has such a direct to indirect transition at $x \approx 0.30$. Following this suggestion, DX centerlike behavior would be expected at aluminum compositions between $x=0$ and $x=0.45$.

This prediction is difficult to verify because of a lack of a suitable substrate at all alloy compositions. At one composition $x=0.48$, $In_xAl_{1-x}As$ is latticed matched to InP. The availability of high quality material has allowed extensive work on this issue at this important composition. Although Nakashima *et al.*[4] believed they had found DX centerlike behavior, other researchers,[5–7] using a combination of techniques, concluded that the traps found are not DX centers. Furthermore, Calleja *et al.*,[6] subjected the material to 12 kbar of hydrostatic pressure roughly equivalent to the band structure of $In_{0.40}Al_{0.60}As$. They found no evidence of DX centerlike behavior. For the nonlatticed match case we could only find one reference. In that work, Hong *et al.* examined thick, $In_{0.43}Al_{0.57}As$ layers grown with a dislocation reducing buffer on InP. Measuring the material by deep-level transient spectroscopy (DLTS), a weak signal corresponding to an emission energy of 0.37 eV and strongly dependent on doping was found.

Considering the binary alloys, for Si-doped AlAs, the DX center level seems to merge with the shallow hydrogenic level 57 meV below the X band.[8] At the other extreme, no DX center has been reported for InAs, as would be the case if the DX center followed one of the non-Γ conduction-band minima. To the best of our knowledge, that is the extent of the known data at the moment with regard to DX centers in *n*-type $In_xAl_{1-x}As$. By determining the alloy compositions where these defects might occur, we can find the "useful range" of this material for electronic and optoelectronic applications.

II. EXPERIMENT

Figure 1 shows a cross-sectional schematic of the structure investigated in this article. The $In_{0.29}Al_{0.71}As/In_{0.3}Ga_{0.7}As$ heterostructures were grown by molecular beam epitaxy on a semi-insulating (100) GaAs substrate cut 2° toward the nearest (110) plane. The active layers are a modulation-doped field-effect-transistor (MODFET) type structure δ doped with silicon to 6×10^{12} cm^{-2}. Between the active layers and the substrate is a 0.8 μm step-graded $In_xGa_{1-x}As$ buffer used to relieve the lattice mismatch. Further details of the crystal growth are described elsewhere.[9]

From the material, eight-armed Hall bridges were photographically defined using standard wet chemical techniques. The ohmic contacts were made by thermal evaporation of a AuGe alloy with a Ni overlayer, which was subsequently annealed at 420 °C. The gate was composed of 300 nm of thermally evaporated aluminum. To reduce errors in the electrical measurements, the length to width ratio was 2400 μm/200 μm or 12/1.

The charge transport properties of the these structures have been initially described in Chen *et al.*[10] In that work, the specimens were cooled to either 77 or 1.6 K, whereupon their electrical properties were measured. At 1.6 K, both the low magnetic field ($B<0.15$ T), as well as high-field Shubnikov–de Haas oscillatory magnetoresistance measurements were obtained as a function of gate bias. This enabled the properties of the individual subbands to be characterized separately. In this work, a gate bias is applied at room tem-

[a]Department of Physics.
[b]Department of Electrical and Computer Engineering; Electronic mail: wieder@ece.ucsd.edu

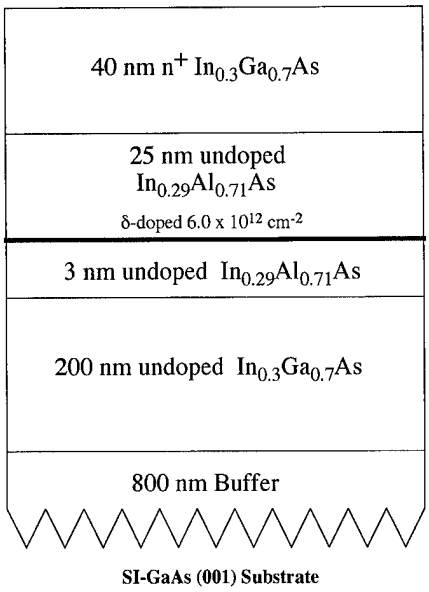


FIG. 1. Schematic diagram of the δ-doped $In_{0.29}Al_{0.71}As/In_{0.3}Ga_{0.7}As$ heterostructure under study.

perature to alter the distribution of free charge when the system is nominally in equilibrium. Then the sample is cooled to $T=1.6$ K, and the charge transport properties are remeasured both at low and high magnetic fields.

III. RESULTS AND DISCUSSION

Figure 2 shows the gate voltage dependence of the free electron density, n_s, of the same specimen for different, fixed values of the cooling bias potential, V_c. Each data point was obtained from a fast Fourier transform of the Shubnikov–de Haas oscillations in a magnetic field up to 7.5 T. While a digital filter could have been used subsequently on the transform, it was decided not to use it in this case due to a slight shifting of the peak of the second subband near zero. The

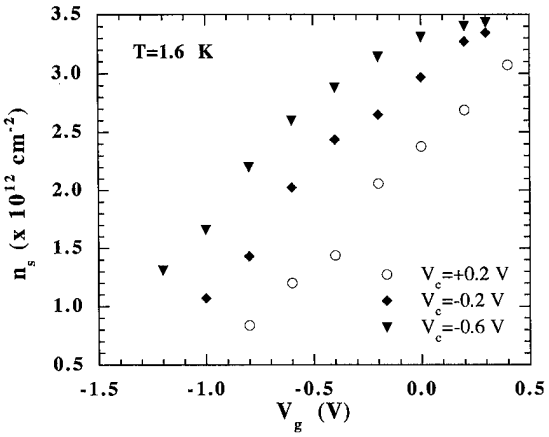


FIG. 2. Free carrier density vs gate bias for several values of the cooling gate bias V_c. Each value of the charge density is the sum of the two subbands as determined from a fast Fourier transform of the Shubnikov–de Haas oscillations.

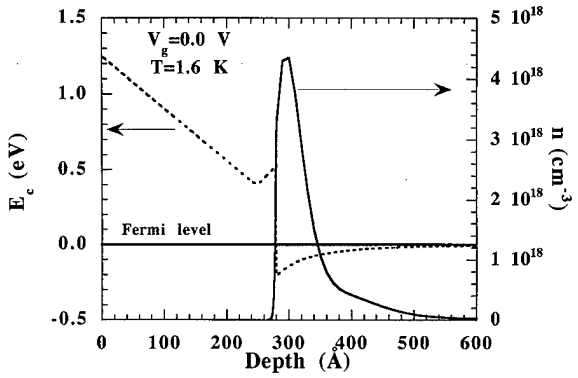


FIG. 3. Conduction-band edge, electron density, and the Fermi level calculated from a self-consistent solution of Schrödinger–Poisson equations. The free charge is measured by integration of the charge density over the whole structure.

peak was considered strong enough that the filter was not necessary. Each of the curves derived from this data has an initially linear $n_s(V_g)$ dependence and a tendency to saturate at higher gate voltages. The gate voltage dependent sum of the electron subbands is in good agreement with the respective charge densities derived from Hall measurements. This indicates that the mobilities of the electrons in the two subbands is similar.

Figure 3 shows a representative plot of the calculated position of the conduction-band edges of the heterostructure obtained by means of Schrödinger–Poisson simulations.[11] The simulated total $n_s(V_g)$ was found to be in fair agreement with the experimentally measured values in the linear regime. However, our simulation does not include possible DX centerlike defects, only the shallow ionized donors. Because our measurements were made at 1.6 K, with a comparatively small displacement of the Fermi level produced by the applied gate voltages, the emission of electrons from DX centers is considered negligible. However, we have introduced explicitly in the simulations other deep levels found in $In_xAl_{1-x}As$.[5,7,12] For modeling purposes, we assume in $In_{0.29}Al_{0.71}As$ a dominant 0.6 eV deep level trap whose density is considered to be 1×10^{16} cm^{-3}.

If there are DX centers in the $In_{0.29}Al_{0.71}As$, there will be a characteristic freeze-in temperature, T_f, where the system will no longer be in equilibrium. As the heterostructure is cooled under the different gate biases, different amounts of free charge would be trapped in the $In_{0.29}Al_{0.71}As$, the result being a change in the ionized donor density, N^+. By extrapolating the charge density in the linear regime of Fig. 2 to zero charge density, the threshold voltage, V_t, expressed as

$$V_t=\phi_b-\Delta E_c+E_{f0}-\frac{qN^+}{C_i} \tag{1}$$

can be obtained. In Eq. (1), $\phi_b=1.25$ eV is the Schottky barrier height of $In_{0.29}Al_{0.71}As$, ΔE_c is the conduction-band offset relative to InGaAs, the parameter $E_{f0}=0.04$ eV is the Fermi level at threshold, which was determined from the experimental data for $V_c=0$ V.[10] The capacitance can be

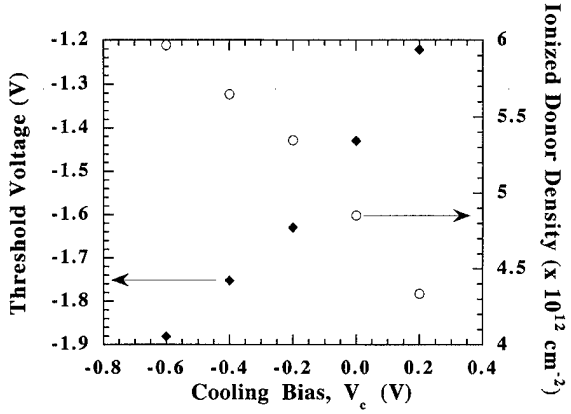


FIG. 4. The threshold voltage and the calculated ionized impurity density as deduced from Eq. (1) as a function of the cooling bias V_c.

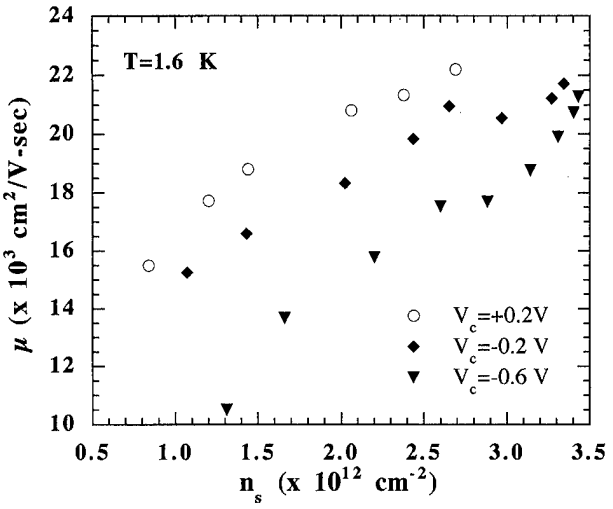


FIG. 5. Mobility vs free-carrier density for several values of the cooling potential bias V_c. As V_c becomes more positive, more electrons become trapped in the $In_{0.29}Al_{0.71}As$ layer as the structure cools. If the scattering sites are correlated in a $DX^- - d^+$ configuration, then the mobility measured at the 2DEG interface will be higher than for uncorrelated defects. If there are correlated DX centers in the $In_{0.29}Al_{0.71}As$, the mobility will increase the higher the cooling bias, as shown in the figure.

computed as the effective distance between the gate and the center of the delta doping layer, then divided by the dielectric constant of $In_{0.29}Al_{0.71}As$ ($\varepsilon = 11.2^* \varepsilon_0$) similar to the analysis in Tian *et al.*[13] This is in reasonable agreement with the capacitance C_i determined from Fig. 2 as the slope of the curves in the linear regime.

Using the capacitance and the threshold voltages determined from Fig. 2, the ionized donor density can be calculated from Eq. (1). Figure 4 shows a plot of both the measured threshold voltage and the calculated ionized donor density as a function of the cooling bias V_c. Because the total impurity density N_d is considered to be the sum of the shallow donors plus N_{DX}, the charges trapped on these DX centers, the decrease in N^+ with increasing V_c, as shown in Fig. 4, is attributed here to the increase in the trapped charge on the DX centers. If $N_d = 6\times10^{12}$ cm^{-2}, then we can see from Fig. 4 that ≈28% of the charge has been "frozen in" by the DX centerlike trap.

The saturation of $n_s(V_g)$ implies Fermi level pinning or, in other words, the Fermi level has risen to the lowest active donor energy in the system. In this case, we believe it is caused by the thermal activation energy of a DX center in a manner similar to that observed in AlGaAs/GaAs heterostructures.[14] From the conduction-band diagram in Fig. 3, we can define this energy E_d as

$$E_d = \Delta E_c - E_F(n_s) - V_p, \tag{2}$$

where V_p is the potential difference between the bottom of the δ-doping well and the heterointerface. In order to calculate this quantity, we need a relation between the Fermi level and the charge density. In the work by Chen *et al.*,[10] just such a relation was found from the linear regime of a figure similar to Fig. 3. They found $E_F(\text{eV}) = 5.84 \times 10^{-14} * n_s + 0.04$. Inserting the measured saturation value of 3.4×10^{12} cm^{-2}, we obtain a value of $E_f = 240$ meV.

Now for the calculation of V_p, the normal component of the displacement vector is continuous across the heterojunction

$$\varepsilon_1 E_{(\text{InGaAs})}(280^+) = \varepsilon_2 E_{(\text{InAlAs})}(280^-) \tag{3}$$

and the electric field on the InGaAs side of the heterojunction is $E_{(\text{InGaAs})}(280^+) = q n_s / \varepsilon_{(\text{InGaAs})}\varepsilon_0$, therefore the electric field on the InAlAs side of the junction is, in this approximation:

$$E_{(\text{InAlAs})}(280^-) = \frac{q n_s}{\varepsilon_{(\text{InAlAs})}\varepsilon_0}. \tag{4}$$

By inserting the saturated charge density $n_{s,\text{sat}}$, and integrating the electric field over the thickness of the spacer=3 nm, we obtain for V_p, 0.165 eV. Now placing the value for V_p and E_F back into Eq. (2), we obtain for the thermal ionization energy E_d a value of 0.32±25% eV.

To further strengthen the argument that we are seeing Fermi level pinning due to a DX centerlike defect, it is useful to consider the arguments in Buks *et al.*[14] in support of the negative U behavior of DX centers in AlGaAs/GaAs heterostructures. This behavior is related to the capture of two electrons by the center in its relaxed state. If such a state exists, then, upon cooling the heterostructure through the freeze-in temperature, the correlation between the negatively charged DX^- state and the positively charged substitutional donor d^+ state would result in an increase in overall mobility for a given charge density as the cooling potential increases. This is precisely what they measured in Fig. 4(a) of that work. In Fig. 5, we plot the mobility in our InAlAs/InGaAs versus the charge density in an identical fashion, and we too see an increase in the mobility as V_c becomes more positive. By reversing this argument, we believe the mobility dependence implies a negative U behavior of this defect, i.e., a DX center.

Figure 6 is a plot of the different conduction-band minima of $In_xAl_{1-x}As$, as a function of alloy composition. On this plot, we have put the thermal ionization level of the DX

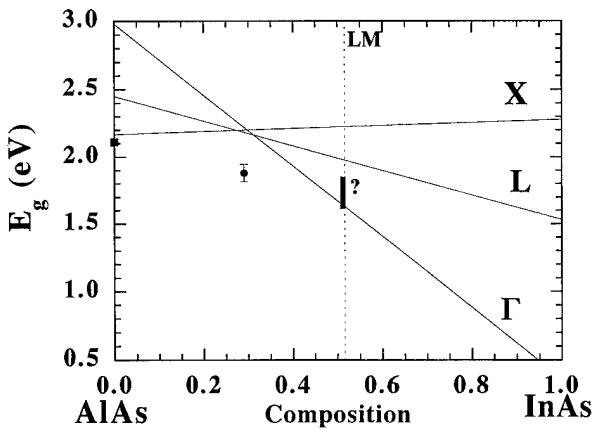


FIG. 6. Conduction-band diagram of the $In_xAl_{1-x}As$ alloy system along with the thermal ionization energies of the DX centers.

center in AlAs, the alloy region where Calleja *et al.*[6] claim not to find DX centers, and the depth of the level measured in this work. While Hong *et al.*[5] measured an emission energy of 0.37 eV and a photocapacitance turn on at 0.61 eV, it is unclear where to place their data, since they did not publish a thermal ionization energy. If the DX center activation energy for $In_{0.29}Al_{0.71}As$ is, in accordance with our measurements, located 0.32 eV below the L valley minimum, then the extrapolated linear relation in Fig. 6 suggests that no DX centers would be occupied in lattice-matched InAlAs/InP, because the energy level would be resonant with the conduction band. That no DX centers were found under pressure by Calleja *et al.*[6] could be due to the need to dope the material more heavily than 1×10^{18} cm^{-3}.

Within the context of a theory about the universality of the DX center is the work by Walukiewicz[15] regarding the position of native deep levels in III–V compounds all being referenced to the Fermi level pinning energies, which are themselves referenced relative to the vacuum level. In the case of $In_xAl_{1-x}As$ in the indirect region of the alloy system, the conduction-band (L band) edge moves up, while the valence band moves down as the amount of indium increases. Therefore, while the overall band gap of $In_xAl_{1-x}As$ decreases relative to AlAs, the conduction-band edge moves higher relative to the Fermi level pinning energy. If you consider the DX center as a kind of "native defect" of the doped material, the free energy of formation of the defect remains constant relative the vacuum level. It is then the prediction of this theory, that as the L band moves higher, the DX center energy level will move into the gap. Once the alloy material moves into the direct region, the conduction-band edge (Γ band) quickly drops relative to the Fermi pinning energy, and the DX center should move back into the gap. Therefore, this theory predicts DX centers should be their deepest at the direct–indirect transition ($x=0.3$). $In_xAl_{1-x}As$ latticed matched to InP ($x=0.52$) would not show any evidence of DX centers, because the energy level moves back into the conduction band. Further work on precisely these issues, including the optical response of the heterostructures, is currently in progress.

IV. CONCLUSION

In summary, from the cooling bias dependence of the threshold voltage and the mobility, we conclude a DX centerlike defect has been found in heavily Si-doped $In_{0.29}Al_{0.71}As$. From a measure of the ionized donor concentration and the known doping density, we find up to 28% of the electrons are trapped in this defect. From the pinning of the charge density in the two-dimensional electron gas (2DEG), we derive a thermal ionization energy for the DX center of $0.32\pm25\%$ eV below the conduction-band edge, twice the depth of the DX center in AlGaAs. At this time, further studies of the doping dependence and the photoconductive properties of new heterostructures are in progress to determine other characteristics of this DX-like energy level.

ACKNOWLEDGMENTS

The authors would like to thank Jianhui Chen for help with the growth of the heterostructures. One author (A.P.Y.) acknowledges the support of the Office of Naval Research under an AASERT fellowship.

[1]P. M. Mooney, J. Appl. Phys. **67**, R1 (1990).
[2]P. M. Mooney, Comments Condensed Mater. Phys. **16**, 167 (1992).
[3]M. Tachikawa, M. Mizuta, H. Kukimoto, and S. Minomura, Jpn. J. Appl. Phys. **24**, L821 (1985).
[4]K. Nakashima, S. Nojima, Y. Kawamura, and H. Asahi, Phys. Status Solidi A **103**, 511 (1987).
[5]W. P. Hong, S. Dhar, P. K. Bhattacharya, and A. Chin, J. Cryst. Growth **16**, 271 (1987).
[6]E. Calleja, A. L. Romero, S. Fernandez de Avila, E. Monuz, and J. Castagne, Semicond. Sci. Technol. 206 (1993).
[7]J. K. Luo, H. Thomas, S. A. Clark, and R. H. Williams, J. Appl. Phys. **74**, 6276 (1993).
[8]N. Chand, T. Henderson, J. Klem, W. T. Masselink, R. Fisher, Y. Chang, and H. Morkoc, Phys. Rev. B **30**, 4481 (1984).
[9]J. Chen, J. M. Fernandez, J. C. P. Chang, K. L. Kavanagh, and H. H. Wieder, Semicond. Sci. Technol. **7**, 601 (1992).
[10]J. Chen, H. H. Wieder, and A. P. Young, J. Appl. Phys. **76**, 4743 (1994).
[11]G. Snider, *Program: Schrodinger-Poisson Solver* (Cornell University, Ithaca, NY, 1992).
[12]P. S. Whitney, W. Lee, and C. F. Fonstad, J. Vac. Sci. Technol. B **5**, 796 (1987).
[13]H. Tian, K. W. Kim, and M. A. Littlejohn, J. Appl. Phys. **69**, 4123 (1991).
[14]E. Buks, M. Heiblum, and Y. Levinson, and H. Shtrikman, Semicond. Sci. Technol. **9**, 2031 (1994).
[15]W. Walukiewicz, Mater. Sci. Forum **519**, 143 (1994).

Microscopic processes during electron cyclotron resonance microwave nitrogen plasma-assisted molecular beam epitaxial growth of GaN/GaAs heterostructures: Experiments and kinetic modeling

Z. Z. Bandić[a] and T. C. McGill
Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125

R. J. Hauenstein[b] and M. L. O'Steen
Department of Physics, Oklahoma State University, Stillwater, Oklahoma 74078

(Received 22 January 1996; accepted 16 April 1996)

A set of δ-GaN_yAs_{1-y}/GaAs strained-layer superlattices grown on GaAs (001) substrates by electron cyclotron resonance (ECR) microwave plasma-assisted molecular beam epitaxy (MBE) was characterized by *ex situ* high resolution X-ray diffraction (HRXRD) to determine nitrogen content y in the nitrided GaAs monolayers as a function of growth temperature T. A first order kinetic model is introduced to quantitatively explain this $y(T)$ dependence in terms of an energetically favorable N for As anion exchange and thermally activated N-surface desorption and surface segregation processes. The nitrogen surface segregation process, with an estimated activation energy $E_s \sim 0.9$ eV appears to be significant during the GaAs overgrowth of GaN_yAs_{1-y} layers, and is shown to be responsible for strong $y(T)$ dependence. © *1996 American Vacuum Society.*

I. INTRODUCTION

In the past few years we have seen a flurry of activity in the development of visible light emitters based on wide band gap refractory semiconductor GaN, related alloys, and heterostructures.[1–4] Most of the research was concentrated on epitaxial growth techniques of nitrides,[5] and growth of device structures (LEDs and lasers) on various substrates,[6] while comparatively little work was done on other materials of potential technological applications, such as mixed anion nitride/arsenide system.[7–9] Although limited by a small solubility,[10] even small amounts of N in GaAs can cause unexpected band gap bowing to infrared,[10,11] while the small amounts of As in GaN can substantially increase valence band spin-orbital splitting which can significantly influence operation of the active layer in the laser structure due to the higher recombination rate.[12] One of the main difficulties in growing such nitride/arsenide structure, for example, in the electron cyclotron resonance plasma-assisted molecular beam epitaxy (ECR-MBE), is controlling the structural and chemical properties of the interface, which is only possible if important microscopic growth processes are fully understood. We have recently shown[7,8] that it is possible to produce high quality, fully commensurate GaN_yAs_{1-y}/GaAs strained-layer superlattices (SLS), grown on the GaAs substrate. Initial high resolution X-ray diffraction (HRXRD) and reflection high energy electron diffraction (RHEED) measurements done on such structures revealed existence of several distinct thermally activated processes, but limited by an unavailability of a suitable quantitative model against which these separate experimental observations could be tested consistently, only semi-quantitative considerations of kinetics and tentative identification of processes have been possible to date.[8]

In this article, we introduce a quantitative model, based on the first order kinetic theory, to examine the initial nitridation (surface N for As exchange), N surface desorption, and N surface segregation processes, observed in ECR-MBE grown GaN_yAs_{1-y}/GaAs heterostructures, characterized by *ex situ* HRXRD. Through application of our model to an unusually strong growth temperature dependence of y, obtained from X-ray experiments, the existence of a thermally activated N surface segregation mechanism is conclusively established and quantitatively confirmed. Although previously suggested,[8] the thermally activated N surface segregation mechanism could not be conclusively established without adequate analytical model, which is necessary to describe the concurrent, thermally activated N surface desorption and segregation processes. Activation energies of both surface segregation and desorption processes are numerically estimated, and the impact of the compositional-profile smearing of N, due to the strong N surface segregation effect, to GaN/GaAs heteroepitaxy and interface quality is analyzed and discussed.

II. EXPERIMENTAL DETAILS

A set of 36-period "δ-GaN_yAs_{1-y}/GaAs" superlattices, with a superlattice period consisting of one GaN_yAs_{1-y} monolayer (ML) and 75 GaAs monolayers, was grown on a GaAs (100) substrate as a function of temperature (540–580 °C) in an ECR-MBE system. GaN_yAs_{1-y} monolayers were produced through brief (4 s) N_2 plasma exposure (nitridation) of an As-stabilized GaAs surface. Nitridation was immediately followed by GaAs overgrowth at a fixed

[a]Electronic mail: zzbssdp.caltech.edu
[b]Electronic mail: rjh@hyper.phy.okstate.edu

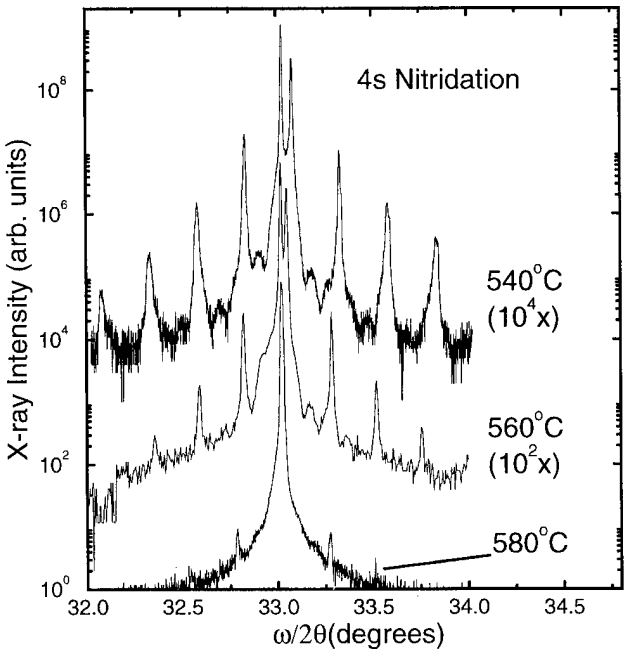


FIG. 1. HRXRD around (004) substrate reflections for δ-GaN_yAs_{1-y} /GaAs superlattices. Results indicate strong dependence of the rocking curves profiles on growth temperature.

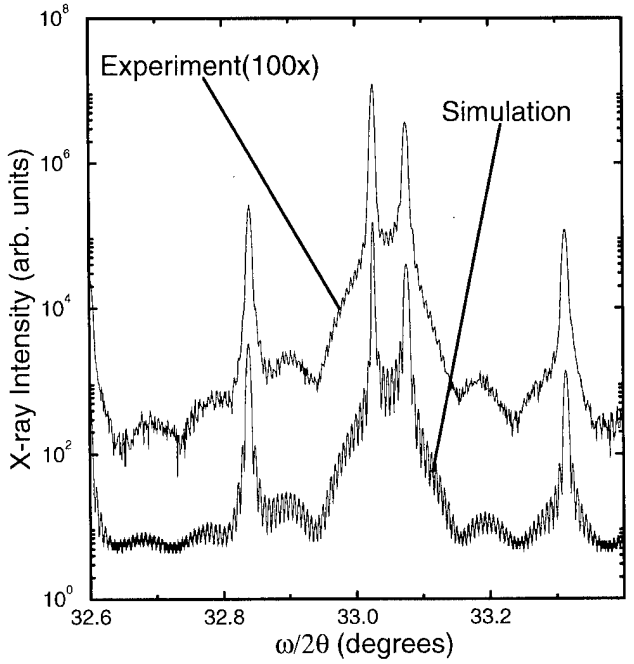


FIG. 2. Typical agreement between experimental rocking curve and HRS siumulated rocking curve.

growth rate (0.75 ML/s) for 100 s. The surface of GaAs is then As-stabilized through exposing to As_2 flux for only 30 s (As-soak). In addition, one superlattice was grown with 6 s nitridation at 550 °C. Further details about growth system and procedures can be found elsewhere.[7,8]

To determine strain and concentration profile, the entire sample set is then characterized by *ex situ* HRXRD, with the use of Philips Materials Research Diffractometer (MRD). The standard $\omega/2\theta$ scans are performed in the 4-crystal mode using Ge (220) reflections, around the substrate (004) reflections, while the reciprocal space scans (area scans) around the (115) reflections are done in the same mode, with additional use of a Bonse–Hart collimator in front of the detector. The area scans confirmed that all samples are coherently strained. The resultant $\omega/2\theta$ rocking curves for 4 s nitridations SLS are presented on Fig. 1 (550 °C and 570 °C scans have been omitted for clarity). Results reveal a strong dependence on growth temperature T of the epitaxial structure, indicating a presence of thermally activated microscopic processes. Two-dimensional equivalent alloy compositions (y) are determined from peak positions, and are confirmed through dynamical simulations using Philips High Resolution Simulation (HRS) software package. The agreement typically obtained between experimental rocking curve and simulation is presented on Fig. 2 on the example of the superlattice grown on 550 °C. Figure 3 shows an Arhennius plot of the resultant compositional dependence on growth temperature, $y(T)$, for 4 s (circles) and 6 s (diamond) nitridations, including experimental uncertainties in temperature, which represent measured variation over a given wafer (i.e., temperature nonuniformity) due to unintended differences in sample mounting. Although not included in the plot, the absolute errors associated with GaN composition can be estimated to $\pm 0.5\%$. These errors are a result of a Vegard's law approximation and uncertainties in numerical values of elastic properties of the GaN_yAs_{1-y} layer.

These results suggest existence of two regimes: (*i*) dose-limited at low temperatures; (*ii*) kinetically limited at high

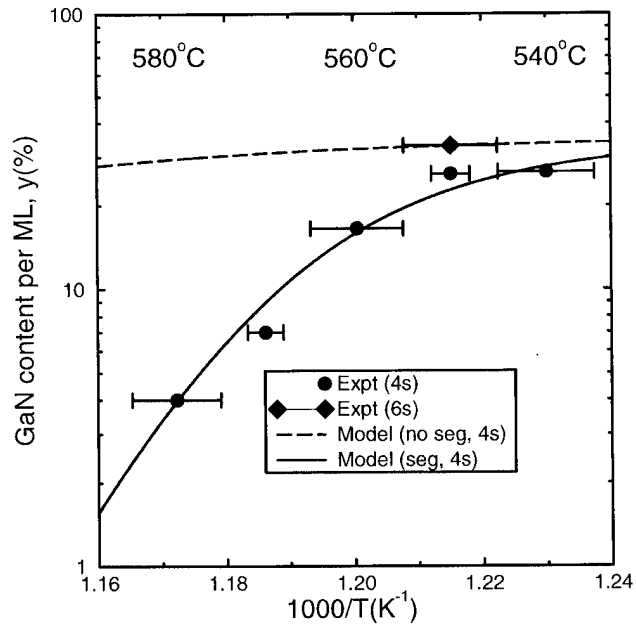


FIG. 3. Growth-kinetics models to explain the strong $y(T)$ dependence observed for ECR-MBE grown δ-GaN_yAs_{1-y}/GaAs strained-layer superlattices. Experimental data points are obtained through HRXRD measurements for 4 s (circles) and 6 s (diamond) nitridations. The dashed curve represents the model neglecting N surface segregation [Eq. (2)], while the solid curve represents the model including segregation effects [Eq. (8)].

temperatures, with unexpectedly strong fall-off in y with increasing growth temperature.

III. KINETICAL MODELING

For the purpose of understanding the physical origin of this $y(T)$ dependence, we have developed a first order kinetic model which explicitly incorporates three phases of our particular SLS growth sequence: *(i)* nitridation; *(ii)* overgrowth, and *(iii)* As-soak, as described in Sec. II.

A. Nitridation

The net rate of surface nitridation is determined by two physical processes which have been previously identified: *(i)* N-for-As surface anion exchange (N gain), and *(ii)* surface N desorption (N loss). These processes can be analyticaly modeled, in the first order in N surface compostion y as

$$\frac{dy}{dt}=s\Phi(1-y)-\frac{y}{\tau_d}. \tag{1}$$

The first term on the right hand side of Eq. (1) corresponds to the N gain on the As-stabilized GaAs surface due to the energetically favorable anion exchange. This process is modeled as being proportional, in the first order, to the flux Φ of incident physically activated N; to the fraction of available As surface sites to be exchanged with N, and to an overall efficiency factor s (units of area), which in general might depend on surface chemistry, strain, and temperature, but is approximated as constant over the temperature and composition range of interest, for the purpose of our simplified model. The second term in this equation, $-y/\tau_d$, represents the N loss due to surface desorption. It can be simply modeled in terms of thermally activated rate constant $\tau_d^{-1}=\tau_{0d}^{-1}\exp(-E_d/kT)$, with activation energy E_d. Such a process has been directly observed in our RHEED measurements of nitrided GaAs surfaces, and estimates for E_d (~2.1 eV) have been previously reported.[7,8]

If we define an "effective dosing rate" $r\equiv s\Phi$, the solution to Eq. (1) for an initially N-free (As-stabilized) surface can be written as

$$y(t_{\exp})=y_{ss}[1-\exp(-t_{\exp}/\tau)], \tag{2}$$

where $y_{ss}\equiv r\tau$, $\tau\equiv(\tau_d^{-1}+r)^{-1}$, and $t_{\exp}$ is the nitrogen plasma exposure time. If Eq. (2) is considered in the limit of low temperatures so that desorption rate τ_d^{-1} is negligible, then $y\to 1$ would be obtained. This reflects the assumption that As atoms from the As-enriched (100) GaAs surface, when exposed to N plasma, will be completely exchanged for N atoms in the steady state, since N terminated surface has lower surface free energy, due to a much stronger Ga–N than Ga–As bond (heats of formation 6.81 eV and 5.55 eV, respectively[13]). Although the essential nitridation kinetics is well described by Eq. (2) without unnecessary mathematical complexity, it ignores the subsequent stages of growth of our superlattices, and therefore cannot be used to model directly our experimental results. This can be seen in Fig. 3, where the dashed curve represents the plot of Eq. (2) taking $t_{\exp}=4$ s and using our RHEED based experimental desorption parameters, $E_d=2.1$ eV and $\tau_d=7$ s at 592 °C.[7,8] In addition, the dosing rate r is estimated to be approximately 0.1 Hz from experimental values of y at 550 °C for 4 s and 6 s nitridations, such that Eq. (2) describes properly the dose-limited (low temperature) regime of growth.

B. GaAs overgrowth—N segregation

Inspection of Fig. 3 clearly reveals that the processes of anion exchange and N surface desorption *alone* result in the weak fall-off of the model (dashed) curve, and thus cannot account for the strong $y(T)$ dependence in the kinetically limited growth regime. Also, even N surface desorption during the short monolayer deposition time (t_{ML}) required to bury (and freeze-in) the nitrided GaN_yAs_{1-y} layer cannot result in significant N loss. To overcome this, we hypothesize that thermally activated N surface segregation process, described with the rate constant $\tau_s^{-1}=\tau_{0s}^{-1}\exp(-E_s/kT)$, is occurring concurrently with N desorption, during GaAs overgrowth.[8] This assumption is supported qualitatively by our experimental observation that GaN-related (3×3) surface reconstruction proceeds faster to GaAs (2×4) reconstruction at a lower substrate temperature during GaAs overgrowth, which is consistent with freeze-out of a thermally activated N segregation process.

Let us assume that (1) segregation occurs from the first subsurface layer upward to the current surface layer; (2) desorption occurs from all currently exposed surfaces, i.e., the current growth surface and the current partially exposed first sub-surface layers; and (3) the segregation and desorption processes are statistically independent, and are occurring concurrently. Also, let S_n represent a relative GaN mole fraction in monolayer n, where by definition the content of the nitrided layer at the end of nitridation process is normalized to unity, $S_0(t=0)\equiv 1$. Under these assumptions, we can relate the N content of monolayers n and $n+1$ through

$$S_{n+1}(0)=[S_n(0)-S_n(t_{ML})]\frac{\tau_s^{-1}}{\tau_s^{-1}+\tau_d^{-1}}. \tag{3}$$

Equation (3) describes the fraction of lost N $[S_n(0)-S_n(t_{ML})]$, which segregates rather then desorbs, and becomes the nitrogen content of the next surface layer. Since layer n loses its N content due to both N desorption and segregation, during the time interval $0\leq t\leq t_{ML}$, $S_n(t)$ decays exponentially as $\exp(-t/\tau_\parallel)$, where $\tau_\parallel^{-1}\equiv(\tau_s^{-1}+\tau_d^{-1})$. Using this expression in Eq. (3) leads us to

$$S_{n+1}(0)=S_n(0)[1-\exp(-t_{ML}/\tau_\parallel)]\frac{\tau_d}{\tau_{tot}}, \tag{4}$$

where $\tau_{tot}\equiv\tau_s+\tau_d$. It is obvious from Eq. (4) that $S_n(0)$ represents geometrical progression. Now, similarly to Eq. (3) we can define δ_n as the amount of initial material lost through desorption from layer n:

$$\delta_n\equiv[S_n(0)-S_n(t_{ML})]\frac{\tau_s}{\tau_{tot}}. \tag{5}$$

Therefore, the total loss of N from one superlattice period during overgrowth would be

$$L=\sum_{n=0}\delta_n, \tag{6}$$

which is simply a sum of a geometric progression. Finally, the total amount of the original GaN content retained within one period of a superlattice is $R=1-L$, or straightforwardly

$$R=\frac{\tau_{\mathrm{tot}}}{\tau_d+\tau_s e^{t_{\mathrm{ML}}/\tau_{\|}}}. \tag{7}$$

It is important to note that the expression for R given in Eq. (7) gives the correct limiting behaviour in describing the loss of N during the GaAs overgrowth, in the limits of no segregation, or no desorption. If there is desorption, but no segregation, then the only opportunity for N loss would be before nitrided layer is completely buried in GaAs; i.e., $R=\exp(-t_{\mathrm{ML}}/\tau_d)$, which is obtained directly from Eq. (7) in the limit $\tau_s\rightarrow\infty$. Similarly, if there is segregation, but no desorption, then R must be unity, since the only possible mechanism for N loss is desorption process. In that case, where $\tau_d\rightarrow\infty$, one would obtain $R\rightarrow 1$, as it should be.

IV. RESULTS AND DISCUSSION

Using the "segregation correction factor" from Eq. (7), it is clear that in order to consistently incorporate N surface segregation into our model, Eq. (2) must be replaced by

$$y(t_{\mathrm{exp}})=y_{\mathrm{ss}}[1-\exp(-t_{\mathrm{exp}}/\tau)]\times\left[\frac{\tau_{\mathrm{tot}}}{\tau_d+\tau_s e^{t_{\mathrm{ML}}/\tau_{\|}}}\right]. \tag{8}$$

Thus, Eq. (8) represents the final result of our model and quantitatively describes the growth sequence (nitridation/overgrowth/soak) of our δ-GaN_yAs_{1-y}/GaAs superlattices, taking into account microscopic processes of anion exchange, desorption, and segregation, and is plotted as the solid curve in Fig. 3. This curve is calculated using identical dosing (r) and desorption (E_d and τ_{0d}) parameter values which were used in Eq. (2) to obtain the dashed curve, while adjusting at the same time two segregation parameters, E_s and τ_{0s}. Nonlinear least-squares fit to the experimental data points for 4 s nitridations finally results in the estimate for N surface segregation energy equal to 0.9 eV$\pm$30%. Obviously, from Fig. 3, it is seen that with the corrected Eq. (8) very good agreement with experimental data is obtained, which confirms our N surface segregation hypothesis.

Additionally, the compositional-profile smearing of N along the growth direction would be expected as a natural consequence of the segregation process, which is confirmed in preliminary cross sectional scanning tunneling microscopy (STM) images of one of our superlattices.[14] Unfortunately, such profile smearing effects would tend to limit the abruptness of the GaN/GaAs interface. However, it is possible that smearing effects might be controlled through some of the MBE techniques which have been succesfully used in the problem of dopant-profile smearing in MBE-grown Si.[15] Also, the cross-sectional STM as a direct imaging technique might provide another way to determine E_s, since our model predicts distribution of nitrogen in the superlattice period ($S_n(0)$), as a function of N surface activation energy E_s and growth temperature T.

V. SUMMARY

In conclusion, a simple, but plausible kinetic model have been developed to conclusively establish and quantitatively explain some of the microscopic processes observed in ECR-MBE grown δ-GaN_yAs_{1-y}/GaAs strained layer superlattices. Particularly, the strong $y(T)$ dependence, obtained through *ex situ* HRXRD of our superlattices, is fully understood in terms of energetically favorable surface anion exchange and combined, thermally activated N surface-segregation/surface-desorption. Finally, our model predicts for the first time numerical estimates of the kinetic parameters associated with N surface-segregation process, which appears to be significant under typical GaN/GaAs ECR-MBE growth conditions.

ACKNOWLEDGMENTS

This work was supported by Advanced Research Project Agency, and monitored by Office of Naval Research under Grant No. N00014-92-J-1845. Additionally, two of us (R.J.H. and M.L.O.) wish to acknowledge the support of Air Force Office of Scientific Research under Contract Nos. F49620-93-1-0211 and F49620-93-1-0389, and the support of the Advanced Research Projects Agency, monitored through the Army Research Office under Contract No. DAAH04-94-G-0393.

[1]S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku and Y. Sugimoto, Jpn. J. Appl. Phys. **35**, 74 (1996).
[2]S. Nakamura, M. Senoh, N. Iwasa, and S. Nagahama, Appl. Phys. Lett. **67**, 1868 (1995).
[3]I. Akasaki, H. Amano, S. Sota, H. Sakai, T. Tanaka, and M. Koike, Jpn. J. Appl. Phys. **34**, 1517 (1995).
[4]H. Sakai, T. Koide, H. Suzuki, M. Yamaguchi, S. Yamasaki, M. Koike, H. Amano, and I. Akasaki, Jpn. J. Appl. Phys. **34**, L1429 (1995).
[5]W. Gotz, N. M. Johnson, J. Walker, D. P. Bour, H. Amano, and I. Akasaki, Appl. Phys. Lett. **67**, 2666 (1995).
[6]S. N. Mohammad, A. A. Salvador, and H. Morkoç, Proc. IEEE **83**, 1306 (1995).
[7]R. J. Hauenstein, D. A. Collins, X. P. Cai, M. L. O'Steen, and T. C. McGill, Appl. Phys. Lett. **66**, 2861 (1995).
[8]R. J. Hauenstein, D. A. Collins, M. L. O'Steen, Z. Z. Bandić, and T. C. McGill, Mater. Res. Soc. Symp. Proc. **388**, 259 (1995).
[9]M. Sato, Jpn. J. Appl. Phys. **34**, 1080 (1994).
[10]J. Neugebauer and C. G. Van de Walle, Phys. Rev. B **51**, 10568 (1995).
[11]S. Sakai, Y. Ueta, and Y. Terauchi, Jpn. J. Appl. Phys. **32**, 4413 (1993).
[12]Z. Z. Bandić and T. C. McGill (unpublished).
[13]*CRC Handbook of Chemistry and Physics,* edited by R. C. West (Chemical Rubber, Boca Raton, FL, 1987), p. E102.
[14]R. M. Feenstra (unpublished).
[15]H. Jorke, Surf. Sci. **193**, 569 (1988).

Tensile strain relaxation in GaN_xP_{1-x} ($x \leq 0.1$) grown by chemical beam epitaxy

N. Y. Li,[a] W. S. Wong,[b] D. H. Tomich, K. L. Kavanagh, and C. W. Tu
Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, California 92093-0407

(Received 23 January 1996; accepted 31 March 1996)

A study of strain relaxation in GaN_xP_{1-x} epilayers grown by chemical beam epitaxy, using a rf-plasma nitrogen radical beam source, tertiarybutylphosphine, and triethylgallium, is reported. Microcracks are observed in GaN_xP_{1-x} epilayers grown on GaP when the nitrogen composition is greater than 1.6%. Transmission electron microscopy results show that the tensile-strain relaxation in GaN_xP_{1-x} epilayers is initially relieved by microcrack formation without misfit dislocations. These microcracks penetrate through the interface, degrading the crystallinity of the GaP substrate. Microcracks formation can not be alleviated by adjusting the growth rate, growth temperature, V/III ratios, and forward plasma power, but they can be eliminated by reducing the growth area of the GaP substrate. © *1996 American Vacuum Society.*

I. INTRODUCTION

Recently, high-brightness blue-light-emitting diodes[1] and high-temperature electronics,[2] have been successfully demonstrated, using the group III nitrides. However, since their lattice constants are much smaller than that of Si, their integration with high-speed Si electronics is difficult. GaN_xP_{1-x} alloys, which can be lattice matched to Si substrates, are promising in potential optoelectronic applications,[3–6] but problems including antiphase domains, pregrowth preparation, and differences in the linear thermal expansion coefficient also make direct growth on Si difficult. To overcome some of these problems, GaP/Si can be used as a substrate. Therefore, we are investigating the growth of GaN_xP_{1-x} on GaP.

A variety of methods have been implemented to grow GaN_xP_{1-x} alloys on GaP using high-pressure ammonia as the N source, co-injected with phosphine[4] or tertiarybutylphosphine (TBP).[7,8] In this study, the GaN_xP_{1-x} samples were grown using a rf-plasma N radical beam source by chemical beam epitaxy (CBE). This N source is highly reactive and has been successfully used to grow GaN,[9] GaN_xAs_{1-x},[3] and to dope ZnSe p-type for blue-light-emitting laser diodes.[10] Recently, we have also demonstrated its use for the growth of GaN_xP_{1-x}, where up to 16%[11] N was incorporated. This is the highest N content reported to date for this alloy.

The lattice constant of GaN_xP_{1-x} alloys is smaller than that of the GaP substrate, therefore, GaN_xP_{1-x} epilayers are under tensile stress. Whereas compressively strained material systems have been extensively studied, there are few reports on tensile-strain relaxation. In this study we report strain relaxation and interface microstructure in GaN_xP_{1-x}/GaP. We found that the tensile misfit strain between GaN_xP_{1-x} and GaP is accommodated by microcrack formation rather than dislocations. Our investigations of the effects of the growth parameters on the surface morphology of GaN_xP_{1-x} epilayers are also reported.

II. EXPERIMENT

The CBE system used in this work was modified from a Perkin-Elmer 425B MBE system. Triethylgallium (TEGa), nitrogen (N_2), and TBP were used as the gallium, nitrogen, and phosphorus source, respectively. TEGa and TBP were introduced into the chamber without any carrier gas, and the flow rates were controlled by mass-flow controllers. The temperature of TBP cracker was 800 °C. Ultrahigh purity nitrogen was introduced into an Oxford Applied Research rf-activated plasma source with the flow rate controlled by a mass-flow controller. The rf-plasma N radical beam source was operated at 13.56 MHz with a forward power of 300 W. The N_2 flow rate was varied from 0.375–1.5 sccm, and the chamber pressure was about 4×10^{-5} Torr during growth. The substrates used in this study were nominally undoped (100) GaP. Before being loaded into the CBE chamber, they were chemically etched in $HCl:HNO_3:H_2O$ (4:4:5) for 5 min and finally given deionized (DI) water rinse.

In this study, GaN_xP_{1-x} ($0.007 \leq x \leq 0.102$) samples, 0.5–0.8 μm thick, were grown at different growth parameters[12] to study the strain relaxation under tensile stress. Typically, the growth rate and the growth temperature were varied from 0.56–0.85 monolayers per second (ML/s) and from 400–690 °C, respectively. The gallium and phosphorus incorporation rates were determined by group III- and group V-induced reflection high-energy electron diffraction (RHEED) intensity oscillations, respectively. The V/III (P/Ga) incorporation ratio was about 1.3. The (P+N)-induced incorporation rate cannot be collected from group V-induced RHEED intensity oscillations because the N radical beam cannot be completely blocked by the shutter. RHEED shows a (2×1) streaky pattern during growth. The surface morphology and strain relaxation of GaN_xP_{1-x} epilayers were characterized by atomic force microscopy (AFM), scanning elec-

[a]Electronic mail: nli@sdcc3.ucsd.edu
[b]Present address: Department of Material Science and Mineral Engineering, University of California, Berkely, CA 94720.

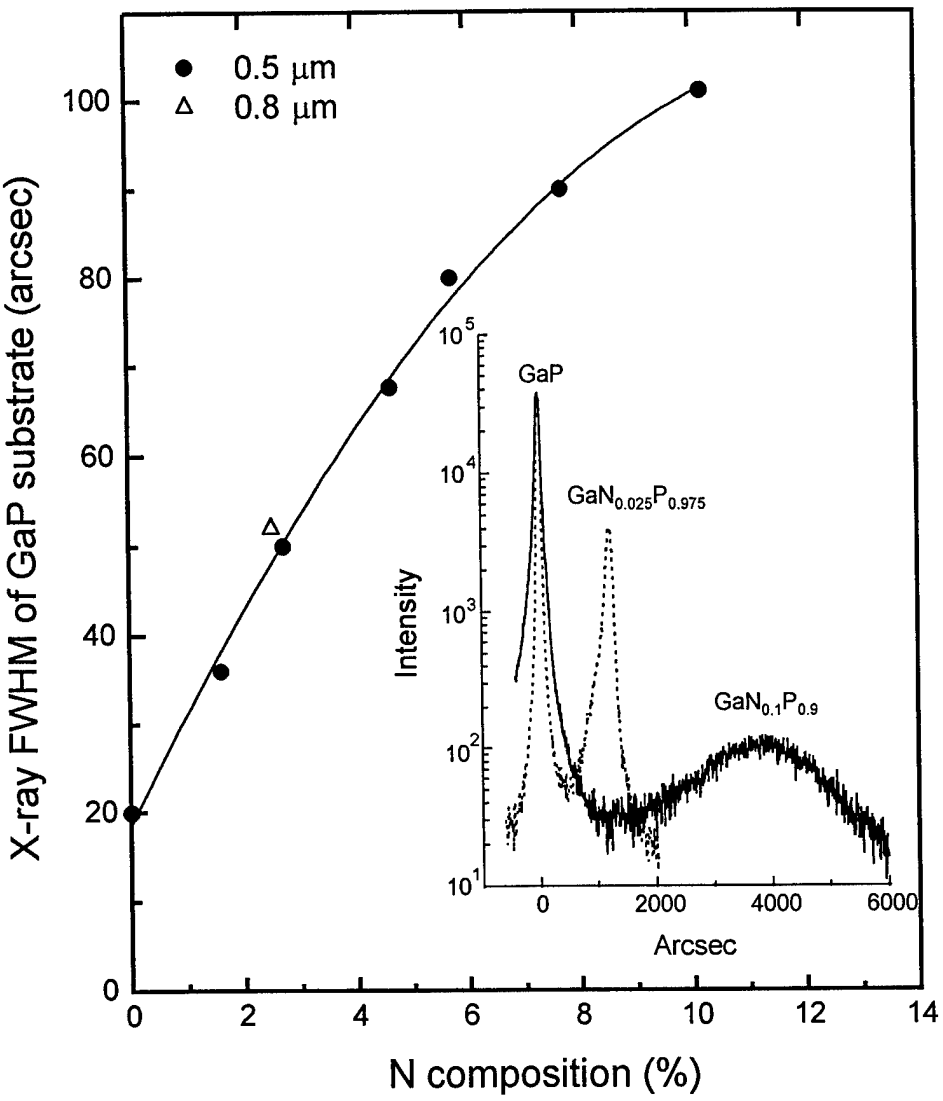


FIG. 1. The x-ray FWHM of the GaP substrates as a function of the N composition. The inset shows (400) x-ray rocking curves of a 0.8-μm-thick $GaN_{0.025}P_{0.975}$ and a 0.5-μm-thick $GaN_{0.1}P_{0.9}$ on GaP using a rf-activated plasma N radical beam source.

tron microscopy (SEM), Nomarski microscopy, high-resolution double-crystal x-ray diffraction, and transmission electron microscopy (TEM).

III. RESULTS AND DISCUSSION

(400) x-ray rocking curves of GaN_xP_{1-x} samples are shown in the inset of Fig. 1. If the film is completely relaxed, the (400) peak splitting (dash line) indicates a N composition of 4.9% using literature values for the GaP[13] and cubic GaN[14,15] elastic constants and Vegard's law. However, analysis of additional asymmetric {511} x-ray rocking curves shows that the epilayer is still highly strained, only 2.5 $\pm$0.2% relaxed in both [011] and [01$\bar{1}$] in-plane directions. Therefore, the N content is determined to be 2.5% (misfit= $-$0.47%). Similar analysis indicates that the film of the solid line in the inset of Fig. 1 is relaxed 58.3% (misfit=$-$1.8%) and has a N composition of 10.2%.

Figure 1 shows the full width at half-maximum (FWHM) of the GaP substrate (400) x-ray peak as a function of the N composition. Compared to a normal GaP substrate, FWHM of 20 arcsec, a much broader substrate peak with a FWHM of 102 arcsec is observed for 0.5-μm-thick $GaN_{0.1}P_{0.9}$ grown on GaP. Figure 2 shows a cross-section TEM image of a 0.8 μm-thick $GaN_{0.025}P_{0.975}$ on GaP. A microcrack is observed in the epilayer that propagates into the substrate, ending in a stacking fault. A similar TEM result was also observed for the sample with a N composition of 4.6% (misfit=$-$0.81%). No misfit dislocations were observed, at least up to 4.6% N. This is surprising since based on the Matthews and Blakeslee model,[16] the critical layer thickness is about 200 Å. With increasing lattice mismatch between GaN_xP_{1-x} and GaP, microcracks apparently generate and propagate into the substrate, degrading the crystalline quality of both the GaP substrate as well as the epilayer. We found that the higher the N

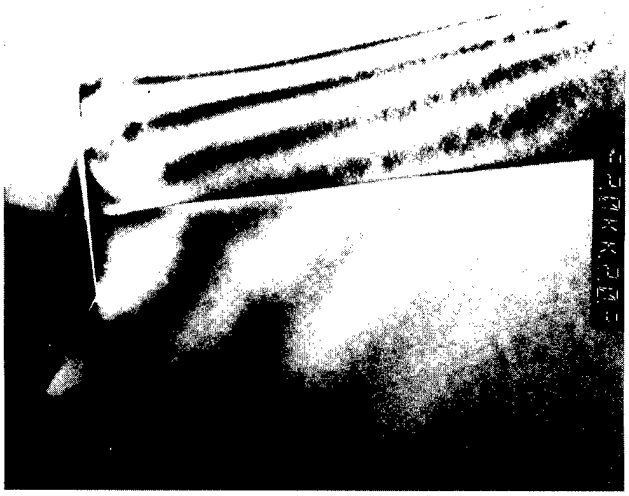


FIG. 2. Bring field cross-section TEM micrograph of $GaN_{0.025}P_{0.975}$ grown at 640 °C. A microcrack extending from the surface through the epilayer into the substrate is visible.

composition or the thicker the epilayer, the broader the GaP substrate x-ray peak.

The surface morphology of the GaN_xP_{1-x} epilayers is optically specular, but consistent with TEM and x-ray diffraction microcracks are visible by eye or with an optical microscope when the N composition exceeds 1.6%. Shown in Fig. 3 is an AFM image of the $GaN_{0.025}P_{0.975}$ layer. Microcracks with narrow openings (50 nm) are observed along both [011] and [01$\bar{1}$] directions, and one with a wider opening of 350 nm is also observed. In between the microcracks, the surface roughness of the $GaN_{0.025}P_{0.975}$ layer is about 15 Å, indicating a smooth surface morphology. Using SEM, wide microcracks in the range 150–550 nm are observed with a density of <1000/cm. We also noticed large cracks forming during growth through a viewport of the CBE when the thickness of the epilayer was about 0.3 μm. Figure 4 shows the percentage of strain relaxation in 0.5-μm-thick GaN_xP_{1-x} epilayers, as a function of the N composition. The GaN_xP_{1-x} epilayers with a N composition of 2.5%, 4.6%, 5.7%, 7.8%, and 10.2% have relaxed 2.4%, 3.2%, 11.9%, 33.8%, and 58.3%, respectively. The percentage of strain relaxation in GaN_xP_{1-x} ep-

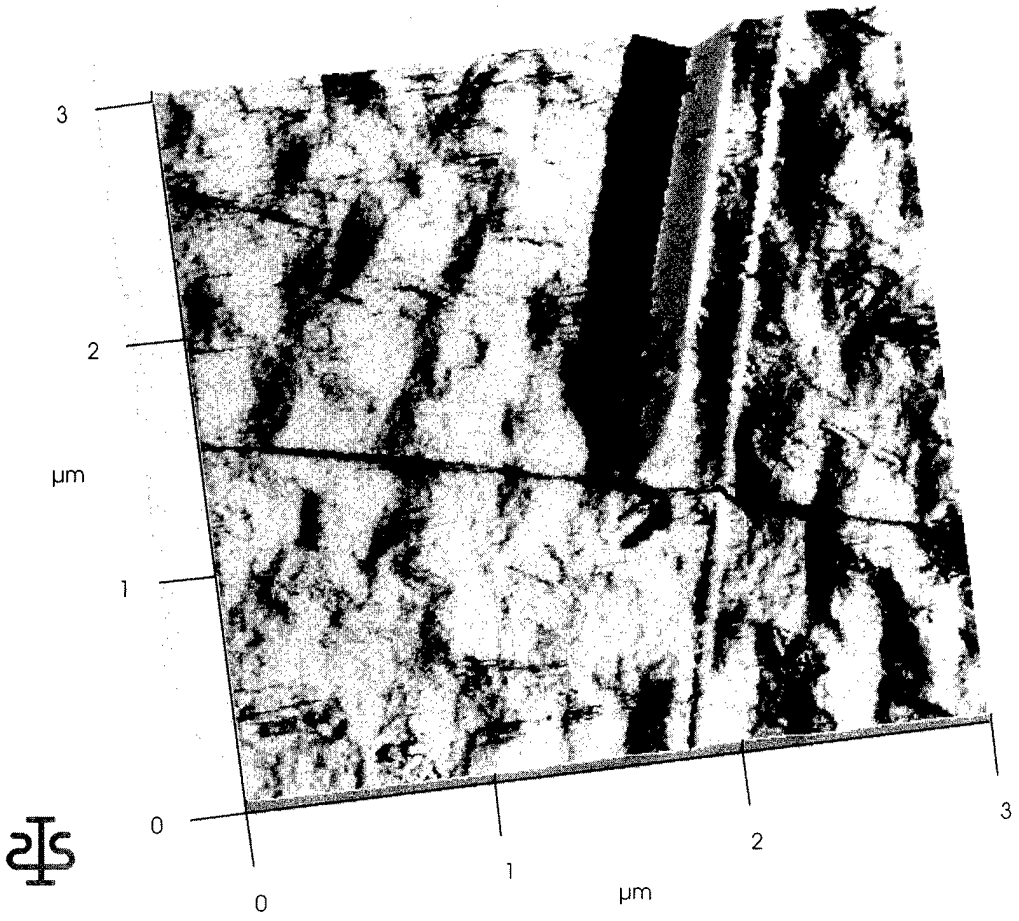


FIG. 3. AFM image of a 0.8-μm-thick $GaN_{0.025}P_{0.975}$ grown at 640 °C. Microcracks are observed in both [011] and [01$\bar{1}$] directions.

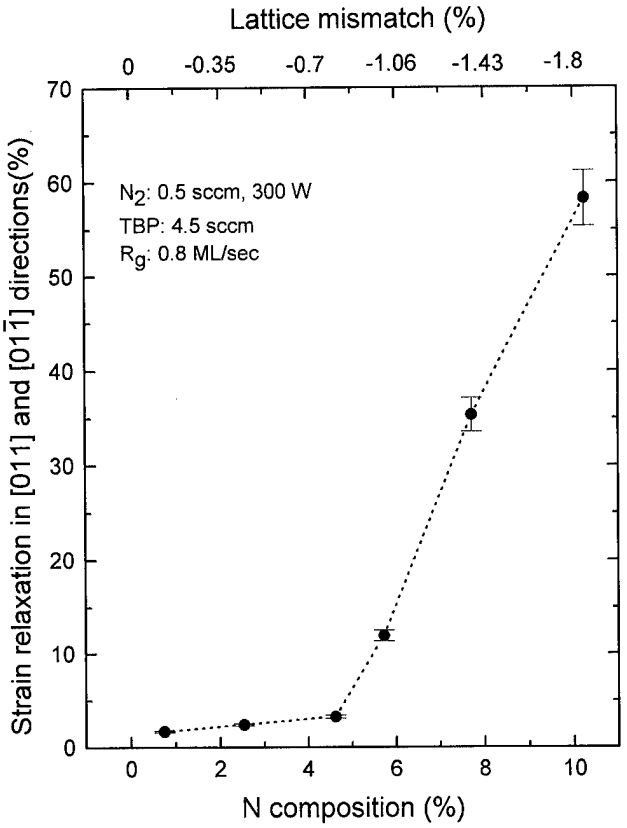


FIG. 4. The percentage of strain relaxation of GaN_xP_{1-x} ($0.007 \leq x \leq 0.102$) in both [011] and [01$\bar{1}$] directions as a function of N composition.

ilayers increases dramatically when the N composition is higher than 4.6%. This may be associated with the onset of misfit dislocation formation, however, TEM still needs to be carried out to confirm this.

Based on the above results, we conclude that when the GaN_xP_{1-x} epilayer thickness exceeds a certain thickness, the strain is relieved initially by microcracks rather than dislocations. By taking asymmetric {511} x-ray rocking curves in the two perpendicular in-plane ⟨110⟩ directions, we found that the strain relaxation via microcracks in GaN_xP_{1-x} ($x \leq 0.1$) epilayers is isotropic. Microcrack densities are symmetric in both [011] and [01$\bar{1}$] directions. This result is different from the anisotropic strain relaxation via dislocations in most III–V compounds grown with compressive strain. Our result is different from other tensile strained material systems. Gao *et al.*[17] reported that the distribution of microcracks in 4.5-μm-thick GaAs/Si is anisotropic. The majority of the microcracks run in the [011] direction and only a few run in the [01$\bar{1}$] direction, which is possibly related to the substrate misorientation. Woodbridge *et al.*,[18] however, found that GaAs/Si has no clear preference for microcracks in either ⟨011⟩ directions, is similar to our result. Olsen *et al.*[19] also reported asymmetric microcracks in epitaxial InGaP on GaAs when the tensile misfit exceeded ~0.15%, and asymmetric cracking was explained by invoking asymmetry of the distribution of dislocations involved in the microcrack formation.

Several groups have reported the growth of GaN_xP_{1-x} by gas-source molecular beam epitaxy[4,11] (GSMBE) and organometallic vapor phase epitaxy[15] (OMVPE). Microcracks, however, are not observed for a 0.5-μm-thick GSMBE-grown GaN_xP_{1-x} samples with N composition up to 16%. Microcracks are observed in our study when the N composition is higher than 1.6%, which is much smaller than 16%. This discrepancy may be due to the different growth mechanism among CBE, GSMBE, and MOVPE. Furthermore, the C impurity concentration might play an important role in the tensile-strain relaxation process. In our previous growth study of GaN_xP_{1-x},[7] secondary-ion mass spectroscopy shows that the C impurity has a concentration of 7×10^{18} cm^{-3} in the epilayer, mainly from the dissociation of TBP cracked at 800 °C. Due to the smaller covalent radius of carbon, a lattice contraction should be expected in GaP, which can introduce an extra tensile stress in the epilayer. Also, the unintentional p-type C doping in GaN_xP_{1-x} probably increases the activation energy for nucleation and/or mobility of dislocations, facilitating tensile-strain relaxation via microcracks. Further studies are necessary to confirm an effect of C impurity on the tensile-strain relaxation in GaN_xP_{1-x}.

To eliminate microcrack formation in the GaN_xP_{1-x} epilayers, growth parameters such as growth rate, V/III ratio, growth temperature, and forward plasma power, were varied. We found that microcracks in GaN_xP_{1-x} epilayers can not be alleviated by adjusting these growth parameters, but the surface morphology can be improved significantly by reducing the growth area. A patterned GaP substrate with SiO_2 as a mask was used to reduce the growth-substrate area for GaN_xP_{1-x} growth. Our preliminary results show that microcracks are not observed in the 0.5-μm-thick GaN_xP_{1-x} epilayers with the N composition up to 10.2% grown in the small area (300 μm×300 μm) of a patterned GaP substrate. This result implies that strain relaxation via the microcrack formation in GaN_xP_{1-x} epilayers can be effectively eliminated by reducing the growth area. This effect is still under investigation, and further results will be reported in details elsewhere.

IV. SUMMARY

AFM, Nomarski microscope, SEM, double crystal x-ray diffraction, and TEM were used to study the microstructure of GaN_xP_{1-x} epilayers on GaP grown by CBE using a rf-plasma N radical beam source. We found that even a small N composition of 1.6% in the GaN_xP_{1-x} epilayer leads to the formation of microcracks in this tensile strained material system. From TEM investigations, for N composition up to at least 4.6%, no threading or misfit dislocations are found in GaN_xP_{1-x}/GaP epilayer. The tensile misfit strain between GaP and GaN_xP_{1-x} grown by CBE is first accommodated by microcrack formation rather than misfit dislocations. The strain relaxation in GaN_xP_{1-x}/GaP up to 60% is isotropic in both [011] and [01$\bar{1}$] directions.

ACKNOWLEDGMENTS

This work is supported by Texas Instruments System Components Laboratory. The authors would like to thank Professor J. C. P. Chang of Purdue University and F. Deng of UCSD for valuable assistance and discussions in the experiments.

[1]S. Nakamura, T. Mukai, and M. Senoh, Appl. Phys. Lett. **64**, 1687 (1994).
[2]M. Rubin, N. Newman, J. S. Chan, T. C. Fu, and J. T. Ross, Appl. Phys. Lett. **64**, 64 (1994).
[3]M. Kondow, K. Uomi, K. Hosomi, and T. Mozume, Jpn. J. Appl. Phys. **33**, L1056 (1994).
[4]J. N. Baillargeon, K. Y. Cheng, G. E. Hofler, P. J. Pearah, and K. C. Hsieh, Appl. Phys. Lett. **60**, 2540 (1992).
[5]J. N. Baillargeon and K. Y. Cheng, J. Appl. Phys. **70**, 1841 (1991).
[6]X. Yang, Z. Lin, Z. Li, L. Wu, and C. Mao, J. Appl. Phys. **77**, 5533 (1995).
[7]N. Y. Li, W. S. Wong, D. H. Tomich, H. K. Dong, J. S. Solomon, J. T. Grant, and C. W. Tu, J. Cryst. Growth (to be published).
[8]A. Wakahara, K. Hirano, X. L. Wang, and A. Sasaki, Inst. Phys. Conf. Ser. **136**, 607 (1994).
[9]C. T. Foxon, T. S. Cheng, S. V. Novikov, D. E. Lacklison, L. C. Jenkins, D. Johnston, J. W. Orton, S. E. Hooper, N. Baba-Ali, T. L. Tansley, and V. V. Tret'yakov, J. Cryst. Growth **150**, 892 (1995).
[10]K. Ohkawa, A. Ueno, and T. Mitsuya, J. Cryst. Growth **117**, 375 (1992).
[11]W. G. Bi and C. W. Tu (unpublished).
[12]N. Y. Li, D. H. Tomich, W. S. Wong, J. S. Solomon, and C. W. Tu, Mater. Res. Soc. Symp. Proc. (to be published).
[13]J. Hornstra and W. J. Bartels, J. Cryst. Growth **44**, 513 (1978).
[14]E. Sherwin and T. J. Drummond, J. Appl. Phys. **69**, 8423 (1991).
[15]Miyoshi, K. Onabe, N. Ohkouchi, H. Yaguchi, R. Ito, S. Fukatsu, and Y. Shiraki, J. Cryst. Growth **124**, 439 (1992).
[16]J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).
[17]Gao, A. W. Stevenson, A. W. Wilkins, and G. N. Pain, J. Cryst. Growth **129**, 134 (1992).
[18]Woodbridge, P. Barnes, R. Murray, C. Roberts, and G. Parry, J. Cryst. Growth **127**, 112 (1993).
[19]H. Olsen, M. S. Abrahams, and T. J. Zamerowski, J. Electrochem. Soc. **121**, 1650 (1974).

Nickel doping of boron carbide grown by plasma enhanced chemical vapor deposition

Seong-Don Hwang, N. B. Remmes, and P. A. Dowben[a)]
Center for Materials Research and Analysis, Department of Physics, Behlen Laboratory of Physics, University of Nebraska, Lincoln, Nebraska 68588-0111

D. N. McIlroy
Department of Physics, University of Idaho, Moscow, Idaho 83844-0903

(Received 22 January 1996; accepted 15 March 1996)

We have nickel doped boron carbide grown by plasma enhanced chemical vapor deposition. The source gas closo-1,2-dicarbadodecaborane (ortho-carborane) was used to grow the boron carbide, while nickelocene [$Ni(C_5H_5)_2$] was used to introduce nickel into the growing film. The doping of nickel transformed a *p*-type, B_5C material, relative to lightly doped *n*-type silicon, to an *n*-type material. Both *n-n* heterojunction diodes and *n-p* heterojunction diodes were constructed, using as substrates *n*- and *p*-type Si(111), respectively. With sufficient partial pressures of nickelocene in the plasma reactor, diodes with characteristic tunnel diode behavior can be successfully fabricated.
© *1996 American Vacuum Society.*

I. INTRODUCTION

Through the decomposition of cluster borane molecules by plasma-enhanced[1–4] and synchrotron radiation-induced[4,5] chemical vapor deposition, heterojunction devices of boron and boron carbide have been successfully fabricated. These techniques have not only been used to construct simple diodes, but have also successfully yielded field effect transistors.[3] Until now, no attempts have been made to intentionally dope the intrinsic films of boron carbide using the technique of plasma enhanced chemical vapor deposition (PECVD).

The introduction of metal dopants into boron carbide is not a trivial process. Initial attempts to dope molecular films of the precursor cluster molecule closo-1,2-dicarbadodecaborane ($C_2B_{10}H_{12}$), otherwise known as orthocarborane, with the common dopant mercury were unsuccessful.[6] Mercury was found to segregate to the molecular film-substrate interface and is indicative of the weak interaction between Hg and orthocarborane. While initially disappointing, subsequent attempts to dope molecular films of orthocarborane did prove to be successful.[6,7] Such molecular films can be doped with sodium.[6,7] The possibility of doping boron carbide with Hg cannot be excluded based on these results. This work, nonetheless, did suggest that doping of this material may be a complex process. This is particularly true since the suitability of orthocarborane ($C_2B_{10}H_{12}$) for the chemical vapor deposition of a B_5C films has been established[3–5] and because this molecule is very similar to the "building block" of boron carbide.

Nickel, however, is a very promising dopant for the boron rich solids. A molecular nickel carborane complex has been synthesized by inorganic chemists[8] and the inclusion of nickel in other boron rich solids is well established. Nickel is a common component in the boron carbide superconductors[9] and the reactions of nickel with boron phosphide have been investigated.[10–11] Nickelocene, $Ni(C_5H_5)_2$, has been shown to be a suitable source compound for the deposition of nickel containing thin films.[12] Nickelocene is volatile and far less toxic than nickel-tetracarbonyl, though a number of other nickel containing organometallic compounds may be suitable.[13] Since both orthocarborane and nickelocene are easily sublimed from the solid, introduction of suitable mixtures into the plasma reactor can be readily accomplished. In this paper we present the device characteristics of diodes constructed with nickel doped boron carbide grown by PECVD. To the best of our knowledge, this is the first successful demonstration of metal doping of boron carbide films grown by PECVD. Both high and low levels of Ni doping of boron carbide have been investigated. The diode devices have markedly different behaviors depending upon the nickel doping.

II. EXPERIMENT

The *p-n* and *n-p* heterojunctions were formed by depositing boron carbide thin films on *n*-type and *p*-type Si(111) substrates, respectively, following procedures described in detail elsewhere.[1,2] Deposition of films was performed in a custom designed parallel plate 13.56 MHz radio-frequency PECVD reactor used in previous studies.[1,2] The silicon substrates were doped to $7\times10^{14}/cm^3$. The Si(111) substrates surfaces were prepared by Ar^+ ion sputtering in the plasma reactor. The source molecule gas closo-1,2-dicarbadodecaborane (*ortho*-carborane) was used as the source compound for growing the boron carbide, while nickelocene [$Ni(C_5H_5)_2$] was used to introduce nickel into the growing film. As discussed above, nickelocene [$Ni(C_5H_5)_2$] was simultaneously introduced into the plasma reactor with orthocarborane [closo-1,2-dicarbadodecaborane ($C_2B_{10}H_{12}$)].

[a)]Electronic mail: pdowben@unlinfo.unl.edu

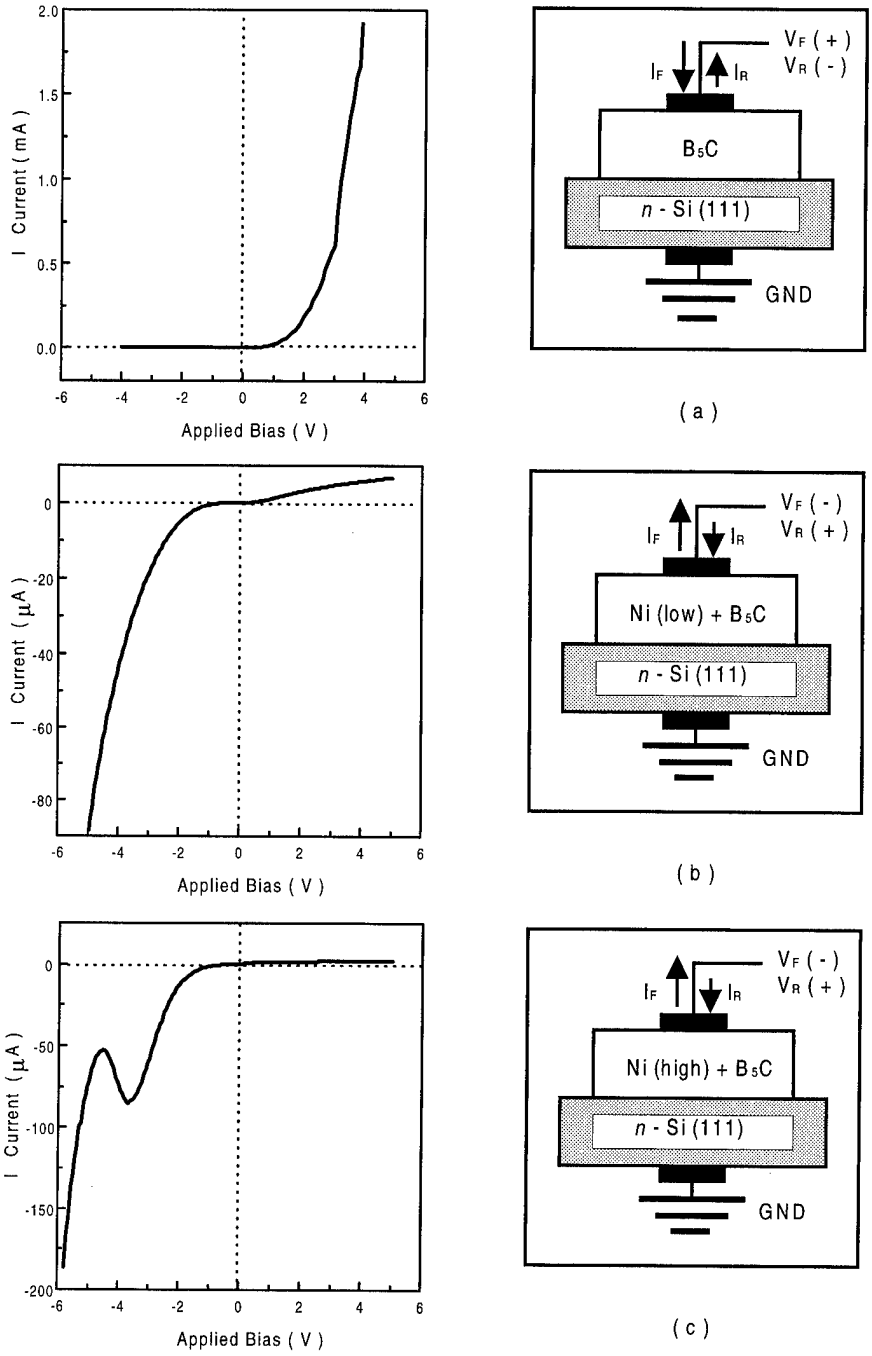


FIG. 1. The $I-V$ characteristics and the corresponding schematic diagrams of the diodes from (a) B_5C/n-Si(111), (b) lightly Ni doped B_5C on n-Si(111), and (c) heavily Ni doped B_5C on n-Si(111). The shift in polarity of the diode demonstrates the relative n-type behavior following nickel doping and the characteristic behavior with sufficiently high levels of nickel doping (see the text).

III. RESULTS AND DISCUSSION

Typical B_5C/n-type silicon heterojunctions have been routinely formed by this technique.[1–4] An example of one such diode device is presented in Fig. 1(a). An examination of the $I-V$ curve in Fig. 1(a) shows the excellent diode characteristics of devices built using the technique of PECVD. These devices typically have onsets of 1 eV with very little leakage current (less than 5 μA at 25 °C). Figure 1(a) also demonstrates the p-type character of undoped PECVD boron carbide in this geometry.

The inclusion of Ni into the boron carbide films with the introduction of nickelocene into the plasma reactor has been verified with Auger electron spectroscopy (AES). The signature of Ni in the Auger spectra suggest that the Ni uptake is large and that these films are highly doped ($\geqslant 1\times10^{21}$). The

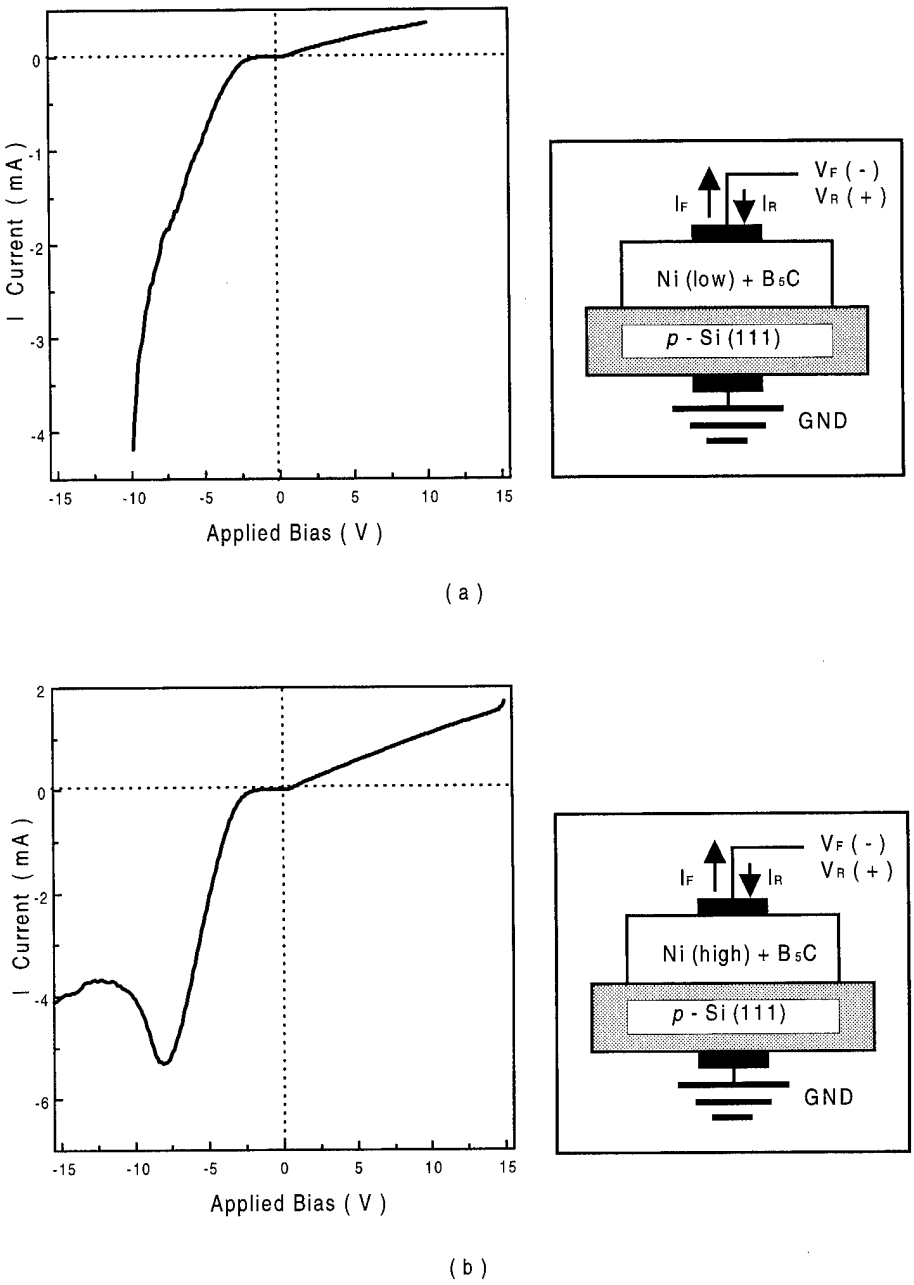


FIG. 2. The $I-V$ characteristics and corresponding schematic diagrams of the diodes from (a) lightly Ni doped B_5C on p-Si(111) and (b) heavily Ni doped B_5C on p-Si(111).

$I-V$ curves of two diodes constructed with Ni-doped boron carbide grown on n-type Si(111) are displayed in panels (b) and (c) in Fig. 1. The diode constructed with a "low" level of Ni doping corresponded to a partial pressure of nickelocene to orthocarborane of <0.1 during film growth. The "high" doping corresponded to a relative partial pressure ratio ≈ 9, respectively.

With the inclusion of Ni, the boron carbide films which are normally p-type relative to n-type silicon, evolve to an n-type material. This is evident from the device characteristics of the diode shown in Fig. 1(b). This results in the formation of rectifying diodes with reverse bias. Thus the nickel doped boron carbide heterojunction diodes appears n-type relative to the lightly doped n-type silicon substrate. This is consistent with the fabrication of n-p heterojunction diodes on p-type silicon, again by including nickelocene with the orthocarborane as an additional source gas, as seen in Fig. 2, for the nickel doped boron carbide.

With the "higher" nickel doping, a negative differential resistance, or a valley in the current, occurs in the effective forward bias direction for diodes formed on both n-type silicon and p-type silicon substrates, as seen in Figs. 1(c) and 2(b). This behavior is characteristic of a tunnel diode[14] and is consistent with degenerative doping of a pinned state relative

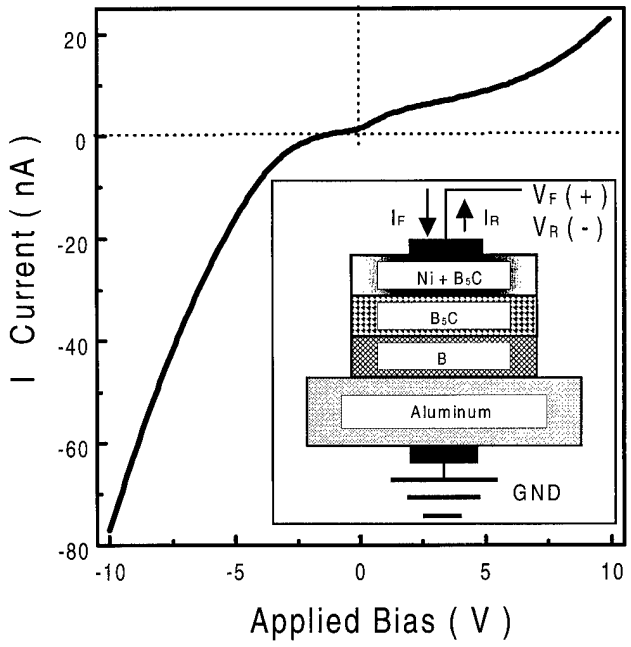


FIG. 3. The $I-V$ characteristics and corresponding schematic diagram of a trilayer diode fabricated from a PECVD Ni doped B_5C layer deposited on a PECVD undoped B_5C layer deposited on a PECVD rhombohedral boron layer placed on an aluminum substrate.

to the conduction band edge. Certainly at the higher doping levels, sufficient nickel from nickelocene is incorporated into the boron carbide during film growth to provide for degenerative doping concentrations.

The hump in the current occurs at a larger bias voltage for the tunnel diodes fabricated on the p-type silicon. This is consistent with an n-type layer [$Ni_x n(B_5C)$] and the formation of heterojunction n-p diodes. In particular, this behavior suggests that the nickel states are pinned to the conduction band edge. States pinned to one band edge have been proposed for boron carbide[2] and have been identified pinned to conduction band edge.[15,16] It may be that such states are occupied or filled through nickel doping. Due to the location of the gap states near the conduction band edge, much larger applied voltages should be needed to observe the negative differential resistance for diodes constructed with p-type Si(111), as compared with diodes constructed with n-type Si(111). This is exactly what is observed from the $I-V$ curves of the n-n and n-p diodes in Figs. 1(c) and 2(b), respectively.

PECVD fabricated B_5C-boron heterojunctions can also form diodes on substrates other than silicon, such as aluminum.[3] Since the construction of diodes through the doping of a layer of B_5C is now potentially possible, we feel that this advance in the doping of PECVD boron carbide holds the prospect that rhombohedral boron rich semiconductor homojunctions can be formed from this material, without the need of a silicon substrate. A preliminary diode of this type has been fabricated by doping a layer B_5C with nickel as seen in Fig. 3.

IV. SUMMARY

In summary, nickelocene can be used to dope PECVD boron carbide. The mildly p-type boron carbide, B_5C, relative to n-type silicon, is strongly n-type following the doping with nickel. With sufficient doping levels, the resulting heterojunction diode acts like a tunnel diode. All evidence suggests that nickel populates states within the gap of this normally highly resistive material.[1,2]

ACKNOWLEDGMENTS

This work was supported by the AFOSR through Grant No. F49620-94-1-0433, the W. M. Keck Center for Molecular Electronics, and the Center for Materials Research and Analysis. One author (N.B.R.) was supported by the NSF-REU program.

[1]S. Lee, J. Mazurowski, G. Ramseyer, and P. A. Dowben, J. Appl. Phys. **72**, 4925 (1992).
[2]S. Lee and P. A. Dowben, Appl. Phys. A **58**, 223 (1994); S. Lee, T. Ton, D. Zych, and P. A. Dowben, Mater. Res. Soc. Symp. Proc. **283**, 483 (1993).
[3]S.-D. Hwang, D. Byun, N. J. Ianno, P. A. Dowben, and H. R. Kim, Appl. Phys. Lett. **68**, 1495 (1996).
[4]D. Byun, B. R. Spady, N. J. Ianno, and P. A. Dowben, Nanostruct. Mater. **5**, 465 (1995).
[5]D. Byun, S.-d. Hwang, P. A. Dowben, F. K. Perkins, F. Filips, and N. J. Ianno, Appl. Phys. Lett. **64**, 1968 (1994).
[6]D. N. McIlroy, Jiandi Zhang, P. A. Dowben, P. Xu, and D. Heskett, Surf. Sci. **328**, 47 (1995).
[7]D. N. McIlroy, J. Zhang, P. A. Dowben, and D. Heskett, Mater. Sci. Eng. (in press).
[8]L. F. Warren and M. F. Hawthorne, J. Am. Chem. Soc. **90**, 4823 (1968); **92**, 1157 (1970); M. F. Hawthorne, D. C. Young, T. D. Andrews, D. V. Howe, R. L. Pillings, A. D. Pitts, M. Reintjes, L. F. Warren, and P. A. Wegner, *ibid.* **90**, 879 (1968); K. P. Callahan and M. F. Hawthorne, Adv. Organomet. Chem. **14**, 145 (1976).
[9]K. Widder, D. Berner, A. Zibold, H. P. Geserich, M. Knupper, M. Kielwein, M. Buchgeister, and J. Fink, Europhys. Lett. **30**, 55 (1995); R. J. Cava, Physica C **226**, 170 (1994); T. Siegrist, H. W. Zandbergen, R. J. Cava, J. J. Krajewski, and W. F. Peck, Nature **367**, 254 (1994); R. J. Cava, Nature **367**, 252 (1994); R. Nagarajan, C. Mazumdar, Z. Hossain, S. K. Dhar, K. V. Gopalakrishnan, L. C. Gupta, C. Godart, B. D. Padalia, and R. Vijayaraghavan, Phys. Rev. Lett. **72**, 274 (1994).
[10]N. Kobayashi, Y. Kumashiro, P. Revesz, J. Li, and J. W. Mayer, Mater. Res. Soc. Symp. Proc. **162**, 595 (1990).
[11]N. Kobayashi, Y. Kumashiro, P. Revesz, J. Li, and J. W. Mayer, Appl. Phys. Lett. **54**, 1914 (1989).
[12]Y.-G. Kim, D. Byun, C. Hutchings, P. A. Dowben, H. Hejase, and K. Schroeder, J. Appl. Phys. **70**, 6062 (1991); G. T. Stauf and P. A. Dowben, Thin Solid Films **156**, L31 (1988); G. T. Stauf, D. C. Driscoll, P. A. Dowben, S. Barfuss, and M. Grade, **153**, 421 (1987).
[13]N. M. Boag and P. A. Dowben, *Metallized Plastics 4: Fundamental and Applied Aspects*, edited by K. L. Mittal (Plenum, New York, 1996); D. Zych, A. Patwa, S. S. Kher, J. T. Spencer, J. Kushneir, A. J. Goodby, N. M. Boag, and P. A. Dowben, J. Appl. Phys. **76**, 3684 (1994).
[14]E. O. Kane, J. Appl. Phys. **32**, 83 (1961); E. O. Kane, Phys. Chem. Solids **2**, 181 (1960); A. G. Chynoweth, W. L. Feldmann, and R. A. Logan, Phys. Rev. **121**, 684 (1961); R. A. Logan and A. G. Chynoweth, *ibid.* **131**, 89 (1963); R. S. Claassen, J. Appl. Phys. **32**, 2372 (1961); A. S. Epstein and J. F. Caldwell, *ibid.* **35**, 3050 (1964); D. Meyerhofer, G. A. Brown, and H. S. Sommers, Jr., Phys. Rev. **126**, 1329 (1962); R. M. Minton and R. Glicksman, Solid-State Electron. **7**, 491 (1964); J. Karlovsky, *ibid.* **10**, 1109 (1967).
[15]H. Werheit, U. Kuhlmann, M. Laux, and T. Lundstrøm, Phys. Status Solidi B **179**, 489 (1993).
[16]H. Werheit, M. Laux, and U. Kuhlmann, Phys. Status Solidi B **176**, 415 (1993).

Evolution of deep levels and internal photoemission with annealing temperature at ZnSe/GaAs interfaces

X. Yang,[a] L. J. Brillson, and A. D. Raisanen
Xerox Wilson Center for Research and Technology, 114-41D, Webster, New York 14580

L. Vanzetti,[b] A. Bonanni, and A. Franciosi[c]
Laboratorio Nazionale Tecnologie Avanzate Superfici e Catalisi dell' Istituto Nazionale per la Fisica della Materia, Area di Ricerca, Padriciano 99, I-34012 Trieste, Italy and Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

M. Grundmann and D. Bimberg
Department of Physics, Technical University of Berlin, Berlin 10623, Germany

(Received 15 March 1996; accepted 7 May 1996)

We have measured charge states deep within the ZnSe band gap and localized near the ZnSe/GaAs interface as a function of annealing temperature by means of photoluminescence spectroscopy, as well as cathodoluminescence from the cross section of a ZnSSe/GaAs heterostructure with a scanning electron microscope. Annealing produces new emissions at 1.9 eV and 2.25 eV which increase with annealing temperature and which vary with different growth conditions. Cathodoluminescence spectra show two similar features which originate within ZnSe and whose relative emission intensities vary with depth. Since the 2.25 eV feature appears only under Zn-rich growth conditions, these results indicate a role of both interface diffusion and crystal growth in the formation of this defect. We observe a proportional increase of the 1.9 eV and 0.9 eV peak intensities with annealing, indicating their complementary nature with respect to a common deep level. Finally, a detailed analysis of the relative intensities of ZnSe and GaAs features due to internal photoemission across the heterointerface suggests that the heterojunction band offset changes with formation of these new deep levels. All these results emphasize that multiple deep levels form near the buried interface with annealing which can dominate the heterojunction electronic properties.
© *1996 American Vacuum Society.*

I. INTRODUCTION

The ZnSe/GaAs heterojunction is a prototypical system for studying II-VI/III-V compound semiconductor interfaces due to its small lattice mismatch (0.27%), which permits high quality epitaxial growth.[1] An understanding of its electronic and chemical properties is valuable for fundamental research and technological applications. Specifically, modification of the corresponding band offsets could have a substantial impact on the performance of II-VI based blue-green laser, and several authors have argued that the band alignment across the junction may depend on the local atomic structure of the interface,[2–6] which, in turn, affects free carrier transport across heterojunction interfaces and their device applications. Our focus is on the presence of charge states localized near the heterojunction interface, their change with annealing, and their effect on carrier transfer across the junction.[1,7,8]

Nicolini *et al.*[9] have demonstrated a connection between atomic interface structure and band alignment. They showed that variations in the beam pressure ratio (BPR) of Zn to Se during the first a few monolayers growth of ZnSe epilayer on GaAs in a molecular beam epitaxy (MBE) system could be used to modify the ZnSe relative concentration at the interface. As a result, the local atomic bonding alters ZnSe/GaAs heterojunction band offsets over a wide energy range from below 0.6 eV to above 1.2 eV.[10,11] Their photoemission measurements showed that those specimens with BPR>1 (Zn-rich) have large valence band offsets between GaAs and ZnSe, while specimens with BPR<1 (Se-rich) exhibit valence band offset values more than 0.6–0.8 eV lower.[12,13]

Deep levels also appear to play a major role in determining the behavior of ZnSe/GaAs interfaces. We have extended the low energy cathodoluminescence and photoluminescence spectroscopy techniques developed earlier for studying metal–semiconductor interfaces[14,15] to heterojunction interfaces. These techniques have revealed optical emissions related to dislocation-induced defects deep within the band gap,[15] injection and recombination of free carriers across heterointerfaces, and deep level luminescence near the ZnSe/GaAs heterointerface.[16–18] Recently, we have measured near band edge (NBE) emissions of GaAs and ZnSe as a function of the Zn to Se BPR used to grow the ZnSe epilayers.[19] We found systematic variations in the NBE relative intensity with BPR due to internal photoemission, which are consistent with the band offset variations reported earlier.[9]

We have now performed photoluminescence spectroscopy studies of the ZnSe/GaAs heterojunction interfaces by using HeCd laser excitation as a function of annealing temperature.

[a]Electronic mail: xinxing_yang@wb.xerox.com
[b]Present address: Centro Materiali Biofisica Medica, I-38050 Povo, Trento, Italy.
[c]Also at Dipartimento di Fisica, Universita' di Trieste, I-34014 Trieste, Italy.

With post growth annealing, new deep level emissions appear which depend on initial growth conditions. Our results provide additional evidence of internal photoemission across the interface from GaAs to ZnSe. We identify complementary transitions due to formation of new states deep within the ZnSe band gap. We also find changes in internal photoemission and heterojunction band offset with annealing processes which correlate with deep level formation.

II. EXPERIMENTS

The specimens in our study were ZnSe/GaAs heterostructures which were grown by MBE. The MBE system included interconnected growth chambers and an analysis chamber with monochromatic x-ray photoemission spectroscopy (XPS) capabilities. The ZnSe/GaAs growth has been described previously.[9] The Zn to Se BPR was monitored by means of a nude ion gauge, and the corresponding near-interface composition was determined by means of XPS using the integrated emission intensity of Zn and Se core levels as a function of ZnSe thickness for the first few monolayers of ZnSe growth. In this article, the specific specimen growth from substrate to the top layer includes a GaAs (100) wafer (Si-doped, $n = 7-9 \times 10^{17}$ cm^{-3}, a GaAs buffer layer (0.5 μm thick, Si-doped, $n = 10^{17}$ cm^{-3}) grown at 580 °C, a ZnSe epilayer (0.15 μm thick, nominally undoped) grown at 290 °C, and an amorphous Se cap layer (0.5–1 μm thick) grown at 25 °C. The latter protects the epitaxial ZnSe film from ambient contamination during transfer in air from growth chamber to other experimental chambers. We have used specimens with various BPR, i.e., BPR=0.1, 0.2, 0.5, 1, 2, 10. Overlayers grown with various BPR exhibited different reflection high energy electron diffraction (RHEED) patterns,[16,17] which depended upon the initial growth conditions and which were consistent with the literature.[20]

We performed luminescence experiments under ultrahigh vacuum (UHV) conditions in order to minimize any effects of surface contamination or surface recombination and to permit low temperature (at 90 K) photoluminescence measurements. The optical excitation source for the emission spectra was a HeCd laser with photon energy 2.808 eV (λ=4416 Å). Our photodetectors were an S-1 photomultiplier and a cooled North Coast Ge photodiode. A CaF_2 lens collects the emitted light and focuses it through a sapphire viewport into a Leiss prism monochromator with a spectral resolution better than 0.05 eV. Both CaF_2 and sapphire are transparent in our measuring energy range. We annealed ZnSe/GaAs heterojunction specimens with different Zn-to-Se BPR in 5 min annealing cycles at increasing temperatures from 250 °C to as high as 500 °C. After each anneal, the specimens were cooled to 90 K for each annealing stage of the photoluminescence measurement.[21] Using low energy electron diffraction (LEED) to detect the surface reconstruction, we found that the temperature for completely desorbing the Se cap layer was between 200 and 290 °C. All specimens displayed a predominant (2×1) mixed with a $c(2 \times 2)$ LEED pattern with the BPR=10 specimen showing the sharpest pattern.

We have also used a scanning electron microscope (SEM) operating at 3 keV with a nominal electron beam diameter of 10–20 nm to produce cross-sectional luminescence spectra from a ZnSe/ZnSeS/GaAs superlattice at 5 K. This particular specimen was grown by metal-organic vapor phase epitaxy (MOVPE) at 340 °C and consists of a GaAs (001) wafer as a substrate, 120 periods of ZnSe layer (10 nm thick) and ZnSeS layer (10 nm thick), and a 10 nm thick ZnSe cap layer. It was subsequently annealed for 2.5 h at 580 °C. The SEM measurements were made with the electron beam either near the GaAs substrate or close to the ZnSe/ZnSeS free surface. In both cases, the signals are averages from multiple ZnSe/ZnSeS layers since the beam exceeds the superlattice period, even without blooming.

III. RESULTS

Figures 1 and 2 display photoluminescence spectra of ZnSe/GaAs heterojunctions with HeCd laser excitation. Figure 1 shows a set of photoluminescence spectra measured at 90 K from three ZnSe/GaAs heterojunction interfaces with various BPR as a function of annealing temperature with an S-1 photomultiplier. The S-1 photomultiplier response covered the energy range of both ZnSe NBE (2.79 eV) and GaAs NBE (1.5 eV). Figure 1(a) shows spectra for a specimen with BPR=1 (i.e., the beam pressure of Zn and Se are balanced). There are four major features (1.3, 1.5, 2.79 eV, and the laser line at 2.808 eV) in these spectra at temperatures lower than 350 °C, which is the thermal threshold for this ZnSe/GaAs heterostructure (the temperature at which the new deep level emission starts to increase, is called the thermal threshold). These low temperature spectra suggest that some ZnSe defects exist near the heterojunction even before annealing. The ZnSe emission features have relatively high intensities for the HeCd laser excitation employed here (see Sec. IV). We also observed two new deep levels (1.9 eV and 2.25 eV) form in the ZnSe band gap for annealing temperatures 350 °C and above. Their emission intensities increase with increasing annealing temperature, dominating the spectra at the highest temperature investigated.

Figure 1(b) presents analogous spectra for a specimen with BPR=10. These spectra exhibit deep level emission and NBE features similar to those in Fig. 1(a). However, the intensity of the new feature at 2.25 eV is more pronounced relative to the GaAs NBE. Figure 1(c) shows the spectra for a specimen with BPR=0.1 (i.e., Se-rich growth conditions). Note that only the 1.9 eV feature is evident in these spectra with increasing annealing temperature. The feature at 2.25 eV is undetectable. We also measured other specimens with BPR=0.2, and 0.5 (not shown). All samples exhibited similar spectral features, with relative intensities that depended strongly on the BPR employed during ZnSe growth. Specifically the feature at 2.25 eV was detected and found to increase with high temperature annealing only in samples grown in Zn-rich conditions, indicating a role of both interface diffusion and crystal growth in formation of this defect.

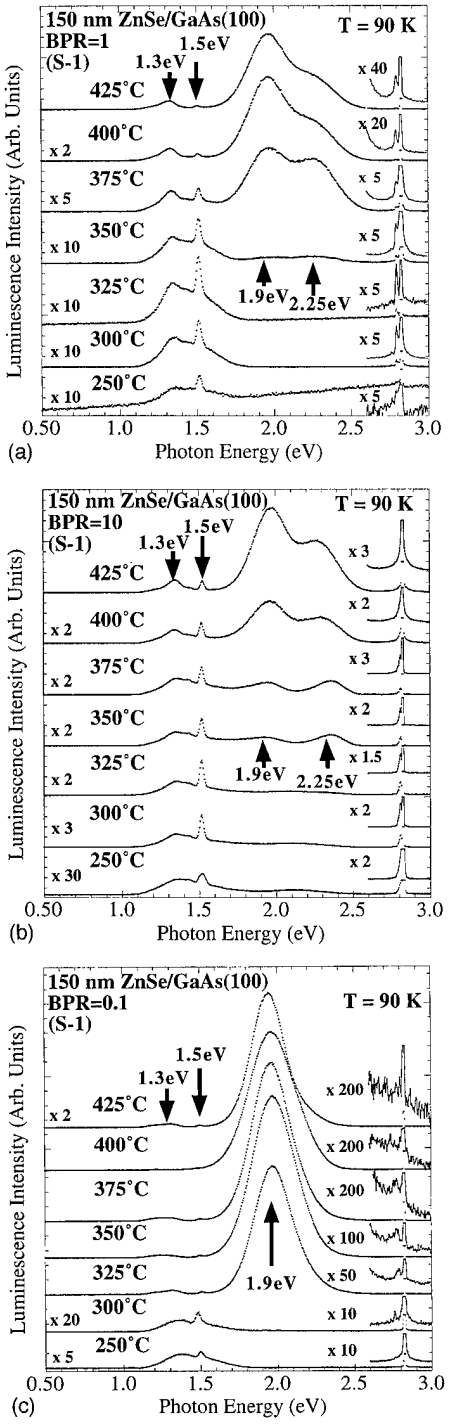


FIG. 1. Deep level emission spectra for 150 nm ZnSe/GaAs heterojunction structure with a HeCd laser excitation and S-1 photomultiplier detection as a function of annealing temperature for specimens with (a) BPR=1, (b) BPR=10, and (c) BPR=0.1. Pronounced new deep level emissions appear at higher annealing temperatures.

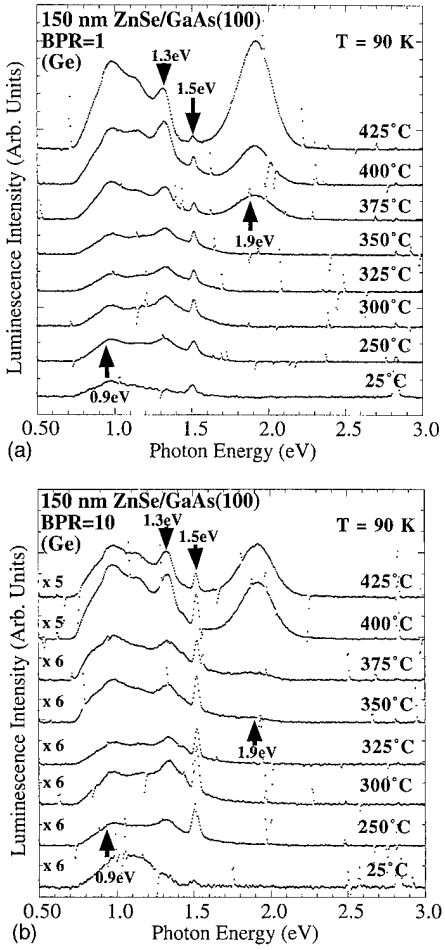


FIG. 2. Deep level emission spectra for the same specimens as in Fig. 1 [(a) BPR=1, (b) BPR=10] with a HeCd laser excitation and Ge photodiode detection as a function of annealing temperature. The intensities of both features at 0.9 eV and 1.9 eV increase with annealing temperature.

We also acquired lower energy photoluminescence spectra with a Ge photodiode to complement the S-1 spectra. The S-1 photomultiplier is sensitive in the range of 1.1 eV to 4 eV, while the Ge photodiode responds within a narrower and lower energy range of spectral sensitivity, i.e., 0.7 eV to ~2.0 eV. The response of this Ge detector is heavily weighted toward the near infrared, which allows us to detect emissions from states deep within the semiconductor band gaps. Figure 2 shows a set of photoluminescence spectra measured at 90 K by the Ge photodiode from the ZnSe/GaAs heterojunction interfaces with various BPR as a function of annealing temperature. We observed multiple deep level transitions with energies less than the band gap energy of GaAs (1.5 eV) as shown in Fig. 2. The spectral features resemble those of similar specimens reported previously.[18]

Evidence of new defect creation appears at the annealing temperature of 350 °C and above, but the feature at 2.25 eV shown in the S-1 spectra of Fig. 1 did not appear in these spectra because of the much lower response of the Ge photodiode at energies above 2 eV. Figure 2(a) shows spectra for the same BPR=1 specimen measured in Fig. 1(a) but with Ge photodiode detection. Among low energy features, the emission intensity of a peak at 0.9 eV grows with annealing temperature and its intensity increases proportionally with that of the 1.9 eV peak after completely desorbing the Se cap. These intensities increase 2–7 times for 0.9 eV peak, and 5–10 times for 1.9 eV peak, from the threshold temperature to 425 °C. Figure 2(b) shows spectra for the same

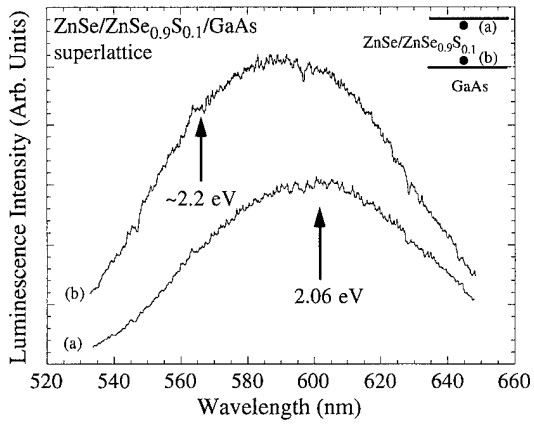


FIG. 3. The cathodoluminescence spectra from ZnSSe/GaAs structures in cross section using an SEM. The top spectrum corresponds to a spot near the buried $ZnSe_{0.9}S_{0.1}$/GaAs interface (the lower spot in the inset). The bottom spectrum corresponds to a spot near the free ZnSe surface (the upper spot in the inset). The higher energy peak component appears to increase near the GaAs interface.

BPR=10 specimen measured in Fig. 1(b). The spectral features are similar to those shown in Fig. 2(a), but the intensity of the 1.9 eV feature relative to the GaAs NBE emission increases less than that of the specimen with BPR=1. In general, the emission intensities of the lower energy peaks relative to the 1.9 eV peak are larger for Zn-rich growth conditions, consistent with previous results.

In order to characterize further the nature of the 1.9 eV and 2.25 eV features, we obtained luminescence spectra from similar heterojunctions in cross section using an SEM at 10 K. Figure 3 shows cathodoluminescence from a $ZnSe/ZnSe_{0.9}S_{0.1}$/GaAs superlattice at two spatial locations relative to the interface (see inset). Two similar features originate within the ZnSeS, whose relative emission intensities vary with depth. The bottom spectrum is measured near the free surface (upper spot in the inset), while the top spectrum corresponds to a position near the buried $ZnSe_{0.9}S_{0.1}$/GaAs interface (lower spot in the inset). The SEM-induced luminescence from the GaAs substrate alone exhibits no such emission in the energy range from 1.8 to 2.4 eV. While the spectral features are largely unresolved, the differences between the two spectra indicate at least two peaks. Subtraction of one spectrum from the other minimizes the background and indicates peak positions at roughly 2.06 eV and 2.16 eV. The relative intensity of the 2.16 eV feature appears to increase near the GaAs interface.

IV. DISCUSSION

Previous work of deep level formation near the ZnSe/GaAs interface with post growth annealing have shown a dramatic dependence on initial epitaxial growth conditions.[21] We now find that annealing produces new pronounced emission features at several energies. Most prominent are features at both 1.9 eV and 2.25 eV, whose intensities increase rapidly with annealing temperature and which differ for structures fabricated with Se- versus Zn-rich growth. The data in Figs. 1–3 shed light on the nature of the thermally-induced deep levels. The spectra in Fig. 1 show that the 2.25 eV peak is most intense for Zn-rich (BPR=10) conditions. In general, the 2.25 eV feature only appears under the BPR-balanced or Zn-rich growth conditions with increasing annealing temperature. Overall, previous studies provide no well-accepted identification of the 2.25 eV luminescence peak feature. Originally, this feature was termed "Cu-green" luminescence,[22] but subsequent studies[23] provide evidence for such emission in the absence of Cu impurities. Similarly, our results show the variation of this emission depends on the balance of Zn to Se during epitaxial overlayer growth. Alternatively, the emission may arise from a Ga-Se interfacial complex. Zahn *et al.*[24] have reported various luminescence peaks in the 2–2.5 eV range for different Ga-Se bulk material. Likewise, Skromme *et al.*[25] found that a similar 2.27 eV peak is stronger when the Ga is deposited in a sheet on a Se-rich ZnSe surface than when bulk- or sheet-doped on a Zn-rich surface. Such Ga-Se compound emission is counter to our observations, since Zn-rich interfacial conditions are likely to suppress any Ga-Se bonding.

More is known about the 1.9 eV peak emission, which is the most intense feature in all of our specimens at elevated annealing temperature. This feature can be associated with self-activated emission due to Zn vacancies (V_{Zn}) or a complex defect center involving a Zn vacancy and a substitutional Ga atom on a Zn site[21,23,26] (i.e., Ga_{Zn}-V_{Zn}). The nature of this defect is consistent with the emission increase with increasing annealing temperature: as described previously,[18,21] Ga outdiffusion occurs at elevated temperatures from GaAs to ZnSe, resulting in formation of defect complexes such as Ga_{Zn}, especially in samples fabricated under Se-rich growth conditions.

The SEM cathodoluminescence spectra from the $ZnSe_{0.9}S_{0.1}$/GaAs cross section in Fig. 3 are consistent with our ZnSe/GaAs photoluminescence results. The peak energies of ~2.06 eV and ~2.16 eV for this $ZnSe_{0.9}S_{0.1}$/GaAs heterojunction appear related to the 1.9 eV ZnSe feature but shifted 0.16 and 0.26 eV, respectively to higher energy. The lower temperature can account for ~0.1 eV of this shift of both peaks.[23] The ternary alloy's larger band gap can account for the 0.1 eV difference in energy of the two unresolved peaks. Thus the two peaks in $ZnSe/ZnSe_{0.9}S_{0.1}$/GaAs and the 1.9 eV peak in ZnSe/GaAs display a good correspondence. The broad features indicate at least two peaks, whose intensities appear to vary with depth. The relative variation of the component intensities within the broad feature certainly indicates that more than one deep level is involved. Whether these variations are due to defect generation proceeding from the buried interface or from the free ZnSe surface remains to be established. Recombination and defect states at the cross sectional surface may well affect both spectra. However, the primary effect appears to be the proximity of the GaAs layer. The absence of such emission within the GaAs alone demonstrates the II-VI origin of these emissions. This latter result also demonstrates that electrons in the GaAs do not diffuse laterally near the cross sectional surface and recombine hun-

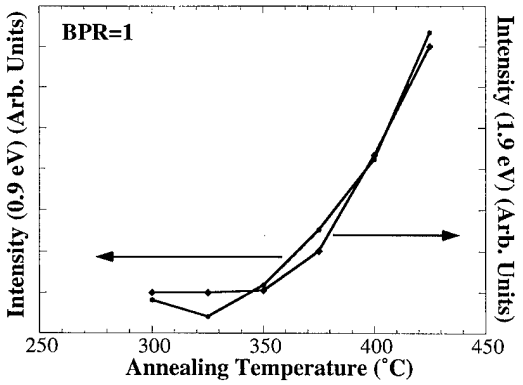


FIG. 4. Intensity increase deep level emissions at 0.9 eV and 1.9 eV as a function of annealing temperature for 150 nm thick ZnSe/GaAs structures (BPR=1). The two peaks exhibit a threshold for defect creation and increase proportionately.

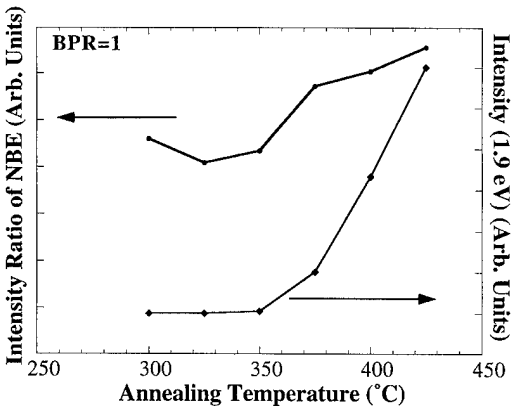


FIG. 5. Intensity ratio of ZnSe to GaAs NBE emissions and intensity of new deep level emission at 1.9 eV vs annealing temperature (BPR=1). The bottom curve represents the intensity of the 1.9 eV feature which increases with annealing. At the threshold temperature, when the new deep level emission starts to increase, the ZnSe/GaAs NBE emission intensity ratio also shows a pronounced change.

dreds of nanometers away in the neighboring ZnSe. This result in fact emphasizes the importance of near-interface excitation for gauging internal photoemission.

The coupled growth of deep level features with annealing underscores the growth of deep rather than shallow level densities with annealing. From Fig. 1 and Fig. 2, we observed that the intensities of both 0.9 eV and 1.9 eV features increase proportionally with annealing temperature. We extracted their intensities from the spectra by using Gaussian distributions to deconvolve the multiple peak structures. Figure 4 shows the intensities of 0.9 eV and 1.9 eV peaks versus annealing temperature for the specimen with BPR=1. Both emission intensities are low and constant below the annealing threshold, but increase rapidly above the threshold temperature. Similar behavior appears in specimens grown at all BPR. This result is consistent with our previous data, which also shows these two features increasing together with annealing.[18] Furthermore, the 0.9 eV and 1.9 eV peaks energies add up to the ZnSe band gap, indicating that these two transitions are complementary, i.e., the emissions correspond to transitions into and out of the same deep level in the band gap. This supports the contention that the increase in the intensity of the 0.9 and 1.9 eV features upon annealing reflects a variation in the spatial concentration of the same deep level, as opposed to unrelated shallow levels.

We have also used the relative intensities of ZnSe and GaAs NBE emissions in Fig. 1 to infer changes in the ZnSe/GaAs heterojunction band offset with annealing. Previously, we showed that the unusually large ZnSe recombination with HeCd excitation at 90 K was the result of free carrier injection across the interface due to internal photoemission.[19] If we take into account the relative absorption of HeCd radiation of ZnSe versus GaAs and the relative response of the S-1 detector at their NBE emission energies,[19] the intensity of the ZnSe NBE (2.79 eV) emission should be 300–600 times smaller (depending on the precise temperature) than the GaAs NBE (1.5 eV) emission in Fig. 1(a) for free carrier recombination without diffusion. Internal photoemission can explain the large intensity of the ZnSe NBE. In such a process, free carriers are photoexcited in GaAs near the metallurgical interface at energies below the ZnSe band gap, and cross the junction to recombine in the ZnSe. We have already shown[19] that specimens known to have different band offsets displayed strong changes in the ratio of ZnSe to GaAs NBE emission, indicating the sensitivity of this internal photoemission to the interfacial barriers.

Annealing causes additional changes in internal photoemission. Figure 5 shows a comparison of the intensity ratio of ZnSe to GaAs NBE emissions (top curve) and the new deep level emission at 1.9 eV (bottom curve) as a function of annealing temperature. At the threshold temperature, the ZnSe/GaAs NBE emission intensity ratio also exhibits a pronounced change. Specimens grown under different BPR conditions also show pronounced changes at the threshold temperature for 1.9 eV defect creation. Therefore, this pronounced change in the relative intensities of NBE emissions at the defect threshold temperature suggests that deep level creation affects the heterojunction band offset. When new deep levels start to form the valence band offset appears to increase, so that there is a lower barrier to movement of free electrons across the interface from GaAs to ZnSe. Thus, the intensity of internal photoemission increases, as evidenced by free carrier recombination not only across the ZnSe band gap, but also into the ZnSe deep levels. These results are a strong indication that near-interface defects and epitaxial growth conditions are critical factors in forming the heterojunction band offsets.

Besides internal photoemission, other processes may affect the ZnSe/GaAs emission ratio. For several of the specimens investigated, including BPR=1, the increase in ZnSe/GaAs NBE ratio is due primarily to a decrease in GaAs NBE intensity (not shown). Such a decrease could be due to: (a) the transfer of photoexcited carriers out of GaAs into ZnSe (i.e., internal photoemission), (b) an increase in sub-band gap recombination within the GaAs (i.e., "non-radiative" recom-

bination), or (c) a change in GaAs band bending which increases the "dead" layer within the depletion width. (This "dead" layer separates free electrons and holes and reduces NBE recombination near the GaAs interface.)

No increase in GaAs deep level recombination [case (b)] is evident in the near-IR spectra of Figs. 2(a) and 2(b), at least for energies above 0.7 eV. A change of band bending [case (c)] due to creation of new near-interface states could indeed affect GaAs NBE since the HeCd absorption depth of 50 nm in GaAs is within the buffer layer's depletion width (ca. 140 nm for 10^{17} cm^{-3} doping and 0.7 eV nominal band bending). "Dead" layer effects may therefore account for part of the ZnSe/GaAs NBE changes with defect formation. However, previous photoluminescence measurements of similar ZnSe/GaAs specimens comparing HeCd versus HeNe spectra show only a factor of 2 increase for excitation well within versus well beyond (HeNe absorption depth of 200 nm) the depletion region.[17] Therefore these effects do not account for the orders-of-magnitude enhancement of ZnSe NBE emission relative to GaAs before annealing based on absorption coefficients. Further investigation of possible band bending effects on ZnSe/GaAs intensity ratios with annealing is in progress. In any case, the thermally induced diffusion and bonding which produces new electronic states leads to changes in the electric potential gradients at the ZnSe/GaAs heterojunction on a microscopic scale.

V. CONCLUSIONS

We have performed photoluminescence spectroscopy studies of ZnSe/GaAs heterojunctions in a wide energy range by means of HeCd laser excitation and complementary photodetection by two detectors with different response sensitivities. Using this technique, we have observed new deep level emissions which appear within the ZnSe energy band gap (1.9 eV and 2.25 eV) with post growth annealing. The thermally induced defect formation depends on specimen growth conditions. Annealing permitted us to identify the complementary nature of the deep level emissions at 0.9 eV and 1.9 eV into and out of the same mid-gap state. The dramatic variations in internal photoemission with thermally-induced defect formation emphasize that annealing processes lead to the creation of multiple deep levels near the buried interface, and that such deep level formation can affect heterojunction band offsets.

ACKNOWLEDGMENTS

Several authors (L. A., A. B., and A. F.) acknowledge support by the U. S. Army Research Office under Grant No. DAAL03H04-93-G-0319 and Grant No. DAAH04-93-G0206, and by the Center for Interfacial Engineering of the University of Minnesota under National Science Foundation Grant No. CDR8721551.

[1]L. J. Brillson, in *Handbook on Semiconductors*, edited by P.T. Landsberg (North-Holland, Amsterdam, 1992), Vol. 1, p. 281.
[2]J. McKinley and G. Margaritondo, in *Contacts to Semiconductors: Fundamentals and Technology,* edited by L. J. Brillson (Noyes, Park Ridge, NJ, 1993), p. 600.
[3]W. Harrison, J. Vac. Sci. Technol. **16**, 1492 (1979).
[4]K. Kunc and R. M. Martin, Phys. Rev. B **24**, 3445 (1981).
[5]R. G. Dandrea, S. Froyen, and A. Zunger, Phys. Rev. B **42**, 3213 (1990).
[6]A. Baldereschi, R. Resta, M. Peressi, S. Baroni, and K. Mader, in *Proceedings of the NATO Advanced Research Workshop Physical Properties of Semiconductor Interfaces at Sub-Nanometer Scale,* edited by H. W. Salemink (Kluwer, Dordrecht, 1993).
[7]*Heterojunction Band Discontinuities: Physics and Device Applications,* edited by F. Capasso and G. Margaritondo (North Holland, Amsterdam, 1987).
[8]A. Franciosi and C. G. Van de Walle, Surf. Sci. Repts. **214**, 1 (1996).
[9]R. Nicolini, L. Vanzetti, G. Bratina, L. Sorba, and A. Franciosi, Phys. Rev. Lett. **72**, 294 (1994).
[10]L. Vanzetti, A. Bonanni, G. Bratina, L. Sorba, A. Franciosi, M. Lomascolo, D. Greco, and R. Cingolani, J. Cryst. Growth **150**, 765 (1995).
[11]A. Bonanni, L. Vanzetti, L. Sorba, A. Franciosi, M. Lomascolo, P. Prete, and R. Cingolani, Appl. Phys. Lett. **66**, 1092 (1995).
[12]R. Nicolini, L. Vanzetti, G. Mula, G. Bratina, L. Sorba, A. Mura, J. E. Angelo, W. W. Gerberich, and A. Franciosi, Mater. Res. Soc. Symp. Proc. **326**, 3 (1994).
[13]A. Franciosi, L. Vanzetti, A. Bonanni, L. Sorba, G. Bratina, and G. Biasol, Proc. SPIE **2346**, 100 (1994).
[14]L. J. Brillson, R. E. Viturro, J. L. Shaw, and H. W. Richter, J. Vac. Sci. Technol. A **6**, 487 (1988).
[15]A. Raisanen, L. J. Brillson, R. S. Goldman, K. L. Kavanagh, and H. H. Wieder, Appl. Phys. Lett. **64**, 3572 (1994).
[16]A. Raisanen, L. J. Brillson, A. Franciosi, R. Nicolini, L. Vanzetti, and L. Sorba, J. Electron. Mater. **24**, 163 (1995).
[17]A. Raisanen, L. J. Brillson, L. Vanzetti, L. Sorba, and A. Franciosi, J. Vac. Sci. Technol. A **13**, 690 (1995).
[18]A. Raisanen, L. J. Brillson, L. Vanzetti, A. Bonanni, and A. Franciosi, Appl. Phys. Lett. **66**, 3301 (1995).
[19]X. Yang, L. J. Brillson, A. D. Raisanen, L. Vanzetti, A. Bonanni, and A. Franciosi, J. Vac. Sci. Technol. A **14**, 867 (1996).
[20]M. C. Tamargo, J. L. de Miguel, D. M. Hwang, and H. H. Ferrell, J. Vac. Sci. Technol. B **6**, 784 (1988).
[21]A. Raisanen, L. J. Brillson, L. Vanzetti, A. Bonanni, and A. Franciosi, J. Vac. Sci. Technol. B **13**, 1705 (1995).
[22]Y. Shirakawa and H. Kukimoto, J. Appl. Phys. **51**, 2014 (1980).
[23]J. Gutowski, N. Presser, and G. Kudlek, Phys. Status Solidi A **120**, 11 (1990).
[24]D. R. T. Zahn, A. Krost, M. Kolodziejczyk, and W. Richter, J. Vac. Sci. Technol B **10**, 2077 (1992) and unpublished.
[25]B. J. Skromme, S. M. Shibli, J. L. de Miguel, and M. C. Tamargo, J. Appl. Phys. **65**, 3999 (1989).
[26]M. Yamaguchi, A. Yamamoto, and M. Kondo, J. Appl. Phys. **48**, 5237 (1977).

Interfacial chemical bonds, reactions, and band alignment in ZnSe/GaAs(001) heterojunctions

G. Bratina,[a] T. Ozzello,[b] and A. Franciosi[c]
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455 and Laboratorio TASC-INFM, Area di Ricerca, Padriciano 99, I-34012 Trieste, Italy

(Received 21 January 1996; accepted 30 April 1996)

ZnSe/GaAs(001) interfaces were fabricated through molecular beam epitaxy by predosing the GaAs(001)2×4 surface with ordered monolayers of Se or Zn, prior to ZnSe growth in stoichiometric growth conditions. We performed the predosing at room temperature (RT) and achieved ZnSe epitaxial overgrowth at RT to minimize desorption of the adsorbed species and minimize interdiffusion. *In situ* synchrotron radiation photoemission studies indicate that Se-predosing results in valence band offsets as low as 0.40±0.05 eV, and that Zn-predosing yields valence band offsets as high as 1.01±0.05 eV. © *1996 American Vacuum Society.*

I. INTRODUCTION

The possibility of controlling the band discontinuities at semiconductor heterojunctions by modifying the local interface environment has attracted much attention in the past ten years.[1–3] Methods to change the band offsets would implicitly clarify the microscopic mechanism that determine the band alignment, and would have a substantial technological impact. Semiconductor heterojunctions are extensively employed in modern optoelectronic and high speed technology, and tuning the band offset to a desired application could optimize carrier confinement or injection in devices ranging from lasers to heterojunction bipolar transistors.

Recent theoretical developments have identified heterovalent heterojunctions with polar orientation as the most promising candidates for offset tuning through modification of the local interface environment.[4–7] In particular, linear response theory (LRT) results have shown that while in isovalent heterojunction the band offsets for abrupt interfaces should be a property of the bulk semiconductor constituents, in heterovalent heterojunctions with polar orientation, the band alignment should depend on the detail of the atomic structure of the interface.[5] Furthermore, ideally abrupt heterovalent heterojunctions with polar orientation would be charged, so that for these interfaces there is a substantial thermodynamic driving force toward the formation of mixed local atomic configurations to yield neutral interfaces.

Attempts to exploit molecular beam epitaxy (MBE) to achieve different local atomic configurations and different band offsets have met with success in the case of the ZnSe/GaAs(001) heterojunction.[8,9] This prototypical, lattice-matched (0.27% in-plane mismatch) heterovalent heterojunction is a crucial component of most blue-green solid state lasers demonstrated to date,[10–13] and changes in the corresponding band offsets could have an important impact on device performance.

The growth parameter successfully employed to modify the ZnSe/GaAs band offset has been the Zn/Se beam pressure ratio (BPR) employed during the early stages of ZnSe MBE on GaAs surfaces.[8,9] Heterojunctions grown with low BPR (BPR~0.1, Se-rich growth conditions) have shown relatively low valence band offsets ($\Delta E_v \sim 0.6$ eV) as compared to junctions grown with high BPR (BPR~10, $\Delta E_v \sim 1.2$ eV), in samples probed *in situ* by photoemission spectroscopy,[8,9] and *ex situ* by optical techniques,[14] and in fully functional heterojunction diodes examined by low-temperature tunneling measurements.[15]

The mechanism through which the BPR change affects the band alignment is more difficult to probe experimentally. Monochromatic x-ray photoemission (XPS) studies of the Zn 3*d*/Se 3*d* intensity ratio from ZnSe/GaAs interfaces as a function of ZnSe thickness indicate that high (low) BPRs yield a Zn-rich (Se-rich) composition near the interface.[8] The local nature of the mechanism is further supported by evidence that the offset change is also observed when the non-stoichiometric growth conditions (i.e., high or low BPR) are confined to a thin (~2 nm thick) ZnSe layer near the interface,[9] while the rest of the II–VI epilayer is grown with BPR=1. Calculations examining some of the simplest possible mixed atomic configurations that would lead to neutral interfaces predict $\Delta E_v = 0.62$ eV for interfaces comprised of a single Se–As mixed plane, and $\Delta E_v = 1.59$ eV for interfaces comprised of a single Zn–Ga mixed plane,[8,16] but no studies relating the BPR employed to the formation of different chemical bonds at the interface have been reported to date.

Here we take a different approach to band offset modification in ZnSe/GaAs(001) interfaces with the goal of ascertaining if the presence of Zn or Se just at the interface is sufficient to induce the band offset changes, and probing more closely the interfacial bonds that are established in the two cases. In the MBE growth procedure employed, the initial GaAs(001)2×4 surface was exposed to an elemental flux of Zn or Se prior to ZnSe growth. The surface predosing was used to achieve an initial ordered Zn or Se coverage of the surface. Since the formation of a Zn ordered monolayer on

[a]On leave from Fotona, Stegne 9, 61000 Ljubljana, Slovenia.
[b]Present address: Intel Corporation, Santa Clara, CA.
[c]Also with Dipartimento di Fisica, Università di Trieste, I-34012 Trieste, Italy.

GaAs in our experimental conditions required lower sample temperatures than normally employed for ZnSe MBE, we elected to perform both types of predosing at room temperature (RT). To avoid Zn desorption, and minimize atomic interdiffusion, we successfully implemented epitaxial regrowth of ZnSe at RT on the predosed surface. We examined the resulting heterojunction band offsets and interfacial chemical bond formation *in situ* by means of synchrotron radiation, soft-x-ray photoemission spectroscopy (SXPS), which yields substantially higher energy resolution and surface sensitivity relative to conventional XPS.

II. EXPERIMENTAL DETAILS

Si-doped ($N_D \sim 1\times10^{18}$ cm^{-3}), 500-nm thick GaAs(001)2×4 buffer layers were grown at 580 °C on n^+-GaAs(001) wafers in a solid source MBE facility following the procedures described elsewhere.[17] A micron-thick As cap layer was deposited *in situ* onto the samples to protect them during transfer in air to the Synchrotron Radiation Center of the University of Wisconsin-Madison. The samples were introduced into a miniature II–VI MBE chamber connected to the photoelectron spectrometer, and the As-cap layer was removed by heating the samples to 450 °C in about 30 min, resulting in a sharp 2×4 reflection high energy electron diffraction (RHEED) pattern, and a surface rms roughness comparable to that of as-grown surfaces.[18]

Predosing of the GaAs(001)2×4 surface with Se or Zn was performed using elemental solid sources, and atomic fluxes of the order of 3×10^{13} atoms/cm^2 s. In these conditions it was found that monolayer or multilayer adsorption of Se could be readily achieved at the typical ZnSe growth temperature of 290 °C, but that only traces of Zn could be detected on the surface. Although a systematic studies of Zn condensation on GaAs as a function of temperature and exposure time is beyond the scope of this paper, and will be presented elsewhere,[19] we mention that submonolayer to monolayer Zn saturation coverages were observed at 290 °C on GaAs only following hour-long exposures at Zn fluxes 25 times as high as those available here.

To increase the Zn sticking coefficient, we performed the Zn predosing—and for consistency the Se predosing—at RT. In these conditions monolayer coverages of Zn could be achieved in a few minutes, and gave rise to the removal of the initial 2×4 surface reconstruction, and the emergence of a sharp 1×1 RHEED pattern. The coverage was estimated using the reading of a quartz thickness monitor located near the sample. Estimates performed from the SXPS-derived attenuation of the Ga 3*d* and As 3*d* core levels were within 20%–30% of the thickness monitor reading. Se exposures also removed the initial 2×4 reconstruction giving rise to a 1×1 pattern. However, for coverages above 1 Å the pattern became increasingly diffused, and disappeared at about 2 Å. We therefore restricted our studies to a coverage of 1 Å for both Zn and Se at RT.

To minimize desorption and redistribution of the preadsorbed species, as well as reduce interdiffusion across the interface, ZnSe growth after predosing was also implemented at RT, using a compound ZnSe cell. For relatively low growth rates of 74 Å/h, epitaxy could also be achieved at such a relatively low temperature, admittedly with some degree of three-dimensional growth (3D), especially in the case of Zn-predosed surfaces, as indicated by the RHEED patterns (see next section).

In what follows the Zn, Se, or ZnSe coverage is given in angstroms or monolayers. The former values reflect the thickness monitor readings, and assume an overlayer density equal to that of bulk Zn, Se, or ZnSe.

At selected stages, the growth was interrupted, and the samples were cooled to RT, and transferred to the photoelectron spectrometer. We used synchrotron radiation from the 800 MeV electron storage ring Aladdin dispersed by a 6-m toroidal grating monochromator, and a commercial hemispherical analyzer to record angle-integrated photoelectron energy distribution curves EDCs from the Ga 3*d*, As 3*d*, Zn 3*d*, and Se 3*d* core levels with an overall energy resolution (electrons+photons) of 0.2 eV for a photon energy of 120 eV.

Most measurements reported here were obtained by collecting the photoelectrons at an average emission angle of ~20° from the sample normal. On occasions, to further enhance the surface sensitivity, photoelectrons were collected at about ~80° from the sample normal.[20]

When necessary, a line shape analysis of the different core levels was performed in terms of the superposition of up to three Voigt doublets, using a standard nonlinear least-squares minimization routine.[20–22]

III. RESULTS AND DISCUSSION

A. Epitaxy

In Fig. 1(a) we show the RHEED pattern (10 keV primary energy, [1$\bar{1}$0] sample azimuth) from a ZnSe/Se/GaAs(001) heterostructure fabricated by MBE at RT. The initial Se predosing of the surface (1 Å) was followed by the deposition of 70 Å of ZnSe at RT. The streaky pattern suggests the presence of a highly ordered overlayer. The 2×1 reconstruction is that observed during ZnSe growth at 290 °C with a single compound effusion cell, corresponding to a BPR~1 (Ref. 20). As discussed in the previous section, the initial 2×4 GaAs reconstruction (not shown) is rapidly replaced by a sharp 1×1 pattern (not shown) upon Se predosing, and then by the streaky 2×1 at a ZnSe coverage of about 5 Å.

In Fig. 1(b) we show the RHEED pattern (10 keV primary energy, [1$\bar{1}$0] sample azimuth) from a ZnSe/Zn/GaAs(001) heterostructure deposited by MBE at RT. The initial Zn predosing of the surface (1 Å) was followed by the deposition of 18 Å of ZnSe at RT. A well-defined diffraction pattern is still visible in Fig. 1(b), indicating crystallinity of the deposits, although the oval-shaped features in the pattern suggest a gradual transition to the spotty pattern indicative of 3D growth. The presence of the oval features complicates the identification of the surface symmetry which, however, shows several analogies with initial 1×1 pattern observed after Zn predosing.

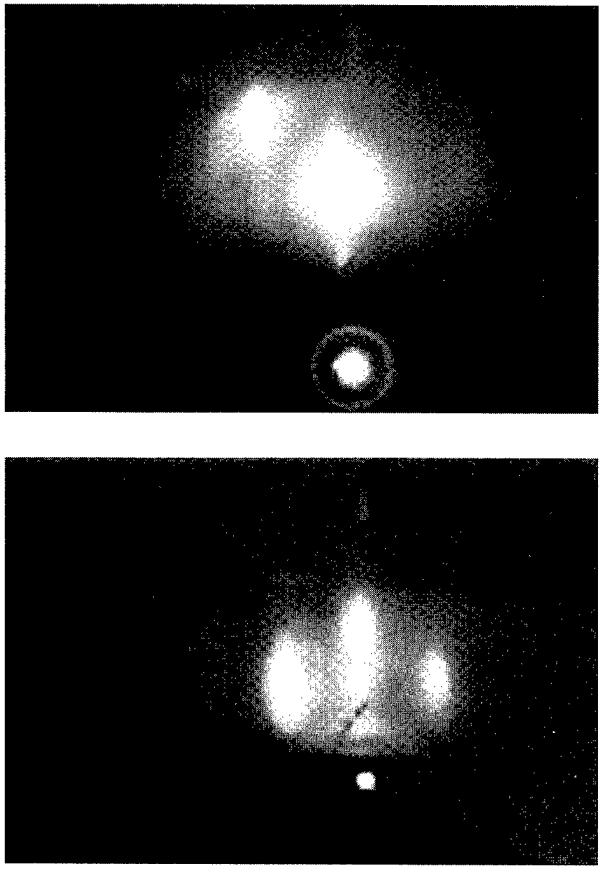


FIG. 1. Top: RHEED pattern (10 keV electrons) from a 70-Å-thick ZnSe layer deposited at room temperature (RT) on a GaAs(001)2×4 surface predosed with 1 Å of Se at RT to yield a 1×1 pattern. Sample azimuth: [1$\bar{1}$0]. Bottom: RHEED pattern from an 18-Å-thick ZnSe layer deposited at room temperature on GaAs(001)2×4 predosed with 1 Å of Zn at RT to yield a 1×1 pattern. Sample azimuth: [1$\bar{1}$0].

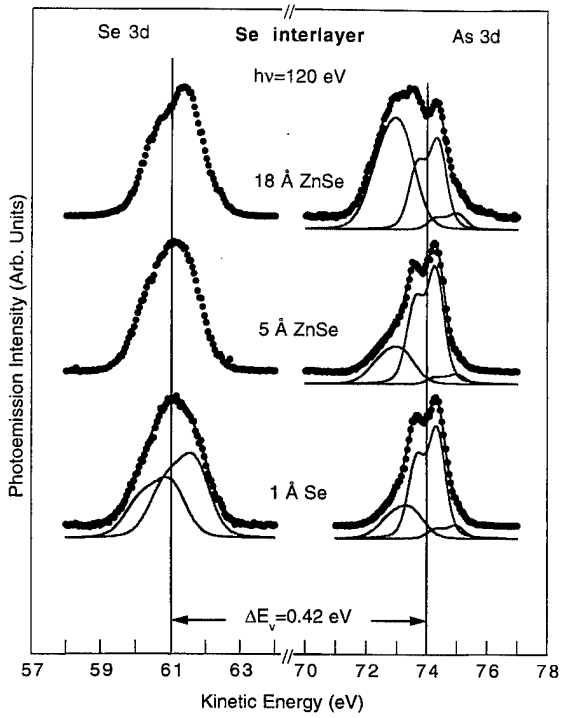


FIG. 2. As 3*d* and Se 3*d* core level spectra from ZnSe epilayers of increasing thickness (center and top) deposited at RT onto a GaAs(001) surface predosed with 1 Å of Se (bottom). The photon energy employed was 120 eV. The vertical lines mark the positions of the centroids of the bulk-related As 3*d* and Se 3*d* doublets. The resulting interface core level separation is consistent with a valence band offset $\Delta E_v = 0.48 \pm 0.06$ eV.

B. Se predosing

Representative SXPS results showing the effects of predosing and ZnSe overgrowth are summarized in Figs. 2 and 3 for Se predosing, and in Figs. 4 and 5 for Zn predosing. In Fig. 2 we show the evolution of the Se 3*d* (left) and As 3*d* (right) emission following predosing of the GaAs surface with 1 Å of Se (bottom), and upon deposition of 5 Å (center), and 18 Å (top) of ZnSe, all at RT. For selected spectra we also show superimposed to the data (solid circles) the result of a best-fit (solid line) of the overall line shape in terms of the superposition of two or three Voigt doublets (also shown, solid line). The corresponding results for the Ga 3*d* and Zn 3*d* core levels are shown in Fig. 3.

After Se predosing, the Se 3*d* emission in Fig. 2 (bottommost spectrum) is comprised of two inequivalent doublets, shifted by 0.69 eV relative to each other. This type of line shape is consistent with that reported following Se passivation of GaAs surfaces at RT.[23] In this passivation study the high-kinetic-energy component has been associated with the formation of As–Se bonds, and the low-kinetic-energy component has been associated with Se that overlays Se bonded to As. Note that Se predosing seems to induce little change in the Ga 3*d* emission in Fig. 3 (bottom), while the As 3*d* line shape in Fig. 2 (bottom-right) clearly shows the emergence of a broad reacted component, shifted by 0.80 eV to lower kinetic energy. A smaller high-kinetic energy contribution to the As 3*d* line might be—at least initially—related to some residual surface emission from regions not yet saturated with Se (Se coverage is only 1 Å in Figs. 2 and 3).

The lack of any Se-induced change in the Ga 3*d* line shape, and the formation of a chemically shifted, high binding energy contribution to the As 3*d* emission are also consistent with the results of Ref. 23, and has been tentatively explained with a preferential Se–As bond formation. The direction and order of magnitude of the As 3*d* chemical shift is consistent with previous reports on the formation of As–Se bonds. Simeca *et al.*[23] observed a component at 1.23 eV higher binding energy relative to the bulk As 3*d* doublet upon Se exposure of GaAs(001) at RT. Since the intensity of the reacted component decreased relative to the bulk component upon annealing at 250 °C, the authors attributed the shifted doublet to As segregated on the Se-terminated GaAs(001) surface.

We use the same interpretation for the low-kinetic energy, reacted As 3*d* component in Fig. 2 based on three considerations. First, upon ZnSe growth on top of the Se-predosed surface, the reacted component rapidly becomes the dominant As 3*d* feature when the ZnSe overlayer thickness become comparable with the photoemission sampling depth, and shifts to 1.18 eV below the substrate line. Second, mea-

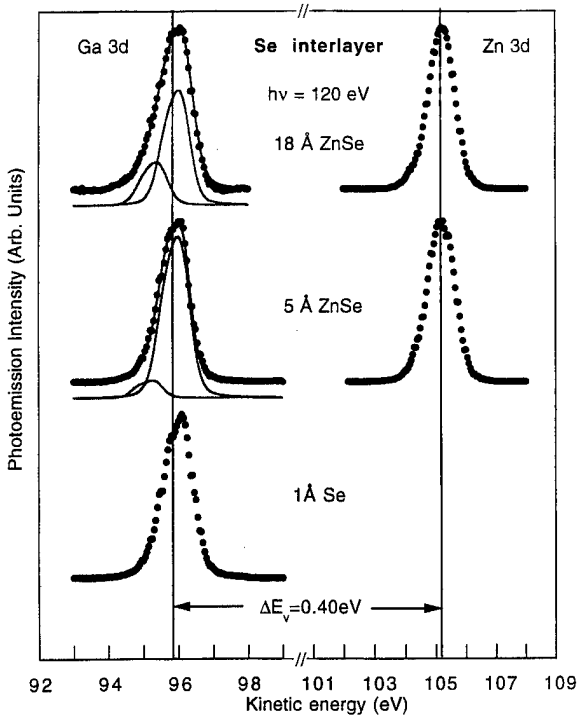


FIG. 3. Ga $3d$ and Zn $3d$ core level spectra from ZnSe epilayers of increasing thickness (center and top) deposited at RT onto a GaAs(001) surface predosed with 1 Å of Se (bottom). The photon energy employed was 120 eV. The vertical lines mark the positions of the centroids of the bulk-related Ga $3d$ and Zn $3d$ doublets. The resulting interface core level separation is consistent with a valence band offset $\Delta E_v = 0.40 \pm 0.06$ eV.

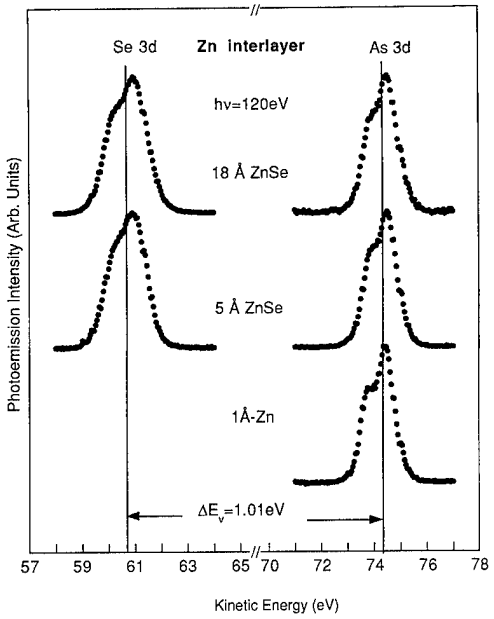


FIG. 4. As $3d$ and Se $3d$ core level spectra from ZnSe epilayers of increasing thickness (center and top) deposited at RT onto a GaAs(001) surface predosed with 1 Å of Zn (bottom). The photon energy employed was 120 eV. The vertical lines mark the positions of the centroids of the bulk-related As $3d$ and Se $3d$ doublets. The resulting interface core level separation is consistent with a valence band offset $\Delta E_v = 0.97 \pm 0.06$ eV.

surements at grazing emission angles (not shown) yield a dramatic increase of the intensity of this component relative to all other As $3d$ features. Third, upon annealing the overall ZnSe/Se/GaAs heterostructures at 400 °C (not shown), the component is selectively removed leaving the other As $3d$ features unchanged. We therefore suggest that Se predosing displaces a number of As atoms from GaAs which segregate on top of the Se-terminated overlayer surface, and that such atoms account for the high-binding energy reacted As feature in Fig. 2. Having extracted the segregated As–Se contribution from the As $3d$ line shape in Fig. 2, we associate the remaining main contribution at 74 eV of kinetic energy with emission from the GaAs substrate.

With increasing ZnSe deposition, the initially complex Se $3d$ line shape in Fig. 2 converges rapidly to the single doublet expected for ZnSe (see also Fig. 4). Therefore it seems feasible to use the main Se $3d$ doublet at a ZnSe coverage of 5–18 Å, and the substrate related As $3d$ feature in Fig. 2 to estimate the valence band offset. From the measured separation of the centroids in Fig. 2 (13.00 ± 0.05 eV, vertical solid lines), and the known position[24] of each doublet relative to valence band maximum in bulk ZnSe and GaAs, we calculated $\Delta E_v = 0.48 \pm 0.05$ eV.

In Fig. 3 we show the evolution of the Ga $3d$ (left) and Zn $3d$ (right) emission following predosing of the GaAs surface with 1 Å of Se (bottom), and upon deposition of 5 Å (center) and 18 Å (top) of ZnSe, all at RT. The Ga $3d$ shows no line shape change following Se predosing, but upon deposition of ZnSe onto the predosed surface, a low kinetic energy doublet becomes increasingly visible 0.75 eV below the main line. The direction and order of magnitude of the shift would be consistent with the formation of Ga–Se bonds, however, its origin is more ambiguous. Since the intensity of this feature become gradually more important relative to the main Ga $3d$ line with increasing ZnSe coverage, it is likely to reflect Ga atoms located at the overlayer surface or within the overlayer, rather than at the interface. Measurements at grazing emission (not shown) also tend to emphasize this feature. Upon annealing at 400 °C, however, this feature is *not* selectively removed as its As-related counterpart in Fig. 2, but rather it is rapidly enhanced until it becomes the dominant feature.

We therefore propose that the reacted Ga $3d$ component derives from interdiffusion across the ZnSe/GaAs interface, i.e., from Ga atoms outdiffusing from the substrate into the overlayer. Such outdiffusion is enhanced by annealing, and possibly by the presence of Zn vacancies, which is likely to accompany ZnSe regrowth on a Se-rich substrate surface. Enhanced Ga outdiffusion in Se-rich ZnSe/GaAs(001) interfaces has indeed been observed upon annealing, with the formation of characteristic deep level emission at 2.0 eV which reflects Ga_{Zn}–V_{Zn} complexes.[25,26] We emphasize that this is quite unlike the mechanism proposed to explain the reacted As component in Fig. 2, which derives from a fixed number of As atoms displaced during the initial Se predosing, and that segregate at the overlayer surface. Upon annealing, the segregated As atoms simply desorb from the surface, while Ga atoms continue to outdiffuse in the overlayer.

The Zn $3d$ line shape in Fig. 3 remains unchanged

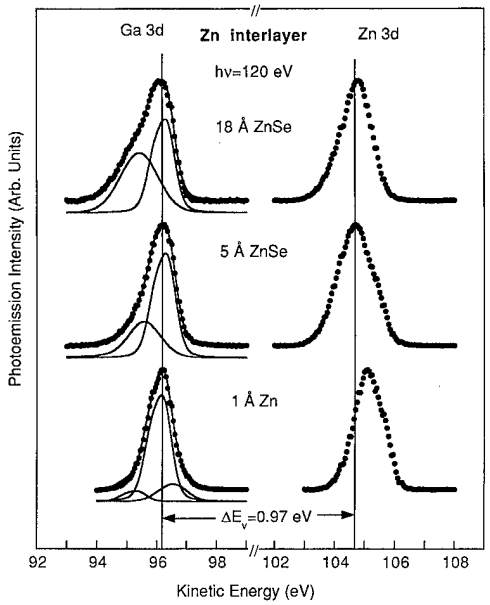


FIG. 5. Ga $3d$ and Zn $3d$ core level spectra from ZnSe epilayers of increasing thickness (center and top) deposited at RT onto a GaAs(001) surface predosed with 1 Å of Zn (bottom). The photon energy employed was 120 eV. The vertical lines mark the positions of the centroids of the bulk-related Ga $3d$ and Zn $3d$ doublets. The resulting interface core level separation is consistent with a valence band offset $\Delta E_v = 0.97 \pm 0.06$ eV.

throughout the coverage range explored, so that one can use the Zn $3d$ line, and the substrate related Ga $3d$ feature in Fig. 3 to estimate the valence band offset. From the measured separation of the centroids in Fig. 3 (9.30 ± 0.04 eV), and the known position[24] of each doublet relative to the valence band maximum in bulk ZnSe and GaAs, we found $\Delta E_v = 0.40 \pm 0.05$ eV, in good agreement with the value derived from the anion core levels in Fig. 2. This low value for the valence band offset is quantitatively consistent with that determined in ZnSe/GaAs(001) heterojunctions grown by MBE at 290 °C (BPR~0.1).[8,9,14,15] This compelling similarity has two important implications. First, it supports the contention that the effect of using a low BPR in the early growth stages of ZnSe is to drive the system toward the formation of a Se-terminated interface. Second, since interdiffusion at the interface is likely to be minimized at RT relative to 290 °C—as indicated by the increase in Ga outdiffusion which we observe in our sample upon annealing—and we still measure a comparable offset in the two cases, it argues against a substantial role of interdiffusion-related doping dipoles in determining the experimental band offsets.

C. Zn predosing

In Fig. 4 we show the evolution of the Se $3d$ (left) and As $3d$ (right) emission following predosing of the GaAs surface with 1 Å of Zn (bottom), and upon deposition of 5 Å (center) and 18 Å (top) of ZnSe, all at RT. The corresponding results for the Ga $3d$ and Zn $3d$ core levels are shown in Fig. 5, where we also show, superimposed to the data (solid circles) for selected Ga $3d$ spectra the result of a best-fit (solid line) of the overall line shape in terms of the superposition of two or three Voigt doublets (also shown, solid line).

The line shape of the Se $3d$ and As $3d$ doublet remains remarkably constant in Fig. 4. In particular, we observe no qualitative changes in the As $3d$ line shape after Zn predosing and ZnSe regrowth, in sharp contrast with Se predosing. We can therefore directly use the position of the two centroids to gauge the valence band offset. From the measured separation of the centroids in Fig. 4 (13.53 ± 0.02 eV), and the known position[24] of each doublet relative to valence band maximum in bulk ZnSe and GaAs, we obtained $\Delta E_v = 1.01 \pm 0.05$ eV, i.e., a value which is 0.6 eV higher than that observed for interfaces fabricated with Se predosing.

The Ga $3d$ and Zn $3d$ core levels in Fig. 5 exhibit substantial changes following predosing and ZnSe regrowth. Upon Zn predosing (bottom), the Ga $3d$ line shape shows a definite high kinetic energy contribution, together with a smaller low kinetic contribution that could derive from residual surface-related emission. Although for the moment our attempt at obtaining a consistent line shape analysis for the Zn $3d$ doublet at different emission angles and photon energy has been less than satisfactory, qualitatively it seems clear that a Zn $3d$ line shape following predosing also shows an unresolved high kinetic energy contribution relative to the main line.

The formation of Ga–Zn bonds at the surface might explain the presence of high kinetic energy contributions to both the Ga $3d$ and Zn $3d$ line shape following predosing, in view of the electronegativity difference between Ga, Zn, and As. However, upon deposition of 5 Å of ZnSe the high kinetic energy component to the Ga $3d$ lines is not detected anymore, so that it is not clear if Ga–Zn bonds at the interface play a role in the final heterojunction. The Zn $3d$ line shape at 5 Å ZnSe coverage (center, Fig. 5) still shows an unresolved high kinetic energy contribution, suggesting that this might be the case, but the same contribution might be due to segregated Zn atoms displaced from the interface. Since no reacted As $3d$ contribution is visible in Fig. 4, as opposed to Fig. 2, i.e., no As segregation is observed, we are forced to conclude that any Zn–Ga bond formation occurs either without displacing the As from the interface region, leading to the formation of a complex ternary phase, or that the displaced As atoms are desorbed from the surface. The latter would be somewhat surprising in view of the high stability of Zn–As bonds, and of the low temperature at which the experiment was performed.

The emergence of an important low kinetic energy contribution 0.81 eV below the main Ga $3d$ line in Fig. 5 upon deposition of 5 and 18 Å of ZnSe also raises some puzzling issues. This feature is substantially broader and has a somewhat larger chemical shift as compared to the corresponding feature in Fig. 3. In the absence of data concerning the dependence of the intensity of this feature on the escape depth and annealing time and temperature, at this stage we cannot discriminate between a reacted interface contribution, or a surface segregated contribution. In both cases, the direction

and order of magnitude of the corresponding chemical shift points to Ga–Se bond formation.

From the measured separation of the centroids of the Zn $3d$ line at high ZnSe coverage in Fig. 5 (vertical solid lines), and the substrate-related Ga $3d$ doublet (8.73 ± 0.05 eV), and the known position[24] of each doublet relative to valence band maximum in bulk ZnSe and GaAs, we found $\Delta E_v = 0.97\pm0.06$ eV, in very good agreement with the value derived from the anion core levels in Fig. 4.

IV. CONCLUSIONS

Ordered monolayer coverages of Zn and Se were obtained on the GaAs(001)2×4 surface at room temperature prior to ZnSe MBE. Adsorption removed the initial 2×4 surface reconstruction and yielded a 1×1 surface periodicity in both cases. On top of these adsorbed layers, we achieved ZnSe epitaxial overgrowth at room temperature. Low temperature was required to minimize Zn desorption, and suppress the Zn indiffusion and the Ga outdiffusion which are known to take place across the ZnSe/GaAs(001) interface at higher growth temperatures. Arsenic segregation on top of the ZnSe/Se/GaAs(001) structure was nevertheless observed, as a result of the displacement of As atoms from the interface during Se predosing. This type of interface reaction was not observed following Zn predosing.

Although ZnSe epitaxial growth was performed with identical Zn/Se beam pressure ratios for both types of interfaces, widely different band offsets were observed for the two interfaces. Se predosing resulted in valence band offsets as low as 0.40 ± 0.05 eV, while Zn predosing yielded valence band offsets as high as 1.01 ± 0.05 eV. Such values are compellingly similar to those observed for ZnSe/GaAs(001) heterojunctions grown with low and high Zn/Se BPR, respectively.

Our results support the contention that the effect of using a low (high) BPR in the early growth stages of ZnSe is to drive the system toward the formation of a Se-terminated (Zn-terminated) interface. Furthermore, since interdiffusion across the interface is minimized at RT relative to typical growth temperatures, and comparable band offsets are observed in the two cases, interdiffusion-related doping dipoles are unlikely to have a substantial role in determining the experimental band offset.

ACKNOWLEDGMENTS

This material was based on work supported in part by the U.S. Army Research Office under Grant Nos. DAAH04-93-G-0206 and DAAH04-93-G-0319, and by the Center for Interfacial Engineering of the University of Minnesota under NSF Grant CDR 87211551. The Synchrotron Radiation Center of the University of Wisconsin-Madison is supported by NSF, and the authors gratefully acknowledge the cheerful assistance of its staff.

[1] L. J. Brillson, in *Handbook on Semiconductors*, edited by P. T. Landsberg (North-Holland, Amsterdam, 1992), Vol. 1, p. 281.
[2] G. Margaritondo and P. Perfetti, in *Heterojunction Band Discontinuities: Physics and Device Applications*, edited by F. Capasso and G. Margaritondo (North-Holland, Amsterdam, 1987), Chap. 2.
[3] A. Franciosi and C. G. Van de Walle, Surf. Sci. Rep. **214**, 1 (1996).
[4] W. A. Harrison, J. Vac. Sci. Technol. **14**, 1016 (1977); W. A. Harrison, E. A. Kraut, J. R. Waldrop, and R. W. Grant, Phys. Rev. B **18**, 4402 (1978).
[5] See S. Baroni, R. Resta, A. Baldereschi, and M. Peressi, in *Spectroscopy of Semiconductor Microstructures*, edited by G. Fasol, A. Fasolino, and P. Lugli (Plenum, London, 1989); A. Baldereschi, R. Resta, M. Peressi, S. Baroni, and K. Mäder, in *Proceedings of the NATO Advanced Research Workshop Physical Properties of Semiconductor Interfaces at Sub-Nanometer Scale*, edited by H. W. Salemink (Kluwer Academic, Dordrecht, 1993).
[6] A. Muñoz, N. Chetty, and R. M. Martin, Phys. Rev. B **41**, 2976 (1990).
[7] R. G. Dandrea, S. Froyen, and A. Zunger, Phys. Rev. B **42**, 3213 (1990).
[8] R. Nicolini, L. Vanzetti, G. Mula, G. Bratina, L. Sorba, A. Franciosi, M. Peressi, S. Baroni, R. Resta, A. Baldereschi, J. E. Angelo, and W. W. Gerberich, Phys. Rev. Lett. **72**, 294 (1994).
[9] A. Bonanni, L. Vanzetti, L. Sorba, A. Franciosi, M. Lomascolo, P. Prete, and R. Cingolani, Appl. Phys. Lett. **66**, 1092 (1995).
[10] M. A. Haase, J. Qiu, J. M. DePuydt, and H. Cheng, Appl. Phys. Lett. **59**, 1272 (1991).
[11] H. Jeon, J. Ding, W. Patterson, A. Nurmikko, W. Xie, D. Grillo, M. Kobayashi, and R. L. Gunshor, Appl. Phys. Lett. **59**, 3619 (1991).
[12] N. Nakayama, S. Itoh, T. Ohata, K. Nakano, H. Okuyama, M. Ozawa, A. Ishibashi, M. Ikeda, and Y. Mori, Electron. Lett. **29**, 1488 (1993).
[13] D. Herve', R. Accomo, E. Molva, L. Vanzetti, J. J. Paggel, L. Sorba, and A. Franciosi, Appl. Phys. Lett. **67**, 2144 (1995).
[14] X. Yang, L. J. Brillson, L. Vanzetti, A. Bonanni, and A. Franciosi (unpublished).
[15] V. Pellegrini, M. Börger, F. Beltram, J. J. Paggel, L. Sorba, S. Rubini, and A. Franciosi (unpublished).
[16] A. Kley and J. Neugebauer, Phys. Rev. B **50**, 8616 (1994).
[17] L. Sorba, G. Bratina, A. Antonini, A. Franciosi, L. Tapfer, A. Migliori, and P. Merli, Phys. Rev. B **46**, 6834 (1992).
[18] Y. Fan, I. Karpov, G. Bratina, L. Sorba, W. Gladfelter, and A. Franciosi, J. Vac. Sci. Technol. B **14**, 623 (1996).
[19] See S. Heun, J. J. Paggel, S. Rubini, and A. Franciosi, J. Vac. Sci. Technol. B **14**, 2980 (1996).
[20] G. Bratina, T. Ozzello, A. Bonanni, J. J. Paggel, L. Vanzetti, L. Sorba, and A. Franciosi (unpublished).
[21] P. H. Citrin, G. K. Wertheim, and Y. Baer, Phys. Rev. B **27**, 3160 (1983).
[22] A. Raisanen, G. Haugstad, X. Yu, and A. Franciosi, J. Appl. Phys. **72**, 1104 (1992); **70**, 3115 (1991).
[23] T. Simeca, Y. Watanabe, R. Berrigan, and M. Oshima, Phys. Rev. B **46**, 10201 (1992).
[24] The observed binding energies of the Ga $3d$, As $3d$, Zn $3d$, and Se $3d$ cores relative to the valence band maximum in bulk GaAs(001) and ZnSe(001) were: 18.83 ± 0.05 eV, 40.72 ± 0.05 eV, 9.13 ± 0.05 eV, and 53.24 ± 0.05 eV, respectively.
[25] A. Raisanen, L. J. Brillson, L. Vanzetti, L. Sorba, and A. Franciosi, J. Vac. Sci. Technol. A **13**, 690 (1995).
[26] A. Raisanen, L. J. Brillson, L. Vanzetti, L. Sorba, and A. Franciosi, Appl. Phys. Lett. **66**, 3301 (1995).

Interface states of ZnSe/GaAs interface

Z. Yang,[a] I. K. Sou, Y. H. Yeung, and G. K. L. Wong
Department of Physics, The Hong Kong University of Science and Technology, Clearwater Bay, Kowloon, Hong Kong

Jie Wang, Cai-xia Jin, and Xiao-yuan Hou
Surface Physics Laboratory, Fudan University, People's Republic of China

(Received 22 January 1996; accepted 26 April 1996)

The in-plane anisotropy ($\epsilon_x \neq \epsilon_y$) and the off-plane anisotropy ($\epsilon_x \neq \epsilon_z$) of ZnSe/GaAs interfaces formed under conditions that promote the formation of either Zn–As or Ga–Se bonds are studied by reflectance difference spectroscopy. Two resonance lines, one at 2.70 eV and the other around 3.0 eV, have been observed. The in-plane anisotropy is along the [110] and the [1$\underline{1}$0] principal axes. The dependence of the resonances on interface formation conditions, the results of the photoreflectance spectroscopy, and the annealing experiments all suggest that the anisotropy is not due to the electro-optic effect resulting from an interface electric field. The reflectance difference spectroscopy results are consistent with the assumption that the anisotropy is the intrinsic properties of ordered ZnSe/GaAs heterovalent interface, that the resonance at 2.70 eV is associated with the interface state of Zn–As bonds and the resonance near 3.0 eV is associated with the interface state of Ga–Se bonds. The presence of a thin layer of S at the ZnSe/GaAs interfaces results in a third resonance at 3.2 eV, probably due to the Ga–S bonds. The resonance line shape changes as the top layer thickness varies. Such change can be well explained by a three-layer system with the conventional Fresnel optics and is understood as due to the ZnSe top layer. It is our hope that our results will stimulate more interest in the theoretical study of this interface. © *1996 American Vacuum Society.*

I. INTRODUCTION

The properties of semiconductor interfaces are an important topic of research for both their fundamental and practical significance. As the technology of epitaxial thin film growth advances, it becomes clear that the quality of a thin film grown on a substrate depends critically on the formation of film/substrate interface. Experimentally, however, a buried solid interface is difficult to probe. Current surface diagnostic techniques such as electron beam diffraction have limited penetration depth and can only probe interfaces covered by a very thin layer of several nanometers thick. However, the interfaces may change as the epilayers grow thicker such as in the case of strain relaxation. Interfaces buried hundreds of nanometers beneath the surface are out of reach by the current surface diagnostic techniques, and the development of new techniques becomes necessary. Optic spectroscopy has much larger penetration depth, and so far the most successful one applied to surface/interface has been the reflectance difference/anisotropy spectroscopy (RDS/RAS).[1] RDS measures the difference between the reflectances of light polarized along two perpendicular directions as a function of photon energy. When light is normally incident on a surface/interface sample and if the bulk optical response of the sample is isotropic, then the only contribution to the reflectance difference (RD) comes from the surface/interface. RDS is therefore a surface/interface sensitive technique, and has been widely employed in the study of surface electronic properties of III–V semiconductors.[2–4] RDS has also been applied to study the Si/SiO_2 interfaces[5] buried under a layer of transparent SiO_2, and sulfur-passivated GaAs surface.[6]

ZnSe is an important material for light emitting devices in the blue color region. Continuous wave lasing at room temperature has been demonstrated and the life span of the devices has been extended to over 100 h. At present, most of the ZnSe based devices are grown on GaAs substrates because of their high quality and relatively low cost. The quality of the ZnSe/GaAs interface is one of the key factors, if not the most important factor, that limits the device life span. In the meantime, a perfect ZnSe/GaAs interface is expected to contain a single layer of Zn–As or Ga–Se bonds, depending on how the substrate surface is treated prior to growth, and these bonds are pointing at either the [110] or [1$\underline{1}$0] direction for a (001) oriented substrate. As a result, the ZnSe/GaAs interface is expected to exhibit optical anisotropy of the same magnitude as a clean GaAs surface.

One problem to be solved in studying the ZnSe/GaAs interface using optical techniques such as RDS is to separate the interface signal from that of the ZnSe surface. Although this has not been clearly revealed, a well ordered ZnSe surface is nevertheless expected to exhibit sizable dielectric anisotropy and contribute significantly to RDS signal. To study the ZnSe/GaAs interface, it is therefore necessary to *passivate* the ZnSe surface to make it optically isotropic.

In our previous work, we have reported the RDS study of ZnSe/GaAs interfaces buried hundreds of nanometers below the ZnSe surface.[7] Sharp resonances in RDS spectra at 2.70 and 3.0 eV had been observed. Efforts were made to identify the origin of the resonances, and it was concluded that RD due to the electro-optic effect was not the likely cause be-

[a]Electronic mail: phyang@usthk.ust.hk

cause no interface electric field had been observed. It was suggested that the Ga–Se and/or the Zn–As bonds might be related to the resonances. In this article, we report systematic study of the optical anisotropy of ZnSe/GaAs (001) interfaces prepared under different conditions. The surface of the GaAs substrates were pretreated to promote the formation of either Zn–As bonds or Ga–Se bonds. More evidence is obtained that suggests the absence of the electro-optic effect in our samples. Within the interface plane the anisotropy occurs along the [110] and [1$\underline{1}$0] principal axes ($\epsilon_x \neq \epsilon_y$). In addition, with the employment of a special RDS configuration, the optical anisotropy off the interface plane ($\epsilon_x \neq \epsilon_z$) is also studied.[7] Both the in-plane and the off-plane anisotropy vary for interfaces formed under different conditions. Efforts to identify the origin of the interface optical anisotropy are made and results will be reported.

II. EXPERIMENT

The RDS optical setup is essentially the same as described in Ref. 1 except that the polarization of the incident light now is at 45° to the incidence plane. In such configuration, both the anisotropic and isotropic interface states can be observed.[7] The samples can be rotated within the surface plane. The incidence angle is 7° off the normal direction of a sample surface. The difference between the complex reflectances of light polarized in the two directions ($r_s - r_p$), where s and p denote the conventional polarization directions in the oblique incidence case, is measured. The samples are undoped ZnSe single crystal films grown at 250° to 300 °C by molecular beam epitaxy (MBE) on (001) GaAs semi-insulating substrates. Planer and cross section transmission electron micrographs show that the interfaces are sharp and flat, with no sign of corrugated quantum-wire type structures. Stacking faults dislocation density is in the range of 10^8 to low 10^4 per cm^2. The surfaces of the ZnSe layers are covered with a layer of Se at the end of growth to prevent oxidation when exposed to air. Typical full width at half-maximum (FWHM) of x-ray diffraction peak of the films is a few hundred arc second. Epilayers with thickness ranging from 18 Å to more than 1 μm have been studied. All RDS measurement are done in air at room temperature.

The measured RD spectrum, $\Delta\rho \equiv 2^*(r_s - r_p)/(r_s + r_p)$, contains three parts, namely the anisotropic part $\Delta\rho_a$ which is due to the interface dielectric anisotropy, the isotropic part $\Delta\rho_i$ which is due to the off-plane anisotropic of the interface, and the featureless isotropic part $\Delta\rho_f$ of the conventional Fresnell RD due to the nonzero incidence angle. $\Delta\rho_a$ has a $\sin 2\theta$ dependence on the angle θ between the incidence polarization and the [110] principal axis, while $\Delta\rho_i$ and $\Delta\rho_f$ has no such angular dependence. By measuring $\Delta\rho(\theta)$ as a function of θ, one can then separate $\Delta\rho_a$ from the isotropic parts $\Delta\rho_i + \Delta\rho_f$, i.e., $\Delta\rho_a = [\Delta\rho(0°) - \Delta\rho(90°)]/2$, and $\Delta\rho_i + \Delta\rho_f = [\Delta\rho(0°) + \Delta\rho(90°)]/2$. The isotropic parts can further be confirmed by the $\Delta\rho(45°)$ spectra.

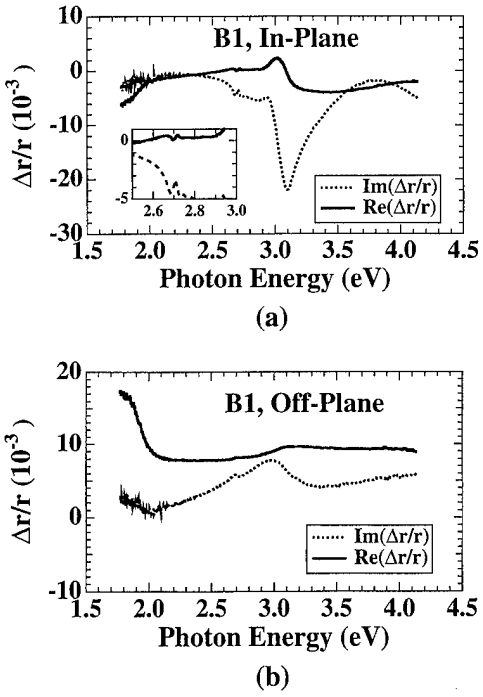


FIG. 1. (a) The complex in-plane RD spectra of sample B1. The solid curve represents the real part of RD, while the dashed curve represents the imaginary part. The thickness of the ZnSe epilayer on GaAs(100) substrate is about 1000 Å. (b) The complex off-plane RD spectra of sample B1. The solid curve represents the real part of RD, while the dashed curve represents the imaginary part.

III. RESULTS AND DISCUSSIONS

An undoped ZnSe single crystal (100) substrate of about 1 mm thick was measured first to test the RDS apparatus. The obtained $\Delta\rho$ is featureless and is of the order of 10^{-3}, largely due to the Fresnel RD of p and s polarization. Bulk ZnSe contribution to RD can therefore be neglected for our purposes.

A. Effects of surface treatments

The five samples, all with the same nominal ZnSe layer thickness of 1000 Å but with different GaAs substrate surface treatment, are presented first. The samples were all grown by molecular beam epitaxy (MBE) at 250 °C on (001) GaAs substrates. The substrate for sample B1 was first treated by $5H_2SO_4 + H_2O_2 + H_2O$ and heated in vacuum at 600 °C for 10 min. Growth of ZnSe was then initiated by simultaneous exposure of Zn and Se flux. Figs. 1(a) and 1(b) show the in-plane and the off-plane RD spectra of sample B1. A strong resonance line around 3.1 eV is obviously seen in the in-plane RD spectrum. A much sharper but much weaker resonance line is also observed at 2.70 eV [see the insert of Fig. 1(a)], with a FWHM of only 20 mV. The off-plane RD spectrum shows a pronounced resonance near 3.0 eV in the imaginary part of $\Delta\rho$ ($=\Delta r/r$) and a derivative-like feature near the same energy in the real part of $\Delta\rho$.

The substrates for samples B2 and B4 were first treated with $(NH_3)S_x$ and heated to 380 °C for 4 min in vacuum. The

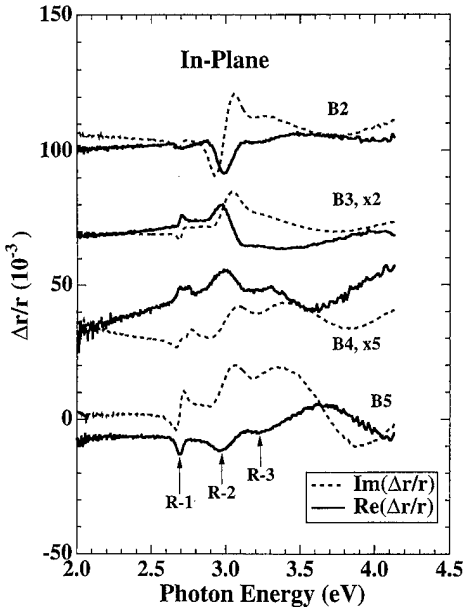


FIG. 2. The complex in-plane RD spectra of samples B2 through B5. The solid curves represent the real parts of RD, while the dashed curves represent the imaginary parts of RD.

substrates for samples B3 and B5 were first treated with S_2Cl_2 and heated to 380 °C for 4 min. The chemical pretreatment left a layer of S on the GaAs surface.[8] The substrates of B2 and B3 were exposed to Se flux (partial pressure 3.2×10^{-4} Pa) for 1 min prior to growth, while those of B4 and B5 were exposed to Zn flux (partial pressure 1.5×10^{-4} Pa) for 1 min before growth. Their in-plane and off-plane RD spectra are shown in Figs. 2 and 3, respectively. Several

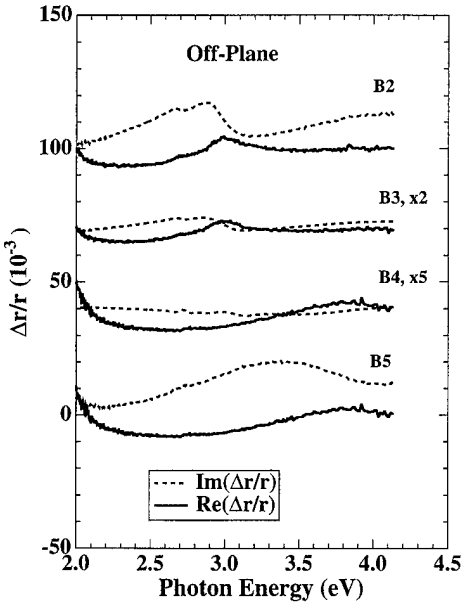


FIG. 3. The complex off-plane RD spectra of samples B2 through B5. The solid curves represent the real parts of RD, while the dashed curves represent the imaginary parts of RD.

important points can be made by examining these spectra.

The in-plane RD spectra of all five samples are qualitatively the same. Two resonance lines, one at 2.70 eV (R-1), and the other near 3.0 eV (R-2), are present in all the spectra. The energy of 2.70 eV is slightly above the E_0 band gap energy of 2.67 eV at room temperature, while the energy of 3.0 eV is near the $E_0+\Delta$ gap of ZnSe and the E_1 gap of GaAs. The fact that the resonance energies are close to the critical energy points of bulk ZnSe and GaAs suggests that the resonances are caused by perturbation of bulk electronic states. The relative intensities of the two resonances, however, depend on the substrate surface exposure to molecular beams prior to growth. For sample B1 which was grown without any pregrowth beam exposure, the intensity of resonance line R-2 at 3.0 eV is about 60 times the intensity of R-1 at 2.7 eV. For B2 and B3 which were grown on Se-rich substrate surfaces the ratio of R-2 and R-1 intensities is about 15, and for B4 and B5 grown on Zn-rich substrate surfaces the ratio is only about two. The off-plane RD spectra of samples grown on Zn-rich surfaces (B4 and B5) are also distinctively different from those grown on Se-rich surfaces (B2 and B3) or bare GaAs surface (B1). A resonance in the off-plane RD spectra at nearly the same energy of R-2 of in-plane RD spectra is observed for samples B1, B2, and B3, while there are no resonances for B4 and B5. Resonances in off-plane RD spectra indicate the presence of interface states because only states localized at the interface, i.e., of two-dimensional nature can cause off-plane anisotropy $\epsilon_x \neq \epsilon_z$.

A third resonance R-3 are observed for samples B2, B4, and B5 (although barely seen in Fig. 2, the resonance is well defined when proper vertical scale is used). The resonance, however, is not observed in samples without S coverage.

Different chemical treatments of substrate surface seem to have different effects on the overall magnitude of the RD spectra. The RD magnitude of B5 is about five times of that of B4. Both samples were grown on Zn-exposed surfaces. The substrate surface of B5 was pretreated with S_2Cl_2 while that of B4 was pretreated with $(NH_3)S_x$. Both chemical treatments leave an S covered GaAs surface with much reduced surface states with S_2Cl_2 treatment resulting in stronger Ga–S bonds.[8] For samples B2 and B3 grown on Se-exposed surfaces the effect of chemical pretreatment is opposite. The RD magnitude of B2 which was $(NH_3)S_x$ treated is about two times that of B3.

Other samples such as ZnSe-7, ZnSe-8, ZnSe-41, and ZnSe-66 were grown on $5H_2SO_4+H_2O_2+H_2O$ treated GaAs surface. ZnSe-7, ZnSe-8, and ZnSe-66 were grown on a Zn-rich GaAs surface, and ZnSe-41 was grown on a Se covered 2×3 reconstructed GaAs surface as a result of prolonged Se beam exposure. The interface defects density of ZnSe-41 is much lower than the others grown on 2×1 reconstructed surface. For example the stacking faults density of ZnSe-41 is in the low 10^4 cm^2, close to the high quality samples grown in two-chamber MBE systems. The preparation procedures of all the samples along with their RDS results are summarized in Table I.

TABLE I. Substrate surface pretreatments, RDS resonance intensity ratio, RDS resonance peak-to-peak values, and the thickness of ZnSe overlayer of the ZnSe/GaAs samples.

	Chemical dip	Preheat	Flux exposure	RHEED pattern	R2/R1 ratio	RDS value peak-to-peak	ZnSe layer thickness (Å)
B1	$5H_2SO_4$ $+H_2O_2+H_2O$	10 min at 600 °C	No	None	60	0.02	1000
B2	$(NH_3)S_x$	4 min at 380 °C	Se	None	15	0.03	1000
B3	S_2Cl_2	4 min at 380 °C	Se	None	15	0.015	1000
B4	$(NH_3)S_x$	4 min at 380 °C	Zn	None	2	0.004	1000
B5	S_2Cl_2	4 min at 380 °C	Zn	None	2	0.02	1000
ZnSe-7	$5H_2SO_4$ $+H_2O_2+H_2O$	3 min at 380 °C	Zn	2×1	6	0.005	1200
ZnSe-8	$5H_2SO_4$ $+H_2O_2+H_2O$	3 min at 380 °C	Zn	2×1	6	0.005	1000
ZnSe-41	$5H_2SO_4$ $+H_2O_2+H_2O$	3 min at 380 °C	Se prolonged	2×3	60	0.04	325
ZnSe-66	$5H_2SO_4$ $+H_2O_2+H_2O$	3 min at 380 °C	Zn	2×1	6	0.04	18

B. Line shape

It is noted in the spectra shown in Figs. 1–3 that the real part of $\Delta\rho$ is closely related to the imaginary part of $\Delta\rho$, namely that if the real part of the resonance is in line shape then the corresponding imaginary part is derivativelike, or vise versa, such as in the case of R-1 resonance of B4 and B5, and R-2 resonance of B2 and B4. In other cases, the resonances are neither line nor derivative like, but a mixture of the two, such as in the case of R-1 and R-2 resonances of B3. As will be seen in the following text, the line shape of the resonances depends on the ZnSe epilayer thickness, and a single interface anisotropic dielectric function can produce RD resonance which is in line shape, derivativelike, or a mixture of the two depending on the ZnSe overlayer thickness. The difference in line shape of the resonances of these five samples indicates that there is a small variation of ZnSe thickness from sample to sample, even though nominally they are all the same. The effect of finite overlayer thickness on interface RD spectra has been effectively treated in Ref. 5. Here we take a slightly different approach to demonstrate how drastic the resonance line shape is affected by the top ZnSe layer, in the energy range where ZnSe is highly absorbing.

Rather than showing $\Delta\rho$ ($d=0$) for zero overlayer thickness from the measured $\Delta\rho(d)$ using the approach in Ref. 5, we simulate the interface with an anisotropic thin layer with a Lorentz-type anisotropic dielectric function resonant at 2.72 eV, and show the calculated RD spectra for various ZnSe overlayer thickness. For clarity, only the R-1 portion of the spectra are shown. The three-layer optic system consisting of a ZnSe top layer, an interface layer of 50 Å, followed by an infinitely thick GaAs substrate. The reflectance of the system is calculated by applying the electromagnetic fields boundary conditions at each interface. The dielectric functions of bulk ZnSe and GaAs are taken from Refs. 9 and 10, respectively. Only the results at normal incidence are presented here, as a small tilt angle of 7° in this model merely results in a featureless RD background of the order of 10^{-3}. The upper curves of Figs. 4(a) and 4(b) show the calculated complex $\Delta\rho$ of two ZnSe thicknesses, respectively. For a 950-Å-thick ZnSe top layer there is a sharp line in Re($\Delta\rho$)

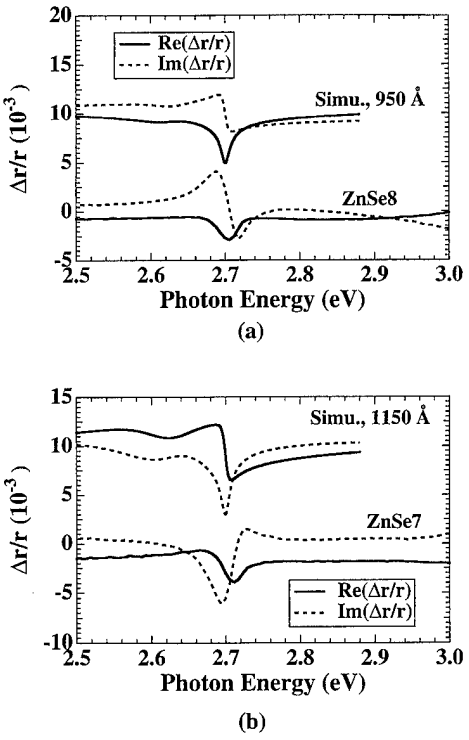


FIG. 4. Simulation spectra of RD with (a) 950 Å ZnSe layer, and (b) 1150 Å ZnSe layer. The experimental RD spectra of sample ZnSe-8 with an 1000 Å ZnSe layer, and sample ZnSe-7 with a 1200 Å ZnSe layer are shown for comparison.

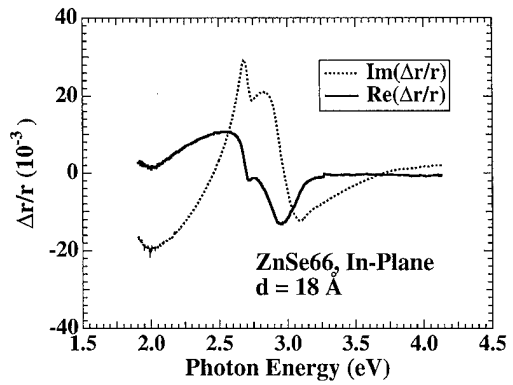


FIG. 5. In-plane RD spectra of sample ZnSe-66 with an 18-Å-thick ZnSe overlayer. The solid curve represents the real parts of RD, while the dashed curve represents the imaginary parts of RD.

and a derivative-like feature in $\mathrm{Im}(\Delta\rho)$ spectrum. The experimental RD spectra of ZnSe-8 with an 1000 Å top layer is shown in the lower part of the figure for comparison. The thickness of ZnSe-8 was determined from the interference fringe of the x-ray diffraction curve. For an 1150-Å-thick ZnSe layer the calculated line shape of $\mathrm{Re}(\Delta\rho)$ and $\mathrm{Im}(\Delta\rho)$ are reversed, i.e., $\mathrm{Im}(\Delta\rho)$ is in line shape and $\mathrm{Re}(\Delta\rho)$ is derivative-like, and they resemble those of sample ZnSe-7 with 1200 Å ZnSe layer. Simulation RD spectra at layer thicknesses between 1000 and 1200 Å are a mixture of the two shown in Fig. 4, which explains the shape of the RD spectra of sample B3. In a wider thickness range the simulation spectra evolve periodically from line shape to derivativelike, and back to line shape, in consistency with the RD spectra of other ZnSe/GaAs samples with corresponding overlayer thicknesses. It is important to count for the epilayer phase factor in the RD spectra. Otherwise, one would come to the wrong conclusion that the ZnSe/GaAs interface anisotropy changes with the thickness of the ZnSe layer.

C. Origin of anisotropy

We now turn to the discussions of possible causes of the observed resonances. First, the observed features are not from the ZnSe surface layer, because the RD spectra of ZnSe bulk samples show no resonance features, and the RD spectra of thick (>5000 Å) samples are featureless in the energy range above the band gap of ZnSe. This is consistent with the fact that the surface oxide layer is amorphous and isotropic. Second, the observed RD is not a bulk effect. Apart from the evidence from the results of thick samples it is also because no obvious thickness dependence of RD resonance intensity has been found from the dozens of samples investigated so far, if the thickness is below 2000 Å. RDS resonance intensity varies from sample to sample with nearly the same ZnSe thickness, the cause of which will be discussed later in this article. Strong resonances comparable to that of B5 were observed for samples with ZnSe layers as thin as 18 Å, as is shown in Fig. 5. Nor is it due to the GaAs bulk effect because the observed resonance energies shifted as alloy composition in the same way as the energy gap E_0 for ZnCdSe and ZnSSe alloy films. The only remaining source of anisotropy is therefore the epilayer/substrate interface.

One possible origin of the anisotropy is the electro-optic effect due to the presence of an interface electric field.[6,11] It is an attractive assumption because the observed resonance energies are close to the critical energies of bulk ZnSe. The field could be due to extended interdiffusion of Ga and Zn atoms which serve as dopants in the adjacent layers[12] over dozens of atomic layers at the interface, due to Fermi level pinning at an interface state, or due to an abrupt ZnSe/GaAs interface.[13] The electric field due to the first two cases extends over dozens to hundreds of atomic layers (extended field), while in the last case, the field exists only within several atomic layers from the interface (confined field). However, according to the following observation, the extended electric field is not strong enough to produce anisotropy of such magnitude. Photoreflection PR experiments were performed but the Franz–Keldysh oscillation,[14] which signifies the presence of an extended interface electric field, is absent in the PR spectra of all the samples. To produce RD resonance of the order of 10^{-2} as observed in our samples, the electric field must be comparable to the one produced by a doping level of the order of $10^{18}/\mathrm{cm}^2$,[16] and should easily be observed as the Franz–Keldysh oscillation in the PR spectra. Furthermore, the electro-optic effect transfers oscillation strength between E_1 and $E_1+\Delta$ transitions,[11] and is likely to do the same to the E_0 and $E_0+\Delta$ transitions, so that the resonances R-1 and R-2 should have opposite signs but comparable amplitudes. It is therefore impossible for a single electric field, extended or confined, to cause such a drastic change of the relative intensities of R-1 and R-2 from 2 to 60. Furthermore, in most of the samples studied so far including the ones with ZnSe layers so thin that the ZnSe layer plays no role in the resultant RD spectra, the two resonances all have the same sign. Since the electro-optic effect is more or less a bulk effect, the RD intensity should depend on the layer thickness for overlayers thinner than 1000 Å, which is contradictory to what has been observed. Extended interdiffusion can be ruled out as RD spectra of samples after 10 h of annealing at 300 °C are virtually the same as the ones before annealing. We therefore conclude that neither the extended nor the confined interface electric field can be counted for the observed two resonances in the RD spectra.

As the possibility of interface electric field is ruled out there left two possible causes for the observed resonances, namely that (i) they are a direct consequence of interface defects such as stacking faults, or (ii) they reflect the intrinsic properties of well ordered II–VI/III–V heterovalent semiconductor interfaces, because bonds at the interface are between Group II and Group V or Group III and Group VI elements which are different from the II–VI bonds in bulk ZnSe or the III–V bonds in bulk GaAs. From the few samples examined by planar-view TEM, we have found no obvious trend between stacking faults or other defects densities and RDS resonance intensity. The mechanism of anisotropic stacking faults or other interface defects is also an open question. However, based on the available data, we

cannot rule out this possibility either. Further work is on the way to examine planar TEM graphs of all the samples, and the results will be duly reported. Here we concentrate our discussion on the second possibility that the two resonances reflect the intrinsic properties of ZnSe/GaAs interface.

We first consider an ideal abrupt As (Ga) terminated interface that only Zn–As (Ga–Se) bonds are present. Due to the presence of these bonds the interface loses its D_{2d} symmetry and becomes anisotropic along the two principal axes [110] and [1$\bar{1}$0].[15] In reality the surface is partially Ga and partially As terminated after the oxide has been removed. With certain treatment prior to ZnSe growth, such as preexposure of Zn (Se) flux, one can create, to a large extent, a Zn(Se)-rich surface, resulting in a Zn–As or Ga–Se bonds dominated interface. Intermixing within several monolayers, which may occur because it is energetically more favored,[13] will change the numbers of Zn–As (Ga–Se) bonds. This is shown:

Ga–As–Ga–As–**Ga**–As–**Zn**–**Se**–Zn–Se–Zn–Se.

After the location of the boldfaced **Zn** atom at the As–Zn interface and the boldfaced **Ga** atom to the left of the As atom are exchanged, the resultant Zn–As bond and the As–Zn bonds cancel each other in RDS because the projection of the two bonds onto the interface plane is perpendicular to each other, while a new Ga–Se bond is created. The same result holds if the intermixing extends to several monolayers. The occurrence of the few-monolayer intermixing is mainly due to its reduction of the interface formation energy,[13] so that once the interface is formed there is no further mixing. This is probably the reason why postgrowth annealing at the growth temperature did not change the RDS spectra. Stacking faults of odd number of monolayers will reduce the total number of effective II–V and III–VI bonds because the bonds in the two adjacent atomic planes separated by a stacking fault are perpendicular to each other so that the contribution to RDS of the bonds in the atomic plane of smaller area cancels part of the bonds in the larger plane, resulting in a reduction of RDS intensity. Interfaces with high stacking faults density will then posses diminishing optical anisotropy, while well ordered interfaces will have large optical anisotropy.

Based on the above argument we then assign the R-1 resonance to a Zn–As bond related interface state, and the R-2 resonance to a Ga–Se bond related interface state. The intensity ratio of these two resonances reflects the relative number of Ga–Se bonds and Zn–As bonds near the interface, and their overall intensity reflects the degree of order, or quality, of the interface. With such assignment we can explain the resonance intensities change with the pretreatment of GaAs substrates.

(1) Samples B1 was kept at 600 °C for 10 min, resulting in an As deficient surface and a Ga–Se bonds rich interface after the over layer was grown while there were few Zn–As bonds. This is consistent with the fact that the R-2 to R-1 intensities ratio is 60 for B1.

(2) Samples B2, B3, and ZnSe-41 were exposed to Se flux prior to growth, resulting in a Ga–Se bond-rich interface. The expose time for ZnSe-41 was much longer than that of B2 and B3, resulting in an almost complete Se covered surface, and an R-2/R-1 intensity ration of 60. The ratio of R-2 to R-1 intensities for B2 and B3 is 15. That the R-1 resonance can still be observed for Se-rich interfaces indicates that there were still some Zn–As bonds near the interface, probably due to intermixing, and incomplete coverage of Se in the case of B2 and B3.

(3) The substrates of samples ZnSe-7, ZnSe-8, ZnSe-66, B4 and B5 were exposed to Zn flux prior to growth, resulting in a somewhat Zn-rich interface. The pregrowth exposure, however, was not enough to form a large Zn coverage at 300 °C, since the growth temperature for a complete layer Zn on GaAs is much lower. The sizable R-2 resonance of these samples indicates that only a fraction of the substrate surface was covered with Zn, and it is much easier for Se to stick to a GaAs surface.

(4) The R-2 to R-1 intensity ratio is about six for samples ZnSe-7 and ZnSe-8 which had no S coverage prior to ZnSe growth, as compared to a ratio of two for samples B2 and B3, indicating that the Zn coverage was less in the former group of samples. This suggests that an S terminated GaAs surface is more receptive to Zn than to Se.

(5) Interface of B2 which was $(NH_3)S_x$ treated is better ordered than that of B3 which was S_2Cl_2 treated. Interface of B5 with S_2Cl_2 pretreatment is better ordered than B4 which was $(NH_3)S_x$ pretreated. This suggests that the stronger Ga–S bonding will produce a better ordered Zn-rich interface, and a less ordered Se-rich interface. This is understandable because for Ga–Se to form it must first break a Ga–S bond. It is made harder if the Ga–S bonds are strong, such as in the case of S_2Cl_2 pretreated surface, and not so hard for a weakly bonded S-terminated surface created by $(NH_3)S_x$ pretreatment. The best ordered interfaces, however, are those samples with no S coverage.

(6) A third resonance, R-3, is observed only for samples B2 through B5, i.e., samples with S at the ZnSe/GaAs interface. It is probably due to the Ga–S bonds.

IV. CONCLUSIONS

We have studied in detail the optical properties of ZnSe/GaAs interface by using reflectance difference spectroscopy. Both the in-plane anisotropy ($\epsilon_x \neq \epsilon_y$) and the off-plane anisotropy ($\epsilon_x \neq \epsilon_z$) are studied. Three resonance lines at 2.70, 3.0, and 3.2 eV have been observed in the in-plane RD spectra. It is confirmed that the observed RD spectra are not due to the ZnSe surface in contact with air, or the bulk effect of either ZnSe epilayer or GaAs substrate, but is rather due to the ZnSe/GaAs interface. The in-plane anisotropy is along the [110] and the [1$\bar{1}$0] principal axes. The dependence of the resonances on interface formation conditions, the results of

the photoreflectance spectroscopy, and the annealing experiments all suggest that the anisotropy is not due to the electro-optic effect resulting from an interface electric field. The results are consistent with the assumption that the observed optical anisotropy is the intrinsic properties of ZnSe/GaAs interface, caused by the III–VI and the II–V bonds at the interface which are different from the III–V bonds of the substrate or the II–VI bonds of the epilayer. The resonance R-1 at 2.7 eV is associated with the interface state of Zn–As bonds, the resonance R-2 at 3.0 eV is associated with the interface state of Ga–Se bonds, and the resonance R-3 at 3.2 eV is associated with the interface state of Ga–S bonds, as the intensity of R-1 increases with the Zn coverage, the intensity of R-2 increases with the Se coverage at the ZnSe/GaAs interface, and R-3 exists only in samples with S at the ZnSe/GaAs interface. Theoretical work along this direction is very much desired.

The resonance line shape changes drastically as the top layer thickness varies. This is well explained by a three-layer system with conventional Fresnel optics and is understood as due to the extra phase factor introduced by the ZnSe top layer. Such phase factor changes the line shapes of the measured RD resonances. Both the real and the imaginary parts of the RD spectra must be measured to make the phase correction.

The application of RDS to semiconductor interfaces buried deeply underneath a surface has revealed interface properties that have never been discovered before. When correlated with well documented initial growth conditions, RDS could very well be a powerful tool in studying interface electronic properties, even in the energy region where the top layer is absorbing. It is hoped that the development of this new experimental technique will stimulate more interest in the theoretical study as well.

ACKNOWLEDGMENTS

This work was supported by the Research Grant Council, Grant No. 210/93E from the Hong Kong Government and the experiments were performed in the Zheng Gu Ru Thin Film Physics Laboratory, the Joyce M. Kuok Laser and Photonics Laboratory, and the William Mong's Semiconductor Cluster Laboratory at the Hong Kong University of Science and Technology.

[1]D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[2]M. Wassermeier, I. Kamiya, D. E. Aspnes, L. T. Florez, J. P. Harbison, and P. M. Petroff, J. Vac. Sci. Technol. B **9**, 2263 (1991).
[3]O. Archer, S. M. Koch, F. Omnes, M. Defour, M. Razeghi, and B. Drevillon, J. Appl. Phys. **68**, 3564 (1990).
[4]J. Jonsson, G. Paulsson, and L. Samuelson, J. Appl. Phys. **70**, 1737 (1991).
[5]T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, J. Vac. Sci. Technol. A **12** 1152 (1994).
[6]V. L. Berkovits and D. Paget, Appl. Phys. Lett. **61**, 1835 (1992).
[7]Z. Yang, I. K. Sou, Y. H. Yeung, and G. K. L. Wong, J. Cryst. Growth (in press).
[8]Z. S. Li, W. Z. Cai, R. Z. Su, G. S. Dong, D. M. Huang, X. M. Ding, X. Y. Hou, and Xun Wang, Appl. Phys. Lett. **64**, 3425 (1994).
[9]Sadao Adachi and Tsunemasa Taguchi, Phys. Rev. B **43**, 9569 (1991).
[10]D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).
[11]S. E. Acosta-Ortiz and A. Lastras-Martinez, Phys. Rev. B **40**, 1426 (1989).
[12]M. S. Yeganeh, J. Qi, J. P. Culver, A. G. Yodh, and M. C. Tamargo, Phys. Rev. B **49**, 11196 (1994).
[13]A. Kley and J. Neugebauer, Phys. Rev. B **50**, 8616 (1994).
[14]H. Shen and Fred H. Pollak, Phys. Rev. B **42**, 7097 (1990).
[15]Paulo V. Santos, P. Etchegoin, M. Cardona, B. Brar, and H. Kroemer, Phys. Rev. B **50**, 8746 (1994).

Zn preadsorption on GaAs(100)2×4 prior to ZnSe growth

S. Heun, J. J. Paggel,[a] S. Rubini, and A. Franciosi[b]
Laboratorio Nazionale Tecnologie Avanzate Superfici e Catalisi dell' Istituto Nazionale, per la Fisica della Materia, Area di Ricerca, Padriciano 99, 34012 Trieste, Italy and Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

(Received 22 January 1996; accepted 24 March 1996)

We report here a systematic study of the adsorption of Zn on GaAs(100)2×4 prior to ZnSe growth. At the typical ZnSe growth temperature of 290 °C, the Zn coverage was found to increase only logarithmically with increasing exposure time, so that only submonolayer coverages of Zn can be absorbed and remain stable on the GaAs substrates. At room temperature, and in general at temperatures below 150–200 °C, a drastic increase in the Zn sticking coefficient makes monolayer and multilayer adsorption possible. However, subsequent annealing at the ZnSe growth temperature removes most of the Zn atoms from the surface, so that only submonolayer coverages are found on the initial growth surface. We discuss the possible role of Zn as an interface control layer between ZnSe and GaAs in the light of our results. © *1996 American Vacuum Society.*

I. INTRODUCTION

Blue-green semiconductor laser diodes capable of continuous wave operation at room temperature (RT) potentially are key devices for use in high density optical recording, full color displays, and many other applications. Most blue-green solid state lasers demonstrated to date are comprised of wide-gap II–VI epilayers grown by molecular beam epitaxy (MBE) on III–V substrates.[1–4]

The lifetime of such devices is currently limited primarily by microstructural mechanisms that generate dark-line defects in the active layer. These defects cause nonradiative recombination and rapidly reduce the gain of the active region to the point where stimulated emission ceases. Evidence exists that the degradation originates from the presence of crystalline defects, primarily stacking faults, that nucleate at the II–VI/III–V heterovalent interface, and give rise to dislocation sources at the intercepts with the active layer.[5,6]

Some evidence points to the initial growth mode of the II–VI epilayer on the III–V substrates as an important factor in determining the stacking fault concentration. Two-dimensional (2D) growth of low-roughness wide-gap II–VI epilayers has been associated with the effect of exposing the GaAs(001) growth surface to a Zn flux prior to II–VI MBE. Such Zn predosing reportedly gives rise to reduced stacking fault densities.[7,8] However, the optimum parameters to be used for this pretreatment, the amount of Zn actually present at the interface, and the microscopic mechanism responsible for defect suppression are still unclear, and no detailed *in situ* study has been conducted for the relevant interfaces, which include ZnSe/GaAs(001), ZnSe/InGaAs(001), and ZnMgSeS/GaAs(001).

Recent studies of ZnSe/GaAs interfaces fabricated by MBE in the presence of a Zn overpressure have also shown that a resulting Zn-rich composition near the interface yields changes in the band offsets,[9,10] suppresses atomic interdiffusion across the interface, and the related deep-level formation,[11,12] and modifies the optical and structural quality of the II–VI epilayer.[13,14] These effects have been tentatively associated with the presence of monolayer amounts of Zn at the interface, yielding a Zn-terminated interface, but the mechanism by which the Zn overpressure determines the Zn-rich interface composition, and the very nature of the local deviation from stoichiometry remain to be determined.

For the above reasons, a systematic investigation of Zn adsorption on GaAs(100)2×4 surfaces as a function of substrate temperature and exposure time seems warranted at this stage. We used MBE protocols successfully employed for the fabrication of blue-green lasers, and examined the results of Zn-predosing by means of monochromatic x-ray photoemission spectroscopy (XPS) and reflection high energy electron diffraction (RHEED) *in situ*.

II. EXPERIMENTAL DETAILS

All of the results presented here were obtained in a facility that includes MBE chambers for the growth of III–V and II–VI semiconductors, and an analysis chamber with monochromatic XPS capabilities. All chambers are interconnected via ultrahigh-vacuum (UHV) transfer modules. The sample surface periodicity was monitored *in situ* by RHEED, with a primary electron energy of 10 keV.

Si-doped ($N_D \sim 1\times10^{18}$ cm^{-3}), 500-nm-thick GaAs(001) buffer layers were grown on *n*-GaAs(100) wafers at 580 °C in the III–V MBE growth chamber following the methods described in Refs. 15 and 16. During the growth, the surfaces exhibited an As-stabilized 2×4 surface reconstruction as determined by RHEED. Both Zn predosing and ZnSe growth were performed in the II–VI MBE chamber. Elemental Zn and Se effusion cells with a Zn/Se beam pressure ratio (BPR) of 1 were employed for ZnSe deposition, with a typical growth rate of 0.36 μm/h at 290 °C. For Zn predosing, the GaAs(001)2×4 surface was exposed to the same Zn flux

[a]Present address: Philipps-Universität Marburg, Fachbereich Physik, Renthof 6, D-35032 Marburg, Germany.
[b]Also with Dipartimento di Fisica, Universita' di Trieste, I-34127 Trieste, Italy.

employed for ZnSe growth, i.e., an equivalent beam pressure of 1.6×10^{-6} Torr as measured by means of a nude ion gauge at the sample position. The Zn flux onto the surface was switched on and off in a few tenths of a second using the cell shutter of the Zn source in the MBE system.

Zn predosing, and on occasion ZnSe regrowth, were performed at a variety of substrate temperature in the RT to 400 °C range. Substrate temperatures were measured using an infrared optical pyrometer in the temperature range above 250 °C and by means of a thermocouple calibrated using known melting points in the temperature range below 250 °C. The estimated error in the sample temperature was of the order of ±10 K.

At selected stages, the growth was interrupted, and the samples were quenched to RT and transferred to the analysis chamber. Desorption of Zn was negligible at temperature below 100–200 °C, so that this treatment minimized unintentional desorption prior to XPS measurements of the Zn coverage. Emission from the Zn $3d$, Zn $2p_{3/2}$, Ga $3d$, and As $3d$ core levels was monitored by means of a monochromatic XPS spectrometer that uses Al $K\alpha$ radiation and a crystal monochromator to achieve a typical energy resolution of 0.7 eV. The photoelectron escape depth employed was $\lambda=15$ Å, having taken into consideration the average photoelectron collection angle.

To monitor the Zn coverage, we used the variation with exposure of the integrated intensities of the different core levels. For exposures performed at RT, we observed an identical exponential attenuation of the Ga $3d$ and As $3d$ emission with increasing exposure, and no detectable change in the Ga $3d$/As $3d$ emission ratio. This would support a layer-by-layer growth mode with constant sticking coefficient. Assuming a layer-by-layer growth mode, the ratio of the integrated photoemission intensity $I(\theta)$ of the Zn and Ga core levels can be used to determine the amount of Zn adsorbed on the surface. In the multilayer coverage range, the coverage θ can be approximately derived from:

$$\theta=\lambda \ln\left(\frac{I_{\mathrm{Ga}}(0)}{I_{\mathrm{Zn}}(\infty)}\frac{I_{\mathrm{Zn}}(\theta)}{I_{\mathrm{Ga}}(\theta)}+1\right), \quad (1)$$

where the $I_{\mathrm{Ga}}(0)$ is the integrated intensity of the Ga core level emission prior to any deposition, and $I_{\mathrm{Zn}}(\infty)$ is the integrated intensity of the Zn core level emission from a thick (quasi-infinite) Zn film.

For substrate temperatures higher than RT, and in particular above 200 °C, the attenuation of the substrate core level emission and the increase in the Zn emission were found to be much slower than those encountered at RT, although no change in the Ga $3d$/As $3d$ intensity ratio was observed at any of the temperatures examined. This is likely to reflect a decrease in the sticking coefficient that gives rise to a finite (submonolayer) saturation coverage. In these conditions, Eq. (1) is not strictly appropriate, since the individual core level intensities would vary linearly with coverage,[17] but we still elected to use Eq. (1) to obtain a rough estimate of the coverage. The resulting equivalent thickness of the Zn films is given in Angstroms throughout the article. Assuming the atomic density in the adsorbed film to be equal to that of bulk Zn, one could also express the film thickness in monolayers (ML; 1 ML=2.48 Å). We caution the reader that in terms of the As (or Ga) atomic density on the GaAs(100) surface, 1 monolayer of Zn atoms would rather correspond to about 1 Å coverage.

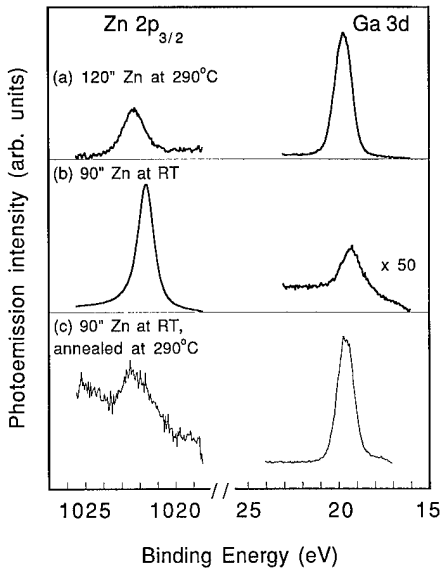


FIG. 1. Angle-integrated photoemission spectra showing the Zn $2p_{3/2}$ (left) and Ga $3d$ (right) emission from GaAs(001)2×4 surfaces exposed to a Zn flux. The zero of the energy scale corresponds to the position of the Fermi level of the spectrometer. Top: spectra recorded following exposure to the Zn flux at 290 °C (typical ZnSe growth temperature) for 120 s, in relative units. Center: exposure to the Zn flux at room temperature (RT) for 90 s. Bottom: exposure to the Zn flux at room temperature (RT) for 90 s, followed by 5 min annealing at 290 °C.

III. RESULTS AND DISCUSSION

Figure 1 illustrates the drastic changes in Zn coverage that correspond to different predosing procedures. In the topmost panel of Fig. 1 we show in relative units angle-integrated photoemission spectra from the Zn $2p_{3/2}$ core level (left) and Ga $3d$ doublet (right) observed after exposure of a GaAs(001)2×4 surface at 290 °C to the Zn flux for 120 s. The relatively small Zn-related emission is in sharp contrast with the dominant Zn $2p_{3/2}$ feature observed following Zn predosing at RT for 90 s (center panel). Multilayer adsorption in this case largely attenuates the substrate emission (right). However, when heating the sample for 5 min at the typical ZnSe growth temperature of 290 °C (bottom-most panel), most of the Zn atoms desorb, and the residual emission intensity in Fig. 1 is similar to that observed following predosing at 290 °C (topmost panel).

A more systematic comparison of the three types of predosing procedures just described is illustrated in Figs. 2–4, using the normalized integrated intensities of the Zn $2p_{3/2}$ and Ga $3d_{5/2}$ core levels, and Eq. (1). In Fig. 2, we show the Zn $2p_{3/2}$ core level photoemission intensity relative to the Ga $3d_{5/2}$ core level photoemission intensity[18] as a function of the Zn exposure time (in a logarithmic scale) at a sample

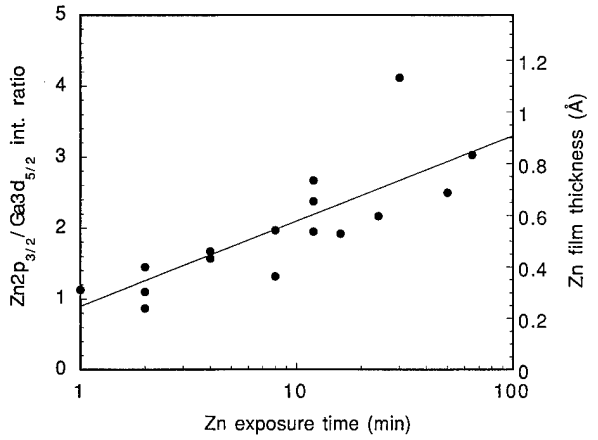


FIG. 2. XPS-derived Zn $2p_{3/2}$/Ga $3d_{5/2}$ core level intensity ratio as a function of the logarithmic Zn exposure time from GaAs(001)2×4 surfaces at 290 °C. The right-hand scale shows the corresponding estimated Zn coverage.

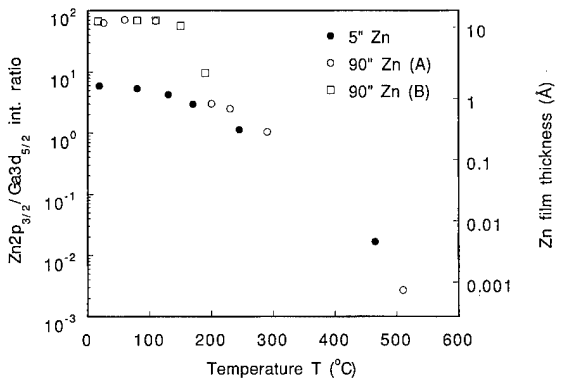


FIG. 4. XPS-derived Zn $2p_{3/2}$/Ga $3d_{5/2}$ intensity ratio from GaAs(001)2×4 surfaces following exposure to a Zn flux at RT followed by 5 min annealing cycles in ultrahigh-vacuum at increasing temperatures. The right-hand scale shows the corresponding estimated Zn coverage. Solid circles: 5 s exposure at RT, yielding an initial Zn coverage of ~0.6 ML (1 ML=2.48 Å). Open symbols: 90 s exposure at RT, yielding an initial Zn coverage of 5 ML. Open circles and squares correspond to two different experimental runs (sample series A and B).

temperature of 290 °C. The right-hand scale shows the corresponding Zn coverage, estimated from Eq. (1). The logarithmic dependence of the Zn film thickness from the exposure time points in the direction of a competition between adsorption and re-evaporation of Zn from the surface during exposure. At the longer exposure times examined (~100 min) θ approaches asymptotically a saturation coverage of the order of 0.7–0.9 Å, i.e., 0.3–0.4 ML in terms of the Zn atomic surface density, or 1 monolayer in terms of the As (or Ga) surface density.

In the exposure range examined in Fig. 2 the RHEED pattern changed from the original 2×4 to a 1×1. First, after 10 min, the 2× streaks faded out and disappeared. Next, after 45 min, the 4× streaks became weaker and disappeared. Finally, after 60 min, a sharp 1×1 pattern was observed. This pattern did not change at least for Zn-exposure times up to 90 min.

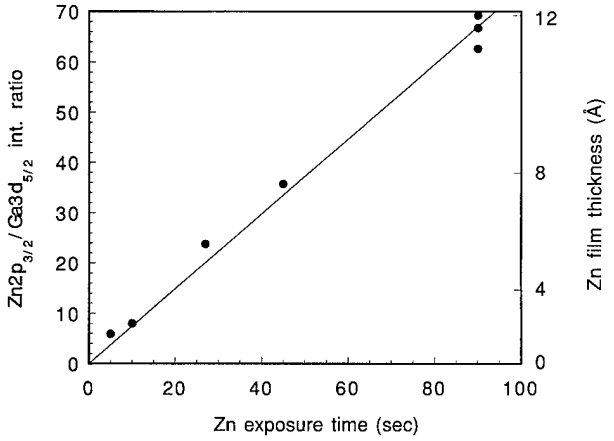


FIG. 3. XPS-derived Zn $2p_{3/2}$/Ga $3d_{5/2}$ core level intensity ratio as a function of the Zn exposure time. The GaAs(001)2×4 substrate was kept at RT. The right-hand scale shows the corresponding estimated Zn coverage.

We emphasize that the relatively large scatter in the data of Fig. 2 is well outside the error margins of the XPS analysis, and reflects true coverage differences from sample to sample. We propose that small, unintended variations of the sample temperature may introduce relatively large variations in coverage. In what follows we will show that for sample temperatures near 290 °C, the Zn coverage indeed depends critically on the sample temperature.

The scatter in the data is dramatically reduced in Fig. 3, that depicts the corresponding exposure-time dependence of the Zn $2p_{3/2}$/Ga $3d_{5/2}$ intensity ratio, following a Zn-predosing at RT. As before, the right-hand scale shows the corresponding Zn coverage, estimated from Eq. (1). At RT, the thickness of the adsorbed Zn layer its found to depend linearly on the exposure time. The implication is that the sticking coefficient of Zn is large and adsorption dominates. Already after a 90 s exposure, a Zn film thickness of 12 Å is obtained, corresponding to a multilayer (~5 ML) coverage.

As explained in the previous section, upon RT exposure the intensity of both the Ga $3d$ and As $3d$ doublets decreases exponentially with identical attenuation length with increasing coverage, suggesting a layer-by-layer growth mode. The RHEED patterns, however, show only mixed indications of the presence of long range order. After a 5 s exposure (θ~0.6 ML), the substrate-related 2×4 reconstruction is still observed, even though it becomes increasingly diffuse. The 2×4 pattern disappears completely following a 10 s exposure (θ~1 ML). For longer exposure times a RHEED pattern comprised of a high background and a superposition of circles and streaks is observed suggesting the presence of epitaxial domains in a polycrystalline matrix.

The monolayer or multilayer Zn adsorbates fabricated at RT on GaAs do not remain stable at typical ZnSe growth temperatures. In Fig. 4 we show that when surfaces on which a 0.6 ML (5 s exposure, solid circles) or 5 ML (90 s exposure, open circles and squares) Zn overlayer was condensed

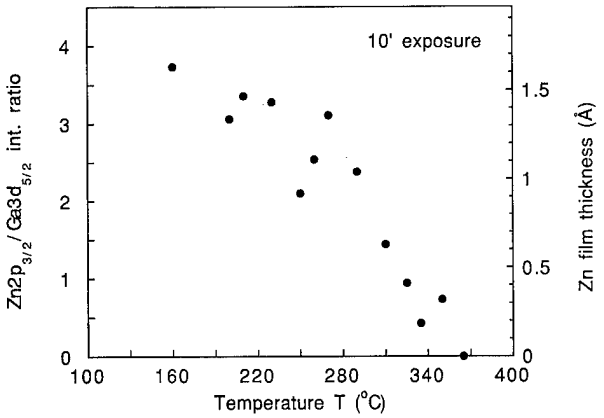


FIG. 5. XPS-derived Zn $2p_{3/2}$/Ga $3d_{5/2}$ intensity ratio from GaAs(001)2×4 surfaces exposed to a Zn flux for the same exposure time of 10 min, but at different temperatures in the 150–400 °C range. The right-hand scale shows the corresponding estimated Zn coverage.

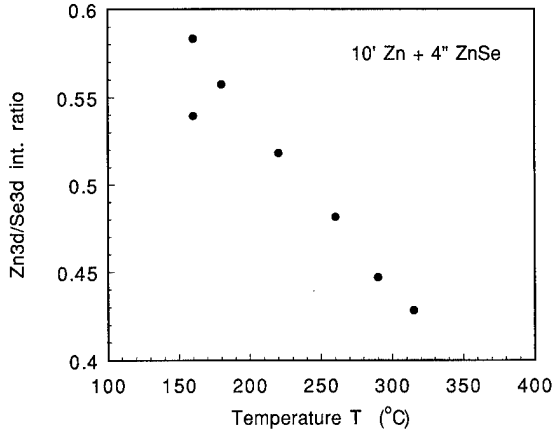


FIG. 6. Zn $3d$/Se $3d$ integrated photoemission intensity ratios from ZnSe/Zn/GaAs structures fabricated by Zn adsorption and ZnSe regrowth on GaAs at the same temperature. The ratio was determined by XPS in the early ZnSe growth stage (3–5 Å) and is plotted in arbitrary units versus the adsorption/regrowth temperature. The apparent intensity ratio reflects the emission from the stoichiometric, thin ZnSe overlayer plus that of the preadsorbed Zn below the overlayer.

at RT are annealed for 5 min cycles in UHV at increasing temperatures, substantial desorption takes place. The evolution of the Zn $2p_{3/2}$/Ga $3d_{5/2}$ intensity ratio for multilayer samples (open symbols) clearly indicate that in the 150–200 °C temperature range the samples revert to the submonolayer coverage stage. The XPS data for the multilayer and submonolayer samples coincide for temperatures above 200 °C, although the corresponding RHEED patterns differ. The initial partially polycrystalline pattern of the multilayer sample is replaced by a 1×1 pattern at about 300 °C, while the diffuse 2×4 initial pattern of the submonolayer sample becomes sharper above 250 °C. Further gradual desorption at higher temperatures makes the Zn $2p$ signal all but disappear at 450–500 °C.

The data in Figs. 2–4 suggests that the fabrication of full monolayers of Zn on GaAs prior to ZnSe regrowth, to explore the effect of Zn-terminated interfaces on the defect density,[7,8] band offsets,[9,10] interdiffusion and deep levels,[11,12] and optical quality of the II–VI overlayers[13,14] will require growth temperatures well below those normally employed for ZnSe. We explored more quantitatively the acceptable growth temperature range by predosing GaAs for a constant exposure time of 10 min at a number of different temperatures in the 150–400 °C range. The results are summarized in Fig. 5. The Zn $2p_{3/2}$/Ga $3d_{5/2}$ intensity ratio in Fig. 5 decreases monotonically with increasing substrate temperature, so that a 10 min exposure time is seen to give rise to a Zn coverage of 0.7 ML at 150 °C, but only ~0.05 ML at 350 °C.[19] We note that the typical ZnSe growth temperature of 290 °C is in a range that corresponds to the strongest dependence of the sticking coefficient on temperature. This may explain the scatter of the data in Fig. 2 as compared to Figs. 3 and 4. Small unintentional temperature variations from sample to sample may be the culprit.

ZnSe regrowth on top of the adsorbed Zn layer was also examined for a variety of temperatures. Elsewhere, we report on studies of ZnSe regrowth at RT, and the effect of the adsorbed Zn layer on the ZnSe–GaAs band offsets.[20] Results for Zn adsorption and ZnSe regrowth in the 150–320 °C temperature range are summarized in Fig. 6. At each temperature, the sample was exposed to the Zn flux for 10 min, and then the Se shutter was opened to start ZnSe growth with BPR=1 at the same temperature. In Fig. 6 we plot the apparent Zn $3d$/Se $3d$ integrated intensity ratio (in arbitrary units) in the early stages of growth (ZnSe coverage 3–5 Å). The intensity ratio reflects the emission from the stoichiometric, thin ZnSe overlayer plus that of the preadsorbed Zn below the overlayer.

The monotonic decrease in the Zn $3d$/Se $3d$ intensity ratio in Fig. 6 with increasing growth temperature primarily reflects the different initial Zn coverage of the predosed surface (see Fig. 5). This figure proves that the adsorbed Zn is not desorbed during the deposition of ZnSe. This is the main reason why increasingly Zn-rich interfaces—in the sense employed in Refs. 9 and 10—are observed with decreasing temperatures in Fig. 6. This illustrates how the results of Figs. 1–5 might be employed to determine the predosing parameters that yield a desired interface composition, and how the same results offer a new approach to engineer ZnSe/GaAs interfaces.

IV. CONCLUSIONS

Adsorption of Zn on GaAs(001)2×4 surfaces prior to ZnSe growth has been examined using different types of predosing procedures. Exposure to Zn was performed at the typical ZnSe growth temperature of 290 °C, at room temperature, at several intermediate temperatures, and through room-temperature deposition followed by annealing. We found that only submonolayer saturation coverages can be obtained at 290 °C, and that in general, full monolayer or multilayer adsorption and stability requires adsorption and processing temperatures below 150–200 °C.

In particular, the predosing procedure (2 min at 280 °C) employed by Kuo *et al.*[7,8] to enhance 2D growth and reduce stacking fault formation at II–VI/III–V interfaces produces Zn coverages of the order of 0.1 ML rather than a full Zn termination of the interface. Any model targeted to explaining the enhanced 2D growth will have to take this into account. In this coverage regime the Zn preadsorbed layer is unlikely to affect atomic interdiffusion,[11,12] or the band offsets.[9,10]

The fabrication of Zn-terminated interfaces using Zn predosing of the III–V substrates requires growth temperatures well below those normally employed to grow high quality II–VI epilayer. Evidence that Zn-rich interface compositions can be achieved instead at 290 °C using high BPR-values in the early growth stages suggests that codeposition of ZnSe and excess Zn is the essential ingredient that makes this alternate procedure work. We speculate that the immediate growth of ZnSe on top of the adsorbed Zn layer is essential to stabilize the Zn atoms at the surface and avoid reevaporation.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the U.S. Army Research Office under Grant Nos. DAAH04-93-G-0319 and DAAH04-93-G-0206, and by the Center for Interfacial Engineering of the University of Minnesota. Two of us (S.H. and J.J.P.) acknowledge support from the Commission of the European Union under the Human Capital and Mobility Programme. We thank L. Sorba for fabricating the GaAs epilayers used in this work. We also acknowledge useful discussions with A. Kahn and R. L. Gunshor.

[1]M. A. Haase, J. Qiu, J. M. DePuydt, and H. Cheng, Appl. Phys. Lett. **59**, 1272 (1991).
[2]H. Jeon, J. Ding, W. Patterson, A. V. Nurmikko, W. Xie, D. C. Grillo, M. Kobayashi, and R. L. Gunshor, Appl. Phys. Lett. **59**, 3619 (1991).
[3]N. Nakayama, S. Itoh, T. Ohata, K. Nakano, H. Okuyama, M. Ozawa, A. Ishibashi, M. Ikeda, and Y. Mori, Electron. Lett. **29**, 1488 (1993).
[4]D. Herve', R. Accomo, E. Molva, L. Vanzetti, J. J. Paggel, L. Sorba, and A. Franciosi, Appl. Phys. Lett. **67**, 2144 (1995).
[5]S. Guha, J. M. DePuydt, J. Qiu, G. E. Höfler, M. A. Haase, B. J. Wu, and H. Cheng, Appl. Phys. Lett. **63**, 3107 (1993).
[6]G. C. Hua, N. Otsuka, S. M. Clark, D. C. Grillo, Y. Fan, J. Han, L. He, M. D. Ringle, R. L. Gunshor, M. Hovinen, and A. V. Nurmikko, Appl. Phys. Lett. **65**, 1331 (1994).
[7]L. H. Kuo, L. Salamanca-Riba, G. Hofler, B. J. Wu, and M. A. Haase, Proc. SPIE **2228**, 144 (1994).
[8]L. H. Kuo, L. Salamanca-Riba, B. J. Wu, J. M. DePuydt, G. M. Haugen, H. Cheng, S. Guha, and M. A. Haase, Appl. Phys. Lett. **65**, 1230 (1994).
[9]R. Nicolini, L. Vanzetti, Guido Mula, G. Bratina, L. Sorba, A. Franciosi, M. Peressi, S. Baroni, R. Resta, A. Baldereschi, J. E. Angelo, and W. W. Gerberich, Phys. Rev. Lett. **72**, 294 (1994).
[10]A. Franciosi, L. Vanzetti, A. Bonanni, L. Sorba, G. Bratina, and G. Biasiol, Proc. SPIE **2346**, 100 (1994).
[11]A. Raisanen, L. J. Brillson, L. Vanzetti, L. Sorba, and A. Franciosi, J. Vac. Sci. Technol. A **13**, 690 (1995).
[12]A. D. Raisanen, L. J. Brillson, L. Vanzetti, A. Bonanni, and A. Franciosi, Appl. Phys. Lett. **66**, 3301 (1995).
[13]A. Bonanni, L. Vanzetti, L. Sorba, A. Franciosi, M. Lomascolo, P. Prete, and R. Cingolani, Appl. Phys. Lett. **66**, 1092 (1995).
[14]L. Vanzetti, A. Bonanni, G. Bratina, L. Sorba, A. Franciosi, M. Lomascolo, D. Greco, and R. Cingolani, J. Cryst. Growth **150**, 765 (1995).
[15]L. Sorba, G. Bratina, A. Antonini, A. Franciosi, L. Tapfer, A. Migliori, and P. Merli, Phys. Rev. B **46**, 6834 (1992).
[16]A. Franciosi, L. Sorba, G. Bratina, and G. Biasiol, J. Vac. Sci. Technol. B **11**, 1628 (1993).
[17]A. Franciosi, D. J. Peterman, J. H. Weaver, and V. L. Moruzzi, Phys. Rev. B **25**, 4981 (1982).
[18]The Ga $3d_{5/2}$ component was routinely obtained from an analysis of the lineshape of the Ga $3d$ emission in terms of two Voigt functions. Analogous results would have been obtained using the overall Ga $3d$ emission.
[19]The RHEED patterns remained 2×4 throughout the temperature range explored in Fig. 5.
[20]See G. Bratina, T. Ozzello, and A. Franciosi, J. Vac. Sci. Technol. B **14**, 2967 (1996).

Electronic properties of ideal and interface-modified metal-semiconductor interfaces

Winfried Mönch[a)]
Laboratorium für Festkörperphysik, Gerhard-Mercator-Universität Duisburg, D-47048 Duisburg, Germany

(Received 22 January 1996; accepted 26 April 1996)

The electronic properties of metal-semiconductor or Schottky contacts are characterized by their barrier heights. At ideal contacts, they are determined by the continuum of metal-induced gap states. Extrinsic interface defects and dipoles as well as interlayers may be present in real contacts and will modify the barrier heights. The zero-charge-transfer barrier height and a slope parameter describe the chemical trends of the barrier heights of Schottky contacts on a specific semiconductor. The zero-charge-transfer barrier heights may be calculated by using the empirical tight-binding method with universal parameters and the width of the dielectric band gaps and the slope parameters are given by the optical dielectric constants of the semiconductors. This is verified by comparison with numerical data from well-established theoretical approaches. Hydrogen doping of metal–diamond and metal–silicon interfaces changes their barrier heights with opposite sign. This is explained by a different orientation of H–C and H–Si interface dipoles. Ag and Pb/Si(111) contacts may be prepared with a "7×7" and a 1×1 interface structure. The difference of their barrier heights is explained by the electric dipole correlated with the stacking fault which is a characteristic of 7×7 reconstructions. Interlayers will also alter the barrier heights of Schottky contacts. Typical examples are SiO_2 and Si_3N_4, i.e., MOS and MNS structures. Previous investigations found their barrier heights to vary as a function of the metals used. The chemical trends of these two data sets are described by the slope parameters predicted from the optical dielectric constants of SiO_2 and Si_3N_4.
© *1996 American Vacuum Society.*

I. INTRODUCTION

The chemical compositions and the atomic arrangements at real semiconductor interfaces may be quite complicated. Therefore, physical explanations of, for example, the barrier heights at metal-semiconductor contacts, which are the main subject of this contribution, have to consider more than just one mechanism.[1] Ideal interfaces, to begin with, are laterally homogeneous, intimate, abrupt, and free of any kind of chemical or structural defects. The band alignment at such ideal interfaces is determined by the continuum of interface-induced gap states.[2,3] Meanwhile, this theoretical conclusion has become widely accepted. These intrinsic interface states at metal-semiconductor interfaces derive from the virtual gap states (ViGS) of the complex semiconductor band structure.[4] At metal–semiconductor interfaces, the metal-induced gap states or, for short, MIGS, are formed by the wave function tails of the metal electrons which decay exponentially into the semiconductor.[2,4,5] This behavior is similar to the exponential tails of the electron wave functions at clean surfaces of solids, i.e., at solid–vacuum interfaces. Below the valence-band maximum of the semiconductor, the wave functions are bulklike on either side of metal–semiconductor contacts.[5]

The continua of interface-induced gap states have branch points where the character of these states changes from mostly valence-band- or donorlike to predominantly conduction-band- or acceptorlike. If the Fermi level is above, coincides with, and drops to below the branch point, the net charge density in interface-induced gap states is negative, vanishes, and becomes positive, respectively. Therefore, these branch points are intuitively called the charge neutrality levels (CNL) of the MIGS. The energy position of the Fermi level at semiconductor interfaces and, what is equivalent, the barrier heights of metal–semiconductor contacts or the band-edge discontinuities in semiconductor heterostructures depend on the occupation of the interface-induced gap states. Space charges are much smaller and may be neglected.

Interfaces are electrically neutral. Therefore, the charge densities on both sides of an interface are of equal size but have opposite sign. A generalization of Pauling's model for small molecules[6] may be used to describe the charge transfer at interfaces by the electronegativity difference of the solids in contact.

The current transport across metal–semiconductor interfaces is governed by their barrier heights.[7] They are the energy distance between the Fermi level W_F and the respective majority-carrier band edge right at the interface, which are the valence-band maximum W_{vi} for p-type and the conduction-band minimum W_{ci} for n-type semiconductors. The MIGS-and-electronegativity model predicts barrier heights to vary as[8]

$$\phi_{Bn}=\phi_{cnl}+S_x(X_m-X_s) \tag{1}$$

and as

$$\phi_{Bp}=\phi^*_{cnl}-S_x(X_m-X_s) \tag{2}$$

a)Electronic mail: w.moench@uni-duisburg.de

on semiconductors doped n- and p-type, respectively. The zero-charge-transfer barrier heights $\phi_{cnl}=W_{ci}-W_{cnl}$ and $\phi^{*}_{cnl}=W_{cnl}-W_{vi}$ result when the difference X_m-X_s between the metal and the semiconductor electronegativities is zero. As a consequence of this, the Fermi level coincides with the charge neutrality level of the MIGS. For one and the same metal–semiconductor pair, the barrier heights ϕ_{Bn} and ϕ_{Bp} add up to the fundamental band gap, i.e.,

$$\phi_{Bn}+\phi_{Bp}=W_g. \tag{3}$$

The MIGS-and-electronegativity model predicts the barrier heights of ideal metal–semiconductor contacts to be determined by the energy position of the charge neutrality level, W_{cnl}, of the MIGS and the slope parameter $S_x=\partial\phi_B/\partial X_m$. Both quantities will be discussed in Sec. II.

At real metal–semiconductor contacts, interface defects, interface dipoles due to foreign atoms or correlated with specific interface structures of substrate atoms, as well as interlayers, have to be considered in addition to the intrinsic MIG states. Examples of such extrinsic mechanisms will be presented and discussed in Secs. III–VI.

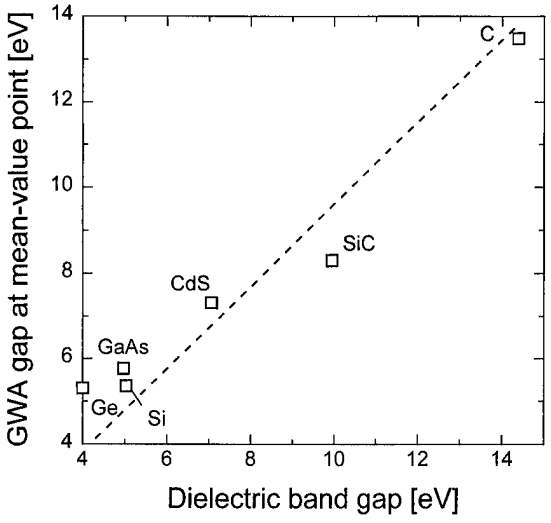


FIG. 1. Width of the band gap at the mean-value point of diamond, Si, Ge, SiC, GaAs, and CdS as calculated in the GW approximation by Rohlfing *et al.* (Refs. 16 and 17) plotted vs the width of the dielectric band gap. The dashed line is a least-squares fit and has a slope parameter of 0.96±0.05.

II. IDEAL INTERFACES: MIGS-AND-ELECTRONEGATIVITY MODEL

A. Zero-charge-transfer barrier height

At their branch point, the MIG states pass over from primarily valence- to predominantly conduction-band character. Tersoff[9,10] calculated the branch points for as many as 15 semiconductors. He defined a cell-averaged real-space Green's function and evaluated it by summing over 152 points in the irreducible wedge of the Brillouin zone. The Green's function changes sign at the branch point of the virtual gap states of the complex band structure. Tersoff calculated the energy bands in the local density approximation (LDA). He adjusted the absolute gaps to the experimental values by rigid shifts of the conduction bands to higher energies. Tersoff's branch point energies $W_{bp}-W_v$ with regard to the valence-band maxima are a most reliable set of data. However, they require a lot of computational effort.

At their branch point the decay length of the virtual gap states is shortest.[9,10] In one dimension, the imaginary part of the ViGS's wave vector passes through its maximum at $V_1^2/4W_1$ below midgap where $2|V_1|$ and W_1 are the widths of the band gap and of the valence band, respectively. With an average or dielectric band gap $W_{dg}\approx 3.7$ eV and an average width $W_1\approx 12$ eV of the valence bands, one estimates $(W_{dg}/2)^2/4W_1\approx 0.04W_{dg}/2$ for three dimensions. The branch point of the ViGS is thus slightly below the midpoint of the dielectric band gap.

Cardona and Christensen[11] took advantage of Baldereschi's concept of mean-value points in the Brillouin zone.[12] For the face-centered cubic lattice, the coordinates of the mean-value point are $\mathbf{k}_{mv}=(2\pi/a_0)(0.6223,0.2953,0)$,[12,13] where a_0 is the lattice parameter. The gap at the mean-value point should have the width of the dielectric band gap in Penn's model.[14] The branch point of the ViGS is then expected to be close to the midpoint of the band gap at the mean-value point of the Brillouin zone. Cardona and Christensen again calculated the energy bands in the local density approximation. Instead of applying the scissors operation, as done by Tersoff, they enlarged the band gap by taking the averages of the top two valence and the bottom two conduction bands at the mean-value point in the Brillouin zone. Their dielectric midgap energies are surprisingly close to Tersoff's branch point energies.

Meanwhile, the LDA band-gap problem has been resolved by using the GW approximation.[15] The width of the band gaps at the mean-value point of GWA band structures and of the dielectric band gap should be equal. The results of respective GWA calculations by Rohlfing *et al.*,[16–18] which are plotted in Fig. 1 as a function of the widths of the respective dielectric band gaps, confirm this conclusion.

GWA band structures need elaborate computations. However, the valence bands are quite well approximated by the empirical tight-binding method (ETB). Figure 2 demonstrates that at the mean-value point, which is of interest here, the GWA[18] and the ETB valence-band energies excellently agree.[19] The valence-band maximum at Γ in the middle of the Brillouin zone was taken as reference. In the ETB calculations, the atomic term values calculated by Fischer[20] and the universal coefficients $\eta_{ll'm}$ of the four matrix elements $V_{ss\sigma}$, $V_{sp\sigma}$, $V_{pp\sigma}$, and $V_{pp\pi}$ as evaluated by Harrison[21] were used. With the group-II chalcogenides, $\eta_{pp\sigma}$ and $\eta_{pp\pi}$ were scaled by a factor of 1.5. These matrix elements vary as

$$V_{ll'm}=\eta_{ll'm}\frac{\hbar^2}{m_0}\frac{1}{d_m^2}, \tag{4}$$

with the nearest-neighbor distance $d_{nn}=a_0\sqrt{3}/4$.

GWA data at the mean-value point are available for the six semiconductors considered in Figs. 1 and 2 only. However, the ETB and the GWA valence-band data are in good agreement. Therefore, the energy position $W_{bp}-W_v(\mathbf{k}_{mv})$ of

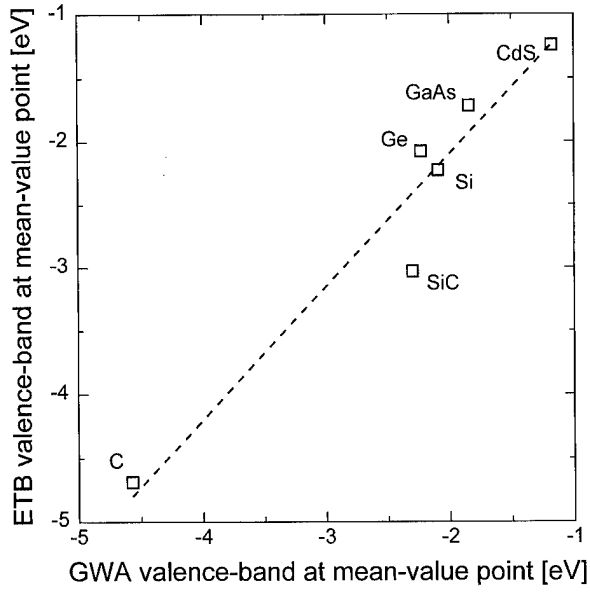


FIG. 2. Position of the top valence band at the mean-value point relative to the valence-band maximum: Comparison of ETB data with GWA data from Rohlfing *et al.* (Refs. 16–18). The dashed line is a least-squares fit and has a slope parameter of 1.05±0.05.

the branch point above the valence band at the mean-value point $\mathbf{k}_{mv}$ may be calculated from Tersoff's $W_{bp}-W_v(\Gamma)$ data[9,10] and the ETB values $W_v(\mathbf{k}_{mv})-W_v(\Gamma)$ at the mean-value point. Figure 3 displays the resulting energy position $W_{bp}-W_v(\mathbf{k}_{mv})$ as a function of the width of the dielectric band gap for all the semiconductors considered by Tersoff. The dashed line is a least-squares fit to the data of the compound semiconductors and has a slope of 0.449±0.008. The branch point in the continuum of the virtual gap states of the

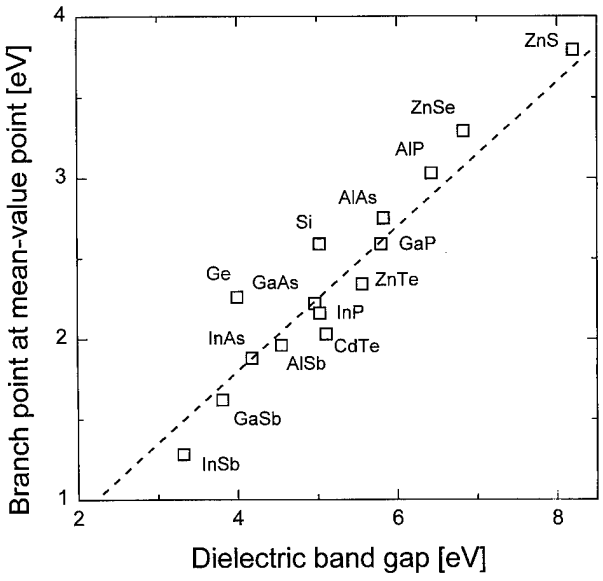


FIG. 3. Position of the ViGS's branch point above the valence band at the mean-value point vs the dielectric band gap. The dashed line is a least-squares fit and has a slope parameter of 0.449±0.008.

complex band structure is thus slightly below midgap position. This finding is explained by the above estimate from the one-dimensional model.

B. Slope parameter

Due to the charge transfer across interfaces, an electric double layer of atomic dimensions exists at metal–semiconductor contacts. The charge density on their semiconductor side is made up of the net charge density in the MIGS continuum

$$Q_{gs}^{mi}=e_0 D_{gs}^{mi}(\mathrm{CNL})(W_{cnl}-W_F)$$
$$=e_0 D_{gs}^{mi}(\mathrm{CNL})(\phi_{Bn}-\phi_{cnl}), \quad (5)$$

since the space charge may be neglected. The density of states $D_{gs}^{mi}(\mathrm{CNL})$ of the MIG states is assumed to be constant around their CNL. The thickness of this intrinsic dipole layer may be estimated as the decay length $1/2q_{gs}^{mi}(\mathrm{CNL})$ of the MIG states at their CNL. The factor 2 in the denominator occurs since the charge distribution rather than the wave functions of the MIGS are considered. The slope parameter $S_x=\partial\phi_B/\partial X_m$ then results from[22,23]

$$A_x/S_x-1=(e_0^2/\epsilon_{ms}\epsilon_0)D_{gs}^{mi}(\mathrm{CNL})/2q_{gs}^{mi}(\mathrm{CNL}), \quad (6)$$

where ϵ_{ms} and ϵ_0 are the interface dielectric constant and the permeability of vacuum, respectively. The parameter $A_x=0.86$ accounts for the use of Miedema's electronegativity scale.[24]

In one dimension, both the density of states and the decay length of the ViGS at their branch point vary inversely proportional to the width of the band gap.[25,26] Penn's nearly free electron model of three-dimensional semiconductors,[14] on the other hand, defines the width of the dielectric band gap W_{dg} by

$$\epsilon_\infty-1=(\hbar\omega_p/W_{dg})^2, \quad (7)$$

where $\hbar\omega_p$ is the plasmon energy of the bulk valence electrons and ϵ_∞ is the optical dielectric constant of the semiconductor. Therefore, the slope parameters S_x were supposed to correlate with the dielectric susceptibility $\epsilon_\infty-1$ of the semiconductors.[26] For several metal-semiconductor contacts the densities of states as well as the charge decay lengths of the MIG states at their charge neutrality level were calculated by using different theoretical approaches.[2,9,10,28,29] According to Eq. (7), Fig. 4 displays a plot of "theoretical slope parameters" $(e_0^2/\epsilon_0)D_{gs}^{mi}(\mathrm{CNL})/2q_{gs}^{mi}(\mathrm{CNL})$ as a function of the experimental susceptibilities $\epsilon_\infty-1$. A least-squares fit to these data gives

$$(e_0^2/\epsilon_0)D_{gs}^{mi}(\mathrm{CNL})/2q_{gs}^{mi}(\mathrm{CNL})=0.29(\epsilon_\infty-1)^{1.91\pm0.24}. \quad (8)$$

The interface dielectric constant ϵ_{ms} was estimated as approximately 2 and 4 for diamond[30] and III–V Schottky contacts.[31] Thus, the theoretical result (8) reproduces the earlier semiempirical rule[29]

$$A_x/S_x-1=0.1(\epsilon_\infty-1)^2. \quad (9)$$

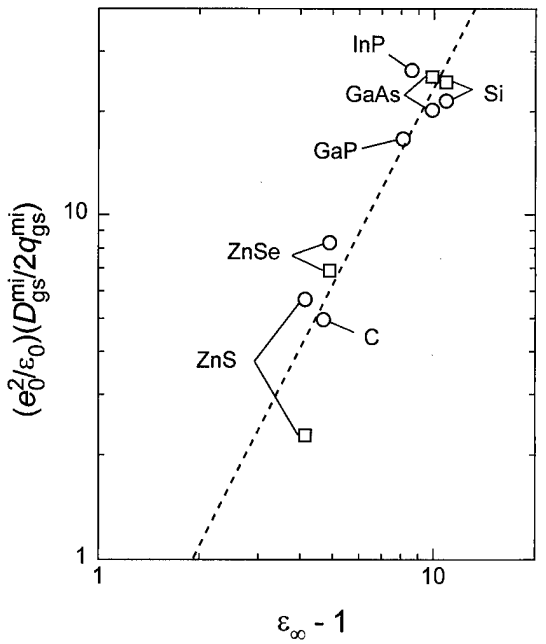


FIG. 4. "Theoretical slope parameters" $(e_0^2/\epsilon_0)D_{gs}^{mi}(\mathrm{CNL})/2q_{gs}^{mi}(\mathrm{CNL})$ as a function of dielectric susceptibilities $\epsilon_\infty - 1$. ○: data from Refs. 2, 9, and 29; □: data from Ref. 28. From Ref. 23.

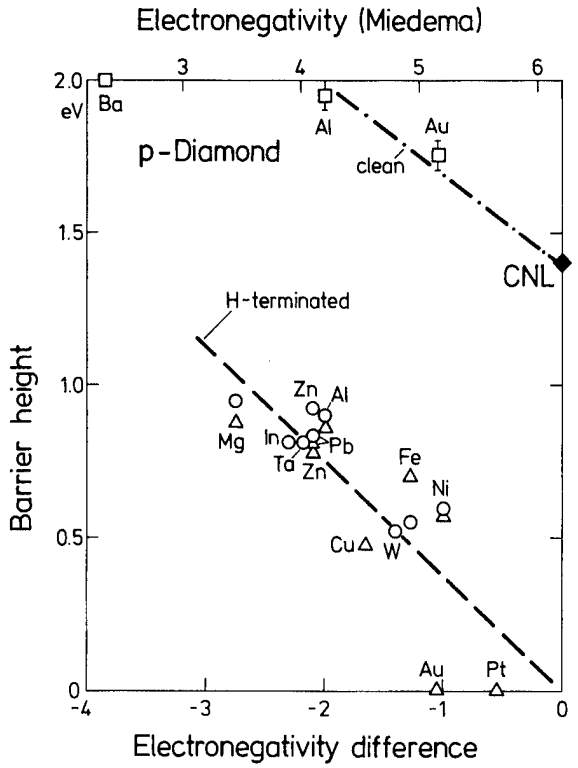


FIG. 5. Barrier heights of p-diamond Schottky contacts as a function of the difference of metal and carbon electronegativities. From Ref. 33.

As will be demonstrated in Sec. IV, relation (9) also predicts the slope parameters for metal contacts on SiO_2 and Si_3N_4.

III. EXTRINSIC INTERFACE DIPOLES

A. Atomic interlayers

Modifications of barrier heights by means of interface doping of Schottky contacts have been frequently tried. As an example, hydrogen interlayers at diamond and silicon Schottky contacts shall be discussed here. Aoki and Kawarada[32] deposited metals on hydrogen-terminated {100}-oriented p-diamond surfaces at room temperature. The barrier heights reported are plotted in Fig. 5 as a function of the difference $X_m - X_C$ between the Miedema electronegativities of the metals used and carbon. The dashed line is a least-squares fit,

$$\phi_{Bp}(\mathrm{H—C}) = 0.017 - 0.37(X_m - X_C) \ \ [\mathrm{eV}],$$

to the experimental data. The dash-dotted line, on the other hand, represents the prediction

$$\phi_{Bp}(\mathrm{C}) = 1.4 - 0.3(X_m - X_C) \ \ [\mathrm{eV}]$$

of the MIGS-and-electronegativity model for ideal diamond Schottky contacts. The CNL position was adopted from Cardona and Christensen[11] and the slope parameter was evaluated by using Eq. (9). This MIGS line reproduces the experimental data for Al and Au contacts with "clean" interfaces while the Ba data point deviates by approximately 0.5 eV.

The MIGS model predicts barrier heights that are larger by approximately 1 eV than what was observed with H-modified diamond Schottky contacts. Mönch[33] explained these findings by H-induced interface dipoles. Recent experimental data obtained with Pb contacts on H-terminated Si(111) surfaces support this interpretation. These latter experimental results will be discussed next.

Figure 6 shows current-voltage (I/V) characteristics of one diode each prepared by evaporation of lead onto a clean n-Si(111)-7×7 and a H-terminated p-Si(111):H-1×1 surface at room temperature.[34] The H-terminated silicon surfaces were prepared by wet chemical etching in buffered hydrofluoric acid with $p\mathrm{H}=9$. Clean 7×7-reconstructed surfaces were obtained by desorption of the hydrogen adatoms from

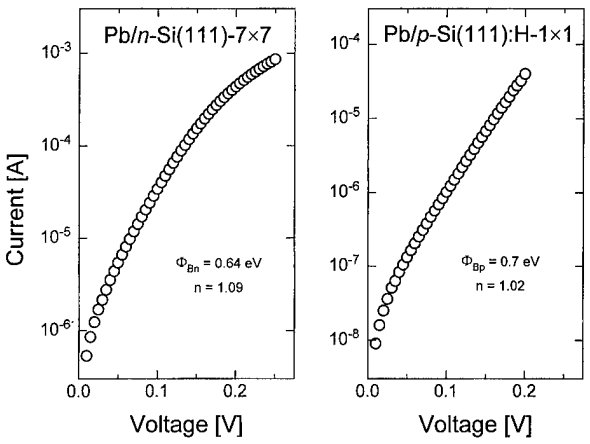


FIG. 6. Current–voltage characteristics of Pb/n-Si(111) and Pb/H/p-Si(111) contacts at room temperature. The diode areas measure 9.3×10^{-3} cm^2. From Ref. 34.

such surfaces at approximately 850 °C. The analysis of the I/V characteristics displayed in Fig. 6 gives barrier heights of 0.64 and 0.7 eV, respectively. By using relation (3) one then obtains barrier heights of 0.48 and 0.42 eV for Pb contacts on p-Si(111)-7×7 and n-Si(111):H-1×1 substrates, respectively. Experimentally, the latter contacts had ohmic I/V curves. Such behavior is expected from simulations for barrier heights lower than 0.5 eV and the increased series resistance which results from the low doping level of the substrates used.

Changes $\delta\phi_B$ of the barrier heights induced by interface doping or, more generally speaking, interface modification, will be of the same size but of different sign on n- and p-type substrates. The barrier heights of interface-modified metal–semiconductor contacts may then be written as

$$\phi_{Bn}^{im}=\phi_{Bn}-\delta\phi_B \tag{10}$$

and

$$\phi_{Bp}^{im}=\phi_{Bp}+\delta\phi_B\,, \tag{11}$$

where ϕ_{Bn} and ϕ_{Bp} denote the barrier heights of the respective "clean" interfaces. The experimental results presented in Figs. 5 and 6 then give

$$\delta\phi_B(\text{H–C})=-1\ \text{eV} \quad \text{but} \quad \delta\phi_B(\text{H–Si})=+0.25\ \text{eV}.$$

The hydrogen-induced changes of the barrier heights are thus of opposite sign at metal–diamond and metal–silicon contacts.

These experimental findings may be easily explained if nearest-neighbor interactions between interlayer (ia) and substrate atoms (s) are considered only. Chemical bonds between monovalent interlayer atoms such as hydrogen and substrate atoms may be described as diatomic "interface molecules." Pauling found a correlation between the partial ionic character of covalent bonds and the difference of the atomic electronegativities. The respective values relevant in the present discussion are $X_{\text{H}}=2.2$, $X_{\text{C}}=2.55$, and $X_{\text{Si}}=1.9$. The H–C and H–Si interface dipoles are thus oppositely oriented. This finding immediately explains the experimentally observed reduction of the barrier heights at H-modified n-silicon but p-diamond contacts.

In a simple point-charge model, the dipole moment of such "interface molecules" may be approximated by

$$p_i=e_0[0.16|X_{ia}-X_s|+0.035(X_{ia}-X_s)^2](r_{ia}+r_s), \tag{12}$$

where r_{ia} and r_s are the covalent radii of the interlayer and the substrate atoms, respectively. The expression in square brackets gives the ionic character of single bonds. It is a revised version[35] of a relation originally proposed by Pauling.[6] The change of barrier heights by such extrinsic dipoles may then be described as the voltage drop across an electric double layer, i.e., by

$$\delta\phi_B=\pm(e_0/\epsilon_{ms}\epsilon_0)p_iN_d\,, \tag{13}$$

where N_d is the area density of the extrinsic interface dipoles. If interfaces are doped with atoms having large covalent radii and then also large polarizabilities, the mutual interaction of the interface dipoles will lower their dipole moment as a function of their area density.[36]

For H-modified diamond and silicon Schottky contacts, relations (12) and (13) give

$$\delta\phi_B(\text{H–C})=-1.86/\epsilon_{ms}\ \text{eV}$$

and

$$\delta\phi_B(\text{H–Si})=+1.00/\epsilon_{ms}\ \text{eV}.$$

As was already mentioned above, the interface dielectric constants were estimated as $\epsilon_{ms}\approx 2$ for diamond[30] and $\epsilon_{ms}\approx 4$ for III–V Schottky contacts.[31] Since the dielectric constant of silicon differs not that much from the values reported for the classical III–V compounds, the interface dielectric constant is also approximated by $\epsilon_{ms}\approx 4$ for silicon Schottky contacts. The H-induced variations of the barrier heights then result as

$$\delta\phi_B(\text{H–C})\approx -1\ \text{eV} \quad \text{and} \quad \delta\phi_B(\text{H–Si})\approx +0.25\ \text{eV}.$$

These estimates are in very good agreement with the experimental data.

Quite recently, Saiz-Pardo *et al.*[37] computed the variations of barrier heights induced by a hydrogen interlayer in K/Si contacts. They used a self-consistent LCAO method supplemented with a local-density prescription to account for many-body effects. The hydrogen atoms were found to form a dipole layer at K/H/Si interfaces and to reduce the barrier height by 0.23 eV with regard to "clean" K/Si contacts. This theoretical result reproduces the experimental observations with Pb/H/Sl(111) contacts and confirms the conclusions from the simple point charge model outlined above.

B. Influence of interface structure on barrier heights

The barrier heights of "clean" metal-semiconductor contacts vary if the interface structures are different. Tung[38] was the first to report on such a correlation. He discovered that the barrier heights of epitaxial $NiSi_2$/Si(111) diodes differ by 140 meV, depending on whether the interfaces are of type A or type B. The lattices are identically aligned on both sides of type-A contacts but they are rotated by 180° around the surface normal for the case of type-B interfaces.[39] Further examples are Ag and Pb contacts on {111}-oriented Si substrates, which will be discussed here.

The atomic structures of buried Ag and Pb/Si interfaces were studied with surface x-ray diffraction.[40,41] First, films evaporated onto clean Si(111)-7×7 surfaces at room temperature will be considered. Each 7×7 unit mesh contains 12 adatoms and a stacking fault in one of its triangular subunits which is surrounded by nine dimers and a corner hole.[42] After evaporation of thick Ag and Pb films, 7th order spots were found to persist. From the diffraction intensities measured with such Ag/Si(111)-"7×7" interfaces, Hong *et al.*[40] concluded that the stacking faults and their associated dimers and corner holes are still present, while the Si-adatoms are

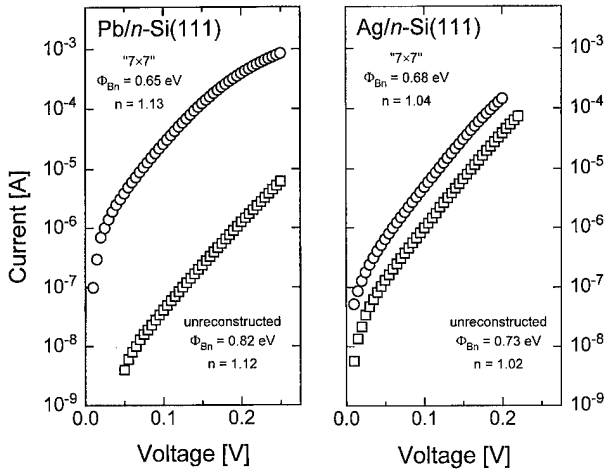


FIG. 7. Current–voltage characteristics of Pb- and Ag/*n*-Si(111) contacts with "7×7" and 1×1 interface structures. The diode areas measure 6.14×10^{-3} cm^2 and 6.6×10^{-3} cm^2, respectively. From Refs. 46 and 45, respectively.

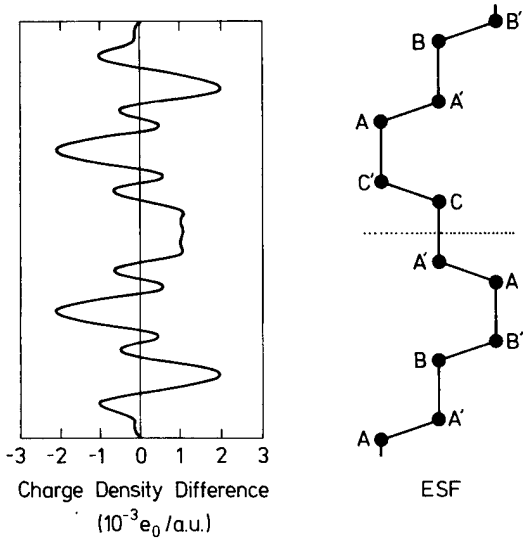


FIG. 8. Integrated difference of the charge density over (111) planes between a Si crystal with a stacking fault and a perfect Si crystal and atomic positions in the (110) plane (1 a.u.$=a_{\mathrm{Bohr}}^3=1.48\times10^{-4}$ nm^3). After Chou *et al.* (Ref. 47).

missing. Similar results were reported by Feidenhans'l *et al.*[43] after the deposition of up to one monolayer of Pb on Si(111)-7×7 surfaces at room temperature.

The Ag/Si(111)-"7×7" interface structure is unstable to annealing at temperatures above 250 °C.[40] With surface x-ray diffraction, no superstructure spots were detected. The resulting contacts are labeled Ag/Si(111)-1×1. Surface x-ray diffraction studies demonstrated that similar Pb/Si(111)-1×1 interfaces exist after the deposition of 100 nm, i.e., thick Pb films onto Si(111):Pb-($\sqrt{3}\times\sqrt{3}$)R30° surfaces at room temperature.[41] Heslinga *et al.*[44] and Schmitsdorf *et al.*[45] measured current–(*I*/*V*) and capacitance–voltage (*C*/*V*) characteristics of Pb and Ag/Si(111) diodes, respectively, with "7×7" and 1×1 interface structures. Irrespective of the metal, the latter interfaces were found to have the larger barrier heights.

Figure 7 displays *I*/*V* characteristics recorded with one Ag and Pb/Si(111) contact each that exhibit a "7×7" and a 1×1 interface structure.[45,46] The clean 7×7-reconstructed surfaces to start with were obtained by hydrogen desorption at approximately 850 °C from Si(111):H-1×1 surfaces prepared by wet chemical etching. Only with such surfaces, thermionic *I*/*V* characteristics as those shown in Fig. 7 were observed. When clean Si(111)-7×7 surfaces were prepared by decomposition of SiO_2 at high temperatures, nonideal *I*/*V* characteristics were observed.[45] Irrespective of the metal used, the barrier heights are larger with diodes exhibiting 1×1 rather than "7×7" interface structures. This observation is attributed to the stacking fault which is a characteristic of the Si(111)-7×7 structure.[45]

Chou *et al.*[47] investigated the electronic properties of intrinsic and extrinsic stacking faults in bulk silicon. The stacking sequences are ...AA′BB′CC′BB′CC′... and ...AA′BB′AA′CC′AA′BB′CC′..., respectively. At such defects the charge densities differ from the charge distribution in the undisturbed bulk. Figure 8 displays the integrated difference of the charge density over the (111) plane between a crystal with an extrinsic stacking fault and a perfect crystal and the atomic positions on the (110) plane. As a whole, stacking faults are electrically neutral but they exhibit electric dipoles.

The stacking sequence ...AA′BB′AA′ is characteristic of the stacking fault in one of the triangular subunits of the 7×7 clean-surface and by this also of the "7×7" interface structure. The computed distribution of the charge density indicates that the electric dipole layer associated with such stacking faults has its negative sign towards the vacuum at clean surfaces and the metal at "7×7" interfaces. These "7×7"-induced interface dipoles are thus oriented the same as the H-induced chemical dipoles considered in the preceding section. This finding immediately explains the experimentally observed increase in barrier height on *n*-type semiconductors when the "7×7"-induced interface dipoles are removed.

The respective change in barrier heights may be estimated by inserting the dipole moment resulting from the theoretical charge density difference plotted in Fig. 8 into relation (13). Numerical integration of these data gives the dipole moment of the ...AA′BB′AA′ stacking fault per unit area as

$$N_{sf}=4.93\times10^{7}e_0 \ \mathrm{m/m^2}.$$

With this value, relation (13) then gives

$$\delta\phi_{Bn}^{sf}=-(e_0/\epsilon_{ms}\epsilon_0)N_{sf}/2=0.446/\epsilon_{ms} \ \mathrm{eV}.$$

The factor $\frac{1}{2}$ accounts for the fact that only one half of the "7×7" unit mesh contains a stacking fault. With an interface dielectric constant $\epsilon_{ms}\approx3$, one obtains

$$\delta\phi_{Bn}^{sf}\approx149 \ \mathrm{meV}.$$

This estimate is close to the difference of 170 meV observed with Pb/Si(111) contacts, but it is larger by approximately

100 meV than what was found with Ag/Si(111) diodes. The latter difference may be attributed to interface defects. This will be discussed next.

The Si(111):Pb-($\sqrt{3}\times\sqrt{3}$)R30° surface structure is due to the arrangement of adsorbed lead atoms on otherwise (almost) ideally terminated Si(111) surfaces.[43,48,49] The removal of this reconstruction and, by this, the formation of a 1×1 interface structure under thick Pb layers thus results from a reordering of the lead atoms at the interface. The coordinates of the Si interface atoms, on the other hand, will experience small shifts parallel and/or normal to the interface only. However, the elimination of the stacking fault and the formation of an ideally terminated 1×1 interface structure during annealing treatments of Ag/Si(111)-"7×7" interfaces needs a supply of eight additional silicon atoms per 7×7 unit mesh. This is equivalent to 16 percent of the atoms in a bulk {111} layer. Residual interface vacancies, for example, might act as interface defects of donor type. They would decrease the barrier height of metal contacts on n-type semiconductors by[50]

$$\delta\phi_B^{id} = -(1 - S_x/A_x)N_{id}/D_{gs}^{mi}(\mathrm{CNL}), \tag{14}$$

where N_{id} is the area density of occupied interface donors. Already approximately 4×10^{13} charged donors per cm^2 suffice to lower the barrier heights of Schottky contacts on n-type silicon by 100 meV. This density corresponds to less than one third of the additional Si atoms needed in the thermally activated conversion of the "7×7" to a 1×1 interface structure at Ag/Si(111) contacts.

As was already mentioned, Heslinga *et al.*[44] discovered that the barrier heights of Pb/Si(111) contacts correlate with the initial reconstruction of the silicon substrate. They deposited thick Pb layers on clean 7×7- and on Pb-($\sqrt{3}\times\sqrt{3}$)R30°-reconstructed n-Si(111) substrates. From the I/V characteristics of their diodes, they evaluated barrier heights of 0.7 and 0.93 eV, respectively. Their and our data agree for Pb/Si(111)-"7×7" interfaces while our values are lower by approximately 60 meV for the Pb/Si(111) interfaces fabricated by evaporation of Pb onto Pb-($\sqrt{3}\times\sqrt{3}$)R30°-reconstructed Si(111) surfaces. This difference is attributed to the nonideal I/V characteristics reported by Heslinga *et al.* Our I/V curves were ideal when we prepared the clean silicon surfaces to start with by hydrogen desorption from Si(111):H-1×1 surfaces processed by HF dips; however, we obtained nonideal characteristics with Si surfaces cleaned by thermal decomposition of SiO_2 at elevated temperatures.[45] Heslinga *et al.* used the latter preparational method in their studies.

IV. METAL-INSULATOR CONTACTS

Metal-insulator-semiconductor (MIS) structures with thin insulating interlayers behave electrically like Schottky contacts. The barrier heights of such devices depend on the nature of the insulator, its thickness, and the specific metal used. Deal *et al.*[51] determined metal–SiO_2 barrier heights in metal–SiO_2–silicon (MOS) structures by using internal photoemission and capacitance-voltage (C/V) measurements. Sobolewski and Helms[52] evaluated Si_3N_4–Si barrier heights

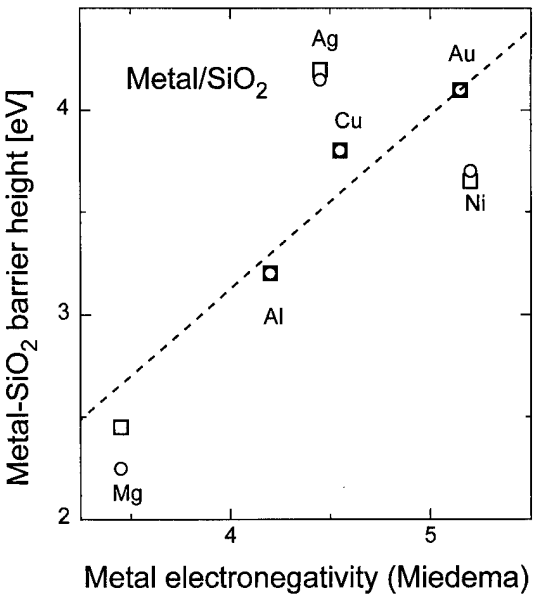


FIG. 9. Metal–SiO_2 barrier heights as a function of the metal electronegativity. The dashed line is a least-squares fit to the data and has a slope parameter of 0.85 eV/Miedema unit. Data from Ref. 51.

from C/V characteristics of metal–Si_3N_4–silicon (MNS) contacts. The nitride thickness varied between 1.6 and 3 nm. Figures 9 and 10 display the MOS and MNS data, respectively. They will be analyzed in the following.

Insulators may be described as semiconductors with large band gaps. Therefore, relation (1) should also apply to metal-insulator contacts. Figure 9 displays the barrier heights at metal–SiO_2 interfaces, which were reported by Deal *et al.*,[51] as a function of the metal electronegativities. The dashed line is a least-squares fit to the data. Its slope amounts to 0.85 ±0.2 eV per Miedema unit. In the infrared, amorphous SiO_2 has an index of refraction $n=1.45$ and one obtains $\epsilon_\infty = n^2 = 2.1$. Relation (9) then predicts a slope parameter $S_x(SiO_2)=0.77$ eV/Miedema unit. This value agrees remarkably well with what resulted from the fit to the experimental data plotted in Fig. 9 although they exhibit a large scatter.

Figure 10(a) displays the barrier heights at the Si_3N_4–Si interfaces of MNS structures as communicated by Sobolewski and Helms.[52] Irrespective of the contact metal, the barrier heights vary linearly as a function of the insulator thickness. The sign of the slope parameters is negative for n and positive for p-type silicon but the least-squares fits to the data give identical absolute values of the slope parameters irrespective of the type of doping and the metal used.

MIS structures may be described by a metal–insulator and an insulator–semiconductor contact in series. Thus, the concept of interface-induced gap states outlined above may be also applied to MIS structures. Figure 11 shows schematically a band diagram of MIS structures in thermal equilibrium. The different energy contributions add up as

$$\phi_{Bn}^{i} = \Delta_i + \Delta W_c^{is} + \phi_{Bn}^{s}. \tag{15}$$

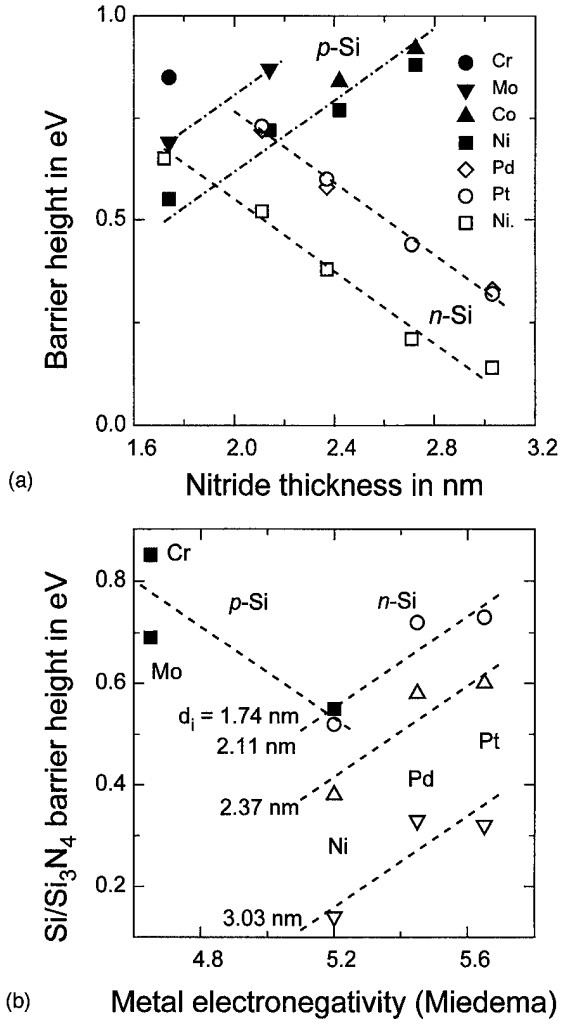


FIG. 10. Barrier heights of metal/Si_3N_4/Si contacts as a function of the nitride thickness for a variety of metals (a) and as a function of metal electronegativity for the same nitride films (b). Data from Ref. 52.

As in semiconductor heterostructures,[3,50] the charge-neutrality levels of an insulator and a semiconductor in contact will align. The conduction-band offsets at such interfaces are then given by

$$\Delta W_c^{si} = \phi_{cnl}^i - \phi_{cnl}^s - D_x, \tag{16}$$

where the last term on the right side accounts for dipole effects due to a nonvanishing difference $X_i - X_s$ of the insulator and semiconductor electronegativities. The potential drop Δ_i across the insulator accounts for a fixed charge:

$$Q_i^f = (\epsilon_i \epsilon_0 / e_0)\Delta_i / d_i, \tag{17}$$

where ϵ_i and d_i are the dielectric constant and the thickness of the insulator. By combining Eq. (15) to Eq. (17) and considering Eq. (1), one obtains

$$\begin{aligned}\phi_{Bn}^s &= (\phi_{Bn}^i - \phi_{cnl}^i) + \phi_{cnl}^s - (e_0/\epsilon_i\epsilon_0)Q_i^f d_i + D_x \\ &= S_{xi}(X_i - X_s) - (e_0/\epsilon_i\epsilon_0)Q_i^f d_i + \phi_{cnl}^s + D_x.\end{aligned} \tag{18}$$

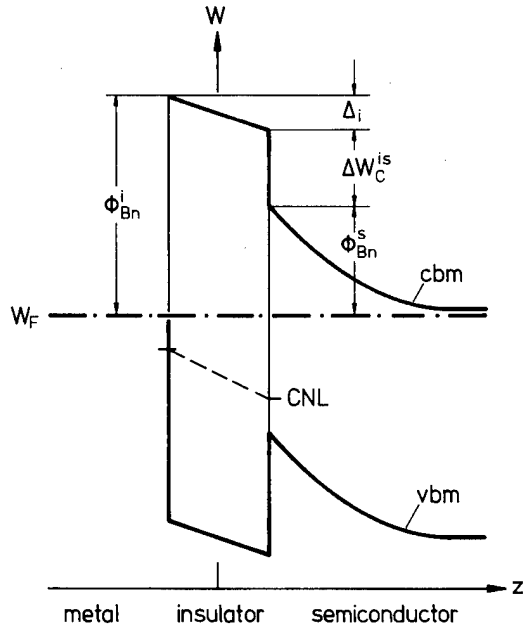


FIG. 11. Band diagram for metal/insulator/semiconductor contacts (schematically).

For p-type semiconductors, the first two terms of (18) have opposite sign.

The barrier height ϕ_{Bn}^s at the insulator-semiconductor interface will thus decrease proportionally to the thickness d_i provided the charge density Q_i^f is fixed. Relation (18) thus explains the linear variation of the barrier height as a function of nitride thickness as plotted in Fig. 10(a). This result is well-known and by no means surprising.

For constant thickness of the insulator, Eq. (18) predicts the barrier heights ϕ_{Bn}^s at the Si_3N_4–Si interfaces to increase proportionally to the electronegativity of the metal. Therefore, the barrier heights displayed in Fig. 10(a) are replotted in Fig. 10(b) versus the electronegativities of the metals with the nitride thickness as a parameter. As to be expected, the barrier heights increase for n- but decrease for p-type doping of the silicon substrates. The straight lines shown are least-squares fits to the data. The average of their slope parameters amounts to $\langle S_{xi}\rangle = 0.45 \pm 0.05$ eV/Miedema unit. The slope parameter S_{xi} is determined by the insulator. In the infrared, Si_3N_4 has an index of refraction $n = 2.05$ so that the optical dielectric constant results as $\epsilon_\infty = n^2 = 4.2$. Relation (9) then predicts a slope parameter $S_x(Si_3N_4) = 0.42$ eV/Miedema unit. This result excellently confirms the value obtained above from the experimental data.

V. SUMMARY

The interface-induced gap states determine the band alignment at ideal semiconductor interfaces. These intrinsic interface states derive from the virtual gap states of the complex band structure of the semiconductor. The chemical bonds at the interfaces establish the charge transfer across the interface and, by this, the net charge in the interface-induced gap states. A generalization of Pauling's concept for

small molecules models the charge transfer by the electronegativity difference of the solids in contact. Two quantities then determine the chemical trends of the barrier heights at metal-semiconductor contacts by the use of the zero-charge-transfer barrier height and a slope parameter. The zero-charge-transfer barrier heights may be estimated from the position of the top valence band at the mean-value point in the Brillouin zone relative to the valence-band maximum and the dielectric band gap. The valence band may be calculated in the empirical tight-binding approximation with universal coefficients for the nearest-neighbor interaction matrix elements. The slope parameters may be directly calculated from the optical dielectric constant of the semiconductor.

In the present article, two types of interface modifications were considered, which are chemical interface doping and differing interface structures. The experimentally observed variations of the barrier heights were explained by extrinsic interface dipoles. They are present in addition to the intrinsic interface dipole layer which results from the metal-induced gap states. Furthermore, interface defects may exist that then absorb charge transferred at the interface. By this, they also effect changes of barrier heights.

Insulators may be regarded as semiconductors with large band gaps. Therefore, the concept of interface-induced gap states not only applies to metal–semiconductor contacts and semiconductor heterostructures but also to metal–insulator and insulator–semiconductor interfaces. Furthermore, layered structures may be described as a sequence of interfaces in series. The thickness of an individual semiconductor or insulator layer has, of course, to exceed the decay lengths of the respective interface-induced gap states.

ACKNOWLEDGMENTS

It is the author's great pleasure to thank Professor J. Pollmann and M. Rohlfing for the computation of the GWA bands at the mean-value point in the Brillouin zone and Dr. A. Mazur for providing the ETB diagonalization program. Without their support some of the concepts presented in this article could not have been pursued.

[1]W. Mönch, Phys. Rev. Lett. **58**, 1260 (1987).
[2]E. Louis, F. Yndurain, and F. Flores, Phys. Rev. B **13**, 4408 (1976).
[3]C. Tejedor and F. Flores, J. Phys. C **11**, L19 (1978).
[4]V. Heine, Phys. Rev. **138**, A1689 (1965).
[5]S. G. Louie and M. L. Cohen, Phys. Rev. B **13**, 2461 (1976).
[6]L. N. Pauling, *The Nature of the Chemical Bond* (Cornell University, Ithaca, NY, 1939/1960).
[7]W. Schottky, Naturwissenschaften **26**, 843 (1938).
[8]W. Mönch, in *Gallium Arsenide Technology*, edited by D. K. Ferry (Sams, Carmel, IN, 1989), Vol. II.
[9]J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[10]J. Tersoff, Surf. Sci. **168**, 275 (1986).
[11]M. Cardona and N. E. Christensen, Phys. Rev. B **35**, 6182 (1987).
[12]A. Baldereschi, Phys. Rev. B **7**, 5212 (1973).
[13]D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).
[14]D. R. Penn, Phys. Rev. **128**, 2093 (1962).
[15]M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[16]M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **48**, 17791 (1993).
[17]M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. Lett. **75**, 3489 (1995).
[18]M. Rohlfing, P. Krüger, and J. Pollmann (private communication).
[19]A. Mazur was so kind as to provide the author with a diagonalization program of the respective 8×8 matrix.
[20]C. Fischer, At. Data **4**, 301 (1972).
[21]W. A. Harrison, Phys. Rev. B **24**, 5835 (1981).
[22]A. W. Cowley and S. M. Sze, J. Appl. Phys. **36**, 3212 (1965).
[23]W. Mönch, Appl. Surf. Sci. (in press).
[24]F. R. de Boer, R. Boom, W. C. M. Mattens, A. R. Miedema, and A. K. Niessen, in *Cohesion in Metals*, edited by F. R. de Boer and D. G. Pettefos (North-Holland, Eindhoven, The Netherlands, 1988), p. 699.
[25]A. Maue, Z. Phys. **94**, 717 (1935).
[26]E. T. Goodwin, Proc. Cambridge Philos. Soc. **35**, 205 (1939).
[27]W. Mönch, in *Festkörperprobleme (Adv. Solid State Phys.)*, edited by P. Grosse (Vieweg, Braunschweig, 1986), Vol. 26, p. 67.
[28]S. G. Louie, J. R. Chelikowski, and M. L. Cohen, Phys. Rev. B **15**, 2154 (1977).
[29]C. Tejedor, F. Flores, and E. Louis, J. Phys. C **10**, 2163 (1977).
[30]J. Ihm, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **40**, 1208 (1978).
[31]R. Ludeke, in *Metallization and Metal-Semiconductor Interfaces*, edited by I. P. Batra (Plenum, New York, 1989), p. 39.
[32]M. Aoki and H. Kawarada, Jpn. J. Appl. Phys. **33**, L708 (1994).
[33]W. Mönch, Europhys. Lett. **27**, 479 (1994).
[34]T. U. Kampen and W. Mönch, Surf. Sci. **331–333**, 490 (1995).
[35]N. B. Hanney and C. P. Smith, J. Am. Chem. Soc. **68**, 171 (1946).
[36]J. Topping, Proc. R. Soc. London Ser. A **114**, 67 (1927).
[37]R. Saiz-Pardo, R. Rincon, and F. Flores, Appl. Surf. Sci. (in press).
[38]R. T. Tung, Phys. Rev. Lett. **52**, 461 (1984).
[39]D. Cherns, G. R. Anstis, J. L. Hutchinson, and J. C. H. Spence, Philos. Mag. A **46**, 842 (1982).
[40]H. Hong, R. D. Aburano, D. S. Lin, H. Chen, and T.-C. Chiang, Phys. Rev. Lett. **68**, 507 (1992).
[41]P. B. Howes, K. A. Edwards, D. J. Hughes, J. E. Mcdonald. T. Hibma, T. Bootsma, and M. A. James, Phys. Rev. B **51**, 17740 (1995).
[42]K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi, Surf. Sci. **164**, 367 (1985).
[43]R. Feidenhans'l, F. Grey, N. Nielsen, and R. L. Johnson, in *Kinetics of Ordering and Growth*, edited by M. Lagally (Plenum, New York, 1990), p. 189.
[44]D. R. Heslinga, H. H. Weitering, D. P. van der Werf, T. M. Klapwijk, and T. Hibma, Phys. Rev. Lett. **64**, 1589 (1990).
[45]R. F. Schmitsdorf, T. U. Kampen, and W. Mönch, Surf. Sci. **324**, 249 (1995).
[46]R. F. Schmitsdorf, M. Gregor, and W. Mönch (to be published).
[47]M. Y. Chou, M. L. Cohen, and S. G. Louie, Phys. Rev. B **32**, 7979 (1985).
[48]E. Ganz, I.-S. Hwang, S. K. Theis, and J. Golovchenko, Surf. Sci. **257**, 259 (1991), and references cited therein.
[49]H. H. Weitering, D. R. Heslinga, and T. Hibma, Phys. Rev. B **45**, 5991 (1992), and references cited therein.
[50]W. Mönch, *Semiconductor Surfaces and Interfaces*, 2nd ed. (Springer, Berlin, 1995).
[51]B. E. Deal, E. H. Snow, and C. A. Mead, J. Phys. Chem. Solids **27**, 1873 (1966).
[52]M. A. Sobolewski and C. R. Helms, Appl. Phys. Lett. **54**, 638 (1989).

Tunable Schottky barriers and the nature of Si interface layers in Al/GaAs(001) diodes

L. Sorba,[a] J. J. Paggel,[b] and A. Franciosi[c]
Laboratorio Nazionale Tecnologie Avanzate Superfici e Catalisi dell'Istituto Nazionale per la Fisica della Materia, Area di Ricerca, Padriciano 99, I-34012 Trieste, Italy and Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

(Received 22 January 1996; accepted 30 April 1996)

Silicon layers grown by molecular beam epitaxy in the interface region of Al/*n*-GaAs(001) Schottky diodes have been shown to tune the Schottky barrier height in the 0.3–1.1 eV range, provided that high enough excess fluxes of As or Al are employed during Si deposition. We studied the incorporation of As and Al in the interface layer for Si growth temperatures of 300 and 400 °C, and for interlayer thicknesses in the 10–150 Å range. Following deposition under an excess As flux we found an As atomic concentration c_{As}=0.4–0.5 at 300 °C and c_{As}=0.2 at 400 °C. The excess Al flux yielded an Al atomic concentration c_{Al}=0.7–0.8 at 300 °C, and c_{Al}=0.6–0.7 at 400 °C. Correspondingly, the lattice structure of the interlayers appears dramatically different from that of Si. We conclude that the 10–100 Å interface layers used by some authors to change the barrier height bear little resemblance to degenerate bulk Si, so that several of the mechanisms proposed to explain Schottky barrier tuning should be reevaluated. © *1996 American Vacuum Society.*

I. INTRODUCTION

Ohmic as well as rectifying contacts have a capital role in solid state electronics. In modern technology, modulation of the properties of metal/semiconductor (MS) junctions is typically achieved by varying the doping of the region of the semiconductor closest to the metallurgical junction to allow a variable degree of tunneling.[1,2] Many authors, however, are continuing the search for reliable methods to engineer the actual Schottky barrier height, with a threefold motivation.[2] First, any procedure that reproducibly changes the barrier height will shed light on the true microscopic mechanism responsible for Fermi level pinning. Second, doping technology is sufficiently advanced only for a very small number of semiconductors. Many new materials (e.g., the wide gap II–VI, III–V, and IV semiconductors employed in short wavelength optoelectronic applications) exhibit doping limitations that severely hinder contact performance. Third, the performance of specific classes of devices may be improved through the incorporation of engineered Schottky barriers. Potential examples include lower leakage currents in metal-semiconductor field-effect transistors, lower dark currents in metal-semiconductor-metal photon detectors, and tunable bandpass for Schottky barrier photon detectors.[3]

In the area of Schottky barrier tuning, several recent reports have emphasized the role of heterovalent semiconductor interface layers, such as Si and Ge in metal/GaAs junctions, fabricated by molecular beam epitaxy (MBE) in the interface region of conventional Schottky diodes. Grant and co-workers initially suggested that the growth of 15–30 Å thick Si interface layers under Ga flux unpinned the Fermi level at Au/*n*-GaAs(001), Cr/*n*-GaAs(001), and Ti/*n*-GaAs(001) contacts, with final Schottky barrier values in the 1.01–1.14 eV range, i.e., considerably higher than those obtained for the same metal/GaAs contacts in the absence of the Si interface layer (0.76–0.89 eV).[4] Studies of the surface position of the Fermi level upon deposition of Si and Ge under As or Ga flux on GaAs(001) surfaces were used to derive Fermi level positions near the bottom of the overlayer conduction bands for overlayers grown under As flux,[5] and predict low Schottky barrier values (~0.3 eV) for hypothetical metal/Si:As/*n*-GaAs MS structures.

Costa *et al.*,[6] fabricated Al/Si/*n*-GaAs(001) MS structures with 6, 30, and 60 Å thick Si layers grown by MBE, and concluded that the barrier decreased from about 1.0 eV to 0.3–0.4 eV with increasing Si thickness. They associated this effect with incorporation of residual As in the Si interface layer.[6] Koyanagi *et al.* also examined Al/Si/*n*-GaAs(001) MS structures grown by MBE, and observed a barrier only some 0.2 eV lower than for the Al/*n*-GaAs(001) case when Si layers 10–40 Å thick were fabricated at the interface under As flux. Conversely, no substantial change of the Schottky barrier relative to the Al/*n*-GaAs(001) case was reportedly observed in Al/Si/*n*-GaAs(001) structures in which Si layers 10–40 Å thick had been grown under Ga flux.[7]

A possible explanation of the above discrepancy was provided by our recent photoemission and transport results from Al/Si/GaAs(001) diodes,[8,9] which showed that the As and Al fluxes required during Si deposition to produce an effect are much larger than those commonly employed for Si doping purposes, i.e., comparable with, or higher than the Si flux, and presumably higher than those employed by Koyanagi *et al.*[7]

Most empirical[6,10] and theoretical models[7,11,12] proposed to explain subsets of the above experimental results consider the electronic structure of the Si interfacial layer as identical to that of bulk Si, require the presence of an As-doped (Ga-doped) n^+ (p^+) degenerate Si layer of sufficient thickness at the interface, e.g., to justify tunneling,[6] depletion charge

[a]Also with Istituto ICMAT del C.N.R., Montelibretti, Italy.
[b]Present address: Philipps-Universität Marburg, Fachbereich Physik, Renthof 6, 35032 Marburg, Germany.
[c]Also with Università di Trieste, Italy.

compensation,[7,10] or metallic screening of the interface states,[12] and assume that the band alignment across the Si–GaAs heterojunction in the Al/Si/GaAs structure is independent of the Si layer thickness. However, the high dopant fluxes required during Si deposition, and evidence that all of the barrier modification is already observed at Si coverages of the order of 1–2 monolayers,[8,9] call into question the role and the very nature of the interface layer. Structural and compositional information about the Si interlayer is woefully scarce. To our knowledge, only in one previous case has the structure of Si epilayers grown on GaAs under As flux been examined by x-ray photoelectron diffraction and shown to be consistent with that of crystalline Si.[10]

Here we focus on the composition and structure of the Si-related species synthesized by MBE on GaAs in conditions homologous to those employed in Refs. 3–11. In particular, we fabricated Si layers at 300 and 400 °C in ultrahigh vacuum (UHV) conditions, as well as under As and Al fluxes comparable with the Si flux. We emphasize that these were the same Si deposition conditions which gave rise to Schottky barrier tunability in Al/Si/GaAs(001) diodes in Refs. 8 and 9. The composition and structure of the layers was examined *in situ* by means of monochromatic x-ray photoemission spectroscopy (XPS) and reflection high energy electron diffraction (RHEED). Preliminary secondary ion mass spectrometry (SIMS) depth profiling studies and transmission electron microscopy (TEM) results lent further support to the surprising conclusions of our *in situ* studies.

II. EXPERIMENTAL DETAILS

All samples were grown by MBE using a facility which includes interconnected growth chambers and an analysis chamber with monochromatic XPS capabilities.[13,14] GaAs buffer layers 5000 Å thick (n-type, $N_D=10^{16}$ cm^{-3}) were first grown at 580 °C on GaAs(100) wafers using the methodology described in Ref. 13. Si was deposited on the buffer layers at 300 or 400 °C in UHV conditions (As partial pressures below 5×10^{-10} Torr), under As flux (As equivalent beam pressures examined were in the $1–5\times10^{-6}$ Torr range), or under Al flux (Al equivalent beam pressures examined were in the $8–10\times10^{-9}$ Torr range). The fluxes were calibrated by exposing a nude ion gauge to the molecular flux at the sample position and no correction was included to take into account the ionization cross sections for the different species. The Si growth rate and atomic flux was determined using the procedure described in Ref. 13. Typical Si growth rate employed was 30 Å/h. The Al growth rate was measured by depositing thick Al films on substrates kept at room temperature, and determining the film thickness with a profilometer. The Al K-cell temperature used in the test was varied between 1000 and 1150 °C. For the growth of Si with Al flux the Al K cell was operated at 950 °C, which corresponds (by linear extrapolation from the measured Al flux values) to an Al flux of $\sim2\times10^{13}$ atoms/cm^2 s. For the growth of Si with As, the As flux employed was $\sim1\times10^{15}$ atoms/cm^2 s which was determined by exposing to the flux a GaAs(001)$c(4\times4)$ surface cooled down to −30 °C.

In principle, RHEED intensity oscillations in MBE can be used to independently determine the Al and As_4 fluxes. The method would involve AlAs growth in cation- or anion-limited conditions, and measurements of the resulting overall compound growth rate from RHEED oscillations. This method yields an estimate of $\sim1\times10^{14}$ atoms/cm^2 s for the Al flux at a K-cell temperature of 1010 °C, and an estimated flux (by linear extrapolation from the measured Al flux values) of 3×10^{13} atoms/cm^2 s at a K-cell temperature of 950 °C. Unfortunately, the same independent flux estimate from RHEED oscillations could not be performed satisfactorily for As.

We observed a $c(4\times4)$ GaAs RHEED pattern at 300 or 400 °C prior to Si deposition. During Si deposition in UHV conditions at 300 and 400 °C, a diffuse 2×1 RHEED pattern became increasingly spotty at Si coverages above the critical thickness (8–10 ML),[13] with an increased background at ~110 ML (150 Å). The nominal Si coverage was determined in all cases from the Si coverage calibration described in Ref. 13. The 2×1 reconstruction is normally observed during Si epitaxy, in the absence of any residual As atmosphere.[15]

During Si deposition at 400 °C under As flux a somewhat diffused 3×1 RHEED pattern became visible at a coverage of about 0.5 ML (1 ML$=6.25\times10^{14}$ atoms/cm^2). The 3×1 RHEED pattern became increasingly diffused at coverages higher than 6 ML, and circles started to appear at ~50 ML. At ~110 ML the dominant presence of circles in the diffraction pattern suggested a polycrystalline structure of the epilayer. The RHEED patterns observed during Si deposition at 300 °C under As flux were similar to those observed during growth at 400 °C, but a polycrystalline-type diffraction pattern was observed at substantially lower coverages (~10 ML).

During Si deposition under Al flux a complex and streaky RHEED pattern was observed for Si growth at 300 and 400 °C and this pattern became relatively spotty with an increased background at 110 ML.

Systematic analyses of the integrated emission intensities from the Si $2p$, Ga $3d$, Ga $3p$, As $3d$, and Al $2p$ core levels and of the valence band line-shape was performed *in situ* by XPS to gauge the nature and composition of the Si-related species. The XPS spectrometer uses Al $K\alpha$ radiation (1486.6 eV) monochromatized and focused by a bent crystal monochromator, and was operated with an overall resolution (electrons and photons) of 0.73 eV and a spot size of 300 μm. The XPS and RHEED measurements were complemented by preliminary cross-sectional TEM measurements, and by SIMS depth profiling studies of thick (110 ML) layers grown under the same conditions.

To compare the nature of the valence states for the Si-related species with that of well-characterized bulk standards, we also fabricated a series of standard samples. A Si reference sample was obtained by depositing a Si epilayer on a Si(001) wafer which was oriented 4° off the [001] along the [110] direction. After chemical removal of the native oxide the substrate was annealed at 900 °C, and a ~10 ML thick Si layer was grown at 500 °C. A double-domain 2×1

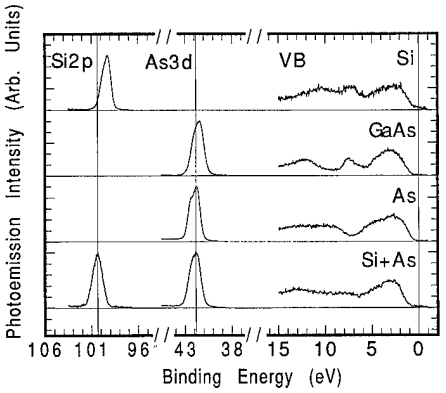


FIG. 1. Si $2p$ (left-hand side), As $3d$ (center) core level, and valence band (right-hand side) emission from bulk Si (top-most panel), GaAs (second top-most panel), elemental As (second bottom-most panel) standards, and from a 110 ML thick Si layer (bottom-most panel) deposited under high As flux at 300 °C onto GaAs(001)$c(4\times4)$. The zero of the energy scale corresponds to the Fermi level of the spectrometer E_F (right-most vertical line). The spectra shown are angle integrated photoelectron energy distribution curves (EDCs) obtained with monochromatic x-ray photoemission spectroscopy (XPS) using Al $K\alpha$ radiation. Left-most and central vertical lines mark the positions of the centroids of the Si $2p$ and As $3d$ doublet in the bottom-most spectra.

RHEED pattern and no detectable contaminants by XPS were observed after growth.

GaAs and AlAs standards were fabricated as epilayers on GaAs buffers. Both the GaAs(001)-2×4 epilayer and the 300 Å thick AlAs(001)3×1 epilayer were grown at 580 °C on a GaAs buffer using the methodology described in Ref. 13.

An elemental As standard was fabricated by condensing a 1 μm thick As layer on a GaAs(001)-$c(4\times4)$ surface cooled down to −30 °C. Diffraction circles in the resulting RHEED pattern suggested a polycrystalline—as opposed to amorphous—structure of the surface, and no contaminants were detected by XPS. The thickness of the resulting adsorbed layer was determined by cross-sectional lateral force microscopy.[16] An elemental Al standard was obtained by depositing 1300 Å of Al on a molybdenum plate at 10 °C. Only a diffuse diffraction intensity was observed by RHEED, suggesting the formation of an amorphous film.

III. RESULTS AND DISCUSSION

In general, the photoelectron valence band and core level emission from the Si-related species deposited under the high As and Al fluxes employed here, and in Refs. 3–11 exhibit large differences as compared to that of bulk Si standards, or of Si epilayers deposited in UHV conditions. In Fig. 1, we compare the valence band and core level emission observed from a 110 ML thick layer deposited under As flux (bottom) at 300 °C with that observed from a variety of standards. The spectra shown correspond to angle-integrated photoelectron energy distribution curves (EDCs) from the Si $2p$ and As $3d$ core level doublets, and the valence band emission. The zero of the energy scale corresponds to the spectrometer Fermi level E_F. The top-most panel in Fig. 1 shows the emission from a bulk Si standard. The three main features in the valence band correspond to the well-known p, sp, and s Si density of states (DOS) features.[17] The position of the centroid of the Si $2p$ core doublet is found 98.94 ±0.08 eV below the linearly extrapolated valence band maximum (E_v), in agreement with the literature.[18] The emission from Si epilayers deposited on GaAs in UHV conditions (not shown) is compellingly similar to that of the bulk standard.

The second topmost panel in Fig. 1 shows the emission from the GaAs standard. The three main features in the valence band, qualitatively similar to those observed for Si, correspond to the well-known As-p, hybrid As-p/Ga-s, and As-s DOS features.[17] The position of the centroid of the As $3d$ doublet is found 40.85±0.08 eV below E_v, in agreement with the literature.[19]

The second bottom-most panel in Fig. 1 shows the emission from an elemental As standard. The valence-band DOS is substantially different from that of Si or GaAs, especially in the region of the p-s hybrid orbitals, and the centroid of the As $3d$ doublet is found 41.45±0.08 eV below the extrapolated leading edge of the valence band. The 0.6 eV shift to higher binding energy relative to the GaAs case is consistent with previous experiments examining monolayer adsorbates of As on GaAs,[20] and reflects the lower ionicity of the chemical bond in going from an As–Ga to an As–As type of coordination. A somewhat larger asymmetry in the core level line shape may reflect an increased metallic character of the epilayer, although we hasten to add that there is no evidence of emission at the E_F in the valence band spectrum.

The emission from the Si-related species formed at 300 °C on GaAs under As flux (bottom-most panel) is dramatically different from that of the bulk Si standard, and qualitatively more similar to what would be expected from a mixed Si–As phase. The centroids of the Si $2p$ and the As $3d$ doublets appear 99.47±0.08 and 41.04±0.08 eV, respectively, below the extrapolated leading edge of the valence band. This would correspond to a 0.53 eV increase in the Si $2p$ binding energy and a 0.41 eV decrease of the As $3d$ binding energy relative to the bulk Si and As standard. The direction and order of magnitude of the shifts supports the formation of a Si–As mixed phase, in view of the relative electronegativities of Si, As, and Ga.

We emphasize that the presence of As impurities incorporated at or below the percent level in a bulk Si crystal would be unlikely to give rise to the large changes in the valence band and core level emission observed when comparing the bottom-most and the top-most spectra in Fig. 1. The comparison argues instead for the formation of a concentrated alloy. The limited width of the core levels also supports the formation of a true alloy, rather than the segregation of separate elemental Si and elemental As phases.

In Fig. 2 we compare the valence band and core level emission observed from a 110 ML thick layer deposited under Al flux (bottom) at 300 °C with that observed from a variety of standards. The spectra shown correspond to angle-integrated EDCs from the Si $2p$ and Al $2p$ core doublets, and the valence band emission. The zero of the energy scale

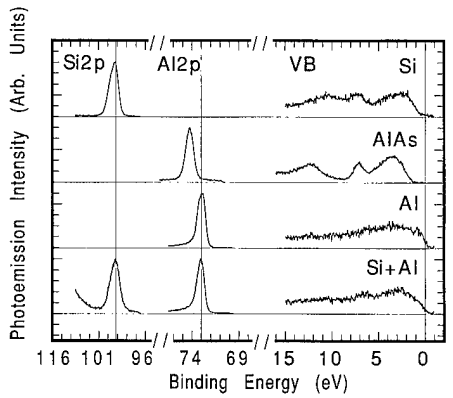


FIG. 2. EDCs for the Si $2p$ (left-hand side), Al $2p$ (center) core level, and valence band (right-hand side) emission from bulk Si (top-most panel), AlAs (second top-most panel), elemental Al (second bottom-most panel) standards, and from a 110 ML thick Si layer (bottom-most panel) deposited under high Al flux at 300 °C onto GaAs(001)$c(4\times4)$. The zero of the energy scale corresponds to the Fermi level of the spectrometer E_F (right-most vertical line). Left-most and central vertical lines mark the positions of the centroids of the Si $2p$ and Al $2p$ doublets in the bottom-most spectra.

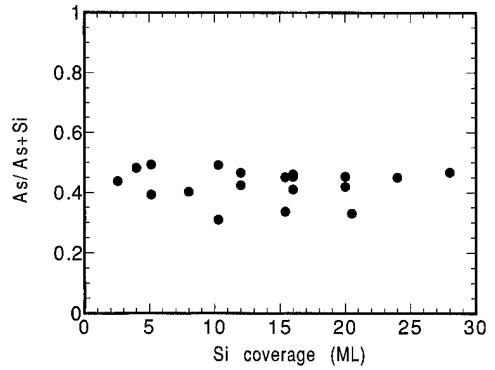


FIG. 3. As/(As+Si) ratio observed during deposition of Si under high As flux at 300 °C onto GaAs(001)$c(4\times4)$ surfaces. The ratio was derived from the integrated As $3d$ emission intensity (and Si $2p$ intensity), subtracting the contribution from the GaAs substrate, so that it is related only to the As atoms incorporated in the overlayer. While the ratio takes into account the relative As $3d$/Si $2p$ relative XPS cross sections, it would correspond to the actual As atomic concentration c_{As} only in the absence of any concentration gradient in the overlayer.

corresponds to E_F. The top-most panel in Fig. 2 shows the emission from a bulk Si standard, as in Fig. 1. The second top-most panel in Fig. 2 shows the emission from an AlAs standard. The three main features in the valence band, qualitatively similar to those observed for Si, correspond to the well-known As-p, hybrid As-p/Al-s, and As-s DOS features.[17] The position of the centroid of the Al $2p$ doublet is found 72.89±0.08 eV below E_v, in agreement with the literature.[19]

The second bottom-most panel in Fig. 2 shows the emission from an elemental Al standard. The well-defined cutoff at E_F and the marked asymmetry of the Al $2p$ core line shape reflect the metallic character of the standard. The centroid of the Al $2p$ doublet is found 72.97±0.08 eV below the Fermi level E_F, consistent with the literature.[8,9]

The emission from the Si-related species formed at 300 °C on GaAs under the Al flux (bottom-most panel) is substantially different from that of bulk Si, as shown by the emission at E_F and the enhanced asymmetry of the Si $2p$.[21] The centroid of the two doublets appear 99.27±0.08 eV and 73.09±0.08 eV, respectively, below the Fermi level. Evidence of hybrid Si–Al states 5–6 eV below E_F in the valence band, and the small increase in the binding energy of the Al $2p$ core levels relative to bulk Al, support the formation of a Si–Al alloys. As was the case for the Si–As mixed phases in Fig. 1, we emphasize that the presence of Al impurities incorporated at or below the percent level in a degenerate Si crystal would not give rise to the large changes in the valence band and core level emission observed when comparing the bottom-most and the topmost spectra in Fig. 2. The comparison argues for the formation of a concentrated alloy.

To follow the formation of the Si-related phases, and gauge their composition, we monitored the attenuation of the As $3d$, Ga $3d$, and Ga $3p$ core level emissions of the GaAs substrate upon Si deposition under the As and Al flux, and the complementary increase of the Si $2p$, Al $2p$, or As $3d$ core level emission as a function of coverage. We used the integrated emission intensities after subtraction of a secondary background multiplied by the tabulated Scofield factors.[22]

Figures 3 and 4 show results for Si deposition under As flux at 300 and 400 °C, respectively. From the experimental overall integrated As $3d$ emission intensity, the contribution associated with the GaAs substrate was subtracted using the experimental Ga $3d$ residual intensity and the known As $3d$/Ga $3d$ intensity ratio from the GaAs substrate. The remaining As $3d$ intensity is assumed to be related to As atoms incorporated in the overlayer. The experimental Si $2p$ intensity and the measured Si $2p$/As $3d$ cross sections were used to derive an As/(As+Si) ratio without any attempt to correct

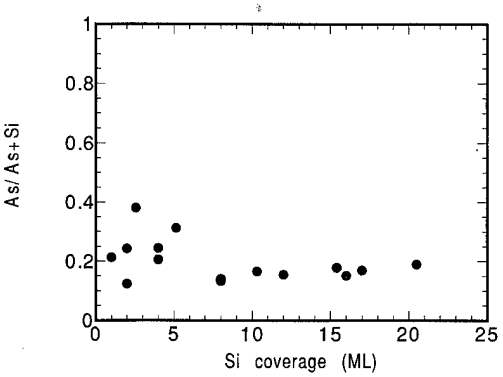


FIG. 4. As/(As+Si) ratio observed during deposition of Si under high As flux at 400 °C onto GaAs(001)$c(4\times4)$ surfaces. The ratio was derived from the integrated As $3d$ emission intensity (and Si $2p$ intensity), subtracting the contribution from the GaAs substrate, so that it is related only to the As atoms incorporated in the overlayer. While the ratio takes into account the relative As $3d$/Si $2p$ relative XPS cross sections, it would correspond to the actual As atomic concentration c_{As} only in the absence of any concentration gradient in the overlayer.

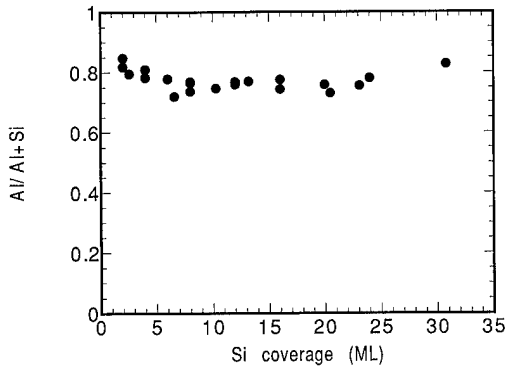


FIG. 5. Al/(Al+Si) ratio observed during deposition of Si under high Al flux at 300 °C onto GaAs(001)$c(4\times4)$ surfaces. The ratio was derived from the experimental overall integrated Al $2p$ and Si $2p$ emission intensity, taking into account the relative Al $2p$/Si $2p$ relative XPS cross sections. However, it would correspond to the actual Al atomic concentration c_{Al} only in the absence of any concentration gradient in the overlayer.

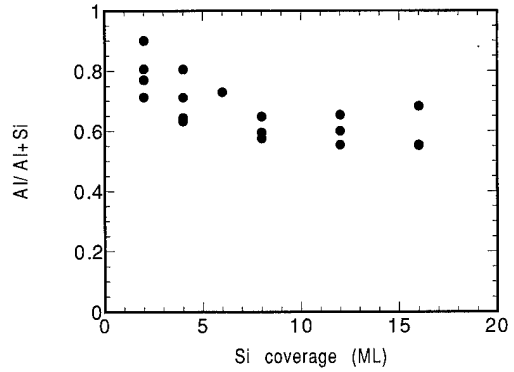


FIG. 6. Al/(Al+Si) ratio observed during deposition of Si under high Al flux at 400 °C onto GaAs(001)$c(4\times4)$ surfaces. The ratio was derived from the experimental overall integrated Al $2p$ and Si $2p$ emission intensity, taking into account the relative Al $2p$/Si $2p$ relative XPS cross sections. However, it would correspond to the actual Al atomic concentration c_{Al} only in the absence of any concentration gradient in the overlayer.

for the actual spatial distribution of Si and As atoms within the overlayer. Therefore this experimental ratio would correspond to the actual As atomic concentration c_{As} only in the absence of any concentration gradient in the overlayer.

The As/(As+Si) ratio in Fig. 3 is seen to remain remarkably constant with increasing coverage, at a value that would correspond to $c_{As}=0.4-0.5$ in a homogeneous alloy. In Fig. 4 the ratio also remains relatively constant with increasing deposition, but at a lower value, consistent with $c_{As}=0.2$. The fact that the As/(As+Si) ratio in Figs. 3 and 4 remains relatively constant throughout the 2–30 monolayer range examined would be consistent with the formation of a relatively homogeneous alloy, and therefore supports the identification between the As/(As+Si) ratio and c_{As}. The formation of an alloy with $c_{As}=0.4-0.5$ would also be quantitatively consistent with the results of Fig. 1. Preliminary SIMS depth profiling results in 110 ML thick samples grown at 300 °C also show a fairly homogeneous As content in the overlayer, with an estimated $c_{As}\sim0.6$. In view of the uncertainties involved in the determination of the XPS and SIMS elemental cross sections, the similarity between the SIMS and the XPS-determined As concentration should be considered very comforting.

We caution the reader that the As equivalent beam pressures employed during Si deposition in the present study and in those of Refs. 8 and 9 ($1-5\times10^{-6}$ Torr) may be substantially higher than those employed in Ref. 10. Chambers and Loebs reportedly employed an As_4 overpressure of $1-5\times10^{-9}$ Torr during Si deposition, and showed by x-ray photoelectron diffraction that the structure of the epilayer was homologous to that of bulk Si. The resulting degenerate Si overlayers were found to change the band bending of GaAs, although no results for Al/Si/GaAs diodes were reported.

Figures 5 and 6 show results for Si deposition under Al flux at 300 and 400 °C, respectively. From the experimental Al $2p$ and Si $2p$ integrated intensity, and the measured Al $2p$/Si $2p$ cross sections, we derived an Al/(Al+Si) ratio without any correction for the actual spatial distribution of Si and Al atoms within the overlayer. As in the previous case, this ratio would correspond to the actual Al atomic concentration c_{Al} only in the absence of any concentration gradient.

The Al/(Al+Si) ratio in Fig. 5 remains constant with increasing coverage above a few monolayers, at a value that would correspond to $c_{Al}=0.7$ in a homogeneous alloy. In Fig. 6 the Al/(Al+Si) shows a somewhat stronger variation with coverage. A saturation value appears to be reached at coverages above 5–6 ML, when the observed value of the ratio would be consistent with $c_{Al}=0.6$ in a homogeneous alloy. The formation of Si–Al alloys with $c_{Al}=0.6-0.7$ would also be quantitatively consistent with the results of Fig. 2.

Unfortunately, the interpretation of SIMS depth profiling results for Si–Al is hindered by substantial knock-in implantation effects, which preclude a quantitative analysis of the Al content in the alloy. We mention, however, that preliminary selected area diffraction results by means of cross-sectional TEM, support a fcc—as opposed to diamond—local crystalline structure for the overlayer, as expected for an Al-rich metallic phase.

Binary equilibrium phase diagrams are only of limited help in the interpretation of the above results. The Si–As bulk equilibrium phase diagram contains at least two stoichiometric compounds (SiAs and $SiAs_2$) in addition to the Si:As solid solution.[23] The equilibrium solubility of As in bulk Si at 300 and 400 °C is much smaller than the values of c_{As} found in Figs. 3 and 4, but can be greatly enhanced through nonequilibrium growth techniques. For example, $c_{As}\sim0.2$ has been observed in Si as a result of ion implantation and subsequent laser annealing.[24] Nonequilibrium solubility might in principle, be consistent with the result of Fig. 4, while the higher value of c_{As} in Fig. 3 could be explained by the presence of a mixed-phase interlayer, possibly involving SiAs and $SiAs_2$ compounds.

The Si–Al bulk equilibrium phase diagram, however, does not contain compounds or solid solutions with c_{Al} large

enough to explain the results of Figs. 5 and 6. The maximum equilibrium solubility of Al in Si is only 0.016 at. % at the eutectic temperature (577 °C).[23] The maximum solubility of Si in Al is 1.5 at. % at the eutectic temperature, and it decreases to 0.05 at. % at 300 °C. Even through nonequilibrium techniques such as splat quenching from the liquid, only c_{Al} values as high as 0.11 were observed.[23] Prior to this study, Al–Si alloys have not been reported to form metastable intermetallics, unlike the case of Al–Ge alloys.[23]

IV. CONCLUSIONS

Growth of Si layers on GaAs at 300 and 400 °C under an As equivalent beam pressure of $1–5\times10^{-6}$ Torr or an Al equivalent beam pressure of $8–10\times10^{-9}$ Torr gives rise to the formation of mixed Si–V and Si–III alloys as opposed to degenerate bulk Si. While RHEED indicates crystallinity of the deposits up to coverages of 10–50 ML, depending on the growth temperature, the atomic structure appears different from that of Si, and group V and III incorporation in the resulting mixed phases may exceed 50%. The trend is towards reduced incorporation with increasing growth temperatures, possibly reflecting a decrease in the sticking coefficient of the groups V and III species relative to Si with increasing temperature, or a decrease in the bulk solubility associated with the formation of phases which are closer to thermodynamic equilibrium. In any case, for both 300° and 400 °C deposition, the resulting intermixed phases bear little resemblance to degenerate bulk Si.

Since fabrication of Si interface layers under an excess group V or group III elemental flux has been used to tune the Schottky barrier height in Al/Si/GaAs(001) diodes, any model which explains the Schottky barrier tunability with the presence of a degenerate Si layer with bulk Si electronic structure, band gap, doping characteristics, as well as equilibrium values of the bulk, unstrained Si–GaAs heterojunction band offsets should be reevaluated taking into account the detail of the growth conditions employed for Si deposition.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the U.S. Army Research Office under Grant Nos. DAAH04-93-G-0319 and No. DAAH04-93-G-0206, by the National Science Foundation under Grant DMR-9116436, and by the Center for Interfacial Engineering of the University of Minnesota. The authors gratefully acknowledge J. E. Angelo for cross-sectional transmission electron microscopy measurements and C. Gerardi for secondary ion mass spectrometry depth profiling studies of our samples. They also acknowledge A. Baldereschi, N. Binggeli, and M. I. Nathan for useful and stimulating discussions.

[1]S. M. Sze, *Physics of Semiconductor Devices* (Wiley-Interscience, New York, 1981).
[2]L. Brillson, in *Basic Properties of Semiconductors*, edited by P. T. Landsberg (North-Holland, Amsterdam, 1992), p. 281.
[3]J. C. Costa, T. J. Miller, F. Williamson, B. A. Bernhardt, and M. I. Nathan, IEEE Trans. Electron. Device Lett. **12**, 324 (1991).
[4]J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **52**, 1794 (1988).
[5]R. W. Grant and J. R. Waldrop, J. Vac. Sci. Technol. B **5**, 1015 (1987).
[6]J. C. Costa, F. Williamson, T. J. Miller, K. Beyzavi, M. I. Nathan, D. S. L. Mui, S. Strite, and H. Morkoç, Appl. Phys. Lett. **58**, 382 (1991).
[7]K. Koyanagi, S. Kasai, and H. Hasegawa, Jpn. J. Appl. Phys. **32**, 502 (1993).
[8]M. Cantile, L. Sorba, S. Yildirim, P. Faraci, G. Biasiol, A. Franciosi, T. J. Miller, and M. I. Nathan, Appl. Phys. Lett. **64**, 988 (1994).
[9]M. Cantile, L. Sorba, P. Faraci, S. Yildirim, P. Faraci, G. Biasiol, A. Franciosi, T. J. Miller, M. I. Nathan, and L. Tapfer, J. Vac. Sci. Technol. B **12**, 2653 (1994).
[10]S. A. Chambers and V. A. Loebs, Phys. Rev. B **47**, 9513 (1993).
[11]H. Hasegawa, M. Akazawa, H. Ishii, and K. Matsuzaki, J. Vac. Sci. Technol. B **7**, 870 (1989).
[12]A. J. Sambell and J. Wood, IEEE Trans. Electron. Devices **37**, 88 (1990).
[13]G. Bratina, L. Sorba, A. Antonini, L. Vanzetti, and A. Franciosi, J. Vac. Sci. Technol. B **9**, 2225 (1991); L. Sorba, G. Bratina, A. Antonini, A. Franciosi, L. Tapfer, A. Migliori, and P. Merli, Phys. Rev. B **46**, 6834 (1992).
[14]G. Bratina, G. Ceccone, A. Franciosi, M. Micovic, F. Tommasini, and J. F. Walker, Vuoto (Italy) **20**, 565 (1990).
[15]S. Heun, J. Falda, and M. Henzler, Surf. Sci. **243**, 132 (1991).
[16]Y. Fan, I. Karpov, G. Bratina, L. Sorba, W. L. Gladfelter, and A. Franciosi, J. Vac. Sci. Technol. B **14**, 623 (1996).
[17]J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1983).
[18]G. P. Schwartz, M. S. Hybertsen, J. Berk, R. G. Nuzzo, J. P. Mannaerts, and G. J. Gualtieri, Phys. Rev. B **39**, 1235 (1989); E. T. Yu, E. T. Croke, D. H. Chow, D. A. Collins, M. C. Phillips, T. C. McGill, J. O. McCaldin, and R. H. Miles, J. Vac. Sci. Technol. B **8**, 908 (1990).
[19]E. A. Kraut, R. W. Grant, J. R. Waldrop, and S. P. Kowalczyk, Phys. Rev. B **28**, 1965 (1983).
[20]G. Le Lay, D. Mao, A. Kahn, Y. Hwu, and G. Margaritondo, Phys. Rev. B **43**, 14301 (1991).
[21]We caution the reader that some of the increased emission on the high binding energy side of the Si $2p$ core line derives from a Ga $3p$ related contribution. Some Ga segregation is always observed during the growth of Si under Al flux, and this effect is enhanced at 400 °C. We associate the segregated Ga with a partial Al–Ga exchange reaction at the surface during the early stages of deposition, that frees Ga atoms.
[22]L. Ley and M. Cardona, *Photoemission in Solids* (Springer, Berlin, 1979).
[23] *Binary Alloy Phase Diagrams*, 2nd ed., edited by H. Okamoto, P. R. Subramanian, and L. Kacprzak (W. W. Scott, Jr., ASM International, Materials Park, OH, 1990), Vol. 1, pp. 211 and 317.
[24]H. D. Geller, G. Gotz, K. D. Klinge, and N. Triem, Phys. Status Solidi A **41**, K171 (1977).

Schottky barrier tuning at Al/GaAs(100) junctions

C. Berthod, J. Bardi, N. Binggeli,[a] and A. Baldereschi
Institut de Physique Appliquée, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

(Received 22 January 1996; accepted 30 April 1996)

Using an *ab initio* pseudopotential approach, we have investigated the electronic structure of ideal Al/GaAs(100) and Al/$Ga_{1-x}Al_xAs$(100) junctions, and the change of the corresponding Schottky barrier height versus the alloy composition x and in the presence of ultrathin group-IV atom interlayers. We find large changes in the Schottky barrier height which agree well with the experimental data. In order to interpret the observed trends we have analyzed the charge density response to chemical substitutions near the junction. This allowed us to extend to metal/semiconductor interfaces a microscopic linear-response theory approach previously employed to interpret band-offset trends at semiconductor heterojunctions. © *1996 American Vacuum Society.*

I. INTRODUCTION

Understanding the microscopic mechanisms which determine the Schottky barrier height in metal/semiconductor junctions is a key issue in solid state physics and technology.[1] It is by now generally accepted that "interfacial" states such as defects[2,3] or metal-induced gap states[4–8] (MIGS) play an important role in the barrier formation. The relative importance, however, of MIGS and defect-related states in determining Schottky barrier trends and the barrier dependence on the interface geometry and composition are issues which are still widely debated, especially in the case of metal contacts to III–V semiconductors such as GaAs.[9–12]

In recent years, there has been a considerable experimental effort targeted to engineering Schottky barriers with the aim of improving the performance or modulating the characteristics of metal/semiconductor devices. One way of engineering Schottky barriers is by altering the bulk semiconductor material, e.g., by using alloys. Schottky barrier heights show, in general, a stronger dependence on the semiconductor material than on the type of metal used for the contact. This type of engineering with bulk perturbations modifies, in principle, also other electronic parameters of the devices such as, e.g., the semiconductor electronic mobility or the band gap. A different approach, which recently yielded encouraging results, is to control the barrier height via artificial modifications of the interface. A number of experimental studies showed, in particular, that the Schottky barrier height at metal/GaAs(100) contacts could be increased/decreased by as much as 0.4 eV by growing thin Si interlayers under an excess group-III/group-V atomic flux.[13,14] The mechanism responsible for these important variations has not yet been established. Various conflicting empirical models based on the *macroscopic* properties of degenerate bulk Si have been proposed to explain different sets of experimental data.[13,14] Very recently, however, Cantile *et al.*[15] have shown that the modified Schottky barrier in Al/GaAs(100) junctions is already established for Si coverages less than 2 monolayers (ML), thus indicating that the tuning has a more *microscopic* origin.

For semiconductor interfaces, a recent microscopic model[16] based on linear-response theory, which treats the interface as a perturbation relative to a reference virtual crystal, successfully explained band-offset chemical trends at semiconductor heterojunctions. In addition, a *microscopic* interface-dipole model[17] explained band-offset variations induced by thin Si or Ge interlayers at III–V/III–V heterojunctions.[18,19] For metal/semiconductor junctions, the complexity of the structure and the lack of precise information on the inhomogeneous screening near the interface hindered the application of similar microscopic models to the analysis of Schottky barrier tuning. The strength of the screening by the metal wave functions of perturbations introduced at or near the interface is a key issue for metal/semiconductor contacts,[5] and a deeper understanding of this screening on the atomic scale is essential to develop models with truly predictive capabilities for Schottky barrier engineering.

In this article we report a first-principles study of the modification of Al/GaAs Schottky barriers with bulk semiconductor $Ga_{1-x}Al_xAs$ alloys and with ultrathin (0–2 ML thick) Si interlayers. Our results for ideal epitaxial junctions show that the experimental alloy composition and Si-doping trends can be fully explained without postulating defect-induced Fermi-level pinning. We explain the observed trends based on an atomic scale analysis of the charge density response to chemical substitutions near the interface. Our study builds upon the linear-response theory work performed for semiconductor heterojunctions,[16] and extends to metal/semiconductor interfaces microscopic models used to interpret band-offset trends with bulk and interfacial perturbations.[16,18]

II. COMPUTATIONAL METHOD

Our calculations were performed within the local-density approximation (LDA) to density functional theory (DFT), using the pseudopotential plane-wave method.[20] We used

[a]Electronic mail: binggeli@irrmasgll.epfl.ch

Troullier–Martins pseudopotentials[21] in the Kleinman–Bylander non-local form.[22] The exchange-correlation functional by Ceperley and Alder was employed.[23] To simulate the isolated Al/GaAs(100) interfaces, we used supercells consisting of 21 layers of GaAs and 7 layers of Al. The polar Al/GaAs(100) interface offers two inequivalent configurations with As-terminated or Ga-terminated GaAs(100), which we both consider in this study. The Al [100] direction was made parallel to the GaAs [100] axis, and the Al fcc lattice was rotated 45° about the [100] axis relative to GaAs cubic lattice in order to satisfy epitaxial relations.[12] The small compressive strain (1%) in the Al in-plane lattice constant was accommodated by a tetragonal elongation (2%) of the Al overlayer, following macroscopic elasticity theory.[24] The interplanar distance at the interface was taken as the average between the interplanar distances in the semiconductor and in the metal, consistently with the result of total energy calculations.

We simulated Al/$Ga_{1-x}Al_xAs$(100) interfaces within the virtual crystal approximation, replacing Ga pseudo-ions with $\langle Ga_{1-x}Al_x\rangle$ pseudo-ions. Si interlayers were introduced at the Al/GaAs(100) junctions by replacing with Si an equal number of Ga and As atoms in the two planes closest to the metal. We thus assumed fully self-compensated Si dopants distributed over two layers, and simulated the doping for coverage $0<x<2$ ML with the virtual crystal approximation using $\langle Ga_{1-x/2}Si_{x/2}\rangle$ and $\langle As_{1-x/2}Si_{x/2}\rangle$ pseudo-ions.

The supercell calculations were performed with a plane-wave kinetic energy cut-off of 16 Ry. For the Brillouin zone integrations we used a (6, 6, 2) Monkhorst–Pack grid,[25] corresponding to 6 special k points, and employed the Gaussian broadening scheme[26] with a full width at half maximum of ~0.1 eV. Convergence tests showed that with the above parameters the numerical accuracy on the barrier height is ~20 meV. Atomic relaxation at the Al/GaAs(100) interfaces was found to have a relatively small effect (~20 meV) on the Schottky barrier height, and was neglected in the present study.[27]

To calculate the Schottky barrier height ϕ_p, we separated the band discontinuity into two contributions:[12,16]

$$\phi_p=\Delta E_p+\Delta V. \tag{1}$$

The band-structure term ΔE_p is the difference between the Fermi energy of the metal, E_F, and the valence-band edge of the semiconductor, E_v, each measured with respect to the average electrostatic potential of the corresponding crystal. This term does not depend on the interface, and was determined from standard bulk band-structure calculations for each of the two materials forming the junction.[28] The second term ΔV is the electrostatic potential lineup across the interface. This term contains all interface-specific features. For Schottky barrier tuning via interfacial perturbations, the change in the potential lineup is the only contribution to the barrier modification. With bulk-semiconductor perturbations such as alloys, instead, both ΔE_p and ΔV contribute to the barrier change. The potential lineup ΔV was derived via

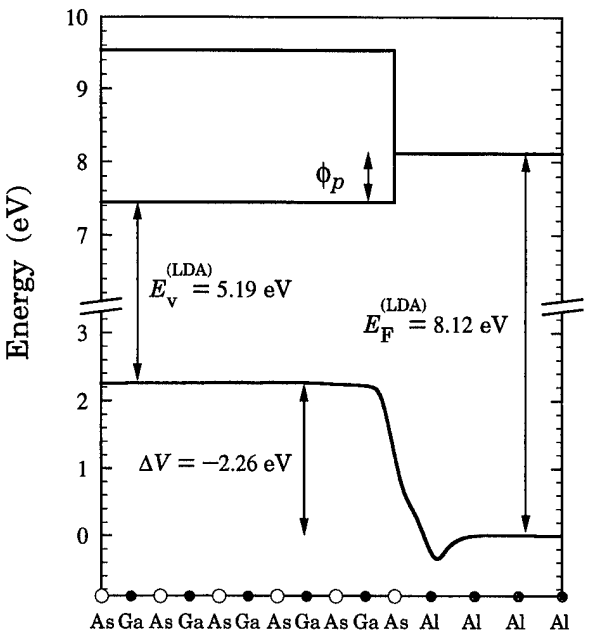


FIG. 1. Macroscopic average of the electrostatic potential in the Al/GaAs(100) As-terminated junction. The potential lineup ΔV and the bulk band energies determining the Schottky barrier are also indicated.

Poisson's equation from the self-consistent supercell charge density using the macroscopic average technique of Ref. 16. In Fig. 1 we show the macroscopic average of the electrostatic potential V and the potential lineup ΔV across the Al/GaAs(100) As-terminated junction. The energy diagram for the calculation of ϕ_p based on Eq. (1) is also shown.[29]

III. RESULTS

A. Unperturbed Al/GaAs(100) interfaces

The calculated LDA p-type Schottky barriers including the GaAs spin–orbit splitting[30] are 0.56 eV for the As-terminated (I) and 0.45 eV for the Ga-terminated (II) Al/GaAs(100) interfaces.[31] These values are about 0.1 eV smaller than the values obtained by Dandrea and Duke in their *ab initio* study of the Al/GaAs(100) junctions.[12] We attribute such differences to the use of different pseudopotentials together with somewhat different atomic geometries.

The potential-lineup contribution to ϕ_p in Eq. (1) is a ground-state property of the junction depending only on its charge density. This term is therefore accurately determined, in principle, within DFT. The LDA band-structure term ΔE_p, instead, should be corrected using quasi-particle energies. The quasi-particle or many-body corrections to the bulk-band structures of Al and GaAs have been evaluated in Ref. 32. Corrections of −0.14 eV and −0.36 eV were reported for the Al Fermi level E_F and for the GaAs valence-band edge E_v, respectively. Using these values to correct the LDA band-structure term ΔE_p, we obtain $\phi_p^{(I)}=0.78$ eV and $\phi_p^{(II)}=0.67$ eV (i.e., $\phi_n^{(I)}=0.64$ eV and $\phi_n^{(II)}=0.75$ eV,

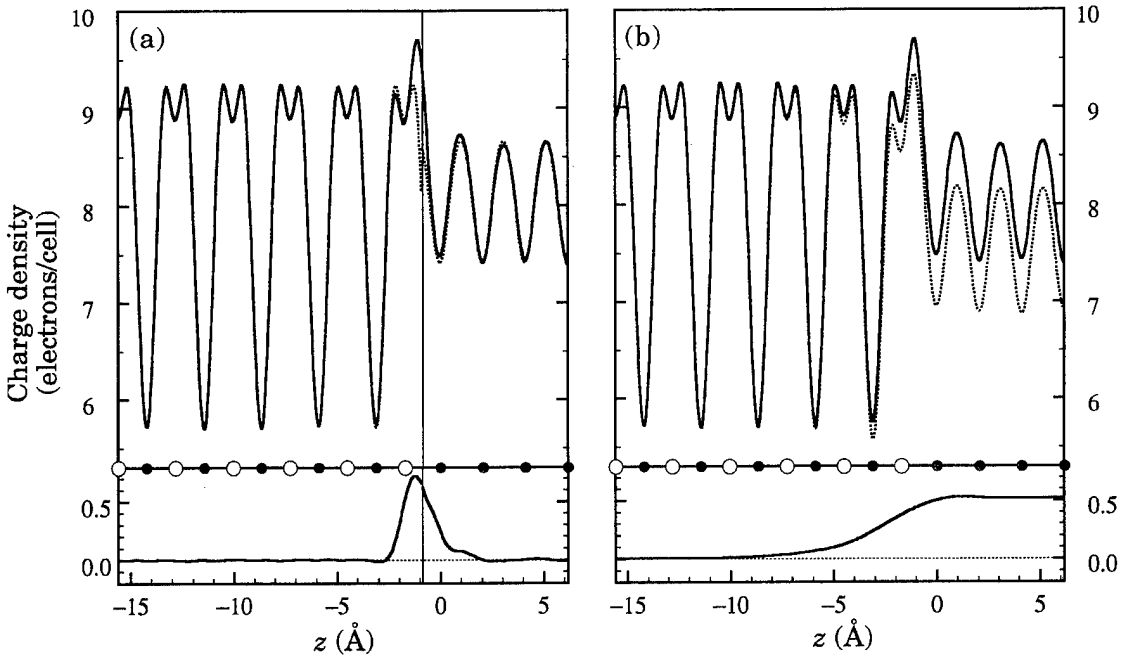


FIG. 2. Planar average of the charge density in the Al/GaAs(100) As-terminated junction (solid line). Comparison with the planar-averaged charge densities of the two bulk materials (dotted line) forming the junction (a). Contribution of the MIGS to the total charge density (b); the dotted line indicates the contribution of the states with energies below the GaAs valence band edge. The lower panels show the macroscopic average of the difference between the interface and bulk charges (a), which somewhat depends on the arbitrary choice of the interface, and the charge density contribution of the MIGS (b).

using the GaAs room-temperature band-gap value of 1.42 eV). The estimated[32] uncertainty on the absolute values of the barriers including many-body corrections is about 0.1 eV.

The higher ϕ_p for the As-terminated interface may be understood in terms of the higher work function of the As- relative to the Ga-terminated GaAs(100) surface. The results for the two interfaces compare well with the measured values of the n-type barriers in epitaxial Al/GaAs(100) junctions with As- and Ga-rich surfaces: 0.72 eV and 0.83 eV in Ref. 33 or 0.66 eV and 0.72 eV in Ref. 34, given the theoretical and experimental uncertainties.

In this work, we will focus on two basic quantities to interpret the Schottky barrier trends with bulk alloy and interlayer perturbations, namely the charge density and the MIGS. Some of their characteristics are illustrated in Fig. 2 for the As-terminated Al/GaAs(100) interface. In Fig. 2(a), we compare the planar average of the charge density in the junction to the planar-averaged charge densities of the bulk GaAs and Al crystals. Inspection of the figure indicates that the charge density in the junction reproduces the bulk semiconductor charge density almost up to the last semiconductor plane. This is an important feature which will allow us to explain (section III C) the "bulk" character of the barrier trend with the semiconductor alloy composition, when the alloying does not affect the chemical composition of the last semiconductor plane. The microscopic bulk behavior of the charge density is related to the local-charge neutrality emphasized by Tersoff[5] in his model of the charge-neutrality level. In Fig. 2(b), we show the contribution of the MIGS to the charge density. These states decay within the semiconductor with a decay length of about 3 Å. The contribution of these states is important to re-establish the bulk semiconductor charge density [see Fig. 2(b)]. The MIGS-related screening will be an important issue in the study of the barrier modification with heterovalent interlayers (section III E).

B. Semiconductor-composition dependence

In Fig. 3, we present the composition dependence of the Al/$Ga_{1-x}Al_xAs$(100) Schottky barrier for the two interfaces

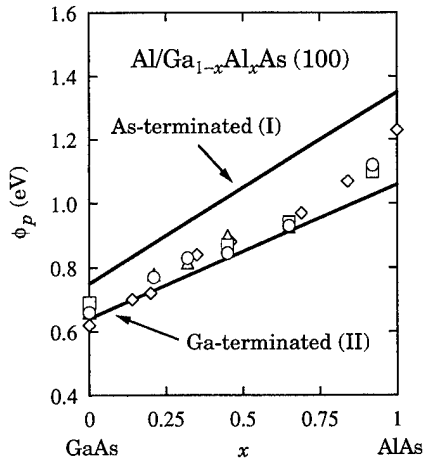


FIG. 3. Composition dependence of the Al/$Ga_{1-x}Al_xAs$(100) Schottky barrier height for the anion- and cation-terminated $Ga_{1-x}Al_xAs$(100) surfaces. The symbols show the experimental data from Ref. 9 [◇ $I(V)$] and Ref. 10 [△ $I(V)$, ○ $C(V)$, □ IPE].

with cation and anion $Ga_{1-x}Al_xAs(100)$ termination. The many-body correction on ϕ_p(Al/AlAs) was included using the quasi-particle correction on ϕ_p(Al/GaAs) of Ref. 32 and the difference in quasi-particle band-edge energies for GaAs and AlAs given in Ref. 35, i.e., ~0.1 eV. A linear interpolation was used for $0<x<1$. The theoretical barriers in Fig. 3 are compared to the experimental current/voltage $I(V)$, capacitance/voltage $C(V)$, and internal photoemission (IPE) data from Refs. 9 and 10. Taking into account the estimated theoretical uncertainly of ~0.1 eV on ϕ_p,[32] and the unknown structural details of the experimental interfaces (with possibly cation- and anion- terminated domains), good general agreement is found between theoretical and experimental trends.

In our calculations, the transitivity rule[36] between the Al/GaAs, Al/AlAs, and GaAs/AlAs band discontinuities is satisfied within 0.03 eV with the type I metal/semiconductor interfaces (As termination), whereas a deviation of −0.21 eV is observed for the type II interfaces (Ga or Al plane closest to the metal). Experimentally the transitivity is approximately verified: the difference ϕ_p(AlAs)−ϕ_p(GaAs) ranges from 0.45 to 0.6 eV, and the valence-band offset VBO(GaAs/AlAs) from 0.45 to 0.55 eV.[16]

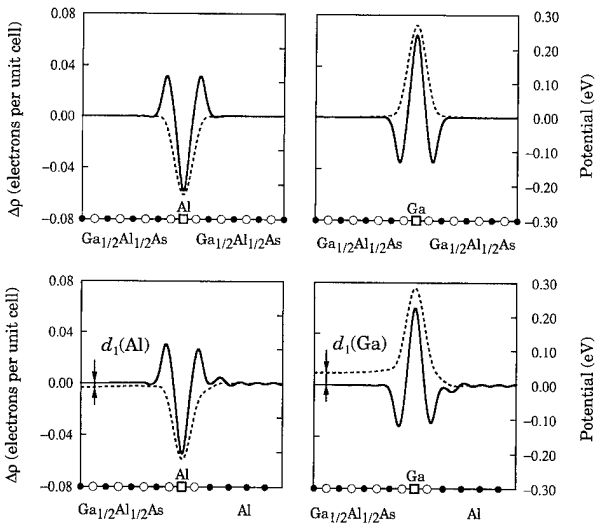


FIG. 4. Charge (solid line) and potential (dashed line) profiles induced by a single Al or Ga plane substitution in the $Al_{1/2}Ga_{1/2}As(100)$ semiconductor homojunction (top panels) and within the semiconductor near the interface in the $Al/Al_{1/2}Ga_{1/2}As(100)$ heterojunction (bottom panels).

C. Interpretation based on linear-response theory

In the case of semiconductor heterojunctions, the linear-response theory model[16] showed that the band offset in isovalent junctions such as GaAs/AlAs is determined only by the bulk properties of the two semiconductor constituents. We will show here that the composition dependence of the Al/GaAlAs Schottky barriers, the transitivity of the Al/GaAs, Al/AlAs, and GaAs/AlAs band discontinuities for the type I interfaces, and the ~0.2 eV deviation from the transitivity rule for the type II interfaces can be explained using a similar approach focusing on charge density response.

Since band terms are bulk properties, our main focus is the potential-lineup term. Within the linear-response approach,[16] the GaAs/AlAs interface is considered as a perturbation with respect to the virtual $Ga_{1/2}Al_{1/2}As$ crystal, and ΔV derives from the charge density $\Delta\rho_{\pm}(z)$ induced by a single cation-plane substitution $v_C=\pm\frac{1}{2}(v_{Al}-v_{Ga})$ changing the virtual ions into real Al (+) or Ga (−) ions, through

$$\Delta V=-4\pi e^2(A_+-A_-)/L, \tag{2}$$

where $A_{\pm}$ is the second moment of the charge,

$$A_{\pm}=\frac{1}{2}\int z^2\Delta\rho_{\pm}(z)dz, \tag{3}$$

and L the distance between consecutive cation planes. The charge profiles $\Delta\rho_{\pm}(z)$ and the corresponding potential profiles, which determine the GaAs/AlAs potential lineup ΔV, are illustrated in Fig. 4 (upper panels). The $\Delta\rho_{\pm}(z)$ reflect the bond polarization associated with the higher electronegativity of Ga relative to Al.

Similarly, starting from a reference $Al/Ga_{1/2}Al_{1/2}As$ system, one may build the Al/GaAs or Al/AlAs interfaces by introducing cation-plane substitutions v_C in the semi-infinite semiconductor region. An ideal linear superposition of semiconductor charge responses $\Delta\rho$ (Fig. 4, upper panel) in the semi-infinite semiconductor region would induce a potential lineup $\Delta V=-4\pi e^2A/L$ across the interface, which is a bulk quantity and satisfies the transitivity rule. In our calculations these bulk terms yield (the quadratic responses cancel out): ΔV_{bulk}(Al/AlAs)−ΔV_{bulk}(Al/GaAs)=0.41 GeV=ΔV(AlAs/GaAs).

The response of the reference $Al/Ga_{1/2}Al_{1/2}As$ system to a substitution of the cation plane closest to the metal, $v_C(1)$, is presented for the type I junction in Fig. 4 (lower panel). The

TABLE I. Dipoles in eV induced by Al (+) and Ga (−) single-plane substitutions in $Al/Ga_{1/2}Al_{1/2}As$ junctions for the first three semiconductor cation planes from the metal (the dipoles are negligible beyond the third plane). The index I (II) refers to the anion- (cation-) terminated interface. ΔV_{corr} is the deviation from the transitivity rule for the band discontinuities ϕ_p(Al/GaAs), ϕ_p(Al/AlAs), and VBO(GaAs/AlAs) as obtained from the *ab initio* calculations.

Interface	d_1^+	d_2^+	d_3^+	d_1^-	d_2^-	d_3^-	$\Sigma_i(d_i^+-d_i^-)$	ΔV_{corr}
I	+0.01	−0.03	+0.00	−0.03	+0.02	+0.01	−0.02	−0.03
II	−0.06	−0.03	−0.02	+0.04	+0.01	+0.01	−0.17	−0.21

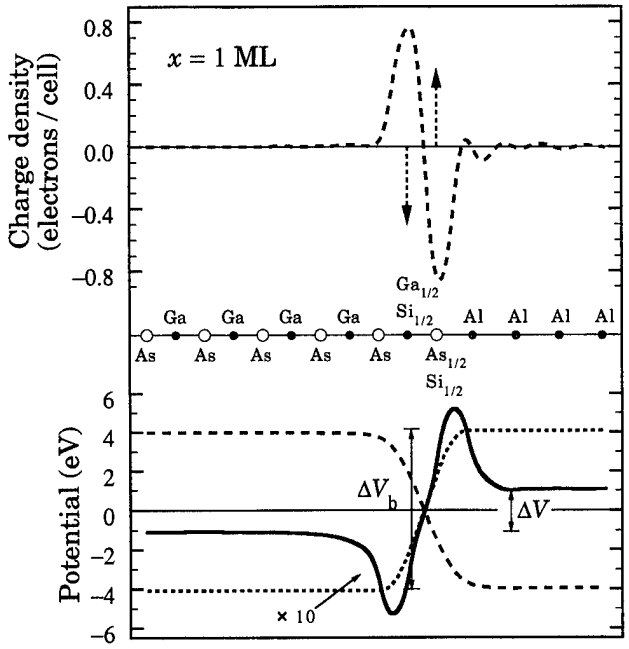


FIG. 5. Planar averages of the ionic (dotted lines) and electronic (dashed lines) charge densities and electrostatic potentials induced by a Si bilayer at 1 ML coverage. The solid line represents the total potential scaled by a factor of 10. The bare ionic, ΔV_b, and screened, ΔV, discontinuities across the Si interlayer are also indicated.

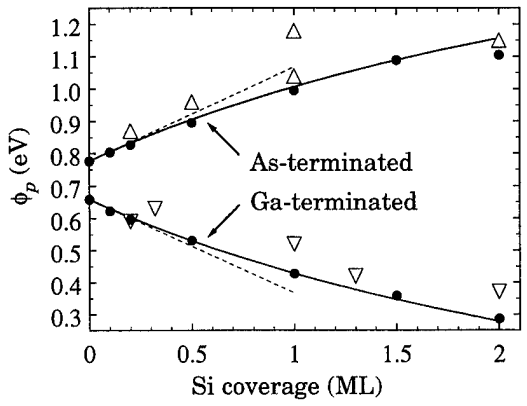


FIG. 6. Schottky barrier height of Al/Si/GaAs(100) diodes as a function of Si coverage. The solid circles are the results of the self-consistent calculations, the open triangles correspond to the experimental data of Ref. 15 for Si grown under As (△) and Al (▽) fluxes. The solid lines give the predictions of our model [see text, Eq. (5)]; the dashed lines indicate the model predictions at low coverage in the linear regime.

presence of the metal can be seen to induce an asymmetry in the response $\Delta\rho_1(z)$ responsible for a dipole d_1. The sign of the dipole d_1 in Fig. 4 may be understood in terms of the higher electronegativity of Ga relative to Al. When the virtual crystal plane is replaced by a Ga (Al) plane the increase (decrease) in electronegativity produces an electronic charge transfer from (to) the metal beyond the second neighboring plane through the MIGS, which accounts for the sign of the observed dipole d_1. To the linear order in the planar perturbations $v_C(i)$, the deviation from the transitivity rule is given by $\Sigma_i(d_i^+ - d_i^-)$, where $d_i^\pm$ are the dipoles induced by the asymmetric $\Delta\rho_i^\pm(z)$ for substitutions $v_C(i)$ in planes close to the interface. These dipoles $d_i^\pm$ are given in Table I.[37] The sum $\Sigma_{i=1}^3(d_i^+ - d_i^-)$ yields -0.02 eV and -0.17 eV for the type I and II interfaces, respectively, and accounts for most of the total deviation ΔV_{corr} from the transitivity rule (-0.03 eV and -0.21 eV).

D. Interfacial perturbation: Heterovalent interlayers

In Fig. 5 we show the ionic and electronic charge densities and electrostatic potentials induced by a Si interlayer (1 ML coverage) on the As-terminated Al/GaAs junction. Si doping of adjacent cation and anion planes in GaAs may be viewed[17] as a proton transfer between substituted As and Ga atoms (indicated by the dashed arrows in Fig. 5). For an arbitrary coverage in the range $0 \leqslant x \leqslant 2$ the bare dipole layer associated with these point charges produces a potential discontinuity $\Delta V_b = \pm(\pi e^2/a)x$ — where a is the GaAs lattice constant[38]—which increases (decreases) ϕ_p at the As- (Ga-) terminated interface. At 1 ML coverage we have $|\Delta V_b| = 8.15$ eV. The electronic contribution, obtained from the self-consistent calculations (see Fig. 5), reduces the discontinuity to $\Delta V = 0.22$ eV for the As-terminated case and $\Delta V = -0.20$ eV for the Ga-terminated case. This corresponds to an effective screening at the junctions $\epsilon_{eff} = \Delta V_b/\Delta V \approx 40$ for 1 Si ML.

In Fig. 6 we present our results for the Schottky barrier height of the two As- and Ga-terminated interfaces as a function of Si coverage together with the measured[15] Al/Si/GaAs Schottky barriers for Si grown under excess As or Al flux. The difference between the barrier heights of the As- and Ga-terminated interfaces increases with Si coverage, and becomes ~ 0.8 eV for $x = 2$ ML, in excellent agreement with the experimental data. The agreement between theory and experiment in Fig. 6 supports the contention by Cantile *et al.*[15] that an excess anion (cation) flux may tend to favor a cation–anion (anion–cation) site sequence for the Si layers. This interpretation is also consistent with expected trends from free energy considerations for the As- and Ga-terminated GaAs(100) surfaces at high As (Ga) chemical potential, i.e., in the presence of excess As (Ga).

The effective screening ϵ_{eff} changes from about 30 at low Si coverages to about 50 at 2 ML coverage. This large variation in ϵ_{eff} accounts for the non-linear coverage dependence of ϕ_p in Fig. 6. The screening of the dipole created by Si interlayers analogously grown at GaAs/GaAs homojunctions has been explained in terms of the GaAs dielectric constant (at low coverages) and of the Si dielectric constant (at ~ 2 ML coverage).[18] Here ϵ_{eff} is 3 to 5 times larger than the GaAs and Si dielectric constants.[39] There is therefore a drastic enhancement of the screening related to the presence of the metal.

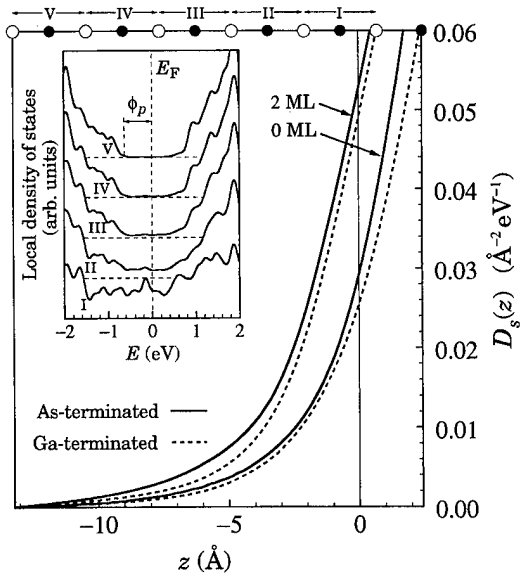


FIG. 7. Surface density of states at the Fermi energy without Si and with 2 ML of Si (see text). The solid line (dashed line) corresponds the As(Ga)-terminated interface. Inset: local density of states within the semiconductor band gap. The curves correspond to supercell regions indicated by the labels in the upper part of the figure.

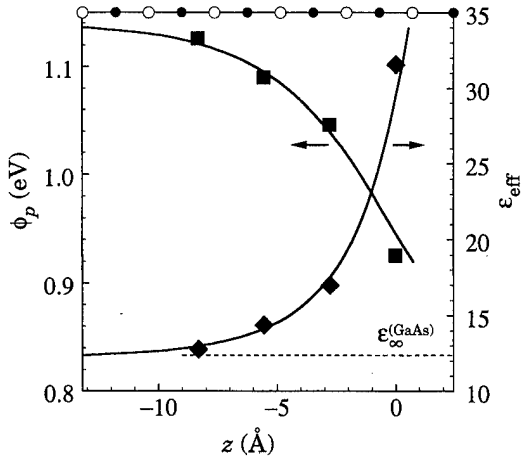


FIG. 8. Schottky barrier (filled squares, left scale) and effective screening (filled diamonds, right scale) as a function of the Si bilayer position within the semiconductor. The symbols give the results of self-consistent calculations for a Si coverage $x=0.5$. Their size indicates the estimated accuracy of the numerical results. The solid lines correspond to the predictions of Eq. (5) in the linear-coverage regime.

E. Model for the dipole screening

At low Si coverage, the additional polarization related to the presence of the metal can be described within a linear-response formulation by

$$\Delta V_b \equiv \epsilon_{eff}\Delta V = (1+4\pi\chi_{sc}+4\pi\chi_m)\Delta V, \quad (4)$$

where $\chi_{sc}=(\epsilon_\infty^{(GaAs)}-1)/4\pi$ is the GaAs electronic dielectric susceptibility, and χ_m accounts for the additional polarization due to the metal on the other side of the junction.

The additional susceptibility χ_m can be explained in terms of the metal-induced gap states (MIGS) at the Fermi energy. The central quantities are the local density of states[7] (LDOS) $N(z)$ and the related surface density of states at the Fermi energy $D_s(z)=\int_{-\infty}^{z}N(z')dz'$, where $-\infty$ indicates a position well inside the semiconductor (where the MIGS vanish) and z is the dipole position. Based on simple electrostatic arguments[40] an estimate for the MIGS contribution $-4\pi\chi_m\Delta v$ to the potential lineup [Eq. (4)] induced by a small change Δv in the Fermi-level position is $-4\pi e^2 D_s(z)\Delta v\,\delta$, where δ is a distance of the order of the exponential decay length λ of the MIGS at the Fermi energy.[4,7] The effective susceptibility describing the MIGS response to a change in potential lineup is thus $\chi_m=e^2D_s(z)\delta$.

In Fig. 7 we present the surface density of states $D_s(z)$ for the As- and Ga-terminated interfaces without Si and with 2 Si ML. In the inset we display the LDOS in the gap region for the As-terminated interface. The figure shows that the $D_s(z)$ for the two terminations are very similar, and both decay exponentially in the semiconductor with a decay length $\lambda\approx 3$ Å. At the position of the Si interlayer ($z=0$) we find $D_s(0)\approx 0.027$ Å^{-2} eV^{-1}. To model the screening effect we will use $\delta=3.2$ Å. These values are consistent with the results of the earlier calculations by Louie, Chelikowsky, and Cohen[7] and Tersoff.[5] At low Si coverage (in the linear regime) the effective screening is $\epsilon_{eff}=\epsilon_\infty^{(GaAs)}+4\pi e^2 D_s(0)\delta\approx 28$, corresponding to the dashed lines in Fig. 6, in perfect agreement with experiment and with the first-principles results.

At high Si coverage x, Eq. (4) can still be used to describe the macroscopic screening of the Si bilayer in the GaAs/GaAs homojunction ($\chi_m=0$) provided the gradual change in the host susceptibility χ_{sc} induced by the Si is taken into account (microscopic capacitor model[18,39]): $\chi_{sc}(x)=\chi^{(GaAs)}+x/2[\chi^{(Si)}-\chi^{(GaAs)}]$. Along the same line, we should also include the change in χ_m induced by the modification of the MIGS with Si doping: $\chi_m(x)=e^2D_s(z)\delta+e^2x/2[D_s'(z)-D_s(z)]\delta$, where $D_s'(z)$ is the surface density of states at the Fermi energy in presence of 2 ML of Si (see Fig. 7). The effective screening for a dipole layer at position z and for a coverage $0\leqslant x\leqslant 2$ is thus

$$\epsilon_{eff}(x,z)=\epsilon_\infty^{(GaAs)}+4\pi e^2D_s(z)\delta+\frac{x}{2}\{\epsilon_\infty^{(Si)}-\epsilon_\infty^{(GaAs)}+4\pi e^2[D_s'(z)-D_s(z)]\delta\}. \quad (5)$$

The perturbed surface density of states at the position of the Si interlayer (Fig. 7) is $D_s'(0)\approx 0.050$ Å^{-2} eV^{-1}. Using this value in Eq. (5) we obtain the coverage dependence indicated by the solid line in Fig. 6, in excellent agreement with the results of the self-consistent calculations. We would like to point out that most of the non-linearity in $\phi_p(x)$ is due to the large difference between $D_s'(0)$ and $D_s(0)$ in Eq. (5); the difference between $\epsilon_\infty^{(GaAs)}$ and $\epsilon_\infty^{(Si)}$ plays only a minor role. Finally, based on Eq. (5), the symmetry in the barrier varia-

tions of the two interfaces in Fig. 6 can be understood from the similarity of their $D_s(0)$ and $D_s'(0)$ values (see Fig. 7).

To probe the correlation between the local-dipole screening and the spatial decay of MIGS, which is explicit in Eq. (5), we moved the Si dipole layer away from the interface to a position $z<0$ into the semiconductor. Given the exponential decay of $D_s(z)$ in the semiconductor, we expect $\epsilon_{\text{eff}}(x=0,z)$ to exponentially converge to $\epsilon_\infty^{(\text{GaAs})}$ with a similar decay length. We investigated the change in the screening as a function of the position of the Si dipole layer for a coverage $x=0.5$. We selected this coverage to minimize the numerical uncertainty on ϵ_{eff} within the linear regime [$x=0$ in Eq. (5)]. The results for $\epsilon_{\text{eff}}(z)$ obtained with first-principles calculations are displayed in Fig. 8 together with the corresponding Schottky barrier values. The exponential decay length of $\epsilon_{\text{eff}}(z)$ towards $\epsilon_\infty^{(\text{GaAs})}$ in Fig. 8 is similar to the decay length of $D_s(z)$ in Fig. 7. The solid line indicates the position dependence predicted by Eq. (5) in the linear coverage regime. The very good agreement, in Fig. 8, between the predictions of the model and the results of the *ab initio* calculations further confirms the soundness of the picture here proposed to explain the screening of local interface dipoles at metal/semiconductor junctions.

IV. CONCLUSION

We have studied the change of Al/GaAs and Al/GaAlAs Schottky barriers versus alloy composition and in the presence of ultrathin (0–2 ML) Si interlayers using first-principles calculations. We presented results for ideal, defect-free, Al/GaAlAs interfaces showing that the experimental composition dependence of the barriers can be explained without postulating defect-induced Fermi-level pinning. We also showed that microscopic-interface dipoles induced by the heterovalent nature of Si interlayers quantitatively explain the experimental barrier tuning at Al/Si/GaAs(100) junctions.

These results were interpreted within a linear-response theory approach focusing on the charge density response to chemical substitutions near the junction. This model approach quantitatively explained the trends with alloy composition obtained from the *ab initio* calculations. We also investigated the screening of the local dipoles induced by interlayers on the atomic scale, and presented a model explaining the effective screening in terms of simple physical parameters such as the semiconductor dielectric constant and the local interface density of states at the Fermi energy.

ACKNOWLEDGMENTS

The authors would like to acknowledge support for this work by the Swiss National Science Foundation under Grant No. 21-40530.94. The authors also thank A. Franciosi for useful discussions.

[1]L. J. Brillson, in *Handbook on Semiconductors*, edited by P. T. Landsberg (North-Holland, Amsterdam, 1992), Vol. 1, p. 281; in *Contacts to Semiconductors*, edited by L. J. Brillson (Noyes, Park Ridge, NJ, 1993).
[2]W. Spicer, Semiconductors Semimetals **38**, 449 (1993).
[3]J. M. Woodall and J. L. Freeouf, J. Vac. Sci. Technol. **21**, 574 (1982).
[4]V. Heine, Phys. Rev. **138**, A1689 (1965).
[5]J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[6]F. Flores and C. Tejedor, J. Phys. C **20**, 145 (1987).
[7]S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B **15**, 2154 (1977).
[8]W. Mönch, Phys. Rev. Lett. **58**, 1260 (1987).
[9]M. Missous, W. S. Truscott, and K. E. Singer, J. Appl. Phys. **68**, 2239 (1990).
[10]P. Revva, J. M. Langer, M. Missous, and A. R. Peaker, J. Appl. Phys. **74**, 416 (1993).
[11]M. van Schilfgaarde, E. R. Weber, and N. Newman, Phys. Rev. Lett. **73**, 581 (1994).
[12]R. G. Dandrea and C. B. Duke, J. Vac. Sci. Technol. A **11**, 848 (1993); B **11**, 1553 (1993).
[13]J. C. Costa, F. Williamson, T. J. Miller, K. Beyzavi, M. I. Nathan, D. S. L. Mui, S. Strite, and H. Morkoç, Appl. Phys. Lett. **58**, 382 (1991), and references therein.
[14]J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **52**, 1794 (1988).
[15]M. Cantile, L. Sorba, S. Yildirim, P. Faraci, G. Biasiol, and A. Franciosi, Appl. Phys. Lett. **64**, 988 (1994); M. Cantile, L. Sorba, P. Faraci, S. Yildirim, G. Biasiol, G. Bratina, A. Franciosi, T. J. Miller, M. I. Nathan, and L. Tapfer, J. Vac. Sci. Technol. B **12**, 2653 (1994).
[16]S. Baroni, R. Resta, A. Baldereschi, and M. Peressi, in *Spectroscopy of Semiconductor Microstructures*, edited by G. Fasol, A. Fasolino, and P. Lugli, NATO ASI Series B (Plenum, New York, 1989), Vol. 206, p. 251.
[17]W. A. Harrison, J. Vac. Sci. Technol. **16**, 1492 (1979).
[18]M. Peressi, S. Baroni, R. Resta, and A. Baldereschi, Phys. Rev. B **43**, 7347 (1991).
[19]G. Biasiol, L. Sorba, G. Bratina, R. Nicolini, A. Franciosi, M. Peressi, S. Baroni, R. Resta, and A. Baldereschi, Phys. Rev. Lett. **69**, 1283 (1992).
[20]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979); **13**, 3095 (1980).
[21]N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[22]L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[23]D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980); J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[24]The calculated equilibrium lattice parameters of bulk GaAs and Al were $a^{(\text{GaAs})}=5.55$ Å and $a^{(\text{Al})}=3.97$ Å. The geometry constraint imposed by epitaxial alignment on the Al in-plane lattice constant is $a_\parallel^{(\text{Al})}=a^{(\text{GaAs})}/\sqrt{2}$. We evaluated the tetragonal distortion of the metal overlayer using the theoretical elastic constants of Al: $C_{11}=120$ GPa and $C_{12}=70$ GPa. Similar results are obtained using the experimental parameters.
[25]H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[26]C. L. Fu and K. M. Ho, Phys. Rev. B **28**, 5480 (1983).
[27]We also neglected atomic relaxations at the interface when Si interlayers were introduced. The effect of atomic relaxation on the band offset of GaAs/Si/AlAs junctions was found to be very small for Si coverages up to 2 ML (Ref. 18).
[28]The band structure terms were calculated with a 20 Ry cutoff, and using a (6,6,6) Monkhorst–Pack grid for the semiconductor and a (12,12,12) Monkhorst–Pack grid with a 0.1 eV electronic level broadening for the metal.
[29]The separation of ϕ_p into a potential-lineup term and a band term [Eq. (1)] is by no means unique. The reference potential V whose macroscopic average defines the lineup ΔV must include the *long-range electrostatic* Hartree and ion point-charge potential terms. The average of the short-range local components of the electronic potential, such as the exchange-correlation potential and the difference between the local component of the pseudopotential and the ion point-charge potential are bulk properties, and can thus be included either in the band term ΔE_p (present work) or in ΔV (as in Ref. 12).
[30]The spin–orbit correction $\frac{1}{3}\Delta$ on the GaAs (AlAs) band edge was included using the experimental value 0.34 eV (0.28 eV) for the spin-orbit splitting Δ: G. Harbeke, O. Madelung, and U. Rössler, in *Semiconductors: Physics of Group IV Elements and III-V Compounds*, edited by O. Madelung, Landolt-Börnstein, New Series, Group III, Vol. 17a (Springer, Berlin, 1982).
[31]We estimated the effect of the semi-core Ga 3*d* orbitals on ϕ_p by including the non-linear core correction. This was found to reduce $\phi_p^{(\text{GaAs})}$ by only 0.03 eV (the band and potential-lineup corrections tend to cancel each other). We therefore neglected the effect in this study.

[32]R. J. Needs, J. P. A. Charlesworth, and R. W. Godby, Europhys. Lett. **25**, 31 (1994).
[33]W. I. Wang, J. Vac. Sci. Technol. B **1**, 574 (1983).
[34]A. Y. Cho and P. D. Dernier, J. Appl. Phys. **49**, 3328 (1978).
[35]S. B. Zhang, D. Tománek, S. G. Louie, M. L. Cohen and M. S. Hybertsen, Solid State Commun. **66**, 585 (1988).
[36]G. Margaritondo and P. Perfetti, in *Heterojunction Band Discontinuities*, edited by F. Capasso and G. Margaritondo (North-Holland, Amsterdam, 1987), p. 59.
[37]The change in sign of d_1 for the type II relative to the type I junctions may be understood (based on the Ga–Al electronegativity difference) by considering the Ga(Al)–As bond polarization which is produced at the type II interfaces when the virtual plane is replaced by a Ga (Al) plane.
[38]We calculated $\epsilon_\infty^{(GaAs)}$ and $\epsilon_\infty^{(Si)}$ by inserting Si bilayers with coverages $x=0.2$ and $x=0.5$ at the GaAs/GaAs homojunction, and using the microscopic capacitor model of Ref. 18. The results are: $\epsilon_\infty^{(GaAs)}=12.37$, $\epsilon_\infty^{(Si)}=13.93$; the experimental values are 10.66 and 11.40 for GaAs and Si, respectively.
[39]In Ref. 18 a weighted average of the inverse dielectric constants was used to interpolate the susceptibility at fractional coverages $0<x<2$. The choice of the average, however, has little effect on the present results since the dielectric constants are close to each other.
[40]A. M. Cowley and S. M. Sze, J. Appl. Phys. **36**, 3212 (1965).

Near-surface dopant passivation after wet-chemical preparation of Si(111):H surfaces

L. Ley,[a] J. Ristein, J. Schäfer, and S. Miyazaki[b]
Universität Erlangen-Nürnberg, Institut für Technische Physik, Erwin-Rommel-Strasse 1, 91058 Erlangen, Germany

(Received 22 January 1996; accepted 17 April 1996)

Recent advances in wet-chemical surface treatment of silicon has resulted in unreconstructed, hydrogen terminated Si(111) surfaces with extremely low surface recombination velocity of only 0.25 cm/s [E. Yablonovitch *et al.*, Phys. Rev. Lett. **57**, 249 (1986)]. This points towards a surface free of surface states that normally pin the Fermi level E_F on Si(111) somewhere near midgap depending on the surface reconstruction. This expectation is not borne out, however, by experiment. Using Si $2p$ core level spectra, we find consistently a surface Fermi level position near midgap for 16 heavily doped n- and p-type samples. Checks for surface photovoltage and ion concentrations sufficient to induce the observed band bending were negative. All samples attain their bulk Fermi level position at the surface after annealing to about 400 °C, a temperature at which the surface hydrogen coverage is unchanged as monitored by ultraviolet excited photoemission. Above 400 °C, hydrogen starts to desorb from the surface and only with the onset of the 7×7 reconstruction does E_F move back towards midgap. We suggest that the initial position of E_F near midgap is due to a hydrogen induced passivation of dopants that leaves an intrinsic surface layer. Annealing at ~400 °C reactivates the dopants by breaking the B–H and P–H complexes as is known from plasma passivation. The necessary passivation depth has been calculated as a function of bulk doping yielding values of the order of one micrometer. © *1996 American Vacuum Society.*

I. INTRODUCTION

Advances in wet-chemical etching processes of silicon have resulted in atomically flat, unreconstructed Si(111) surfaces of unprecedented quality. The IR spectra of these surfaces taken in the attenuated total reflection (ATR) mode show a single sharp Si–H stretching mode at 2084 cm^{-1} with negligible contributions from $Si–H_2$ and $Si–H_3$ configurations.[1] Recent high resolution photoemission measurements show Si $2p$ core line spectra that exhibit linewidths below 130 meV, which was previously unheard of, and a sharp Si–H bonding state about 5.1 eV below the valence band maximum.[2] All these observations point towards an ideally Si–H terminated unreconstructed Si(111) surface with a minimum of steps and imperfections. As a consequence, one would expect a surface essentially free of surface states in the fundamental gap leading thus in contrast to the hydrogen-free, reconstructed Si(111) surfaces (7×7 or 2×1) to an unpinned Fermi level position and a vanishingly small surface recombination velocity. The latter is indeed borne out by the measurements of Yablonovitch *et al.*, who determine a surface recombination velocity of only 0.25 cm/s, which is even lower than that of the best $Si–SiO_2$ interfaces reported so far and corresponds to a density of surface states of only 2.5×10^{-7} cm^{-2}.[3] Here we present data on p- and n-doped hydrogen passivated Si(111) samples that exhibit contrary to expectation a surface Fermi level position close to midgap, corresponding thus to near surface depletion layers in all cases.

II. EXPERIMENT

The surface Fermi level position was determined through the binding energies of the Si $2p$ core level as measured in x-ray excited photoelectron spectroscopy (XPS). Strictly speaking, only differences in Fermi level position are being determined by this method which can, however, be converted into absolute Fermi level positions, as will be shown below.

We investigated three kinds of Czochralski-grown Si samples. P and Sb doped n^+ samples with a resistivity between 0.02 and 0.05 Ω cm and a donor concentration of 1.5 to 2.5×10^{17} cm^{-3}, highly B doped p^+ samples with resistivities of 0.01 to 0.02 Ω cm and acceptor concentrations between 4.5 and 11×10^{18} cm^{-3}, and lightly B-doped p-type samples with acceptor concentrations of 1.3 to 2.3×10^{15} cm^{-3}. The samples were cleaned according to the recipe given by Higashi *et al.* using buffered HF (40% NH_4F in H_2O) as the last etching step, which was followed by a 3 s dip in de-ionized water to remove traces of fluorine at the surface.[1] Infrared absorption spectra of the samples so prepared showed as the dominant absorption band the Si–H stretching mode at 2084 cm^{-1}, with less than 10% contributions from either $Si–H_2$ or $Si–H_3$ modes. Only in a few cases could contaminants in the form of physisorbed C and O be traced just at the detection limit of XPS (~1%). A remarkably sharp low energy electron diffraction (LEED) pattern with very low background confirmed the unreconstructed 1×1 surface unit mesh. XPS was performed with a resolu-

[a]Electronic mail: lothar.ley@physik.uni-edu
[b]Present address: Dept. of Electr. Engineering, Hiroshima University, Hiroshima 724, Higashi, Japan.

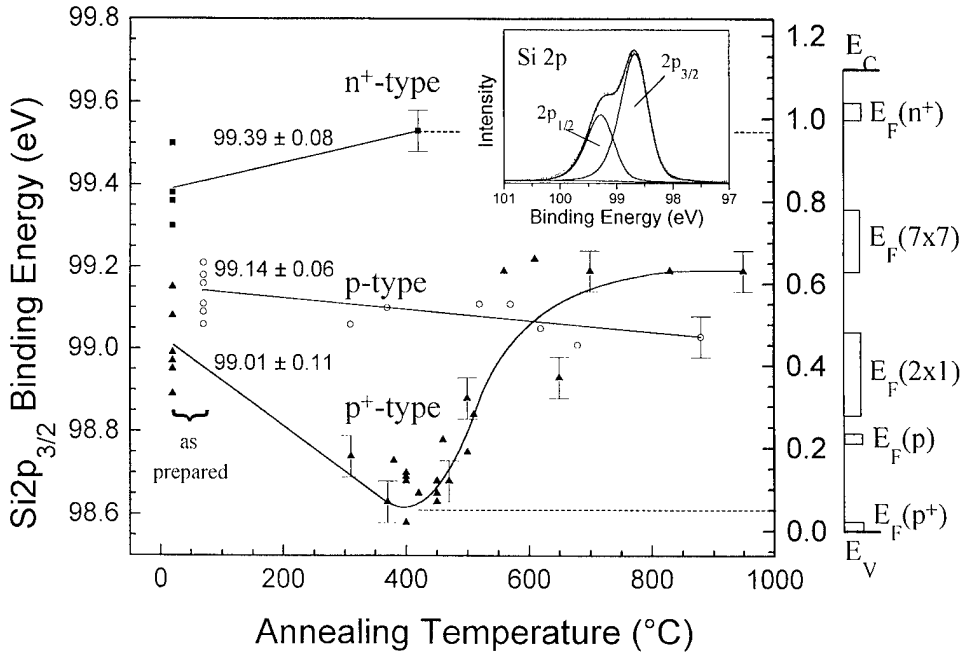


FIG. 1. Si $2p_{3/2}$ binding energies of hydrogen passivated Si(111) surfaces as a function of annealing temperature. The lines are guides to the eye and the average binding energy as measured before annealing is given next to each group belonging to one doping class. On the right ordinate binding energies measured relative to the Fermi level E_F have been converted into Fermi level position with respect to the valence band edge E_V of Si by subtracting 99.59 eV from the former. The hatched regions indicate the bulk Fermi level positions of the samples according to their doping and the open rectangles mark the range of pinning positions of E_F as observed for unhydrogenated and reconstructed Si(111) surfaces according to Ref. 4.

tion of 0.5 eV using monochromatized Al $K\alpha$ radiation (1486 eV), and ultraviolet excited photoemission (UPS) with a resolution of 0.11 eV utilizing HeII radiation (40.8 eV) was used to measure the valence band spectra. In order to exclude band flattening under x-ray or UV illumination, we measured *in situ* the surface photovoltage using a semitransparent Kelvin probe (copper mesh). None of the samples developed photovoltages exceeding the detection limit of 1 meV under the illumination conditions encountered in the photoemission measurements.

III. RESULTS

In Fig. 1 we show the binding energies of the Si $2p_{3/2}$ core level with respect to the Fermi level for different samples as obtained from a fit to the XPS spectra (see insert) in the as deposited state and after various *in situ* annealing steps (5 min held at the annealing temperature, no rapid cooling). Temperatures were measured with an optical pyrometer. From the individual binding energies of the as-prepared samples we derive average binding energies for each group of differently doped samples: 99.39±0.08 eV for the n^+ samples, 99.14±0.06 eV for the p-type samples, and 99.01 ±0.11 eV for the p^+ samples. The scatter of the data is due to the preparation process as well as to the experimental uncertainty in Si $2p$ line position (±50 meV, indicated exemplary on some data points).

On the right side of Fig. 1, the gap of Si and the bulk Fermi level positions of the three kinds of samples are indicated on an energy scale that is shifted by 98.56 eV with respect to the binding energy scale. The bulk Fermi levels are appropriate for 300 K and were calculated from the dopant concentrations given above. It is obvious from this comparison that the variation in surface Fermi level position that amounts to 0.38 eV falls considerably short of the difference in bulk Fermi level of ~1.00 eV between n^+ and p^+ samples. Anticipating the discussion that leads to the alignment of the binding energy and Fermi level scales as given in Fig. 1, we can infer that the Fermi levels tend to lie closer to midgap at the surface than in the bulk. The midgap position corresponds to a binding energy of 99.15 eV and the band bending amounts on average to 0.18 eV for the n^+ samples, 0.37 eV for the p type, and 0.42 eV for the p^+ samples. Note that the surface is never pulled into inversion, i.e., the Fermi level position of the as-prepared samples is always found between its bulk value and midgap.

For annealing temperatures up to 400 °C, the binding energies and thus the surface Fermi levels move in opposite directions depending on the type of doping, i.e., towards their respective bulk position (hatched regions on the right hand scale). The extreme binding energy difference of 0.93 eV that is reached by the p^+ and n^+ samples after annealing at 400 °C agrees with the bulk Fermi level difference of 1.00 eV as well as can be expected. We conclude therefore that the heavily doped samples attain flat band condition after a 400 °C anneal and we have used this fact to align the two energy scales of Fig. 1 as indicated by the dashed horizontal lines. As a by-product, we obtain the Si $2p_{3/2}$ binding energy with respect to the valence band edge E_V as 98.56±0.05 eV.

For annealing temperatures above about 450 °C, the binding energies of the p^+ samples increase again, corresponding

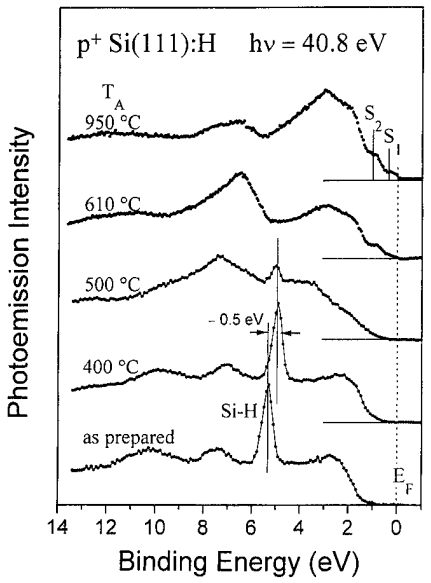


FIG. 2. He II valence band spectra of a hydrogen passivated p^+ Si(111) surface as a function of annealing temperature T_A.

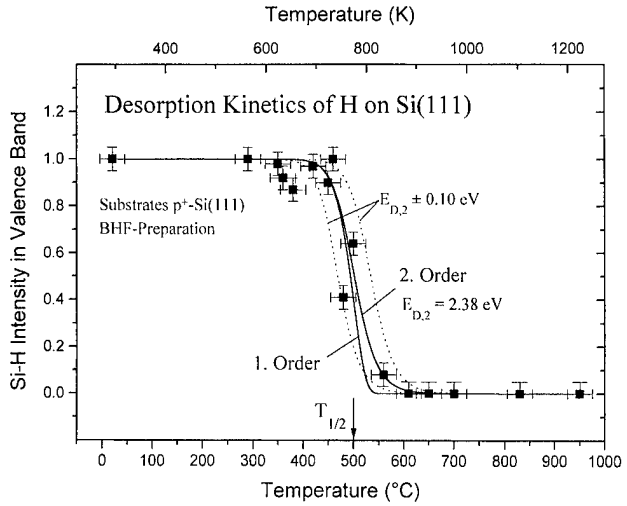


FIG. 3. Hydrogen coverage as a function of annealing temperature. The solid lines have been calculated for first and second order desorption kinetics, respectively, using a common activation energy of 2.38 eV.

to a tendency of E_F to move back towards midgap. Three of the six p-type samples were annealed at various temperatures up to 900 °C. None of them attained the bulk Fermi level position or showed the reversal in band bending observed for the p^+ samples. Instead, we measure a monotonic movement of E_F towards its bulk value albeit by no more than 0.1 eV. Annealing the n^+ samples above 400 °C led to their oxidation, a fact that is currently being investigated. We shall therefore concentrate our discussion on the p^+ samples in what follows.

Connected with annealing above 400 °C is a loss of hydrogen at the surface as witnessed by the UPS valence band spectra of a representative p^+ sample in Fig. 2, where the sharp peak at ~5.3 eV represents Si–H bonding states at the surface. It is evident from the first two spectra that the maximum Fermi level shift of −0.5 eV obtained after annealing at 400 °C occurs without the loss of hydrogen. Only beyond that temperature does the intensity of the Si–H peak drop, signalling a corresponding loss in surface hydrogen that is now indeed accompanied by a movement of E_F back towards midgap. In this state, samples are considerably more susceptible to oxygen contamination as witnessed by the appearance of the O $2p$ signal at ~7 eV. This contamination could be held below 4%, however, in all cases as measured by XPS. The oxygen is reduced again after the 900 °C annealing step and the characteristic surface states S_1 and S_2 of the 7×7 reconstruction[4] can be clearly observed, together with the corresponding LEED pattern.

For one p^+ sample, we have followed the desorption kinetics of hydrogen in somewhat more detail, as illustrated in Fig. 3. Here we plot the area under the Si–H peak at 5.3 eV binding energy relative to the amplitude of bulk derived states at a binding energy of about 4.5 eV as a measurement of the hydrogen coverage as a function of annealing temperature. Also shown as solid lines are coverages calculated according to first and second order desorption kinetics, respectively, calculated according to Ref. 5 using an attempt frequency of 10^{13} s^{-1} and an annealing time at each step of 300 s as stated above. It is obvious that the data are not accurate enough to distinguish unambiguously between first and second order desorption kinetics despite the fact that second order appears to be slightly favored at lower coverages. From a least squares fit to the data, we obtain $E_{D2}=2.4\pm0.1$ eV for the activation energy of the second order process. The first order line was calculated using the same activation energy. Schulz and Henzler[6] deduce first order kinetics for the desorption of H from Si(111) at high coverages and second order at low coverages with activation energies of 2.1 and 2.5 eV, respectively, values that are in as good agreement with our result as can be expected.

IV. DISCUSSION

A common explanation for surface band bending is to assume surface charges in conjunction with a space charge region inside the doped semiconductor. For the p^+ samples, a surface charge density (in units of the elementary charge) of 7×10^{12} cm^{-2} is necessary to account for the observed band bending. These surface charges can be present either in the form of charged surface states or as surface ions with energy levels possibly outside the gap region. Surface concentrations of Na^+ and F^- as the most likely ionic contaminants were proved by XPS to be at least an order of magnitude lower than required to explain the p^+ results. Moreover, sign and concentration of surface ions would have to be just appropriate to pull the Fermi level close to, but never beyond midgap for all three types of samples, which we consider a rather unlikely coincidence.

Surface Fermi level pinning by surface states would thus require a density of surface states on the order of 10^{13} cm^{-2} or more. These conditions are obviously not met at the hydrogen terminated (111) surface as demonstrated by

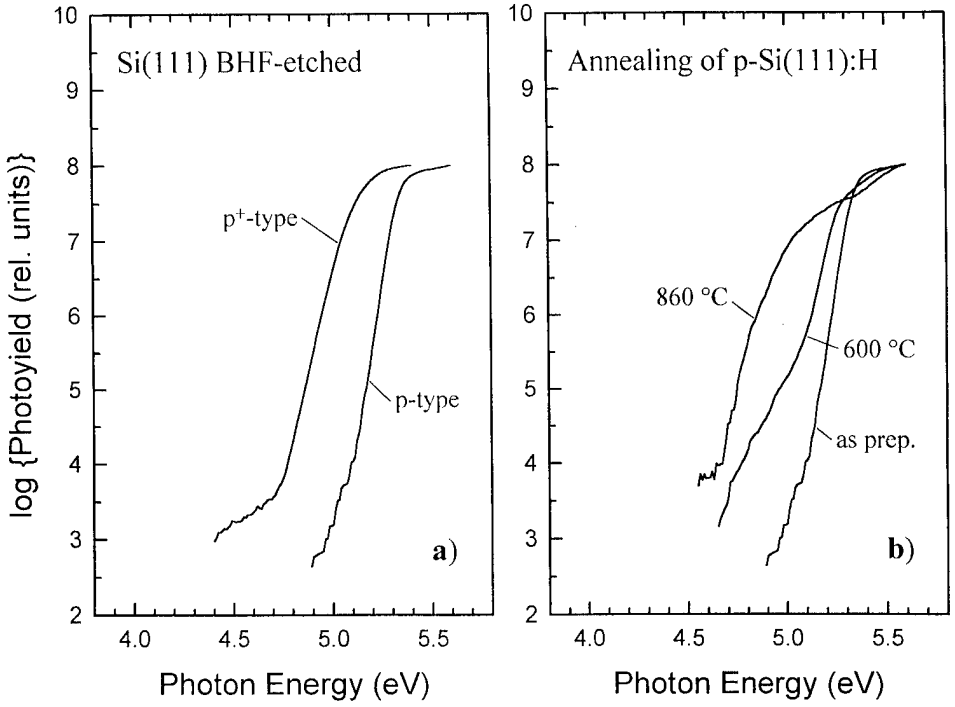


FIG. 4. Yield spectra of two representative Si samples: (a) as-prepared; (b) p-type sample after annealing.

Yablonovitch *et al.*[3] (see above). The lack of deep gap states that could be responsible for the pinning of E_F is further confirmed by electron yield spectroscopy [compare Fig. 4(a)], which provides no evidence whatsoever for emission below the onset of transitions from the valence band edge at threshold energies of 5.1 (n^+,p samples) and 4.8 eV (p^+ samples), respectively. The sensitivity of this technique is sufficient to detect surface densities of states as low as 10^{11} cm^{-2}.[7] As the samples are heated above 450 °C emission from surface states can be detected below the bulk threshold as demonstrated in Fig. 4(b) for temperatures of 600 and 860 °C, respectively.

We therefore suggest that the surface Fermi level position in our Si(111):H samples is the result of a hydrogen passivation of dopant atoms in a surface layer; that is to say, we are really looking at undoped surfaces here. The hydrogen passivation of dopant atoms is a well-known phenomenon that is intentionally performed and studied extensively by exposing silicon at temperatures above ~100 °C to atomic hydrogen created in an H_2 plasma.[8] By comparison, there are only few reports on hydrogen injection and near-surface passivation at room temperature (RT) following wet-etching procedures[9–11] and all apply to boron compensation with the exception of Ref. 11, which suggests that donor passivation might have been observed in one sample as well.

With the assumption that hydrogen passivation is indeed responsible for the depletion layers in our samples the annealing behavior can be understood as follows. For temperatures between 200 and 350 °C the donor- and acceptor-hydrogen complexes dissociate,[12] leading thus to a gradual reactivation of the dopant atoms and a concomitant movement of E_F towards its respective bulk value. A complete reactivation is apparently reached at T_A=400 to 450 °C for both n^+ and the p^+ samples.

Above 450 °C, the desorption of hydrogen from the surface commences in agreement with the temperature controlled desorption study presented above. With the loss of surface hydrogen the now unsatisfied bonds start to rearrange towards the thermodynamically stable 7×7 reconstruction[13] with a corresponding pinning of the Fermi level by surface states. As mentioned earlier, the hydrogen exposition of the n^+ sample above 400 °C led to a surface contamination and the resulting Fermi level positions have therefore not been included in Fig. 2.

The near-surface depletion of doped samples as a result of hydrogen passivation could explain some as yet unaccounted-for observations. Landemark *et al.* commented on the unexpected low Fermi level position 0.68 eV above E_V for an n^+-type Si(111):H sample,[14] whereas the annealed (at 400 °C) Si(111):H surface exhibits the unpinned bulk Fermi level.[15] Ley *et al.*[16] measured a lower than expected Schottky barrier height in the PtSi/Si system prepared on hydrogen terminated n-type Si(111) samples. All these observations are compatible with a (partial) passivation of donors at the surface.

This leaves the question to what degree and what depth the dopant atoms in the as-prepared samples have to be passivated to account for the observed Fermi level positions at the surface. A lower limit for the compensation depth can be obtained under the assumption that all dopants are passivated up to a certain depth. The potential profile calculated for the ensuing i-n and i-p transitions under the proper boundary conditions yields the Fermi level at the surface as a function of compensation depth as shown for three dopant concentra-

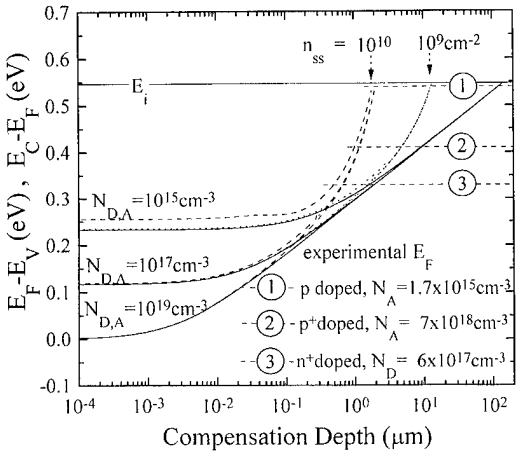


FIG. 5. Surface Fermi level position on Si as a function of dopant compensation depth. The solid lines are calculated for a surface free of surface charge, whereas the dotted lines represent the calculations for fixed densities of charged surface states n_{SS}. Parameter is in either case the dopant concentration. The dashed horizontal lines indicate the average surface Fermi level positions for the three types of samples investigated. Details of the calculation may be found in Ref. 17.

tions in Fig. 5.[17] For simplicity, we did not distinguish between n- and p-type doping in this calculation, which means that we have used the effective valence band density of states for the conduction band as well. Comparing the experimentally determined surface Fermi level positions with the calculation, we obtain compensation depths of 2 and 10 μm for the n^+ and p^+ samples, respectively. These values are compatible with what has been reported by Tavendale *et al.*[11] A depletion depth of about 100 μm as required to account for the surface Fermi position in the p-type samples appears excessive by comparison. However, as is also indicated in Fig. 3, a density of surface charge n_{SS} of 10^{10} cm^{-2} which is below the detection limit of our photoyield experiment substantially reduces the depletion width necessary for all doping levels and leads to values below 2 μm for all samples. We are quick to admit that these surface densities are considerably higher than the results of Yablonovitch *et al.*[3] However, at least for plasma hydrogenated silicon, there is ample evidence for hydrogen induced near-surface defects.[8,18]

V. CONCLUSIONS

In summary, we have observed a Fermi level position near midgap at the surface of hydrogen passivated Si(111):H independent of doping. The annealing behavior of E_F points towards a near surface depletion layer caused by hydrogen passivation during the wet-chemical surface treatment as the likely cause. The bulk Fermi level positions for p^+ and n^+ samples are restored at the surface after a 400 °C anneal, while hydrogen desorption above 450 °C causes a pinning of E_F by surface gap states. These results explain some unaccounted-for observations in the literature concerning the Fermi level position on Si(111):H. A calculation yields passivation depths of the order of 2 μm to account for the observed Fermi level positions.

ACKNOWLEDGMENT

Part of this work was supported by the Deutsche Forschungsgemeinschaft through Grant No. SFB 292.

[1]G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari, Appl. Phys. Lett. **56**, 656 (1990).
[2]K. Hricovini, R. Günther, P. Thiry, A. Taleb-Ibrahimi, G. Indlekofer, J. E. Bonnet, P. Dumas, Y. Petroff, X. Blase, X. Zhu, S. G. Louie, Y. J. Chabal, and P. A. Thiry, Phys. Rev. Lett. **70**, 1992 (1993).
[3]E. Yablonovitch, D. L. Allara, C. C. Chang, T. Gmitter, and T. B. Bright, Phys. Rev. Lett. **57**, 249 (1986).
[4]W. Mönch, in *Semiconductor Surfaces and Interfaces* (Springer, Berlin, 1993).
[5]K. Christmann, Surf. Sci. Rep. **9**, 1 (1988).
[6]G. Schulz and M. Henzler, Surf. Sci. **124**, 336 (1983).
[7]B. A. Orlowski, J. P. Lacharme, C. A. Sebenne, and S. Bensalah, Appl. Surf. Sci. **41/42**, 492 (1989).
[8]See, e.g., *Hydrogen in Semiconductors*, edited by J. I. Pankove and N. M. Johnson (Academic, Boston, 1991).
[9]A. J. Tavendale, A. A. Williams, and S. J. Pearton, Appl. Phys. Lett. **48**, 590 (1986).
[10]C. H. Seager, R. A. Anderson, and J. K. G. Panitz, J. Mater Res. **2**, 96 (1987).
[11]A. J. Tavendale, A. A. Williams, and S. J. Pearton, Mater. Res. Soc. Symp. Proc. **104**, 285 (1988).
[12]N. M. Johnson, in Ref. 8, pp. 113 ff.
[13]Le Thanh Vinh *et al.*, Appl. Phys. Lett **64**, 3308 (1994).
[14]E. Landemark, C. J. Karlsson, and R. I. G. Uhrberg, Phys. Rev. B **44**, 1950 (1991).
[15]C. J. Karlsson, E. Landemark, L. S. O. Johansson, U. O. Karlsson, and R. I. G. Uhrberg, Phys. Rev. B **41**, 1521 (1990).
[16]L. Ley, Y. Wang, V. Nguyen Van, S. Fisson, D. Souche, G. Vuye, and J. Rivory, Thin Solid Films **270**, 561 (1995).
[17]J. Ristein, J. Appl. Phys. **78**, 6083 (1995).
[18]P. Asoka-Kumar, H. J. Stein, and K. G. Lynn, Appl. Phys. Lett. **64**, 1684 (1994).

Correlation between surface structure and ordering in GaInP

H. Murata, S. H. Lee, I. H. Ho, and G. B. Stringfellow[a)]
Department of Materials Science and Engineering, University of Utah, Salt Lake City, Utah 84112

(Received 22 January 1996; accepted 5 April 1996)

Ga and In atoms in $Ga_{0.52}In_{0.48}P$ layers spontaneously segregate to form alternating In- and Ga-rich {111} monolayers during organometallic vapor phase epitaxial (OMVPE) growth on (001) oriented GaAs substrates, thus forming the CuPt ordered structure. This ordering phenomenon is believed to be driven by surface processes, although little direct experimental information is available. This work presents evidence, based on surface photoabsorption data, that $[\bar{1}10]$ oriented P dimers are present on the surface during OMVPE growth using trimethylgallium and ethyldimethylindium combined with tertiarybutylphosphine, suggesting a (2×4)-like surface reconstruction. Furthermore, when the growth temperature is increased above 620 °C, with other parameters constant, both the concentration of these P dimers and the degree of order are observed to decrease. A similar correlation of decreased P-dimer concentration with decreased degree of order is observed for decreases in V/III ratio. Thus, the changes in order parameter for variations in temperature and TBP flow rate are found to be closely correlated with the changes in the order parameter. A third parameter studied was the misorientation of the substrate from (001) toward either the $\{111\}_A$ or $\{111\}_B$ planes. The concentration of P dimers decreased as the misorientation increased in either direction. The degree of order was also observed to generally decrease, supporting the connection between surface reconstruction and ordering. However, the difference in order parameter observed for the two misorientation directions suggests the importance of a second parameter, the step structure, itself. For exactly (001) oriented substrates the surface was observed, using high resolution atomic force microscopy, to consist of islands, elongated in the [110] direction, with heights of 30–60 Å. Monolayer steps are observed for some growth conditions, but for most conditions the boundaries are formed exclusively of bilayer (5.7 Å) steps. Predominantly monolayer steps are formed for low V/III ratios and bilayer steps for high V/III ratios. © *1996 American Vacuum Society.*

I. INTRODUCTION

The spontaneous formation of monolayer superlattices during vapor phase epitaxial growth is a phenomenon of both fundamental and technological importance. The most widely studied example of this phenomenon is the formation of the CuPt ordered structure, with ordering on {111} planes, during organometallic vapor phase epitaxial (OMVPE) growth of $Ga_{0.52}In_{0.48}P$.[1,2] The thermodynamic driving force for CuPt ordering is believed to be related to formation of the (2×4) surface reconstruction,[1–3] since this ordered structure is not stable in the bulk.[2] The ordered structure is formed at the surface during epitaxial growth.[1–4] Evidence of this is the change in ordered structure correlated with the change in surface reconstruction for GaInP grown by molecular beam epitaxy (MBE).[3,4] The initial studies of Murata *et al.*[5] have demonstrated that for GaInP grown by OMVPE a (2×4)-like surface reconstruction can be detected *in situ* using the optical technique of surface photo absorption (SPA). This allows the opportunity to probe the correlation between the degree of CuPt order or order parameter, S, and the extent of formation of the (2×4)-like reconstruction as a function of growth parameters such as temperature, V/III ratio, and substrate misorientation from (001), all of which are known to have marked effects on the degree of order observed in the GaInP epitaxial layers grown by OMVPE.[1,2]

This article will present the results of a study of the effects of the three growth parameters mentioned above on the SPA spectrum and the degree of order determined from the photoluminescence peak energy as well as the step structure of the surface determined using atomic force microscopy (AFM). The results suggest a direct correlation between the (2×4)-like reconstruction determined by SPA and the degree of CuPt order in the GaInP epitaxial layers, although changes in the step structure are also found to be important.

II. EXPERIMENT

The epitaxial layers studied in this work were grown by OMVPE in a horizontal, atmospheric pressure apparatus[6] using the newly developed, less hazardous phosphorus source tertiarybutylphosphine (TBP). The flow rates of the group III precursors trimethylgallium (TMGa) and ethyldimethylindium (EDMIn) were adjusted to give the solid composition lattice matched to the GaAs substrates, i.e., $Ga_{0.52}In_{0.48}P$, and a growth rate of 0.3 μm/h. The TBP flow rate was varied to give the various V/III ratios used. The semiinsulating GaAs substrates were nominally (001), misoriented by angles of 0°, 3°, 6°, 11°, and 15.8° toward either the $[\bar{1}10]$ or the [110] direction in the lattice.

[a)]Electronic mail: stringfellow@ee.utah.edu

For the SPA measurements, P-polarized light from a 150 W Xe lamp irradiates the GaInP layer at an incidence angle of 70° through a polarizer and a chopper.[7] The direction of the incident light is parallel to the direction of gas flow in the reactor. The reflected light is monochromatized and detected by a Si PNN^+ photodiode using standard lock-in amplification techniques. Each run consists of the growth of two epitaxial layers, one where the [110] direction of layer is parallel to the gas flow and one where the substrate is rotated by 90°. The P-stabilized and group III element-stabilized surfaces were formed by switching the TBP flow between reactor and vent, respectively.[5,7]

The characterization of the surface structure was carried out using a Nanoscope III atomic force microscope (AFM) in the tapping mode. Etched single-crystalline Si tips were used with an end radius of about 5 nm, with a sidewall angle of about 35°. Scan rates of 1 to 2 lines per second were used and data were taken at 512 points/line and 512 lines per scan area. This technique has been shown capable of revealing atomic-scale features on the surface.[8] Monolayer steps have been observed *ex situ* for GaInP layers grown by OMVPE.[9]

The solid composition of the GaInP layers was determined, using Vegard's law, from x-ray diffraction measurements using Cu $K\alpha$ radiation. The 20 K PL was excited with the 488 nm line of an Ar^+ laser. The emission was dispersed using a Spex Model 1870 monochromator and detected using a Hamamatsu R1104 head-on photomultiplier tube.

III. RESULTS AND DISCUSSION

The major feature of the SPA spectrum, i.e., the difference in SPA signal with incident light along [110] and $[\bar{1}10]$, is a strong absorption peak observed at approximately 400 nm.[5] This has been attributed to the $[\bar{1}10]$ oriented P dimers on the surface[5] based on the similarity to the SPA results for InP reported by Kobayashi and Kobayashi.[10] These $[\bar{1}10]$ oriented group V dimers are the major characteristic of the (2×4) reconstructed surface.[11] Thus, in what follows, the intensity of this peak is taken as a measure of the degree of (2×4) order on the surface during growth, although SPA is a probe of the local bonding, so it cannot be used to determine the long range order on the surface. For this reason, the reconstruction will be referred to as "(2×4)-like."

The effect of growth temperature on the SPA difference signal at 400 nm is shown in Fig. 1. The SPA signal difference, and hence the concentration of $[\bar{1}10]$ oriented P dimers, $[P^{2-}_{[110]}]$, decreases monotonically as the temperature increases from 620 to 720 °C. This is simply due to a shift in the P adsorption/desorption balance at the surface. The SPA difference signal approaches zero at 720 °C. The SPA signal at 400 nm continues to climb as the temperature is further reduced to 520 °C.[7]

The degree of order is included in Fig. 1. It was determined from the S^2 dependence of the band gap energy on the degree of order,[2] with a band gap energy difference between ordered and disordered $Ga_{0.5}In_{0.5}P$ of 471 meV.[12] The PL peak energy is used to determine the band gap energy. A potential problem with this approach is the possibility of

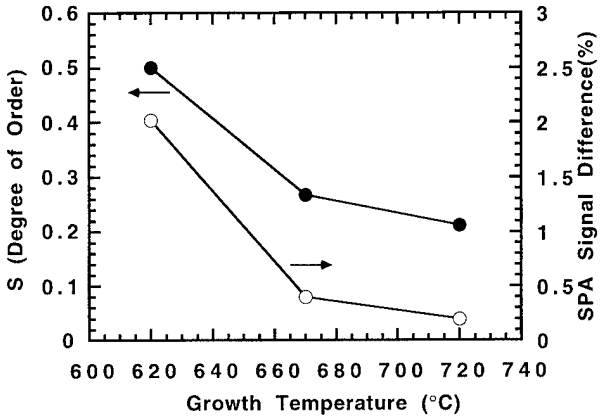


FIG. 1. The effect of growth temperature on the degree of order, S, and the SPA signal difference at 400 nm: (●) degree of order for a V/III ratio of 40 (TBP partial pressure of 50 Pa), (○) SPA signal difference for a TBP partial pressure of 50 Pa.

nonband edge emission dominating the PL spectra. Impurities are one possible source of such low-energy emission peaks, although impurity emission is not normally dominant for the high purity GaInP layers considered here. Impurity effects can normally be minimized by using high excitation energies. A more serious problem for these samples is the possibility of band tailing effects associated with the inhomogeneity of the order parameter or the presence of antiphase boundaries. Such effects have been studied by several workers.[12,13] Again in this case, the best method for minimizing these problems is the use of high excitation intensities. Thus, all order parameters have been deduced from PL spectra measured at a laser power of 10 mW. For the size of the laser spot used here, this corresponds to an excitation power density of 30 W/cm^2.

The data plotted in Fig. 1 indicate that for temperatures of 620 °C and above, both the degree of order and $[P^{2-}_{[110]}]$ decrease with increasing growth temperature. This suggests that loss of the (2×4)-like surface reconstruction results in a reduction of the degree of CuPt ordering. As discussed above, this is consistent with theoretical predictions summarized in Ref. 2. The order parameter is a maximum at approximately 620 °C. A clear decrease is observed with decreasing temperatures below 620 °C.[5,7] Thus, the degree of order decreases with decreasing temperature even though $[P^{2-}_{[110]}]$ increases in this range. This suggests that the degree of order decreases at low temperatures for kinetic reasons. Lower temperatures lead to lower rates of atomic migration on the surface, which may prevent formation of CuPt structure during the time available before the surface layer is covered by the next layer deposited.[7] This mechanism is basically the same as that postulated to account for the decrease in degree of order with increasing growth rate at a constant temperature.[14]

The effect of V/III ratio on both ordering and the surface reconstruction deduced from the SPA measurements is also significant. As seen in Fig. 2, for layers grown using TBP on exactly (001) substrates, a decrease in the TBP partial pres-

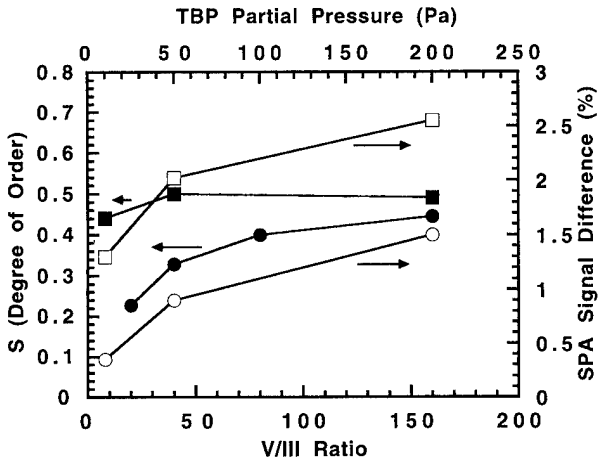


FIG. 2. Influence of input TBP partial pressure on the SPA difference signal at 400 nm and the degree of order, deduced from the 15 K PL peak energy, for GaInP samples grown lattice matched to exactly (001) GaAs substrates. ○, □: 670 °C and ●, ■: 620 °C.

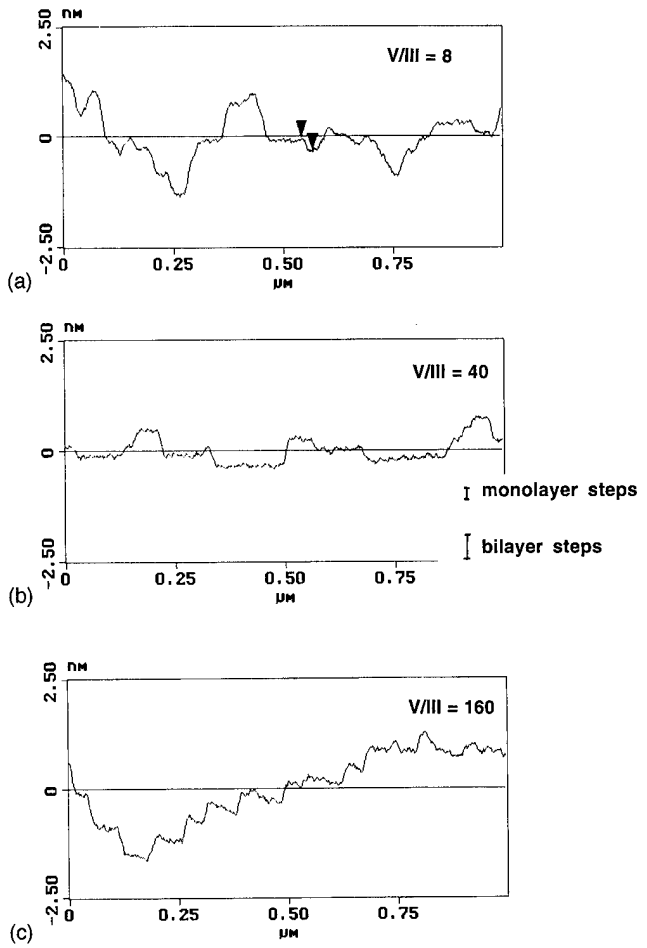


FIG. 3. Cross sectional AFM scans of the surfaces of GaInP layers grown on exactly (001) GaAs substrates at a temperature of 620 °C. (a) V/III=8; (b) V/III=40; and (c) V/III=160.

sure from 200 to 10 Torr results in a significant decrease in the extent of (2×4)-like reconstruction at both 620 and 670 °C. Consistent with the temperature effect discussed above, the SPA signal is always larger at the lower growth temperature. The degree of order deduced from the PL peak energy is also plotted versus input TBP partial pressure in Fig. 2 for growth at both 620 and 670 °C. The monotonic trend of decreasing degree of order for decreasing (2×4)-like reconstruction observed at 670 °C is qualitatively similar to the correlation observed when the growth temperature is changed. At 620 °C, the behavior is more complex. At high V/III ratios at this lower growth temperature, the degree of order is observed to decrease slightly. This is consistent with previous observations for growth using either phosphine or TBP.[15] This behavior at high P surface coverages is not completely understood. However, evidence of formation of an isotropic P dimer on the surface at high group V flow rates and low temperatures observed in InP[10] is also seen in the GaInP layers studied here. This change in surface reconstruction may be responsible for the observed decrease in the degree of order. This also provides an alternate possibility for the reduced ordering observed at lower growth temperatures.

The step structure, another factor that may affect ordering, is also found to be dependent on the V/III ratio. Figure 3 shows AFM scans for layers grown at 620 °C with V/III ratios of 8, 40, and 160. The surfaces for V/III ratios of 40 and 160 are smooth, giving easily interpreted AFM profiles. Growth at the low V/III ratio of 8 is difficult, resulting in rougher surfaces. The scan shown in Fig. 3 is representative of the smoothest areas observed. The [110] steps are seen to be mainly monolayers (2.8 Å) at low V/III ratios and mainly bilayers at high V/III ratios. The average step height is plotted versus input TBP flow rate in Fig. 4 for growth temperatures of both 620 and 670 °C. The formation of bilayer (5.7 Å) steps at high TBP partial pressures has been attributed to the stabilization of the (2×2) structure on the (111)B step edge at high phosphorus partial pressures.[7] Biegelsen *et al.*[16] observed the As-rich (2×2) structure on (111)B GaAs surfaces at high As pressures. The (2×2) unit cell will just fit on the (111)B face of a bilayer step.[7] Recent calculations indicate that bilayer type A ([$\bar{1}$10]) steps become more stable for high anion/cation ratios,[17] although the effect for B type ([110]) steps has not been considered.

The bilayer step structure observed for growth at high V/III ratios apparently explains one type of defect observed in GaInP layers.[9] For growth on exactly (001) oriented substrates, both the ($\bar{1}$11) and (1$\bar{1}$1) variants of the CuPt ordered structure are observed. The domains of the two variants are separated by order twin boundaries having the (001) orientation. The thicknesses of the (001) laminae of the two variants are found experimentally to be always equal to an even number of atomic layers.[18] The results presented here suggest that this is due to the formation of the CuPt structure via propagation of bilayer steps. The order twin boundaries are apparently due to a change in the direction of step propagation due to the undulating nature of the surface as growth proceeds.

The SPA spectra for GaInP layers grown at 620 °C with a V/III ratio of 40 on substrates with misorientation angles of 0°, 6°, 11°, and 15.8° toward the [$\bar{1}$10] direction are super-

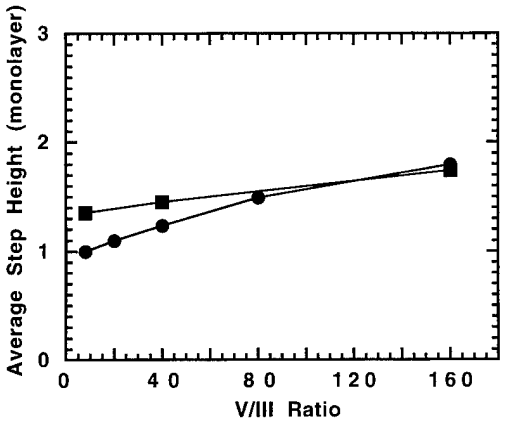


FIG. 4. Average step height (neglecting bunched steps) vs V/III ratio of GaInP layers grown on exactly (001) oriented GaAs substrates at temperatures of 620 (■) and 670 °C (●).

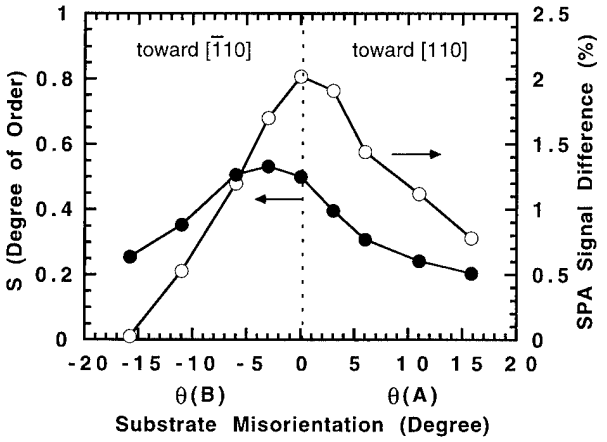


FIG. 6. SPA difference intensity at 400 nm vs misorientation angle, for misorientations in both the A and B directions, to produce [$\bar{1}$10] and [110] steps, respectively. Also plotted are the order parameters determined for these samples.

imposed in Fig. 5. The substrate with a misorientation of 15.8° is, of course, the (115)B surface. Clearly, the feature at approximately 400 nm is strongly dependent on the substrate misorientation. As mentioned above, this strongly anisotropic feature is attributed to [$\bar{1}$10] oriented P dimers on the surface. The SPA difference intensity is plotted in Fig. 6 versus misorientation angle, for misorientation in the [110] direction, to produce A-type steps, and in the [$\bar{1}$10] direction, to produce B-type steps. The data clearly show the decrease in the (2×4)-like reconstruction with increasing misorientation angle. The effect is stronger for misorientation in the B direction. Also included in Fig. 6 are the order parameters calculated as described above from the PL peak energies. The degree of order decreases monotonically with increasing θ_A, as reported previously for GaInP growth using phosphine.[19] For misorientation in the B direction, the degree of order increases with increasing misorientation angle for angles of less than approximately 3°. However, for larger misorientation angles, a decrease in order parameter with increasing misorientation angle is observed, similar to the data for misorientation in the A direction. These data are similar to those obtained earlier using PH_3.[6,19]

By combining these observations, it is clear that the decrease in the (2×4)-like reconstruction induced by increasing the misorientation angle corresponds to a decrease in the degree of order, as for the data obtained for variations in both temperature and V/III ratio. However, a second factor also appears to be important. The [110] steps themselves clearly assist the ordering process and the [$\bar{1}$10] steps have a disordering effect.

The effect of the surface [$\bar{1}$10] P-dimer concentration on the degree of order measured in the GaInP layers can be seen most clearly by directly examining the relation between the degree of order, deduced from the PL peak energy, and the

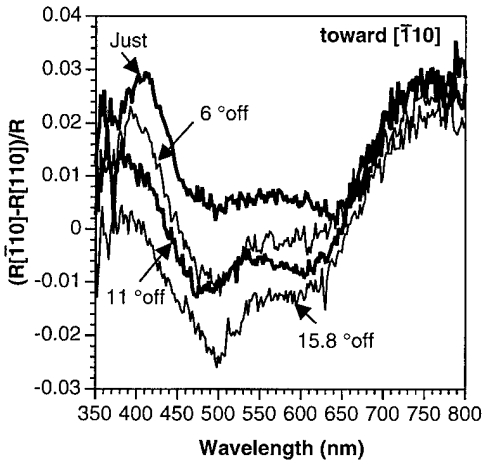


FIG. 5. Superimposed SPA spectra for GaInP samples grown at 620 °C on substrates misoriented from (001) by angles of 0°, 6°, 11°, and 15.8° toward the [$\bar{1}$10] direction in the lattice.

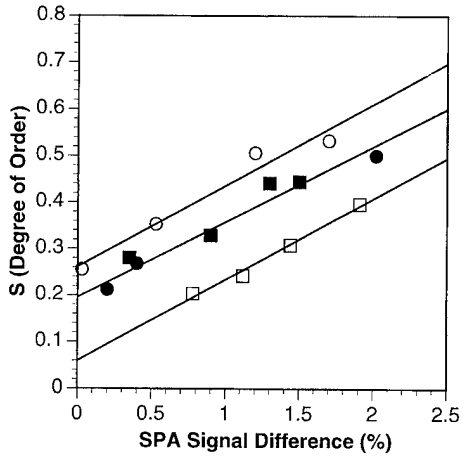


FIG. 7. Degree of order vs SPA difference signal: (●) Ref. 7 for samples grown at several temperatures for exactly (001) substrates; (■) Ref. 7 for samples grown at several values of V/III ratio for exactly (001) substrates; (□) substrates misoriented toward [110]; (○) substrates misoriented toward [$\bar{1}$10].

SPA difference signal measured at 400 nm. The data are plotted in Fig. 7. For samples with the exact (001) orientation, the correlation between the degree of order and the SPA signal intensity induced by changing either the growth temperature or the input TBP flow rate is nearly perfect. The general trends for changes in the misorientation angle are similar, although it is apparent that [110] steps have an additional beneficial effect and the $[\bar{1}10]$ steps have an additional deleterious effect.

The disappearance of the (2×4)-like reconstruction with increasing misorientation angle is thought to relate to the inability to form dimers involving all atoms on the (001) facets between steps. For an exactly (001) oriented surface having the (2×4) reconstruction, all P surface atoms will be able to reduce the energy of their dangling bonds via formation of [110] dimer rows. The introduction of [110] steps produces (001) facets of finite widths. This interferes with the ability of all surface atoms to form dimers. For example, if the (001) facet in a particular location consists of an odd number of [110] rows of P atoms, one row will not be able to participate in the reconstruction. The fraction of surface P atoms not able to participate generally increases with increasing misorientation angle. In terms of the simplest surface structure involving P dimers, the (2×1) reconstruction, one can see qualitatively that a larger fraction of the surface P atoms will be able to dimerize on the surface in the presence of $[\bar{1}10]$ steps, in agreement with the results presented in Fig. 6.

For both the A and B misorientations, the steps on the surface are also expected to affect the reconstruction. Simple electron counting (local charge density) arguments are often used to understand the stabilities of particular surface reconstructions.[16] The presence of steps will certainly affect the electron counting results for large misorientation angles, leading to the expectation that other reconstructions may form.

The general effects of surface steps, in addition to the effect on the surface reconstruction, have been considered qualitatively. For example, simple models to explain the formation of the B variants of the CuPt structure involve [110] steps.[1,20] Additional factors related to the attachment of adatoms at steps were discussed qualitatively by Asahi.[21] For example, a group III adatom approaching a [110] step makes no bonds in addition to the two bonds to the underlying (001) surface. This is expected to lead to a small sticking coefficient which leads, in turn, to a higher probability for ordering, since the lowest energy structure is more likely to form when many adsorption/desorption steps occur at the step edge before the final configuration is attained. On the other hand, for high V/III ratios, the group V adatom approaching a $[\bar{1}10]$ step will be able to form third bond. This is expected to result in a high sticking coefficient, which would not be conducive to formation of an ordered structure. For example, if the sticking coefficient were unity, the step propagation rate would be a maximum, the $[\bar{1}10]$ step would be extremely rough, and the arrangement of the Ga and In atoms would be random. This simple model is consistent with the observation that islands on exactly (001) surfaces of GaInP grown by OMVPE are elongated in the $[\bar{1}10]$ direction.[9] Thus, [110] steps are expected to enhance the formation of the B variants of the CuPt structure and $[\bar{1}10]$ steps are expected to have a deleterious effect.

Examination of the data collected here for variations in growth temperature, V/III ratio, and substrate misorientation suggests that formation of the CuPt ordered structure is intimately related to the formation of the (2×4)-like ordered structure on the surface during growth. However, it is too early to claim a direct mechanistic relationship. Recent efforts to produce order/disorder heterostructures by changing the flow rate of the P precursor during growth have resulted in a clear memory effect. The degree of order in the bottom layer apparently "propagates" into the top layer for a significant distance, i.e., the change in order parameter is not abrupt.[22] The change in the SPA signal is virtually instantaneous. This suggests that the step structure may play an important role in ordering. Investigations of the dynamics of the change is step structure are currently underway.

IV. CONCLUSIONS

The effects of three growth parameters, temperature, V/III ratio, and substrate misorientation, on both the occurrence of CuPt ordering and the surface structure have been studied. Considering first the data for singular (001) substrates, increasing growth temperature above 620 °C is found to decrease the degree of order and to result in a dramatic decrease in the SPA signal at 400 nm attributed to the $[\bar{1}10]$ dimers characteristic of the (2×4)-like surface reconstruction. Decreasing the input TBP flow rate, giving V/III ratios of <160 at 620 and 670 °C is similarly found to decrease both the 400 nm SPA signal and the degree of order measured in the GaInP epitaxial layers. Together, these data strongly suggest a correlation between the concentration of $[\bar{1}10]$ P dimers on the surface, characteristic of the (2×4) reconstruction, and the degree of order produced during epitaxial growth. In addition, the structure of [110] steps is found to change from predominately bilayer at high V/III ratios to monolayer at low V/III ratios. Increasing the concentration of [110] steps by misorientation of the $(00\bar{1})$ substrate is found to decrease the concentration of $[\bar{1}10]$ P dimers on the surface. It is found to increase the degree of order for angles of $<3°$ and to decrease the degree of order for larger angles. Increasing the concentration of $[\bar{1}10]$ steps is found to decrease both the concentration of P dimers and the degree of order. This supports the relationship between the (2×4)-like reconstruction and ordering, but indicates that a second factor is the concentration of steps. [110] steps assist the ordering process and $[\bar{1}10]$ steps lead to disordering.

ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support of the Department of Energy (growth and PL studies) and the National Science Foundation (SPA and AFM studies).

[1]G. B. Stringfellow, in *Common Themes and Mechanisms of Epitaxial Growth*, edited by P. Fuoss, J. Tsao, D. W. Kisker, A. Zangwill, and T. Kuech (Materials Research Society, Pittsburgh, 1993), pp. 35–46.
[2]A. Zunger and S. Mahajan, in *Handbook on Semiconductors*, edited by S. Mahajan (Elsevier Science, Amsterdam, 1994), Vol. 3, Chap. 19.
[3]S. B. Zhang, S. Froyen, and A. Zunger, Appl. Phys. Lett. **67**, 3141 (1995).
[4]A. Goymo, K. Makita, I. Hino, and T. Suzuki, Phys. Rev. Lett. **72**, 673 (1994).
[5]H. Murata, I. H. Ho, T. C. Hsu, and G. B. Stringfellow, Appl. Phys. Lett. **67**, 3747 (1995).
[6]L. C. Su, I. H. Ho, and G. B. Stringfellow, J. Appl. Phys. **75**, 5135 (1994).
[7]H. Murata, I. H. Ho, T. C. Hsu, Y. Hosokawa, and G. B. Stringfellow, J. Appl. Phys. (in press).
[8]C. C. Hsu, J. B. Xu, and I. H. Wilson, Appl. Phys. Lett. **64**, 2105 (1994); H. Bluhm, U. D. Schwarz, F. Herrmann, and P. Paufler, Appl. Phys. A **59**, 23 (1994).
[9]L. C. Su and G. B. Stringfellow, Appl. Phys. Lett. **67**, 3626 (1995).
[10]Y. Kobayashi and N. Kobayashi, J. Electron. Mater. **25**, 691 (1996).
[11]D. K. Biegelsen, R. D. Brignans, J. E. Northrup, and L. E. Swartz, Phys. Rev. B **41**, 5701 (1990).
[12]P. Ernst, C. Geng, F. Scholz, H. Schweizer, Y. Zhang, and A. Mascarenhas, Paper EE2.1, presented at the November 1995 Materials Research Society Meeting, Boston.
[13]M. C. Delong, W. D. Ohlsen, I. Viohl, P. C. Taylor, and J. M. Olson, J. Appl. Phys. **70**, 2780 (1991).
[14]D. S. Cao, E. H. Reihlen, G. S. Chen, A. W. Kimball, and G. B. Stringfellow, J. Cryst. Growth **109**, 279 (1990).
[15]Y. S. Chun, H. Murata, T. C. Hsu, I. H. Ho, L. C. Su, Y. Hosokawa, and G. B. Stringfellow, J. Appl. Phys. **79**, 6900 (1996).
[16]D. K. Biegelsen, R. D. Brignans, J. E. Northrup, and L. E. Swartz, Phys. Rev. Lett. **65**, 452 (1990).
[17]S. B. Zhang and A. Zunger, J. Cryst. Growth **163**, 113 (1996).
[18]C. S. Baxter, W. M. Stobbs, and J. H. Wilkie, J. Cryst. Growth **112**, 373 (1991).
[19]L. C. Su, I. H. Ho, G. B. Stringfellow, Y. Leng, and C. C. Williams, Mater. Res. Soc. Symp. Proc. **340**, 123 (1994).
[20]G. B. Stringfellow and G. S. Chen, J. Vac. Sci. Technol. B **9**, 2182 (1991).
[21]H. Asahi, J. Cryst. Growth **80**, 425 (1987).
[22]G. B. Stringfellow, L. C. Su, H. Murata, Y. S. Chun, I. H. Ho, T. C. Hsu, and Y. Hosokawa, Mater. Res. Soc. Proc. (to be published).

Heteroepitaxy of GaP on Si(100)*

K. J. Bachmann,[a),b),d)] U. Rossow,[c)] N. Sukidi,[b)] H. Castleberry,[b)] and N. Dietz[a),c)]
North Carolina State University, Raleigh, North Carolina 27696-7919

(Received 14 February 1996; accepted 20 April 1996)

In this article, we analyze the kinetics of heteroepitaxial growth of GaP on Si(100) by pulsed chemical beam epitaxy on the basis of results obtained by real-time optical process monitoring. In view of the large barrier to epitaxial growth on oxygen or carbon contaminated silicon surface elements and the low stacking fault energy for GaP, residual contamination of the silicon surface contributes to defect formation in the initial phase of GaP heteroepitaxy on Si, and requires special measures, such as surface structuring, to limit the propagation of defects into the epitaxial film. The control of the supersaturation during the first 10–20 s of film formation is essential for the quality of subsequent epitaxial growth and is limited to a narrow process window between three-dimensional nucleation and overgrowth at low Ga supersaturation and gallium-cluster formation at high Ga supersaturation. Steady state heteroepitaxial growth is described by a four-layer stack substrate/epilayer/surface reaction layer (SRL)/ambient and, depending of the source vapor flux, allows for more than monolayer coverage. Under this special condition of low-temperature CBE, the kinetics of chemical reactions in the SRL is composed of homogeneous reactions creating/consuming intermediates that participate in surface reactions including the incorporation of Ga and P atoms into the epitaxial film. For the homogeneous part of the surface kinetics the dipole approximation provides an adequate correlation of the changes in the dielectric function of the SRL to the activities of randomly distributed reactants and products. No adequate correlation of the dielectric function exists to the concentrations of strongly bonded surface atoms and surface molecules. Therefore, quantitative assessments of the heterogeneous kinetics on the crystal surface cannot use real-time optical monitoring as a reliable basis. © *1996 American Vacuum Society.*

I. INTRODUCTION

Research concerning polar-on-nonpolar heterostructures has been motivated for more than a decade by potential technological applications and scientific questions of general importance. However, progress with regard to both technological and scientific goals has been slow. The primary reasons for this are as follows.

(1) The formation of defects during early stages of nucleation and heteroepitaxial overgrowth and their propagation into the epitaxial film.

(2) The roughening of the surface of the epilayer in the course of epitaxial growth, complicating the modeling and control of the kinetics of heteroepitaxial growth.

(3) Interface broadening related to surface reactions, reconstruction and interdiffusion.

As discussed in Sec. III in more detail, these three problem areas are linked, so that their separation and control provides for a considerable challenge. Here we focus onto GaP heteroepitaxy on Si(100) by pulsed chemical beam epitaxy.

II. EXPERIMENT

Figures 1(a) and 1(b) show schematic representations of the experimental arrangement for pulsed CBE.[1] Pulsed ballistic beams of the source vapors, e.g., tertiarybutylphosphine (TBP) and triethylgallium (TEG) for the growth of GaP, are directed onto a substrate wafer, e.g., a silicon wafer that is heated from the back. For real-time monitoring of the kinetics of heteroepitaxy, we combine several supplementing nonintrusive optical techniques, such as p-polarized reflectance spectroscopy (PRS),[2] reflectance difference spectroscopy (RDS),[3] and laser light scattering (LLS). In addition, reflected high energy electron diffraction (RHEED) and mass spectrometric sampling are available on the system. Here we focus onto results of single-wavelength PRS and LLS. A comparison of these techniques to RDS is provided in our companion article.[4] PRS is based on the changes in the reflectivity during the formation of a heteroepitaxial stack with regard to a chopped beam of p-polarized laser light of below band gap energy that impinges onto the surface at the Brewster angle of the substrate. For monitoring of GaP heteroepitaxy on silicon, we use HeNe laser radiation at 632.8 nm wavelength, polarized by a Glan–Thompson polarizer to a ratio of the s- and p-polarized components of the incident intensity $I_s/I_p \approx 10^{-6}$. This corresponds to a residual reflectivity of the Si surface of the order of 10^{-4}. The reflected laser beam is detected by a photodiode, the output of which is processed through a phase sensitive amplifier and read into a computer. Also, the radiation that is scattered by the front and back interfaces of the growing film into a solid angle of 5° is focused by a lens at a location well removed from the plane of incidence of the impinging light beam onto the entrance

*Published without author corrections.

[a)]Department of Chemical Engineering.
[b)]Department of Materials Science and Engineering.
[c)]Department of Physics.
[d)]Electronic mail: bachmann/@mte.ncsu.edu

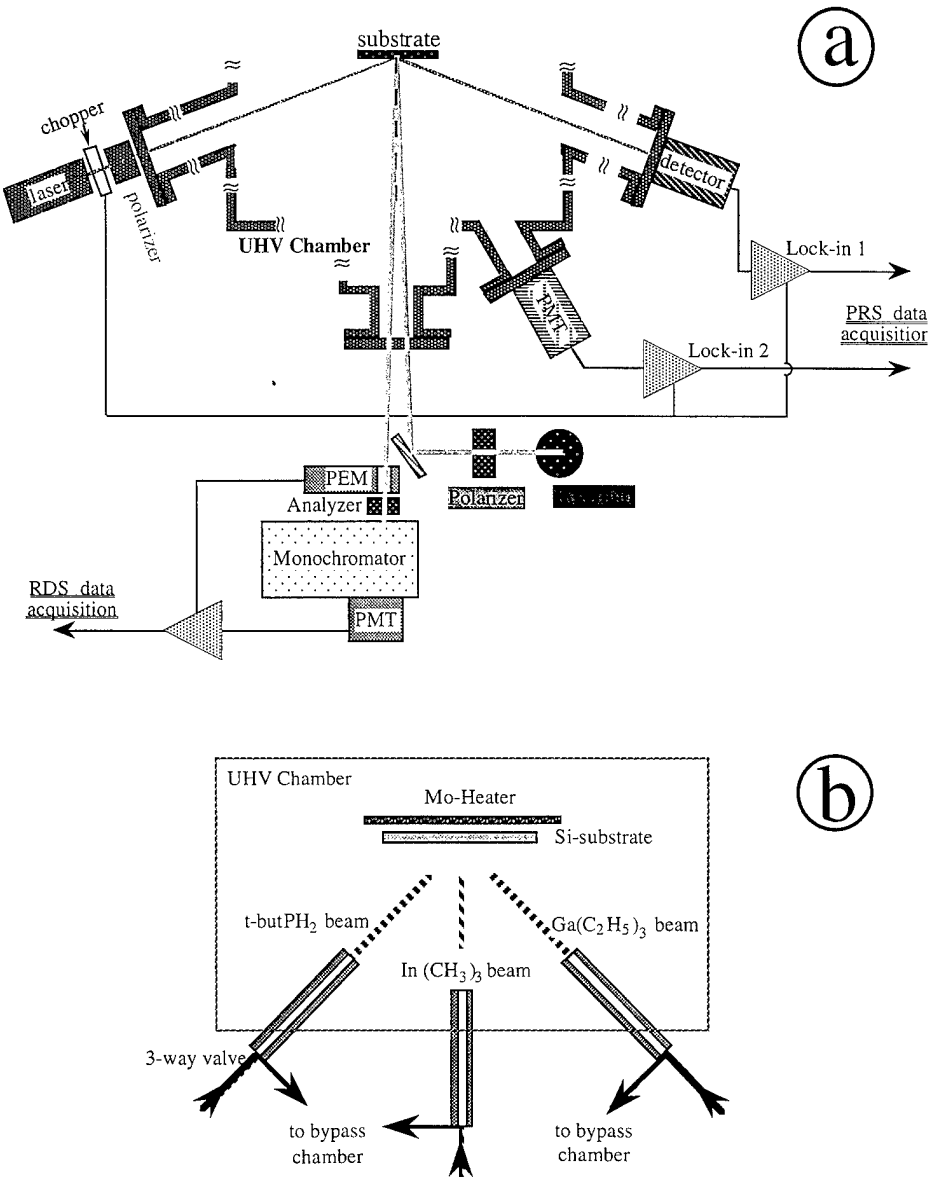


FIG. 1. Schematic representation of (a) the vacuum chamber with real-time optical process monitoring ports for single-wavelength PR, LLS, and RDS. (b) Schematic representation of source vapor switching and substrate stage.

slit of a photomultiplier. Because of its technological significance, most of the experimental data discussed here refer to the Si(100) substrate surface, including vicinal surface orientations, misoriented toward [011]. Upon appropriate surface conditioning, the misorientation for such vicinal surfaces is accommodated by double steps parallel to the $[0\bar{1}1]$ direction, which are energetically favored with regard to single steps and result in large (2×1) domains with parallel dimer rows on the top and bottom surfaces of the step.[5]

The single-wavelength PR intensity shows periodic variations in time due to interference between the partial waves reflected at the top and bottom interfaces of the heteroepitaxial GaP as illustrated in Fig. 2(a). The periodicity of these oscillations is related by analytic expressions to the thickness of the growing epilayer and can be utilized to calculate the growth rate in real time, which is reviewed in detail in Ref. 3. Superimposed to the interference oscillations is a fine structure that varies periodically with the periodicity of the source vapor pulse sequence. Figure 2(a) shows the PR intensity as a function of time over a period of growth corresponding to several interference fringes, and in the inset the magnified fine structure in the region of the second maximum. In the particular experiment of Fig. 2 the LLS intensity exhibits only weak variations, that is, the surface roughness does not undergo a large change from start to finish of the experiment. The top part of Fig. 2(b) shows a schematic representation of the pulse sequence, consisting of a *t*-butPH$_2$ pulse that is followed by pause No. 1, a TEG pulse followed pause No. 2 and so on. The bottom part of Fig. 2(b) shows the associated response in the PR intensity. From the periods in the interference fringes and the precursor cycle the average incremental increase of the GaP epilayer thickness

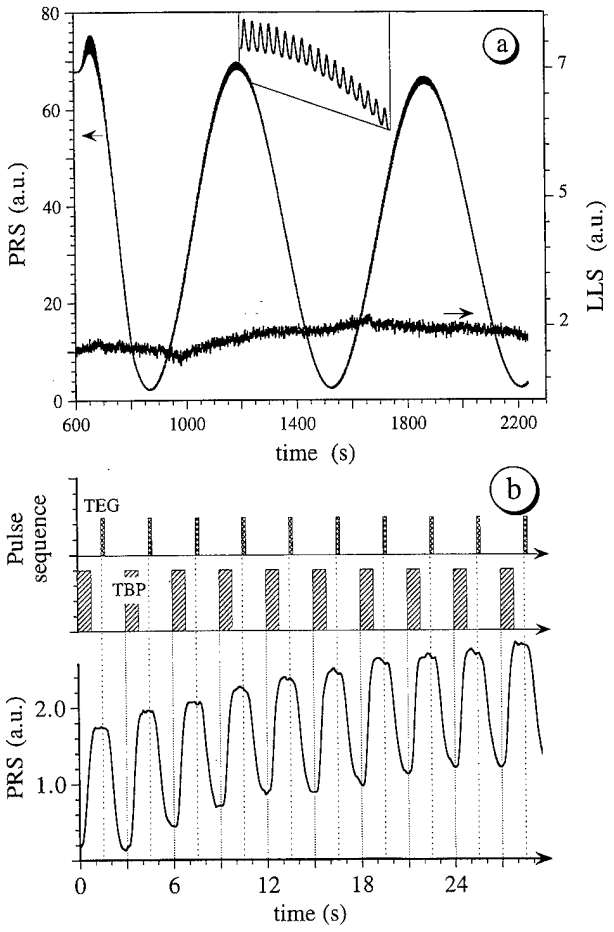


FIG. 2. (a) PR intensity as a function of time for pulsed CBE growth of GaP on Si(100), (b) pulse sequence and the associated PR fine structure.

per cycle can be determined for steady state growth with an accuracy of approximately 5%. The measured time-dependent LLS intensity provides information on the roughness of the top and bottom interfaces of the film.

III. CONTROL OF DEFECT FORMATION/PROPAGATION

At room temperature, GaP has a lattice constant of 5.4512 Å that nearly matches the lattice constant of Si of 5.4301 Å. Exact lattice matching of the III–V epilayer to the silicon substrate is possible at a chosen temperature through alloying, e.g., formation of a GaN_yP_{1-y} solid solution of the appropriate composition.[6] However, the difference in the thermal expansion coefficients for the epilayer and the substrate invariably results in the generation of strain upon thermal cycles in the processing. Also, antiphase domains (APDs) are formed irrespective of misfit strain upon III–V heteroepitaxy on group IV substrate surfaces that exhibit single steps or odd multiples thereof.[7] Fortunately, one APD variant often outgrows the other, so that the APDs remain restricted to regions close to the heterointerface.[8]

At interfaces between singular Si and III–V faces that are polar, such as {100} or {111}, either faceting into nonpolar

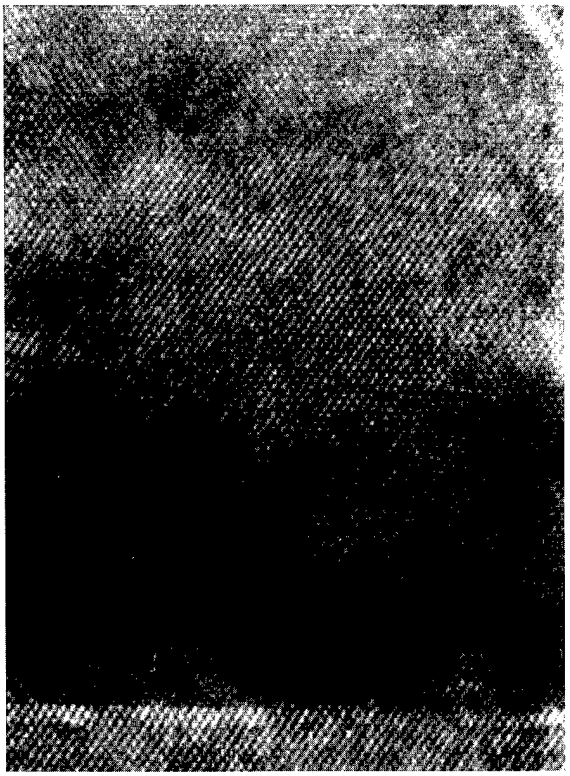


FIG. 3. Cross sectional HREM image of a GaP/Si(100) interface prepared by CBE using TEG and TBP as source materials.

surface elements, such as {110} or {112}, or intermixing have been predicted based on assessments of interface charge.[9] Figure 3 shows a high resolution transmission electron microscopy (HREM) image of a cross section of a GaP/Si(100) heterostructure created by pulsed CBE. No evidence for faceting is observed. However, charge build-up at the polar interface can contribute to interface broadening. In addition, there are other mechanisms of interface broadening such as interdiffusion, which may be enhanced by the formation of native point defects as part of the interfacial reactions. Therefore, it is presently impossible to determine whether or not electrostatic effects play a significant role. Recent STM investigations provide insights into the early stages of the interactions of the Si(100) 2×1 surface with phosphine.[10,11] At low to intermediate dose, the Si dimers on the Si(100) 2×1 surface are completely replaced by Si–P heterodimers, pushing Si surface atoms from their original Si-dimer positions into adatom positions. Since these Si adatoms are likely to become incorporated into the epitaxial film upon further growth, they contribute to interface broadening. On the P-saturated surface, the Si–Si dimers are replaced by P–P homodimers generating surface strain that is relieved, in part, by the generation of shallow trenches parallel to the dimer rows that are bound by elongated {111} facets and can affect the kinetics of subsequent heteroepitaxial overgrowth.[12] In particular, the {111} facets in the Si(100) surface upon overgrowth by GaP can act as nucleation sites for the formation of stacking faults that propagate deeply into the heteroepi-

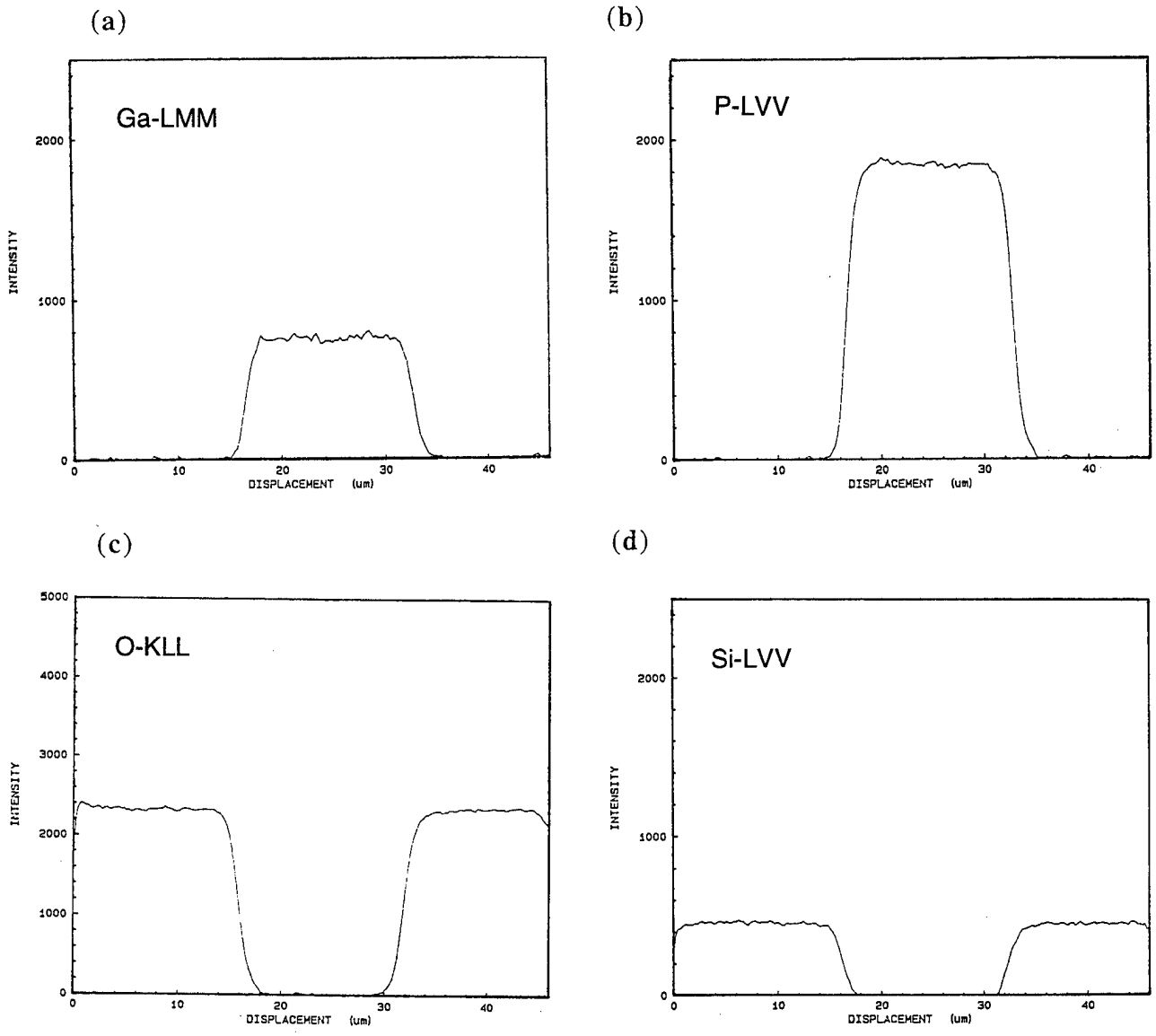


FIG. 4. Auger electron spectroscopy line scans at the energies of the (a) Ga-LMM, (b) P-LVV, (c) O-KLL, and (d) Si-LVV lines over a SiO_2-masked silicon surface with 10×10 μm window areas filled with epitaxial GaP.

taxial film. Driving the phosphorus coverage of the surface to saturation thus is not necessarily a desirable condition of GaP heteroepitaxy on Si, which must be weighed against the control of stoichiometry, requiring large V:III source vapor flux ratio.

An important cause of defect formation is residual contamination of the silicon surface. In our experiments, a modified RCA cleaning procedure ending with an aqueous HF dip was used to create a hydrogen-passivated Si(100) 1×1 surface prior to transfer into the CBE chamber through a load-lock. This initial reconstruction was verified by RHEED. The control of the pH and the concentration of the etching solution, absence of dissolved oxygen, and the conditions of oxidation prior to HF (or, for {111} substrate wafers, NH_4F) etching are all important factors in controlling the surface roughness and chemical integrity of the hydrogen-terminated silicon surface.[13–15] Recent quantitative high resolution XPS studies of carefully prepared hydrogen-terminated Si(001) (1×1) surfaces revealed surface concentrations of residual oxygen and fluorine atoms in the mid-10^{12} cm^{-2} range.[16] Because of the high selectivity of the epitaxial growth of GaP on Si vs SiO_2, this level of residual contamination can make a significant contribution to the generation of defects.

Figure 4 shows Auger electron spectroscopy (AES) line scans across 10×10 μm windows in a SiO_2 mask on a Si(100) substrate after deposition of heteroepitaxial GaP in the window areas at (a) 120, (b) 1070, (c) 510, and (d) 92 eV. They reveal a substantially larger kinetic barrier to nucleation of GaP on the SiO_2-masked surface areas as compared to the bare Si window areas. Although high selectivity is a welcome property in the context of device fabrication, in the context of defect formation, the kinetic barrier to overgrowth of surface areas contaminated by oxygen is a detriment. Upon forced overgrowth of such contaminated surface

patches, defects may form in the epitaxial film due to the generation of strain. This strain may be relieved by the formation of 60° dislocations that dissociate into partials. Also, for amorphous contaminant patches, loss of registry to the underlying substrate lattice can cause the formation of stacking faults. Similar experiments conducted on partially SiC-coated Si surfaces show that the overgrowth by GaP is also highly selective with regard to SiC, that is, carbon contamination of the silicon surface is also of concern. Burying the surface contaminants under a homoepitaxial layer of silicon that exposes a virgin surface at the initiation of heteroepitaxy can result in improvements, because the stacking fault energy for silicon (65 mJ/m^2) is substantially larger than for GaP (41 mJ/m^2).[17] The relatively small stacking fault energy of GaP is a consequence of the decrease of the stacking fault energy in the zincblende structure III–V compounds with increasing ionicity of the bonding. Thus, the structure becomes unstable in the electronegativity difference range between AlP, which has zincblende structure at slightly larger lattice constant than silicon and InN that has wurzite structure at smaller lattice constant than silicon. Therefore, the insistence on nearly lattice-matched materials combination generates conditions, where the gains in minimizing strain are paid for by losses in maximizing the probability of stacking errors. Indeed, high densities of stacking faults have been observed in GaP/Si(100) heterostructures by TEM.[18]

Since in zincblende structure compounds stacking faults nucleate on {111} and, for GaP growth on Si(100), the GaP nuclei are bound by {111} facets,[18] the formation of three-dimensional faceted GaP islands in the initial period of nucleation and overgrowth of the Si surface must be curtailed to control the density of planar defects in the heterostructure. The aspect ratio A of height to width of the nuclei depends on both thermodynamic and kinetic parameters. The assessment of habit of heteroepitaxial nuclei governed by thermodynamics is possible on the basis of the modified Wulff's law

$$\frac{\sigma_{hkl}-\alpha_{hkl}}{r_{hkl}}=\frac{\Delta\mu}{2v}, \tag{1}$$

stating that the habit of the nuclei—formed under the conditions of heteroepitaxy on the substrate surface of orientation (hkl)—is related to the equilibrium habit of homogeneously formed three-dimensional nuclei at same supersaturation $\Delta\mu$ by truncation.[19] This truncation is characterized by central distance r_{hkl} in the heterogeneously formed nucleus to the interface of orientation (hkl), which, for a given supersaturation and volume v of the crystal building blocks, depends on the relative magnitudes of surface free energy σ_{hkl} on the facet (hkl) on the homogeneously formed nucleus and the interfacial attachment energy α_{hkl}. A three-dimensional nucleation mechanism thus ensues for $0<\alpha_{hkl}<2\sigma_{hkl}$, with decreasing aspect ratio A (height/width) of the three-dimensional nuclei for increasing α_{hkl}, approaching a two-dimensional nucleation and growth mechanism in the limit $\alpha_{hkl}\rightarrow 2\sigma_{hkl}$. However, in addition to thermodynamic causes for three-dimensional overgrowth, kinetic enhancements of A results from the catalysis of TBP decomposition on the GaP surface elements.[20] Upon growth, the three-dimensional nuclei thus formed expand laterally and meet with patches of surface contaminants that are randomly distributed on the silicon surface. In view of the large barrier to GaP growth on the surface of the oxide patch, its overgrowth by the GaP nucleus requires a substantial fluctuation in the chemical potential that, given the relatively low stacking fault energy, results in a high probability of stacking fault formation. Although it is thus difficult to avoid the nucleation of stacking faults, their propagation into the III–V epilayer can be minimized by structuring of the Si(100) surface into V groves bound by (111) and $(\bar{1}11)$ facets along $[0\bar{1}1]$. This concept has been applied successfully in controlling stacking fault propagation in InP/Si(100) heterostructures.[21–24]

The blockage of the propagation of stacking faults is explained by interactions between the partial dislocations that support them. Stacking faults form upon dissociation of unit 60° dislocations on (111) and $(\bar{1}11)$ according to

$$\mathbf{b}_1=\frac{a}{2}[\bar{1}10]=\frac{a}{6}[\bar{2}11]+\frac{a}{6}[\bar{1}2\bar{1}]=\mathbf{b}_2+\mathbf{b}_3, \tag{2}$$

$$\mathbf{b}_4=\frac{a}{2}[101]=\frac{a}{6}[211]+\frac{a}{6}[1\bar{1}2]=\mathbf{b}_5+\mathbf{b}_6, \tag{3}$$

respectively, where $\mathbf{b}_1$ and $\mathbf{b}_4$ refer to the Burgers' vectors of the 60° dislocations and $\mathbf{b}_2$, $\mathbf{b}_3$ and $\mathbf{b}_5$, $\mathbf{b}_6$ refer to the Burgers' vectors of the associated partials. This is illustrated in Fig. 5(a). Because of their smaller length as compared to $\mathbf{b}_1$ and $\mathbf{b}_4$, the dissociation reactions are energetically favored and provide a driving force for the generation of the stacking faults suspended between the partials. For InP, which has a stacking fault energy of 18 mJ/m^2, large separations between the partials have been observed.[25] Therefore, for submicrometer groove depth, the faults can expand after nucleation over the entire V-groove facet. Upon interaction between the partials with Burgers' vectors $\mathbf{b}_3$ and $\mathbf{b}_6$ on the intersection of the two glide planes according to

$$\frac{a}{6}[1\bar{1}2]+\frac{a}{6}[\bar{1}2\bar{1}]\rightarrow\frac{a}{6}[011]=\mathbf{b}_7, \tag{4}$$

Lomer–Cottrell sessile dislocations of pure edge character can be formed. Although this dislocation cannot move by glide, in principle, it can move by climb. However, dislocation climb in this case is tagged to motion of the entire stacking fault, which is highly unlikely. In view of the polar nature of the dislocations and possible presence of other defects, that is, APBs and misfit dislocations near the interface, more complex interactions can occur and require further study. Also, the above discussed mechanism of self-termination of stacking fault propagation relies on perfect faceting at a submicrometer scale and the absence of nucleation of stacking faults on {111} variants other than (111) and

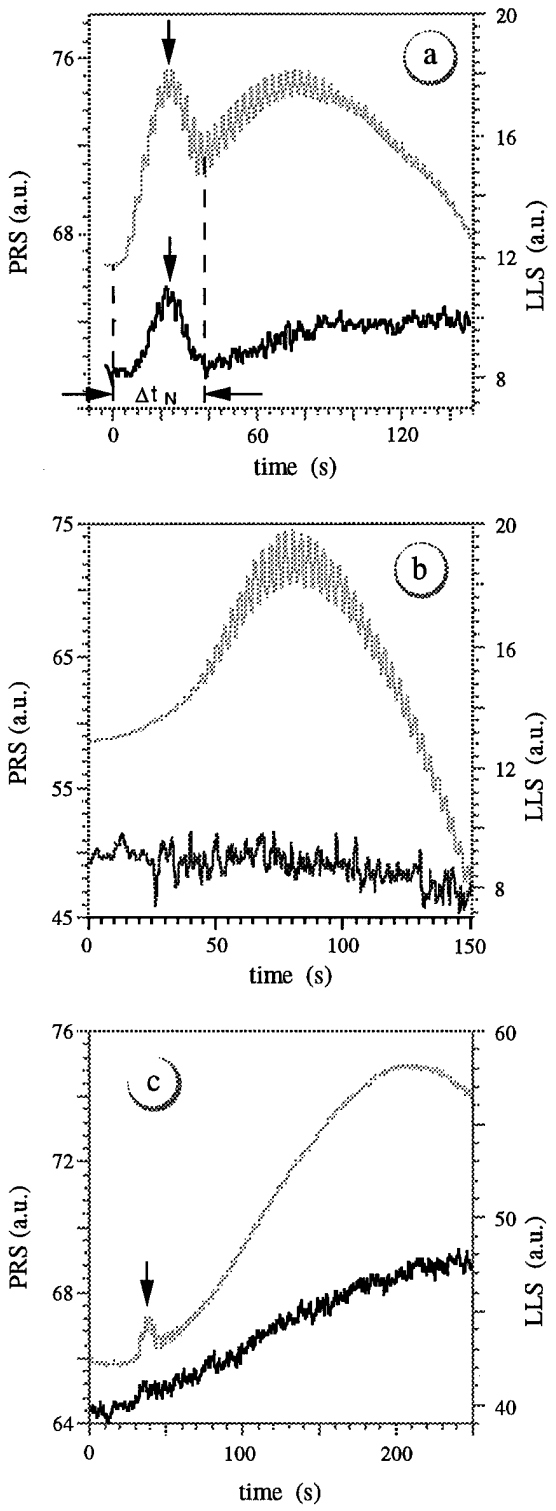


FIG. 5. Initial PR intensity traces for pulsed CBE of GaP on Si(100): (a) at small Ga supersaturation; (b) under optimized processing conditions; (c) at large Ga supersaturation.

$(\bar{1}11)$ due to the faceting of three-dimensional GaP nuclei or uncontrolled formation of microfacets. The first of these two potential problems can be addressed on the basis of optical process monitoring during the initial phase of GaP nucleation and overgrowth of the Si surface by a contiguous GaP film.

IV. INITIAL PHASE OF NUCLEATION AND OVERGROWTH

Since, under the conditions of vapor phase epitaxy, the dielectric function of the ambient is smaller than the dielectric function of the deposited compound, a corrugated incomplete heteroepitaxial film is associated with an effective dielectric function that increases with decreasing pore size upon progressive overgrowth of the substrate surface from initially close to the ambient dielectric function to the dielectric function of the contiguous heteroepitaxial film. Consequently, during this initial phase of heteroepitaxy, distinct features are observed in the time-dependent PR intensity, marked by a vertical arrow in Fig. 5(a). Figures 5(a)–5(c) represent a set of experiments carried out under the same conditions, except for different TEG flux in the initial 20 s of the experiment. At small TEG flux, that is, small supersaturation of the surface with regard to Ga precursors to growth, the low nucleation rate results in the formation of widely spaced nuclei that grow to substantial height before coalescing into a contiguous epilayer. The associated roughening of the surface is observed in Fig. 5(a) in the time-dependent LLS intensity and is also marked by a vertical arrow. At the time of complete coverage of the substrate surface by a contiguous heteroepitaxial film, that is, at the end of the period labeled Δt_N, the reflectance and LLS signals merge into the time dependence for contiguous film growth. The magnitude of the PR and LLS intensity changes in this initial period Δt_N depends on the aspect ration A of the height to width of the GaP islands. *Ex situ* atomic force microscopy investigations reported by us elsewhere confirm the roughening of the surface relative to the initial roughness of the silicon surface in the nucleation period under the conditions of Fig. 5(a), the subsequent smoothening of the epilayer, once a contiguous GaP film has been formed, and the increase of surface roughness upon further growth.[26] This correlation of the initial feature during the nucleation stage to significant roughening in the later stages of steady-state film growth has been observed consistently in several experiments. Therefore, we conclude that improvements in the control of A in the initial nucleation and overgrowth phase can affect the quality of the epilayer during later stages of growth.

Figure 5(b) shows the result of increasing the TEG flux in the first 20 s of the nucleation and overgrowth period at the same TBP flux as in Fig. 5(a), increasing the supersaturation on the surface. Since at higher supersaturation a higher density of nuclei of smaller height is generated and an increased supply of nutrients exists between these islands, they merge into a contiguous film upon a shorter period of lateral growth. In addition, the higher Ga supersaturation on the silicon surface elements between the GaP nuclei should reduce the local differences in the catalytic properties on the partially GaP coated surface. In contrast to Fig. 5(a), no excess intensity in the time-dependent PR and LLS signals is observed during the initial phase of film formation under the conditions of Fig. 5(b). However, there exists a sensitive upper limit in the initial increase of the TEG flux at a given TBP flux, above which gallium droplets form. Also, the tim-

ing of the merging of the initially higher TEG flux into the smaller optimum TEG flux for steady-state growth of the contiguous heteroepitaxial film is critical to avoid phase separation on the surface. Thus, upon exceeding a critical upper limit for the gallium supersaturation, surface roughening due to Ga droplet formation results in a rapid deterioration of the surface morphology, as illustrated in Fig. 5(c), which shows a clearly resolved nucleation peak in the PR intensity—marked by a vertical arrow—and an immediate continuous rise in the LLS intensity that accelerates in later periods of growth not shown in Fig. 5(c). Under extreme conditions, a vapor–liquid–solid (VLS) growth mechanism may take over, leading to very large aspect ratios of the surface features.

V. KINETICS OF STEADY-STATE HETEROEPITAXIAL GROWTH

Details of the growth kinetics after complete sealing of the silicon surface by a contiguous GaP film are revealed by the response in the time-dependent PR intensity to variations in the source vapor cycle and perturbations of the steady-state surface composition. For example, in a preceding publication,[27] we have shown that the growth rate per cycle v_c for constant source vapor pulse heights and widths is insensitive to the duration of pause No. 1 between the trailing edge of the TBP pulse and the leading edge of the following TEG pulse. However, v_c increases with increasing length of pause No. 2 between the trailing edge of the TEG pulse and the leading edge of the following TBP pulse. We interpreted this result as a consequence of a kinetic barrier to the complete dealkylation of the impinging TEG molecules on the GaP surface. Their conversion into surface species that contribute to the growth, for example, Ga surface atoms, represents the time limiting step in the heteroepitaxial growth process. Thus, a reservoir of TEG fragments that do not directly contribute to epitaxial growth coexists on the surface with active fragments that are derived from the inactive fragments over time and feed the GaP growth. Since v_c can correspond to several bilayers of GaP per cycle,[6] in steady state, a store of up to several monolayers of the source vapor fragments must exist on the surface to sustain the growth. Therefore, the modeling of steady-state heteroepitaxial growth must be based on a four-layer stack: (a) ambient; (b) surface reaction layer (SRL); (c) heteroepitaxial film; (d) substrate. The SRL is strictly stabilized by the kinetics of the primary pyrolysis and follow-up reactions. Since only part of the molecules in the SRL are attached to the surface, under this specific condition of pulsed CBE, part of the chemical reactions on the surface proceed homogeneously between molecules that not at all or only weakly bound to the surface atoms of the GaP film and consequently are randomly oriented. Thus, a set of coupled parallel and consecutive homogeneous reactions within the SRL proceed simultaneously with heterogeneous surface reactions and crystallization processes. The latter heterogeneous reactions are coupled to the homogeneous reactions through intermediates. However, there exist also competing reaction pathways, e.g., β elimination of radical ligands by reaction with adsorbed hydrogen atoms in parallel to homogeneous radical chemistry in the SRL. Since the bonding of products to the surface strengthens as the radical ligands are shed, the homogeneous reactions will dominate the initial steps of dealkylation while denuded products, such as Ga and P atoms that bond strongly to the surface and eventually incorporate into the GaP lattice, will probably not be present at large concentration in the random part of the SRL.

The complex reflectivity coefficient for p-polarized light impinging onto the four-layer stack for a specific value of the film thickness can be calculated on the basis of the Fresnel equations.[28] In addition to knowledge of the dielectric functions of the Si substrate, the GaP film, and the ambient, the angle of incidence and wavelength this requires an assessment of the dielectric function of the surface reaction layer and its thickness. Generally, both the dielectric function and thickness of the SRL are functions of time.

For a preliminary assessment, we recently introduced a simplified model that reduces the complex surface reaction scheme discussed above to an exponential build up of t_b upon exposure of the surface to a pulse of TBP, followed by an exponential decay to zero thickness after supply of the pulse of TEG, which is accompanied by an incremental increase in the thickness of the GaP epilayer.[29] This model reveals two important features in the PR intensity:

(i) A periodic increase and decrease in the amplitude of the fine structure that is tagged to the interference oscillations. The relative positions of minima in the fine structure amplitude with respect to the positions in the minima of the interference oscillations depend on the value of the dielectric function of the SRL.

(ii) A change of the signs of the response to the TBP and TEG pulses at the minima in the fine structure amplitude. Because of the turnover in the pulse response at these positions, we refer to them as turning points in the PR fine structure.

Here we expand the model by adding an assessments of the changes in the dielectric function due to changes in the composition of the homogeneous part of the SRL as the growth process progresses. Since in this case the molecules are randomly spaced and oriented the electronic transitions, contributing to the dielectric function refer to localized states. Thus, the dielectric function can be represented by a modification of the Sellmeier equation, summing the contributions of all constituent molecules, that is,

$$\epsilon_b = 1 - \frac{4\pi}{V} \sum_l \sum_k \sum_{Q'_{kl} Q_{kl}} \frac{2|\xi \mu'_{Qkl Qkl}|^2 E'_{Qkl Qkl}}{E_+^2 - E'^2_{Qkl Qkl}}, \qquad (5)$$

where V is the volume of the surface reaction layer, ξ and $\mu'_{Qkl Qkl}$ represent the unit polarization vector and the matrix element of the electric dipole operator for the electronic transition between filled states Q_{kl} and empty states Q'_{kl} associated with the individual molecule labeled by subscript k of the constituent labeled by subscript l, $E'_{Qkl Qkl}$ is the energy of this transition, and $E_+ = \hbar\omega_+ = \hbar(\omega + i\eta)$ is the energy at which ϵ_b is evaluated in the limit of turn-on parameter

$\eta \to 0^+$. Since all molecules of a particular constituent have same molecular orbitals and electric dipole-allowed transitions, the summation over k results simply the number of molecules of each species. Also, the number of molecules of one particular constituent l—divided by the volume—represents its concentration c_l in the surface reaction layer. Therefore, Eq. (5) provides for the desired correlation between the composition of the homogeneous part of the SRL and the dielectric function.

However, Eq. (5) does not account for intermolecular interactions, which in view of the high concentrations of at least some of the constituents must be considered. Therefore, we replace the concentrations of all constituents of the surface reaction layer by activities, that is, $a_l = f_l c_l$, with a specific activity coefficient $f_l \neq 1$ for each constituent that generally depends on both temperature and composition. Then,

$$\epsilon_b = 1 - 4\pi \sum_l a_l \sum_{Q_l' Q_l} \frac{2|\xi \mu'_{Q_l Q_l}|^2 E'_{Q_l Q_l}}{E_+^2 - E'^2_{Q_l Q_l}}, \tag{6}$$

where the first sum extends over all constituents of the surface reaction layer and the second sum extends over the electric dipole allowed transitions from filled states for each of these constituent molecules.

Note that Eqs. (5) and (6) rely on the assumption that the volume of the surface reaction layer is constant in time, which holds only in the limit of small variations in its composition. Therefore, we restrict in the following discussion of specific correlations between surface reaction layer composition and responses in the PRS intensity to experiments carried out at constant temperature with a preceding reproducible surface conditioning step and small variations in composition. Within these restrictions, the activity coefficients can be considered to be constants, which enables the analysis of the homogeneous reactions in the SRL on the basis of analytic expressions for the integrated rate laws.

The pyrolysis of the primary source molecules impinging onto the SRL is represented by

$$C_4H_9\bar{P}H_2 \to C_4H_9^{\bullet} + \bar{P}H_2^{\bullet}, \tag{7}$$

$$Ga(C_2H_5)_3 \to Ga(C_2H_5)_2^{\bullet} + C_2H_5^{\bullet}. \tag{8}$$

In these equations, fragments located in the SRL are represented in italics and molecules that are expelled into the vapor phase and pumped off are represented in plain lettering. The t-butyl-phosphine fragments and ethyl radicals can decompose further under the generation of adsorbed hydrogen atoms according to

$$C_2H_5^{\bullet} \cdot C_2H_4 + H^{\bullet}, \tag{9}$$

$$C_4H_9^{\bullet} + C_4H_9PH_2 \to C_4H_{10} + C_4H_9PH^{\bullet}, \tag{10}$$

$$C_4H_9PH^{\bullet} \to C_4H_9 + |\bar{P}H, \tag{11}$$

$$C_4H_9^{\bullet} \to C_4H_8 + H^{\bullet}, \tag{12}$$

$$\bar{P}H_2^{\bullet} \to |\bar{P}H + H^{\bullet}, \tag{13}$$

$$|\bar{P}H \to |\bar{P}^{\bullet} + H^{\bullet}, \tag{14}$$

that assist the dealkylation of the TEG fragments on the surface according to

$$Ga(C_2H_5)_2^{\bullet} + H^{\bullet} \leftrightarrow \bar{G}a(C_2H_5) + C_2H_6, \tag{15}$$

$$\bar{G}a(C_2H_5) + H^{\bullet} \leftrightarrow \bar{G}a^{\bullet} + C_2H_6, \tag{16}$$

in competition with

$$C_2H_5^{\bullet} + H^{\bullet} \to C_2H_6, \tag{17}$$

$$C_4H_9^{\bullet} + H^{\bullet} \to C_4H_{10}, \tag{18}$$

$$|\bar{P}H + H^{\bullet} \to \bar{P}H_2^{\bullet}, \tag{19}$$

$$\bar{P}H_2^{\bullet} + H^{\bullet} \to \bar{P}H_3, \tag{20}$$

and

$$H^{\bullet} + H^{\bullet} \to H_2. \tag{21}$$

The final step in the overall reaction of TBP and TEG to form GaP is the incorporation of Ga and P atoms into the GaP lattice. As established in a series of experiments using deuterated TBP by Li *et al.*[30] for the growth of GaP using TBP and trimethylgallium (TMG) as source vapors, adsorbed PH_2 and PH radicals play an important direct role in the pyrolysis of both TBP according to reactions (11)–(14) and TMG according to

$$PH_x + Ga(CH_3)_y \to PH_{x-1} + Ga(CH_3)_{y-1} + CH_4, \quad x = 1,2 \text{ and } y = 1,2. \tag{22}$$

Although there exist significant differences in the interactions of TEG and TMG fragments with the GaP surface, we expect that adsorbed PH_x also plays an important role in the heterogeneous dealkylation of TEG and its fragments. All reactants and products shown in the equations in italics contribute to the composition of the surface reaction layer, which exists strictly on the basis of kinetic barriers to the desorption of its constituents, with the exception of gallium, which has at the growth temperature a vapor pressure well below the background pressure in the reactor. We note that the set of equations written above is not exhaustive; that is, other reactions may be added as our knowledge of details grows. Since the heterogeneous reactions that contribute to crystal growth probably entail competition for active surface sites and replacement of inhibitors at such sites, it is reasonable to expect that they will be slower than the homogeneous radical reactions in the SRL. Also, the denuded fragments, e.g., gallium atoms and possibly monoethyl gallium (MEG), are expected to be bonded tighter to the surface than organometallic fragments that retain several ligands, e.g., diethylgallium (DEG). Therefore, we limit the discussion of the homogeneous kinetics onto reactions involving primary and secondary products of the pyrolysis reaction. Under conditions of pulsed growth, the concentrations and activity coefficients of all reactants and products vary with time, and reactions (9)–(22) are coupled through the bimolecular reactions, having surface hydrogen as one of the reactants. No

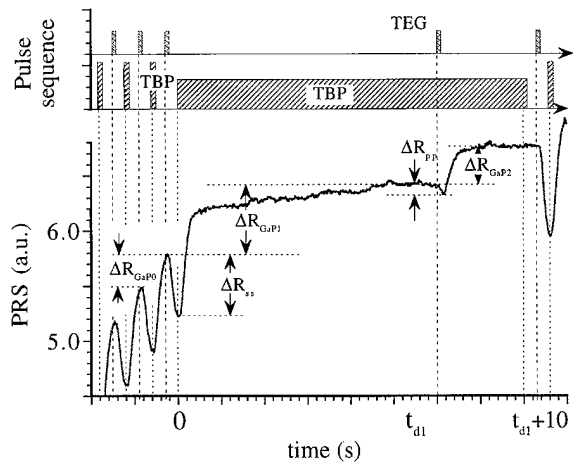


FIG. 6. Schematic representation of the timing of the source gas flow (top) and response in the PR intensity to a TEG probe pulse during interruptions of steady state and switching to continuous TBP flow (0.8 sccm) during the interruption.

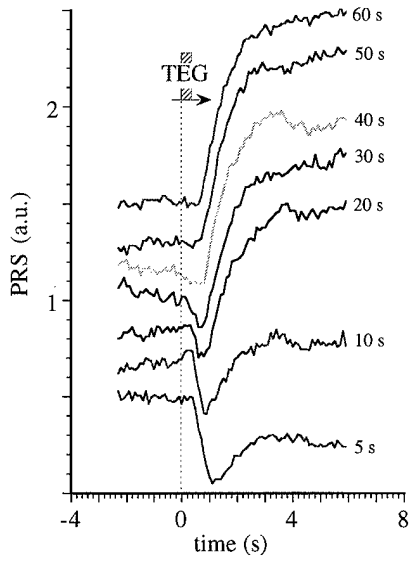


FIG. 7. Single wavelength PRS response to the TEG probe pulse under the conditions of Fig. 9 at λ=632.8 nm for selected values of t_{d1}.

analytical solutions are available for the quantitative analysis of the PR intensity, in this case requiring the application of numerical methods.

Considerable simplifications are gained under the conditions of interrupted steady-state growth, where the TEG pulses are switched off at initiation time t_{int}, continuing the exposure of the surface to a continuous TBP flux, as illustrated in Fig. 6. Since we use in our experiments typically TBP to TEG flux ratios in the range 10:1 to 40:1, at constant TBP flux, the hydrogen and phosphorus activities in the SRL are only marginally perturbed by the TEG probe pulse and can be taken as constants during the delay period t_{d1}. Going back to Fig. 5(b), we note that initially the PR response to the application of TEG pulses after heating the surface of the substrate under TBP exposure to the growth temperature is very small. The development of the fine structure to full swing from low PR intensity after application of the TEG pulse to high PR intensity after the TBP pulse requires several cycles. Also, we note that in Fig. 6, the response in the PR intensity ΔR_{pp} to a TEG probe pulse at a delay t_{d1}, after switching from pulsed to continuous TBP flow and discontinuation of TEG pulse supply at time t_{int}, is substantially smaller than the response to the last TEG pulse under steady-state conditions ΔR_{ss}. We interpret these features in Figs. 5(b) and 6 as being due to the build-up of the surface reaction layer under the conditions of initial nucleation and overgrowth and its destruction upon interruption of steady-state growth, respectively, that is, relate ΔR_{pp} to the deviations of the surface reaction layer from steady state, which includes its coupling to the surface of the substrate. Thus, we can utilize the response to a TEG probe pulse for an assessment of the kinetics of the chemical changes in the reaction layer upon perturbations of steady state, e.g., switching from pulsed to continuous TBP flow and discontinuation of TEG pulse supply.

In the following discussion, we set $t_{\text{int}} \equiv 0$, that is, measure t_{d1} from $t=0$ to the rising edge of the TEG probe pulse. The exposure to the continuous TBP beam is maintained for a period $t_{d1} \leqslant t \leqslant t_{d1}+t_r$. The time interval t_r refers to the recovery period after the application of the TEG probe pulse in the PR intensity before reconditioning the surface to steady-state growth for the next experiment and is typically 10 s. A considerable number of experiments can be conducted within one heteroepitaxial growth run. Comparisons between experiments from several runs were made to establish the reproducibility of our findings. All experiments described below applied the same pulse cycle sequence for conditioning to steady-state growth. Figure 7 shows the decrease in the amplitude of the response in the PR intensity to the TEG prove pulse fort selected values of t_{d1} of increasing in length. Note that, in view of the increase in the amplitude of the PR intensity with increasing distance from the turning points, the response to the TEG probe pulse in a particular experiment depends on the timing of the experiment relative to the extrema in the interference oscillations. In the following evaluations, this effect is eliminated by normalization of ΔR_{pp} to ΔR_{ss}.

Figures 8(a) and 8(b) show plots of the normalized measured response $\Delta R_N \equiv \Delta R_{pp}/\Delta R_{\text{ss}}$ versus t_{d1} for two experiments under the conditions of Fig. 6. The curves labeled exp represents cubic spline fits to the experimental data. While for a given set of experimental parameters, the initial decrease of ΔR_N with time is reproducible, after 5–10 s, variations between different experiments are observed. This is interpreted in terms of the kinetics of the chemical reactions on the GaP surface, starting with the decay of DEG radicals according to Eq. (15). The rate law for this reaction step is of the form

$$-\frac{da_{\text{Ga}(\text{C}_2\text{H}_5)_2}}{dt} = k_{15} a_{\text{H}} a_{\text{Ga}(\text{C}_2\text{H}_5)_2}. \tag{23}$$

Since at large constant flux of TBP, the activity of hydrogen

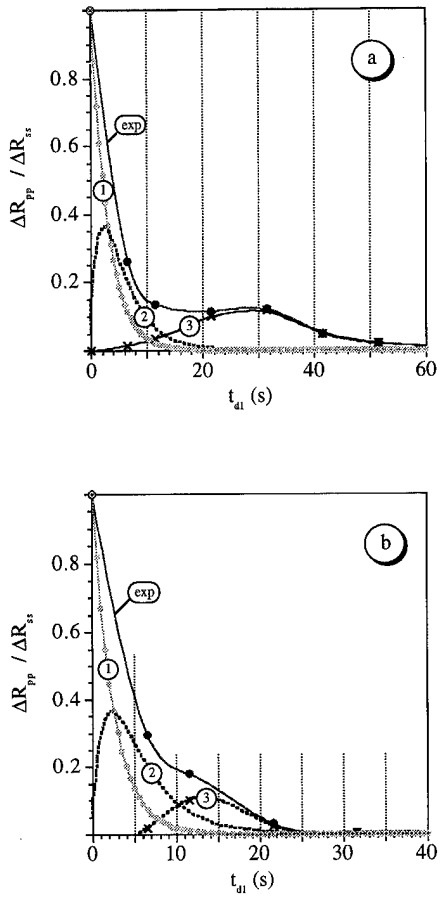


FIG. 8. Plot of $\Delta R_{pp}/\Delta R_{ss}$ vs time delay t_{d1}—measured from last TEG pulse under steady-state conditions of growth—to the application of a single TEG probe pulse: curve 1 based on Eqs. (6) and (25), curve 2 based on Eqs. (6) and (28). (a) and (b) refer to two different experiments run under the same conditions.

in the surface reaction layer a_H at $t \geq 0$ is only slightly perturbed by the ongoing surface reactions, we may in a first approximation treat it as a constant, that is, rewrite Eq. (23) as

$$-\frac{da_{Ga(C_2H_5)_2}}{dt} = k'_{15} a_{Ga(C_2H_5)_2}. \tag{24}$$

With initial condition $a_{Ga(C_2H_5)_2}(t=0) = \bar{a}_{Ga(C_2H_5)_2}$, we thus get the integrated rate law

$$a_{Ga(C_2H_5)_2}(t) = \bar{a}_{Ga(C_2H_5)_2} \exp\{-k'_{15} t\} \approx \bar{a}_{Ga(C_2H_5)_2} \exp\left(-\frac{t}{\tau_1}\right), \tag{25}$$

that is, the DEG activity decays from the value established in steady state $\bar{a}_{Ga(C_2H_5)_2}$ at $t=0$ exponentially with time constant $\tau_1 = 1/k'_{15}$, that is, $k'_{15} = 0.4\ s^{-1}$. We note that under the conditions of the experiment the activities of PH_x are also constants, so that adding Eq. (22) as a parallel reaction path merely changes the interpretation of k'_{15}, but not the rate equation.

The time evolution of the activity of monoethylgallium (MEG) in the surface reaction layer is composed of a contribution to the MEG activity $\Delta a_{Ga(C_2H_5)}$—generated in the course of the decay of the DEG according to reaction (15)—that adds at $t>0$ to the activity of MEG $\bar{a}_{Ga(C_2H_5)}$ measured at $t=0$. By reaction (15), the rate with which this contribution is generated is coupled to the rate of the DEG decay according to

$$\frac{d(\Delta_{Ga(C_2H_5)})}{dt} = -\frac{d(a_{Ga(C_2H_5)_2})}{dt}. \tag{26}$$

Also, with the assumption of constant hydrogen activity, the rate law for reaction (16) can be written as

$$-\frac{da_{Ga(C_2H_5)}}{dt} = k_{r16} a_H a_{Ga(C_2H_5)} = k'_{16} a_{Ga(C_2H_5)}. \tag{27}$$

Therefore, we get

$$\begin{aligned} a_{Ga(C_2H_5)}(t) &= \left[\bar{a}_{Ga(C_2H_5)} + k'_{15}\bar{a}_{Ga(C_2H_5)_2} \times \int_0^t \exp\left(-\frac{t}{\tau_1}\right) dt\right] \exp(-k'_{16}t) \\ &= \left\{\bar{a}_{Ga(C_2H_5)} + \bar{a}_{Ga(C_2H_5)_2}\left[1 - \exp\left(-\frac{t}{\tau_1}\right)\right]\right\} \times \exp\left(-\frac{t}{\tau_2}\right), \end{aligned} \tag{28}$$

where $\tau_2 = 1/k'_{16}$. This function has a single extremum at

$$t_{max} = \tau_1 \ln\left(\frac{\bar{a}_{Ga(C_2H_5)_2}}{\bar{a}_{Ga(C_2H_5)2} + \bar{a}_{Ga(C_2H_5)_2}} \frac{\tau_1 + \tau_2}{\tau_1}\right) \geq 0 \tag{29}$$

and initial slope

$$\left.\frac{da_{Ga(C_2H_5)}}{dt}\right|_{t=0} = \frac{\bar{a}_{Ga(C_2H_5)_2}}{\tau_1} - \frac{\bar{a}_{Ga(C_2H_5)}}{\tau_2} \geq 0, \tag{30}$$

for $\bar{a}_{Ga(C_2H_5)}/\tau_2 \leq \bar{a}_{Ga(C_2H_5)_2}/\tau_1$. Fitting the initial decay of ΔR_N on the basis of $F_1(t_{d1})$ and $F_2(t_{d1})$ can be accomplished and can be extended in time by incorporating a time-dependent loss

$$a_{GaC_2H_5} = \left\{\bar{a}_{Ga(C_2H_5)} + L(t)\bar{a}_{Ga(C_2H_5)_2} \times \left[1 - \exp\left(-\frac{t}{\tau_1}\right)\right]\right\} \exp\left(-\frac{t}{\tau_2}\right) \tag{31}$$

with loss factor $L(t)$ starting from $L(t=0)=1$ and varying with same time dependence as $a_{GaC_2H_5}(t)$, corresponding to a maximum loss of 10% at the maximum in $F_2(t_{d1})$ for both experiments of Fig. 8. The curves labeled 1 and 2 in Figs. 8(a) and 8(b) are modeled with the same set of parameters $\tau_1 = 2.5$ s, $\tau_2 = 4.2$ s, $L(t)$, and $\bar{a}_{Ga(C_2H_5)} = 0.05 \times \bar{a}_{Ga(C_2H_5)_2}$. A possible interpretation of the loss of MEG is its parallel dealkylation through heterogeneous surface reactions at a rate that is tagged to the concentration of MEG in the homo-

geneous part of the SRL. The difference between the predicted behavior according to Eqs. (25) and (28) and the experimental data represented in Figs. 8(a) and 8(b) by the curves labeled 3 in the later stages of the decay of the SRL from its steady-state composition is attributed to surface reactions that become dominant as the TEG fragments become increasingly denuded of radical ligands. In this case, the dipole approximation is no longer a reliable basis for the evaluation of the dielectric function and no interpretation of these surface reactions can be offered on the basis of single pulse PR monitoring.

VI. SUMMARY AND CONCLUSIONS

On the basis of real-time process monitoring by p-polarized reflectance in combination with laser light scattering we have shown that, for a given flux of the phosphorus source vapor, there exists a narrow band of initial gallium supersaturations for the overgrowth of the silicon surface by a contiguous film of GaP with minimum roughening. Also, we have shown that steady-state heteroepitaxial growth must be represented by a four-layer stack: ambient–surface reaction layer–epilayer–substrate. Depending on the source vapor fluxes, the surface reaction layer can have a thickness of several monolayers of fragments of the gallium and phosphorus source vapor molecules. Based on the response in the PR intensity to variations and interruptions in the periodic source vapor pulse sequence—utilized for establishing steady-state conditions—we have show that the dealkylation of the gallium source molecules is the rate limiting step in the low temperature heteroepitaxial growth process. Its initial stages involving molecules and radicals that do not interact strongly with the underlying substrate can be modeled in terms of homogeneous reactions in the SRL. There exists at present no adequate description of the response in the PR intensity due to changes in the concentration of molecules and atoms bound directly to the surface atoms of the heteroepitaxial film. An assessment of the heterogeneous kinetics thus requires further research supplementing optical real-time process monitoring by additional methods of surface analysis.

[1]J. T. Kelliher, A. E. Miller, N. Dietz, S. Habermehl, Y. L. Chen, Z. Lu, G. Lucovsky, and K. J. Bachmann, Appl. Surf. Sci. **86**, 453 (1995).
[2]N. Dietz and K. J. Bachmann, MRS Bull. **20**, 49 (1995).
[3]D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[4]U. Rossow, N. Dietz and K. J. Bachmann, J. Vac. Sci. Technol. B **14**, 3040 (1996).
[5]J. Chadi, Phys. Rev. Lett. **59**, 1691 (1995).
[6]S. Miyoshi, H. Yaguchi, K. Onake, R. Ito, and Y. Shiraki, J. Cryst. Growth **145**, 87 (1994).
[7]H. Kroemer, J. Cryst. Growth **81**, 193 (1987).
[8]S-W Choi, K. J. Bachmann, T. Colpitts, J. B. Posthill, and M. L. Timmons, J. Electrochem. Soc. **139**, 312 (1992).
[9]W. A. Harrison, E. A. Kraut, J. R. Waldrop, and R. W. Kraut, Phys. Rev. B **18**, 4402 (1978).
[10]Y. Wang, M. J. Bonikowski, and R. J. Hamers, J. Phys. Chem. **98**, 5966 (1994).
[11]Y. Wang, X. Chen, and R. J. Hamers, Phys. Rev. B **50**, 4534 (1994).
[12]R. D. Bringans, M. A. Olmstead, R. I. G. Uhrberg, and R. Z. Bachrach, Phys. Rev. B **36**, 9569 (1987).
[13]G. S. Higashi, R. S. Becker, Y. J. Chabal, and A. J. Becker, Appl. Phys. Lett. **58**, 165 (1991).
[14]H. M'saad, J. Michel, A. Reddy, and L. C. Kimmerling, J. Electrochem. Soc. **142**, 2833 (1995).
[15]H. Angermann, K. Kliefoth, W. Füssel, and H. Flietner, Microelectron. Eng. **28**, 51 (1995).
[16]C. H. Bjorkman, J. L. Alay, H. Nishimura, M. Fukuda, T. Yamazaki, and M. Hirose, Appl. Phys. Lett. (to be published).
[17]Landoldt-Börnstein, *Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung and H. Weiss (Springer, Berlin, 1984), Vol. 17, p. 17.
[18]F. Ernst and P. Piroux, Appl. Phys. Lett. **64**, (1988).
[19]W. Kern, in *Current Topics of Materials Science*, edited by E. Kaldis (North-Holland, Amsterdam, 1979), Vol. 3, p. 131.
[20]C. H. Chen, C. A. Larsen, G. B. Stringfellow, D. W. Brown, and A. J. Robertson, J. Cryst. Growth **77**, 11 (1986).
[21]M. Grundmann, A. Krost, D. Bimberg, O. Ehrmann, and H. Cerva, Appl. Phys. Lett. **60**, 3292 (1992).
[22]M. Grundmann, J. Christen, F. Heinrichsdorf, A. Krost, and D. Bimberg, J. Electrochem. Soc. **23**, 201 (1994).
[23]A. Krost, R. F. Schabel, U. Rossow, D. Bimberg, and H. Cerva, J. Cryst. Growth **145**, 314 (1994).
[24]R. F. Schabel, A. Krost, M. Grundmann, D. Bimberg, and H. Cerva, J. Electron. Mater. **24**, 1625 (1995).
[25]S. Mahajan, K. J. Bachmann, D. Brasen, and E. Buehler, J. Appl. Phys. **49**, 245 (1978).
[26]A. E. Miller, N. Dietz, and K. J. Bachmann, Mater. Res. Soc. Symp. Proc. **355**, 197 (1995).
[27]K. J. Bachmann, N. Dietz, and U. Rossow, Mater. Sci. Eng. B **35**, 472 (1995).
[28]R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North Holland, Amsterdam, 1987), p. 340.
[29]N. Dietz and K. J. Bachmann, Vacuum **47**, 133 (1996).
[30]S. H. Li, C. A. Larsen, N. I. Buchan, and G. B. Stringfellow, J. Appl. Phys. **65**, 5161 (1989).

Characterization of $Si/Si_{1-y}C_y$ superlattices grown by surfactant assisted molecular beam epitaxy*

P. O. Pettersson,[a] C. C. Ahn, and T. C. McGill
Department of Applied Physics, California Institute of Technology, Pasadena, California 91125

E. T. Croke and A. T. Hunter
Hughes Research Laboratories, Malibu, California 90265

(Received 22 January 1996; accepted 26 April 1996)

$Si/Si_{0.97}C_{0.03}$ superlattices grown on Si(001) substrates by Sb surfactant assisted molecular beam epitaxy are characterized by *in situ* reflection high energy electron diffraction (RHEED), atomic force microscopy, transmission electron microscopy (TEM), and high resolution x-ray diffraction. The RHEED shows that, in the absence of Sb, the growth front roughens during $Si_{0.97}C_{0.03}$ growth and smooths during subsequent Si growth. In contrast, when Sb is present, the growth front remains smooth throughout the growth. This observation is confirmed by cross-sectional TEM, which reveals that for samples grown without the use of Sb, the $Si/Si_{0.97}C_{0.03}$ interfaces ($Si_{0.97}C_{0.03}$ on Si) are much more abrupt than the $Si_{0.97}C_{0.03}/Si$ interfaces. In the case of Sb assisted growth, there is no observable difference in abruptness between the two types of interfaces. Atomic force microscopy micrographs of the $Si_{0.97}C_{0.03}$ surface reveal features that could be the source of the roughness observed by RHEED and TEM. © *1996 American Vacuum Society.*

I. INTRODUCTION

By introducing carbon into the Si/Ge material system, one could increase flexibility in the design of electronic structure in a Si-based electronics technology. The reason for this is that the SiGeC alloys may provide useful conduction band offsets for layers grown coherently strained to Si. This offset might allow fabrication of novel n-type devices such as n-type resonant tunneling devices (RTDs) or high-electron-mobility transistors (n-HEMTs) that are compatible with VLSI processing lines. In addition, due to the smaller lattice constant of diamond (a_0=0.357 nm vs a_0=0.543 nm for Si and a_0=0.566 nm for Ge), carbon could be used for strain compensation of SiGe structures, resulting in alloy layers that are lattice-matched to Si and therefore in devices that are stable during high temperature processing steps.

When $Si_{1-x}Ge_x$ is grown coherently strained to Si(001), the compressive strain causes the fourfold-degenerate Δ_4 band to become the lowest conduction band state. The combination of strain splitting and compositional shift cancel for this band, resulting in almost no conduction band offset for $Si_{1-x}Ge_x$ alloys grown strained to Si.[1]

While conduction band offsets can be achieved in the $Si/Si_{1-x}Ge_x$ system by growing tensile-strained layers on relaxed $Si_{1-x}Ge_x$ buffer layers, the incorporation of carbon may provide this desirable feature without the substantial defect densities and complicated processing inherent for growth on relaxed buffer layers. Since $Si_{1-x-y}Ge_xC_y$ alloys can be grown lattice matched to Si, the strain splitting of the conduction band can be eliminated and a useful conduction band offset might be provided by a compositional shift in band gap. According to estimates,[1] the band gap for $Si_{1-y}C_y$ alloys strained to Si decreases only slightly with y. Therefore, it is essential to develop growth techniques that allow the introduction of substantially more than 1% carbon to achieve band gap differences and band offsets that are larger than k_BT at room temperature.

Our experiments show that for carbon concentrations in excess of 2%, the normal, two-dimensional layer-by-layer growth of Si(100)[2] is disrupted, resulting in a rough surface. This roughness manifests itself in reflection high energy electron diffraction (RHEED) as "spottiness" in the pattern, rather than the normally streaked, (2×1)+(1×2) pattern associated with growth on atomically smooth, two-domain-reconstructed Si(001) surfaces.

In this article, we report a study of surfactant-mediated growth of $Si/Si_{0.97}C_{0.03}$ superlattices through the use of RHEED, transmission electron microscopy (TEM), atomic force microscopy (AFM), and high resolution x-ray diffraction (HRXRD). The RHEED and TEM results were reported in Ref. 3; here we explain the RHEED analysis in detail and correlate the previous results with AFM measurements. Actual composition and layer thicknesses were confirmed using HRXRD. We demonstrate that through the use of a surfactant, the tendency for the $Si_{1-y}C_y$ surface to roughen during growth can be reduced or eliminated, allowing layers with even higher carbon concentrations to be grown. This technique has been used previously with encouraging results to suppress Stranski–Krastanov islanding during the growth of Si/Ge superlattices[4] and $Ge_{1-y}C_y$ alloy layers[5] on Si substrates.

The article is organized as follows. First, the sample growth is described. Details of the superlattice structure and postgrowth compositional analysis by HRXRD is presented. We then present an analysis of digitized RHEED data taken during growth of the samples and discuss the findings in

*Published without author corrections.
[a]Electronic mail: peo@ssdp.caltech.edu

 0734-211X/96/14(4)/3030/5/$10.00 ©1996 American Vacuum Society

comparison with cross-sectional TEM micrographs and AFM images of the sample surfaces. Finally, we put forward an explanation as to how the surfactant might suppress the islanding.

II. EXPERIMENT

The samples used in this experiment consisted of two 15-period, 26.1 nm Si/4.4 nm $Si_{0.97}C_{0.03}$ superlattices grown at 525 °C on 100-nm-thick Si buffer layers. Prior to growth 2000 Ω cm Si(001) substrates were degreased in trichloroethane and acetone for 2 min each, followed by rinses in methanol for 10 min and de-ionized H_2O. A 15 s dip in 5% HF just prior to loading into the MBE system (Perkin–Elmer Model 430S) was used to hydrogen passivate the wafer surfaces. *In situ*, the wafer surfaces were cleaned at 875 °C under the influence of a slight Si flux (about 0.1 Å/s) until the RHEED pattern consisted of the usual (2×1)+(1×2) streaked pattern indicative of a clean, reconstructed Si(001) surface.

After deposition of a 100 nm undoped Si buffer layer, approximately one monolayer (6.8×10^{14} atoms/cm^3) of Sb was deposited on the surface of sample SL-Sb. Sample SL received no such Sb predeposition. The superlattices were then grown on each sample using growth rates of approximately 1.8 Å/s and 0.015 Å/s for Si and C, respectively, evaporated from electron beam sources. The $Si_{0.97}C_{0.03}$ layer was grown with both shutters opened, so the growth rate for this layer was 1.815 Å/s. Closed-loop control of the flux was accomplished for Si through the use of a Sentinel III deposition controller and for C by monitoring the amplitude of amu 36 (C_3) with a residual gas analyzer (RGA) and adjusting the power to the electron gun to maintain a predetermined signal. Growth rates, layer thicknesses, and compositions were determined later using HRXRD. From the superlattice peak positions, the average carbon concentration and superlattice period were measured and, together with a knowledge of the shutter opening times, used to calculate individual layer thicknesses and the carbon content of the $Si_{1-y}C_y$ layer.

During superlattice growth, images of the RHEED patterns were digitized and captured for later analysis. The setup consisted of a standard RHEED system (Perkin–Elmer), a black and white CCD camera (SONY CCD-IRIS), and an S-VHS VCR (Panasonic AG 7355). Data were recorded on S-VHS videotape and digitized at a resolution of 512×512 pixels with a 256 level grayscale using a computer (Sun Sparc 2) and a video capture board (Data Cell S2200). In order to filter out some of the background due to stray light from the e-gun sources, we inserted a green filter between the camera and the RHEED screen.

The AFM measurements were performed in laboratory ambient without any surface preparation using a Nanoscope III from Digital Instruments. The TEM cross sections were prepared by mechanical polishing and ion milling and the micrographs were acquired at an acceleration voltage of 300 kV on a Philips EM430 electron microscope.

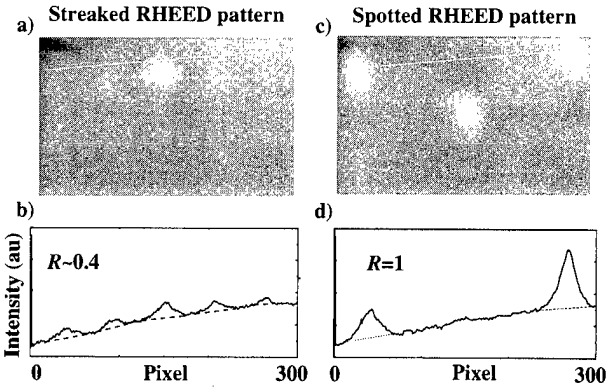


FIG. 1. Analysis of the RHEED pattern. The amplitudes of the intensity along the line [marked in white in (a) and (c)] are shown in (b) and (d). The areas enclosed by the solid curves and the dotted lines in (b) and (d) represent the total intensity associated with that streak or spot. By calculating the ratio of the spot intensity to the total intensity of the pattern, we obtain a qualitative measure of the roughness of the surface R [see Eq. (1)]. This ratio can be used as a qualitative measure of surface roughness since, for a spotty pattern (c), the areas associated with the spots will dominate, making $R\approx1$. On the other hand, when the RHEED pattern is streaked (a), all the areas are comparable, giving an R of about $\frac{2}{5}=0.4$.

III. RESULTS

The RHEED pattern from both samples prior to growth of the Si buffer layer exhibited the usual (2×1)+(1×2) streaked pattern typical of a clean Si(001) surface. During growth of sample SL, immediately upon opening the carbon shutter, the pattern became spotty in appearance, indicative of a rough surface. Each subsequent Si layer caused the pattern to revert back to the (2×1)-reconstructed pattern, suggesting that the Si deposition caused the surface to become smooth again. This alternating behavior of roughening followed by smoothing persisted throughout the growth of this sample. For sample SL-Sb, the half-order streaks originally visible in the pattern diminished in intensity after Sb deposition due to a realignment of the surface reconstruction.[6] The observed (1×1) pattern exhibited no spottiness during the subsequent growth of the superlattice and remained streaked (smooth).

In order to study the differences between the spotted and streaked patterns observed in the growth of sample SL, we employed the following analysis of the digitized RHEED data. First, the intensity along a line [marked in white in Figs. 1(a) and 1(c)] perpendicular to the streak direction and intersecting the $(\bar{1}0)$ and (10) spots was digitized. The amplitude of the intensity along the line is shown in Figs. 1(b) and 1(d). Then, the intensity associated with a certain spot or streak ($n0$) was integrated along the line, to take into account the intensity from the full width of the streak or spot, to give the quantity I_{n0}. The background intensity due to light from the e-gun sources was subtracted, as indicated by the dotted lines in Figs. 1(b) and 1(d). Finally, we calculated the ratio, R, of the spot intensity to the intensity of the whole pattern [see Eq. (1)]. This ratio can then be used as a qualitative measure of surface roughness since, for a spotty pattern, $I_{\bar{1}0}$ and I_{10} will dominate, making $R\approx1$. On the other hand, when the RHEED pattern is streaked [see Figs. 1(a)

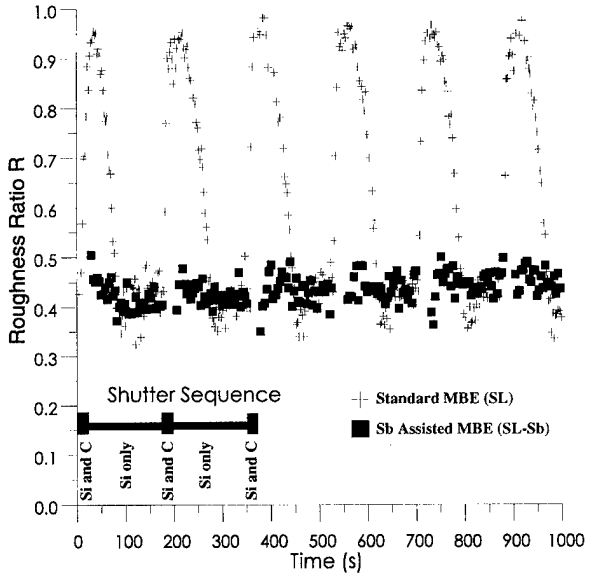


FIG. 2. Roughness ratio R vs growth time. See Fig. 1 for the derivation of R. An R of 1 indicates a spotted RHEED pattern and an R of 0.4 indicates a streaked pattern. For sample SL (no Sb predeposition), immediately upon opening the C shutter the ratio is shown to increase rapidly, saturating at a value near 1. Half ($\approx$13 nm) of the Si layer thickness was required to completely recover the original pattern. During growth of sample SL-Sb, the RHEED pattern remained unchanged from the (1×1), streaked pattern observed immediately following Sb predeposition.

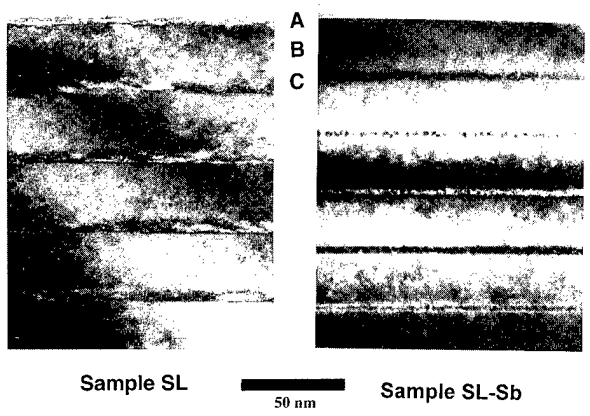


FIG. 3. Cross-sectional TEM images of the superlattice samples. (a) is an image taken from sample SL, showing alternating layers of thick and thin bands corresponding to the Si (B) and $Si_{0.97}C_{0.03}$ (C) layers, respectively. The surface of the superlattice is marked A near the top. Sample SL (a) was terminated with a $Si_{0.97}C_{0.03}$ layer resulting in rough surface morphology (see Fig. 5). Sample SL-Sb (b) was terminated with a Si layer. Clearly, the interfaces that are formed when $Si_{0.97}C_{0.03}$ is grown on Si are much more abrupt than the interfaces formed when Si is grown on $Si_{0.97}C_{0.03}$. In the case of sample SL-Sb (b), both interfaces appear equally abrupt. In comparison with sample SL, they appear more abrupt than the case for which Si is grown on $Si_{0.97}C_{0.03}$ and less abrupt than the case for which $Si_{0.97}C_{0.03}$ is grown on Si.

and 1(b)], all the terms are of comparable magnitude, giving an R of about $\frac{2}{5}=0.4$.

$$R=\frac{I_{\bar{1}0}+I_{10}}{I_{\bar{1}0}+I_{\bar{\frac{1}{2}}0}+I_{00}+I_{\frac{1}{2}0}+I_{10}}. \quad (1)$$

In Fig. 2(d), the ratio, R, is plotted as a function of time during growth of the two samples (C shutter opens at $t=0$ s). For sample SL (no Sb predeposition), immediately upon opening the C shutter the ratio is shown to increase rapidly, saturating at a value near 1. During subsequent growth of the Si layer, the RHEED pattern slowly recovered its original (2×1)+(1×2), streaked pattern and the ratio returned to approximately 0.4 (smooth). From Fig. 2, we note that fully half ($\approx$13 nm) of the Si layer thickness was required to completely recover the original pattern. During growth of sample SL-Sb, the RHEED pattern remained unchanged from the (1×1), streaked pattern observed immediately following Sb predeposition. Data were not available for SL-Sb during $Si_{0.97}C_{0.03}$ deposition because stray light from the e-guns washed out the pattern. Nevertheless, it was possible to view the pattern visually during these periods and no spottiness was observed.

In Fig. 3, we present cross-sectional TEM images of the samples that show features consistent with the RHEED observations. Figure 3(a) is an image taken from sample SL, showing alternating thick and thin layers corresponding to the Si and $Si_{0.97}C_{0.03}$ layers, respectively. The surface of the superlattice is marked A near the top of the figure. Clearly, the interfaces that are formed when $Si_{0.97}C_{0.03}$ is grown on Si are much more abrupt than the interfaces formed when Si is grown on $Si_{0.97}C_{0.03}$. In the case of sample SL-Sb [Fig. 3(b)], both interfaces appear equally abrupt. In comparison with sample SL, they appear more abrupt than the case for which Si is grown on $Si_{0.97}C_{0.03}$ and less abrupt than the case for which $Si_{0.97}C_{0.03}$ is grown on Si. Sample SL was terminated with a $Si_{0.97}C_{0.03}$ layer, resulting in a rough surface morphology evident in the TEM and AFM micrographs (see Fig. 5). Sample SL-Sb was terminated with a Si layer, which gives the superlattice a smooth surface morphology.

In Fig. 4, we display HRXRD of the two samples. The peaks associated with sample SL are larger in magnitude and have narrower widths than those of sample SL-Sb. One reason for this could be that the $Si_{0.97}C_{0.03}$/Si interfaces (when $Si_{0.97}C_{0.03}$ is grown on Si) in sample SL are more abrupt than the $Si/Si_{0.97}C_{0.03}$ interfaces in sample SL-Sb, as seen in the TEM image (see Fig. 3). Another reason might be that there is a higher density of defects that look like stacking faults in sample SL-Sb as compared to sample SL. These defects could reduce the lateral coherence of the superlattices and thus broaden the HRXRD peaks. The defects are discussed in more detail below. As mentioned above, the superlattice peak positions were measured to calculate the layer thicknesses and carbon content of the samples.

In Fig. 5, we show an AFM micrograph taken from the surface of sample SL, terminated with $Si_{0.97}C_{0.03}$. The figure shows features on the order of 1.5 nm peak-to-valley perpendicular to the surface and 40.0 nm laterally. There does not appear to be any directional dependence to the features. The RMS roughness of this surface was measured to be approximately 0.36 nm. These features are apparently responsible for the spotted pattern observed in the RHEED image of Fig. 1(c), since additional samples we have studied, for which Si

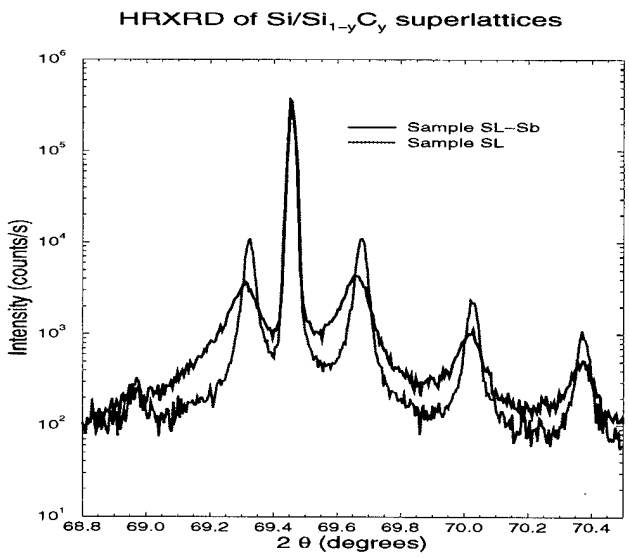


FIG. 4. HRXRD of a 15 period $Si/Si_{0.97}C_{0.03}$ superlattice grown by standard MBE (sample SL) and by Sb assisted MBE (sample SL-Sb). The sample SL superlattice peaks are sharper than those of sample SL-Sb. This is probably due to the $Si/Si_{0.97}C_{0.03}$ interfaces being less sharp in the SL-Sb sample (as observed in the cross-sectional TEM micrographs, Fig. 3) or a higher density of stacking faults in sample SL-Sb as compared to sample SL.

in one sample and Sb-terminated $Si_{0.985}C_{0.015}$ in another were grown on Si(001), lack these features and have streaked RHEED patterns. The feature height observed in the AFM micrograph agrees with the thickness variation seen in the TEM image [see Fig. 3(a)].

In addition to the interface structure, the TEM images also reveal that here are defects that look like stacking faults, possibly originating from point defects such as C dimers or trimers in the SL-Sb sample. These defects are also present in the SL sample, albeit at a lower density. Most of the defects in sample SL-Sb originate in the region between superlattice layers 7 and 10 (counting up from the bufferlayer). Only a small fraction (0.4%) of the volume of the superlattice is enclosed by these defects.

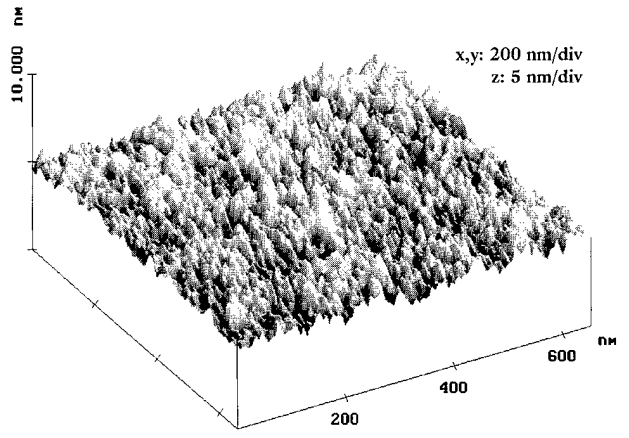


FIG. 5. AFM micrograph of the top layer ($Si_{0.97}C_{0.03}$) of a 15 period $Si/Si_{0.97}C_{0.03}$ superlattice grown by standard MBE (sample SL). The features are on the order of 1.4 nm tall and 40 nm on the side. These surface features apparantly give rise to the spotty RHEED pattern observed [see Fig. 1(c)] and the variation in thickness seen in the TEM (see Fig. 3).

IV. DISCUSSION

The data presented in the above section suggest that one or more of the species in the carbon flux disrupts the epitaxial growth of $Si_{0.97}C_{0.03}$ on Si (sample SL). These species form surface nucleation centers where diffusing adatoms can incorporate in competition with surface steps, resulting in 3D growth. Residual gas analysis of the growth flux shows that it primarily consists of monomers (C), dimers (C_2), and trimers (C_3). In principle, any of the carbon species could act as nucleation centers. We expect the monomers to be the most mobile on the surface, and hence, the most likely to simply be incorporated in step flow growth. Of multiple carbon species the dominant is the dimers.

All of the multiple carbon species, however, are potential candidates for nucleating the rough growth. First, the C–C bond is 1.8 times stronger than the Si–Si bond and based on the observation that the Si dimer is stable up to 600 K,[8] one expects the C dimer and trimers to be stable up to 1100 K or about 800 °C. Since the growth temperature used in this experiment was 525 °C, the carbon dimers and trimers impinging on the surface remain undissociated. Another argument for the stability of the carbon dimers is that the equilibrium constant for sublimation of solid carbon (to monoatomic carbon gas), K_p, is 3×10^{-17} at the growth temperature mentioned above.

Second, the diffusion length of these dimers and trimers is expected to be negligible compared to that of the monomers because of the large activation energy due to the bond bending and stretching required for a dimer or trimer to move on the surface. Thus, while the monomers might diffuse to a step and thus contribute to step flow growth, the dimers or trimers will incorporate at the site of impingement and form nucleation centers.

Third, the nucleation center density generated by the dimers and trimers is large enough for 3D growth to dominate over 2D step flow growth. To show one possibility for how this could occur, we need to consider this probability of an adatom attaching to a dimer or trimer rather than to a step. To assess the relative probability, we follow Mo *et al.*'s argument.[8] Adapting his argument, we consider a square with the side aligned with a step on a slightly miscut substrate. We set the length of the side equal to the average terrace width W of the steps which is given by the degree of miscut of the substrate. Under the conditions of our experiment, this square is the area from which this portion of the step accumulates adatoms. Let us say that the square has a dimer at the center and assume that both the step and the dimer are perfect sinks for adatoms. According to the 2D random walk theory, the number of hops required for an adatom impinging at a random site in the square to reach the dimer is on average $\sim(W/a)^2$, where a is the length of a hop. The number of hops required for the adatom find the step is also on average $\sim(W/a)^2$. In this case then, both island growth and step flow

growth will take place. For both sample SL and SL-Sb, the fluxes of the dimers and trimers were about 10% and 5% of the monomer flux, respectively. These fluxes yield a dimer and trimer density of about 10^{12} molecules/cm^2 in the time required to complete a monolayer of $Si_{0.97}C_{0.03}$ growth. Given a limiting case of a wafer miscut of about 0.5°, the terrace width is $W \approx 2\times10^{-6}$ cm. At the density calculated above, we get four dimers per square, so 3D nucleation should compete effectively with 2D step-flow growth.

We speculate that in the case of sample SL-Sb, where the surface stays smooth throughout the growth, the surfactant assisted growth mechanism prevents the dimers and trimers from forming effective nucleation centers. The primary role of the Sb is to ride as a surface layer, burying the carbon dimer and trimers. An impinging silicon adatom diffuses on the Sb layer until it reaches a proper site. At that site, the adatom undergoes an exchange with a surfactant atom and incorporates. Again, the carbon clusters do not diffuse on the surface; rather, they are incorporated immediately by some exchange mechanism. Since direct contact between Si adatoms and the carbon clusters is reduced, the clusters no longer serve as effective nucleation centers and the growth stays relatively smooth as shown by AFM and the streaked RHEED pattern.

The TEM picture in Fig. 3 shows that the interface between the Si and $Si_{0.97}C_{0.03}$ on the substrate side is slightly rougher on a sample superlattice grown with Sb than it is on a sample grown without Sb. One possible reason for this is that, in our case, a perfectly ordered Sb terminated Si(001) surface was not achieved [RHEED pattern was (1×1) as opposed to (2×1)]. The unordered surface could prompt the exchange to occur at sites other than steps, thus creating a less abrupt growth front. A highly ordered Sb-terminated (2×1)-reconstructed surface could be expected to aid in the formation of perfectly flat interfaces.

In conclusion, the Sb prevents the carbon dimers and trimers from serving as effective nucleation centers for 3D growth and as a consequence the surface stays smooth during growth of $Si_{0.97}C_{0.03}$ on a sample grown with Sb (sample SL-Sb), whereas the surface roughens during $Si_{0.97}C_{0.03}$ growth on the sample grown without Sb (sample SL).

V. SUMMARY

We studied the effect of adding Sb as a surfactant in the MBE growth of $Si/Si_{0.97}C_{0.03}$ superlattices. Our analysis of reflection high energy electron diffraction, transmission electron microscopy, x-ray diffraction, and atomic force microscopy data shows that Sb induces 2D growth of $Si_{0.97}C_{0.03}$ under conditions in which standard MBE yields 3D growth. Epitaxial growth on the 2×1 Si(001) surface could be easily disrupted by carbon dimers and trimers which introduce additional sites for incorporation of Si adatoms. The Sb-terminated Si surface could prevent direct contact between the carbon dimers and the silicon, and hence suppresses the tendency of the surface to roughen. In conclusion, the use of Sb as a surfactant during growth of high-carbon-content $Si_{0.97}C_{0.03}$ alloys was shown to result in sharper film interfaces and appears useful for achieving carbon contents in excess of what would normally be possible for growth on bare Si(001).

ACKNOWLEDGMENTS

This study was supported in part by the Office of Naval Research under Grant No. N00014-93-1-0710. In addition, the authors would like to thank Carol Garland for her assistance in obtaining the TEM images presented in this work.

[1]A. R. Powell, K. Eberl, F. K. LeGoues, B. A. Ek, and S. S. Iyer, J. Vac. Sci. Technol. B **11**, 1064 (1993).
[2]Y.-W. Mo, R. Kariotis, D. E. Savage, and M. G. Lagally, Surf. Sci. **219**, L551 (1989).
[3]P. O. Pettersson, C. C. Ahn, T. C. McGill, E. T. Croke, and A. T. Hunter, Appl. Phys. Lett. **67**, 2530 (1995).
[4]K. Fujita, S. Fukatsu, H. Taguchi, T. Igarashi, Y. Shiraki, and R. Ito, Mater. Res. Soc. Proc. **220**, 193 (1991).
[5]H. J. Osten, E. Bugiel, and P. Zaumsel, J. Cryst. Growth **142**, 322 (1994).
[6]J. Nogami, A. A. Baski, and C. F. Quate, Appl. Phys. Lett. **58**, 475 (1991).
[7]Y. W. Mo and M. G. Lagally, Surf. Sci. **248**, 313 (1991).
[8]Y. W. Mo, J. Kleiner, M. B. Webb, and M. G. Lagally, Surf. Sci. **268**, 275 (1992).

Modulation-doped $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ heterostructures grown on GaAs substrates using step-graded $In_xGa_{1-x}As$ buffers

R. S. Goldman,[a] K. L. Kavanagh, and H. H. Wieder
Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, California 92093-0407

S. N. Ehrlich
School of Materials Engineering, Purdue University, West Lafayette, Indiana 47907

(Received 25 January 1996; accepted 30 April 1996)

We have grown modulation-doped $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ heterostructures on GaAs substrates using compositionally step-graded $In_xGa_{1-x}As$ buffers. Triple-axis x-ray diffraction measurements indicate nearly complete and isotropic strain relaxation in the buffer, lattice matching of the active layers with the top of the buffer, and no significant epilayer tilt. The temperature dependence and the photoresponse of the electron mobility suggest that transport in the heterostructures is limited principally by remote ionized-impurity scattering, with mobility values comparable to those of heterostructures grown lattice-matched to InP. © *1996 American Vacuum Society.*

I. INTRODUCTION

$In_xGa_{1-x}As$ alloys may be employed as quantum wells for the confinement of a two-dimensional electron gas (2DEG).[1] Modulation-doped InGaAs/InAlAs heterostructures are advantageous for field-effect transistor applications due to their high room temperature electron mobilities, high conduction band offsets, and minimal deep-level trapping in InAlAs barrier layers.[2] However, only one heterostructure, $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$, can be grown lattice-matched to a binary substrate (InP). In order to maximize the flexibility in choice of alloy composition without the constraint of a lattice-matched substrate, it is necessary to interpose a strain relaxing buffer between the substrate and heterostructure. The heterostructure which is lattice-matched to InP provides a useful test structure for monitoring the effectiveness of strain relaxing buffers in providing a suitable substrate for optimum electronic properties. A number of buffer layers have been considered for the lattice-mismatched growth of $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ on GaAs ($\Delta a/a \sim 3.8\%$), including compositionally step-graded buffers,[3] linearly graded buffers,[4–7] and a combination of step-grading and strained-layer superlattices.[8] In most of these studies, a variety of complicated layer designs and growth conditions have been implemented. Furthermore, the effectiveness of the buffer in relaxing strain was not correlated with the electronic properties of the active layers. Here, we report the growth of modulation-doped $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ heterostructures on GaAs substrates using simple step-graded $In_xGa_{1-x}As$ buffers with a total thickness of 1 μm. In addition, we present a correlation between nearly complete buffer strain relaxation, lattice matching of the active layers with the top of the buffer, and high electron mobility of the heterostructures.

[a]Present address: Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213; Electronic mail: rgoldman+@andrew.cmu.edu

II. EXPERIMENTAL PROCEDURES

The samples were grown by solid-source molecular beam epitaxy (MBE), with As_4/group III beam equivalent pressure ratios ~35 and growth rates ~0.9 μm/h. The buffers were graded in 200 nm thick steps, from indium mole fraction, $x=0$ to 0.5, in increments of $\Delta x=0.1$, resulting in a composition gradient, 50 at. % In/μm. A cross section of the targeted layer structure is presented in Fig. 1. The GaAs, step-graded $In_xGa_{1-x}As$ buffers, and active layers were grown at 580, 350, and 450 °C, respectively.

Triple-axis x-ray diffraction measurements were performed at Beamline X18A at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory. The wavelength of incident radiation was selected to be 1.653 Å using a double-crystal Si(111) monochromator. In addition, the source point of the synchrotron was focused on the sample by means of a Pt-coated aluminum mirror placed in the beam path, following the monochromator. The samples were mounted onto a six-circle Huber diffractometer and the signal was collected by a Ge(111) analyzer crystal and scintillation counter. With this configuration, the instrumental resolution was 0.012°. Contour maps were performed near the (004) and (224) Bragg peaks. These contour maps consisted of a series of ω–2θ scans, with the same initial value of 2θ (angle between incident x-ray beam and detector) but a slightly different value of ω (angle between the incident beam and sample). The (004) and (224) d spacings were determined from this data.

Electron transport measurements were implemented with six-arm Hall bars (200×800 μm) aligned along both ⟨110⟩ in-plane directions, fabricated by standard contact photolithographic processes, with Ni/AuGe/Ni ohmic contacts deposited in vacuum and annealed at 420 °C. Low magnetic field Hall and resistivity measurements as a function of temperature, from 300 to 1.6 K, and resistivity measurements as a function of magnetic field, from 0 to 7.5 T, were performed in a superconducting NbTi magnet. In addition, the photoionization of electrons trapped in deep levels was monitored

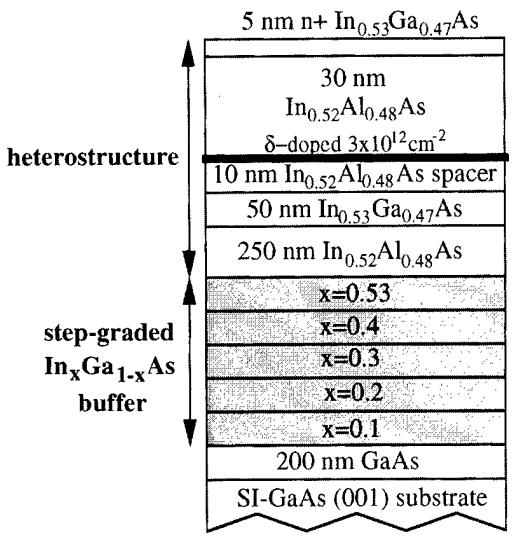


FIG. 1. Cross section of targeted layer structures. The step-graded buffers were graded in 200 nm thick steps. The layers are undoped, except for the δ-doped region.

by illuminating the sample with a GaAs light-emitting diode (LED).

III. RESULTS AND DISCUSSION

Figures 2(a) and 2(b) show symmetric (004) and (224) ω–2θ scans for the sample. In the plots, the epilayer peak positions are referenced with respect to the GaAs substrate peak, arbitrarily set to $\Delta\theta$=0 arcsec, for visual comparison of

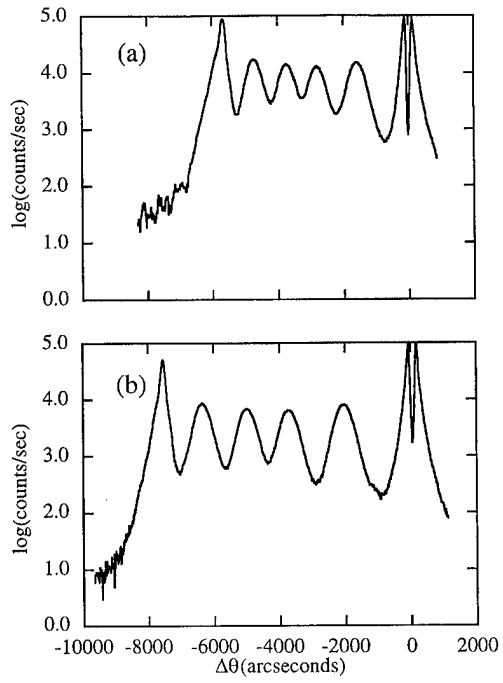


FIG. 2. Symmetric (a) (004) and (b) (224) ω–2θ scans of the five-step compositionally graded $In_xGa_{1-x}As$ buffer layer grown on (001) GaAs.

Bragg angle differences between the substrate and epilayer. The dip in the substrate peak is due to saturation of the detector. It is clear from this data that the top epilayers are lattice-matched to the buffer. Four sets of (004) and (224) reflections were measured and indicated a tetragonal lattice with symmetric strain relaxation in in-plane ⟨110⟩ directions. From the (004) and (224) d spacings, the in-plane and out-of-plane lattice constants, a and c, were obtained.[9] The In mole fraction in each layer was determined, assuming a linear interpolation of elastic and lattice constants between bulk GaAs and InAs.[10]

Analysis of the x-ray data yields layer In mole fractions 0.15, 0.26, 0.35, 0.45, and 0.53, respectively. The strain relaxation in each layer with respect to the substrate ranges from 92 to 95±5%, essentially identical to within experimental error. A more sensitive indication of the effectiveness of strain relaxation is the strain relaxation in each layer with respect to the in-plane lattice constant of the previous layer. This ranges from 94 to 62%, with the expected decreasing trend towards the surface layers, similar to observations in SiGe/Si heterostructures.[11] The overall relaxation of the top epilayer is 92%, with in-plane and out-of-plane lattice constants 5.851 and 5.883 Å, respectively. Assuming that the strain is relaxed by the edge component of 60° misfit dislocations, the density of misfit dislocations at the interface between the fourth and fifth buffer step is $2.6\times10^5/cm^2$.

Figures 3(a) and 3(b) present (004) and (224) contour maps taken with the x-ray beam incident in the [$\bar{1}$10] direction. In the plots, the y axis corresponds to θ and the x axis is ω_{mid}, the center value of ω in each ω–2θ scan. Plotted in this manner, the y axis is a measure of strain, and the x axis is a measure of lattice plane tilt. To compare peak broadening stemming from d spacing variations and microtilting of lattice planes (mosaic spread), the data has been normalized to the top epilayer peak intensity and plotted as contours with lines of equal intensity ranging from $(\frac{1}{8})I_{max}$ to I_{max}, in increments of $(\frac{1}{8})I_{max}$.

In the (004) map, the substrate peak lies at θ=35.76°, and the five steps in the buffer lie at successively smaller θ. The ω_{mid} values of the epilayer peaks are approximately the same as that of the substrate, indicating that there is no measurable macroscopic tilt of the epilayer about the [110] axis. Similar data taken with the sample rotated 90° azimuthally resulted in nominally identical ω_{mid} values of the epilayer and substrate peaks, indicating no measurable macroscopic tilt about the [$\bar{1}$10] axis. In the (224) map, the substrate peak lies at 45.71°, and the five steps in the buffer also lie at successively smaller θ. The ω_{mid} values of the epilayer peaks are the same as that of the substrate, confirming the lack of tilt about the [110] axis, and the nearly complete relaxation of the layers.

For both (004) and (224) reflections, the full width at half-maximum (FWHM) of the ω–2θ scan for the top epilayers was 0.07°, indicating a uniform distribution of (004) and (224) d spacings throughout the thickness. Furthermore, the FWHM of the ω scans of the top epilayers was 0.5° and 0.3° for the (004) and (224) reflections, respectively. We have observed similar amounts of broadening in ω–2θ and ω scans of nominally identical epilayers grown on linearly

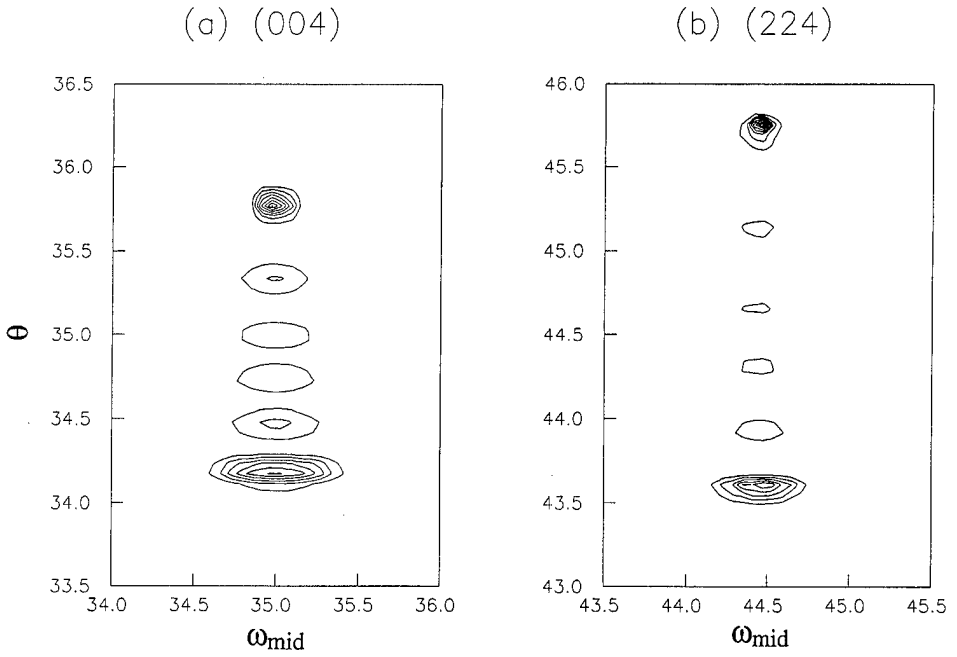


FIG. 3. Triple-axis contour maps (θ_B vs ω_{mid}), performed near the (a) (004) and (b) (224) reflections. The maps consist of a series of ω–2θ scans, each with the same initial value of 2θ but a slightly different value of ω. To compare peak broadening setemming from d spacing variations and microtilting of lattice planes (mosaic spread), the data has been normalized to the top epilayer peak intensity, and plotted as contours with lines of equal intensity ranging from $(\frac{1}{8})I_{max}$ to I_{max}, in increments of $(\frac{1}{8})I_{max}$.

graded $Al_{0.48}(In_zGa_{1-z})_{0.52}As$ buffers (with nominally identical composition gradient) on GaAs.[12] In addition, the smaller ω-broadening for (224) in comparison with (004) is similar to results from studies of step-graded SiGe/Si heterostructures.[13] Assuming that the mosaic spread originates from a Gaussian distribution of misoriented blocks, an estimate of the dislocation density is given by $D \sim (\Delta\omega)^2/9b^2$, where $\Delta\omega$ is the rocking curve width (in radians), and b is the Burger's vector.[14] The resulting estimated dislocation density is $\sim 10^9/cm^2$. Since this density of dislocations cannot be accounted for by interfacial misfit dislocations ($\sim 10^5/cm^2$), a significant density of threading dislocations may exist in the top epilayers. However, plan-view transmission electron microscopy (TEM) studies of the epilayers grown on linearly graded $Al_{0.48}(In_zGa_{1-z})_{0.52}As$ buffers mentioned before indicated a threading dislocation density $\sim 10^7/cm^2$.[6] Therefore, the accuracy of the dislocation density estimates from the misoriented block model needs further investigation.

The 0.5° mosaic spread measured in the top epilayers has not adversely affected the electronic properties of the 2DEG. Figures 4(a) and 4(b) present the temperature dependence of the apparent electron mobility μ_H and the apparent electron concentration $[n]_H$ determined from low-field Hall and resistivity measurements. The solid and open symbols correspond to data taken in the [110] and [$\bar{1}$10] directions, respectively. The peak electron mobility, in the [$\bar{1}$10] direction, is comparable to that of similar heterostructures grown lattice-matched to InP which have room temperature mobilities ranging from 9600 to 13 600 $cm^2/V\,s$.[12,15–23] Similar tem-

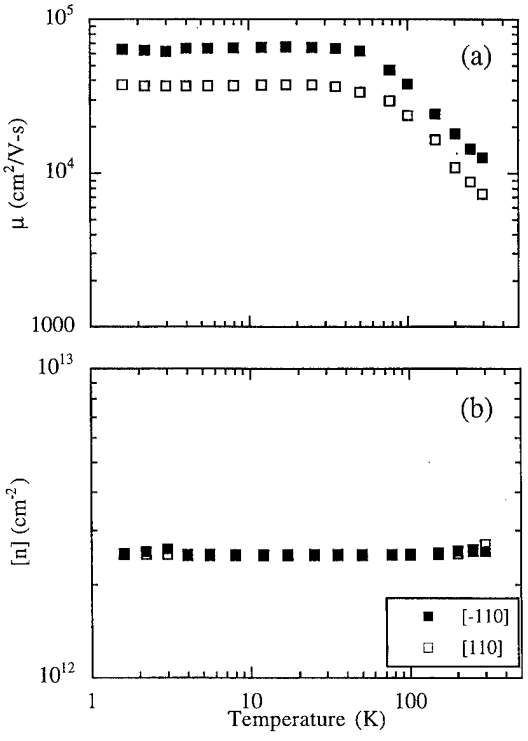


FIG. 4. Temperature dependence of (a) apparent electron mobility μ_H and (b) apparent electron concentration $[n]_H$ determined from low magnetic field Hall and resistivity measurements in ⟨110⟩ in-plane directions.

TABLE I. Summary of 300 and 1.6 K data taken in both ⟨110⟩ in-plane directions, in the dark and under illumination with a GaAs LED. Listed are the apparent electron mobility μ_H and apparent electron concentration $[n]_H$ determined from low-field Hall and resistivity measurements. Also listed are the subbands $[n]_i$ determined from the FFT power spectrum of the ρ_{xx} oscillations in reciprocal magnetic field. The experimental uncertainties are approximately ±5%.

	300 K			1.6 K					
Sample	$\mu[110]$ (cm^2/V s)	$\mu[\bar{1}10]$ (cm^2/V s)	$[n]_H$ (cm^{-2})	$\mu[110]$ (cm^2/V s)	$\mu[\bar{1}10]$ (cm^2/V s)	$[n]_H$ (cm^{-2})	$[n_0$	$[n]_1$ (10^{12} cm^{-2})	$[n]_2$
Dark	7500	12 600	2.7×10^{12}	37 700	64 000	2.5×10^{12}	1.37	0.88	0.40
LED	7500	12 600	2.7×10^{12}	40 500	68 300	3.1×10^{12}	1.66	0.78	0.45

perature dependencies of μ_H and nearly temperature-independent $[n]_H$ are observed in both ⟨110⟩ directions. The temperature independence of the mobility in the range 20–1.6 K indicates that the mobility is limited primarily by remote ionized impurity scattering, as expected for a well-behaved 2DEG. Furthermore, the data indicate anisotropic μ_H at all temperatures. These anisotropies, which are also observed in heterostructures grown lattice-matched to InP, will be discussed elsewhere.[12]

The two-dimensional nature of the electron gas was confirmed from 1.6 K high-field Shubnikov de Haas measurements which indicate oscillations of the resistivity ρ_{xx} and quantum Hall plateaus in the transverse resistivity ρ_{xy}. A slight beating of the ρ_{xx} oscillations is observed, indicating multiple subband occupation of the 2DEG. The electron concentrations $[n]_i$ determined from a fast Fourier transform (FFT) power spectrum of the ρ_{xx} oscillations in reciprocal magnetic field are summarized in Table I. Charge-carrier confinement in the 2DEG is confirmed through a comparison of the total electron concentration $[n]_t=[n]_0+[n]_i$ determined from SdH measurements with $[n]_H$; these agree to within 2%, indicating excellent carrier confinement.

Further confirmation of the high quality of this heterostructure was confirmed by illuminating the sample with a GaAs LED. The photoionization of electrons trapped in deep levels within the fundamental bandgap of InAlAs, was monitored at room temperature and 1.6 K, and the results are listed in Table I. At room temperature, the LED has a negligible effect on μ_H and $[n]_H$. At 1.6 K, the LED has a small effect on the μ_H, $\Delta\mu<5000$ cm^2/V s with a change of $[n]_H$, $\Delta[n]\sim0.5\times10^{12}/cm^2$. The increase in $[n]_H$ leads to an increase in μ_H, suggesting that illumination with the LED has led to the photoionization of electrons trapped in acceptor-like deep levels. The light sensitivity of these heterostructures is comparable to or less than those grown lattice-matched on InP.[12,18] In addition, 1.6 K high-field SdH measurements were performed after illumination for 30 min (the LED had been turned off). The subband electron concentrations $[n]_i$ determined from the FFT power spectrum of the ρ_{xx} oscillations in reciprocal magnetic field are listed in Table I. In this case, the agreement between $[n]_t$ and $[n]_H$ is within 7%.

IV. CONCLUSIONS

In summary, we have demonstrated that a simple, 1 μm thick, compositionally step-graded $In_xGa_{1-x}As$ buffer effectively relaxes the ~3.8% mismatch strain between an $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ heterostructure and a GaAs substrate. Triple axis x-ray diffraction data reveals essentially complete buffer strain relaxation, lattice matching of the active layers with the buffer, and negligible epilayer tilt. High electron mobilities and insignificant light sensitivities of the mobility and carrier concentration of the 2DEG are observed.

ACKNOWLEDGMENTS

The authors thank P.M. Mooney for useful discussions. This work was supported in part by the National Science Foundation (PYI-DMR-9157724) and the Office of Naval Research (ONR) Grant No. N00014-89-J-1147. R.S.G. acknowledges support of the ONR through an AASERT Fellowship. The NSLS is supported by Department of Energy (DOE) Grant No. DE-AC02-76H0016. S.N.E. and Beamline X18A of the NSLS are supported by DOE Grant No. DE-FG02-85ER4518. K.L.K. and R.S.G. also acknowledge support of the NSLS Faculty–Student Research Support Program.

[1] A. Kastalsky, R. Dingle, K. Y. Cheng, and A. Y. Cho, Appl. Phys. Lett. **41**, 274 (1982).
[2] T. Itoh, T. Griem, G. W. Wicks, and L. F. Eastman, Electron. Lett. **21**, 373 (1985).
[3] D. E. Grider, S. E. Swirhun, D. H. Narum, A. I. Akinwande, T. E. Nohava, W. R. Stuart, P. Joslyn, and K. C. Hsieh, J. Vac. Sci. Technol. B **8**, 301 (1990).
[4] C. Harmand, T. Matsuno, and K. Inoue, Jpn. J. Appl. Phys. **28**, L1101 (1989).
[5] K. Hausler, K. Eberl, and W. Sigle, Semicond. Sci. Technol. **10**, 167 (1995).
[6] R. S. Goldman, J. Chen, K. L. Kavanagh, H. H. Wieder, V. M. Robbins, and J. N. Miller, Inst. Phys. Conf. Ser. **141**, 313 (1995).
[7] T. P. Chin and C. W. Tu, Appl. Phys. Lett. **21**, 2708 (1993).
[8] K. Chang, P. Bhattacharya, and R. Lai, J. Appl. Phys. **67**, 3323 (1990).
[9] A. Leiberich and J. Levkoff, J. Vac. Sci. Technol. B **8**, 422 (1990).
[10] S. Adachi, J. Appl. Phys. **53**, 8775 (1982).
[11] P. M. Mooney, J. L. Jordan-Sweet, J. O. Chu, and F. K. LeGoues, Appl. Phys. Lett. **66**, 3642 (1995).
[12] R. S. Goldman, K. L. Kavanagh, H. H. Wieder, S. N. Ehrlich, and V. M. Robbins (unpublished).
[13] E. Koppensteiner, A. Schuh, G. Bauer, V. Holy, G. P. Watson, and E. A. Fitzgerald, J. Phys. D **28**, A114 (1995).
[14] P. B. Hirsch, in *Progress in Metal Physics*, edited by B. Chalmers and R. King (Pergamon, London, 1956), p. 236.
[15] I. Lo, W. C. Mitchel, M. Ahoujja, J. P. Cheng, A. Fathimulla, and H. Mier, Appl. Phys. Lett. **66**, 754 (1995).
[16] W. Kraak, B. Oelze, M. Kunzel, H. G. Bach, J. Bottget, V. Nakov, D. Schulze, and G. Gobsch, Phys. Status Solidi B **183**, 437 (1994).
[17] W. Klein, G. Bohm, M. Sexl, S. Grigull, H. Heiss, G. Trankle, and G. Weimann, J. Vac. Sci. Technol. B **12**, 1306 (1994).

[18]M. A. Tischler and B. D. Parker, Appl. Phys. Lett. **58**, 1614 (1991).
[19]T. Matsuoka, E. Kobayashi, K. Taniguchi, C. Hamaguchi, and S. Sasa, Jpn. J. Appl. Phys. **29**, 2017 (1990).
[20]S. Sasa, Y. Nakata, Y. Sugiyama, T. Fujii, and E. Miyauchi, J. Cryst. Growth **95**, 189 (1989).
[21]W. P. Hong, G. I. Ng, P. K. Bhattacharya, D. Pavlidis, S. Willing, and B. Das, J. Appl. Phys. **64**, 1945 (1988).
[22]H. T. Griem, K. H. Hsieh, I. J. D'Haenens, M. J. Delaney, J. A. Henige, G. W. Wicks, and A. S. Brown, J. Vac. Sci. Technol. B **5**, 785 (1987).
[23]K. H. Hsieh, H. Ohno, G. Wicks, and L. F. Eastman, Electron. Lett. **19**, 160 (1983).

Optical investigations of surface processes in GaP heteroepitaxy on silicon under pulsed chemical beam epitaxy conditions*

U. Rossow,[a] N. Dietz, K. J. Bachmann, and D. E. Aspnes
Physics Department and Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 1 March 1996; accepted 4 May 1996)

Surface processes during the heteroepitaxy of GaP on Si under pulsed chemical beam epitaxy conditions were investigated simultaneously by the optical methods reflectance difference/anisotropy spectroscopy, p-polarized reflectance spectroscopy (PRS), and laser light scattering. Our studies were performed during both cyclic and interrupted growth, where the surface was exposed to individual pulses of the precursors triethylgallium (TEG) and tertiarybutylphosphine (TBP). The data show that the three optical probes provide different perspectives of growth. Several surface processes exhibit time constants of the order of 1 s. One such process is the clustering of Ga atoms, or less likely, of TEG fragments, that occurs with TEG exposure. The optical data also show that TBP dealkylation occurs essentially instantaneously upon arrival at the surface, and that TEG dealkylation is the rate-limiting step. The PRS data exhibit fine structure that shows that heteroepitaxial growth can be described by a four-phase model consisting of the substrate, a GaP layer, a surface reaction layer containing all adsorbed species not yet incorporated in the growing layer, and the ambient. By assuming that this surface layer is very thin we derive approximate equations that allow us to treat the PRS response quantitatively. © *1996 American Vacuum Society.*

I. INTRODUCTION

In this article we discuss real-time monitoring of the pulsed chemical beam heteroepitaxial growth of GaP on Si by three different optical techniques: reflectance difference/anisotropy spectroscopy (RDS/RAS),[1–3] p-polarized reflectance spectroscopy (PRS),[4,5] and laser light scattering (LLS).[6] The objective is a better understanding of the growth process, specifically surface chemistry and reaction kinetics. A discussion of the complicated chemistry of this growth system is given in Ref. 7.

The results are expected to be useful in the engineering of heteroepitaxial growth on silicon. In particular, the quality of the heteroepitaxial layers depends strongly on interfacial properties, and can be optimized only when the deposition process, especially the early stage, is well understood. As discussed earlier,[5,8–10] PRS in the visible range shows a large sinusoidal signal during growth due to interference between front- and back-surface reflections associated with the growing GaP layer, from which the overall growth rate can be determined. Superimposed on this is a fine structure that is linked to the pulsed supply of the precursors. This fine structure is also observed in RDS and, to a lesser extent, in LLS. The three probes provide different perspectives of the growth process, and show in particular that in the growth system studied here, triethylgallium (TEG) dealkylation is the rate-limiting step.

II. EXPERIMENT

Heteroepitaxial growth was performed in a pulsed mode at low substrate temperatures of 360–400 °C as measured by thermocouples and calibrated by a pyrometer. The precursors were tertiarybutylphosphine (TBP) and TEG. The silicon substrates were lightly (1–10 Ω cm) p doped. To determine the possible influence of steps on growth, we used (113) substrates and substrates oriented 6° off (001) toward [110]. Since on-axis (001) Si substrates yielded no net RDS signals as a result of the formation of antiphase domains, (001) substrates were used only for comparison.

Details of the RDS configuration are given in Refs. 1 and 2, and those of PRS and LLS in Ref. 5. For PRS and LLS a He–Ne laser was used as light source (633 nm, 1.96 eV). RDS transients were taken at a fixed photon energy of 3.6 eV, where the optical penetration depth of GaP is low and the detected light intensity is high. The intrinsic RDS time resolution was the RDS sampling period, 1 ms. However, to improve signal-to-noise ratios we averaged 25 and 50 points for interrupted and cyclic growth, respectively, yielding effective time resolutions of 25 and 50 ms, respectively. The time resolution for PRS and LLS was 100 ms. The pneumatic switching valves for TEG and TBP have a small dead volume and a time response faster than 20 ms, as specified by the manufacturer (NUPRO).

III. LINEARIZATION OF THE PRS SIGNAL

A typical PRS transient for a GaP layer heteropitaxially grown on Si is shown in Fig. 1. These data can be described with the so-called four-phase model, which involves the ambient (0), a surface reaction layer (1), the GaP layer (2), and the Si substrate (3), where all interfaces are assumed to be sharp and reflections from the back of the substrate can be ignored.[11] It is necessary to include a surface reaction layer, because unreacted species on the surface of the growing GaP film generally have refractive indices different from that of

*Published without author corrections.
[a]Electronic mail: rossow@unity.ncsu.edu

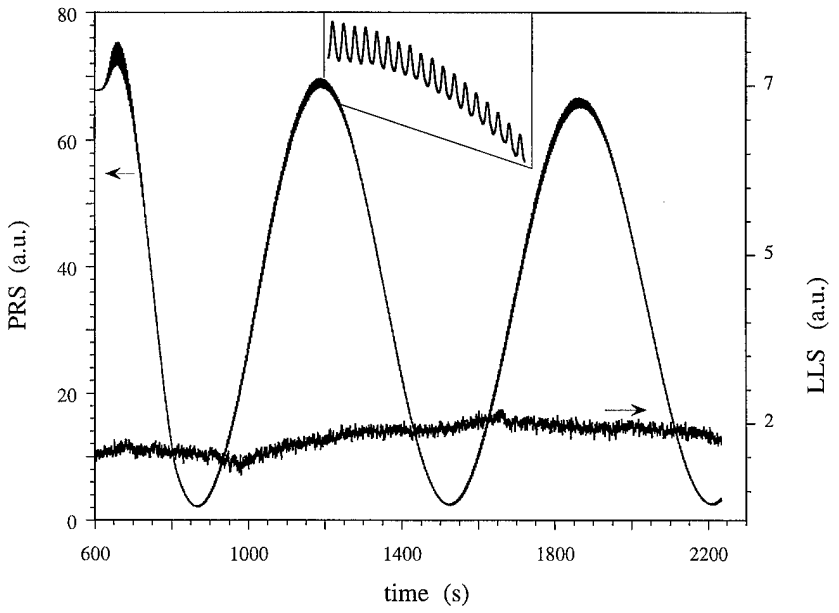


FIG. 1. Typical PRS transient during heteroepitaxial growth on Si. The fine structure is shown on an expanded scale in the inset. The nonzero values of the minima are due to an offset in the electronic signal.

GaP. This is manifest by the data of Fig. 1, which show a fine structure in the PRS data that accompanies cyclic changes in ambient exposure.

In the four-phase model the PRS signal is proportional to the absolute square $R_{4p}=|r_{4p}|^2$ of the complex reflectance coefficient r_{4p} for p-polarized light, which is given by[12]

$$r_{4p}=\frac{(r_{01p}+r_{12p}e^{i2\beta_1})+(r_{01p}r_{12p}+e^{i2\beta_1})r_{23p}e^{i2\beta_2}}{(1+r_{01p}r_{12p}e^{i2\beta_1})+(r_{12p}+r_{01p}e^{i2\beta_1})r_{23p}e^{i2\beta_2}}, \quad (1)$$

where

$$\beta_i=\frac{2\pi d_i}{\lambda}\sqrt{\epsilon_i-\sin^2\theta_0}, \quad i=1,2 \quad (2)$$

and where θ_0 is the angle of incidence. Since the surface reaction layer is thin ($|\beta_1|\ll 1$), we can linearize Eq. (1) with respect to β_1 as

$$r_{4p}(\beta_1)=r_{4p}(\beta_1=0)+\beta_1\left.\frac{\partial r_{4p}}{\partial\beta_1}\right|_{\beta_1=0}$$

$$=r_{4p0}+\beta_1\left.\frac{\partial r_{4p}}{\partial\beta_1}\right|_{\beta_1=0}. \quad (3)$$

The first term on the right side of Eq. (3) describes interference due to the growing GaP layer and the second the fine structure due to the variation of d_1 or ϵ_1 with exposure.

The interference term r_{4p0} is given by

$$r_{4p0}=\frac{r_{02p}+r_{23p}e^{i2\beta_2}}{1+r_{02p}r_{23p}e^{i2\beta_2}}. \quad (4)$$

If the Brewster condition $\theta_0=\theta_B$ is fulfilled exactly, then $r_{4p0}=r_{03p}=0$ when $\exp(i2\beta_2)=1$ and therefore $r_{02p}=-r_{23p}$. Under this condition

$$r_{4p0}=r_{02p}\frac{1-e^{i2\beta_2}}{1-|r_{02p}|^2e^{i2\beta_2}} \quad (5)$$

$$\approx -2ir_{02p}e^{i\beta_2}\sin\beta_2. \quad (6)$$

For real β_2 we therefore have

$$\text{PRS signal}\sim R_{4p0}=|r_{4p0}|^2=4R_{02p}\sin^2\beta_2, \quad (7)$$

where $R_{02p}=|r_{02p}|^2$. Thus for $\theta_0=\theta_B$ and $\beta_1=0$ the PRS signal exhibits a sinusoidal behavior with a period $\Delta d=\lambda/2(\epsilon_2-\sin^2\theta_0)^{-1/2}$. For GaP at 632.8 nm [$\epsilon_{\text{GaP}}=10.96+i0.00$ (see Ref. 13)] and $\theta_0=70°$, Δd is approximately 101 nm.

For growth monitoring the Brewster condition is difficult to realize because the dielectric function, and consequently θ_B, is temperature dependent. For the more general case of $\theta_0\approx\theta_B$, R_{4p0} is given by

$$R_{4p0}\approx R_{03p}+2[R_{23p}-\text{Re}(r^*_{03p}r_{23p})]-2\cos 2\beta_2[R_{23p}$$
$$-\text{Re}(r^*_{03p}r_{23p})]-2\sin 2\beta_2\,\text{Im}(r^*_{03p}r_{23p}), \quad (8)$$

where the R_{ijp} terms have obvious meaning. If the substrate is weakly absorbing, as is the case for Si at 632.8 nm, the last term can be neglected and we obtain

$$R_{4p0}\approx R_{03p}+2[R_{23p}-\text{Re}(r^*_{03p}r_{23p})][1-\cos 2\beta_2], \quad (9)$$

which is basically Eq. (7) with a dc offset $R_{03}\geq 0$, as seen in Fig. 1. If $R_{03}=0$ ($\theta=\theta_B$) then from either Eq. (7) or (9) the PRS signal must begin to rise when GaP growth is initiated. However, Fig. 1 shows that this is not observed: the signal begins near its maximum and then decreases. From Eq. (9) this is possible if $R_{03}\neq 0$ and $r_{23p}<r_{03p}$, which means that $[R_{23p}-\text{Re}(r^*_{03p}r_{23p})] < 0$. Furthermore, with increasing thickness of the GaP layer we observe that the signal ap-

proaches zero. Therefore, $R_{03p} \approx 4|R_{23p} - \mathrm{Re}(r^*_{03p} r_{23p})|$. From this result it follows that $r_{03p} \approx 2 r_{23p}$. Using these conditions and the values for the dielectric functions of Si and GaP given above, we can calculate θ_0. We find this to be about 70.4°, which is slightly less than θ_B.

To determine the fine-structure contribution we must evaluate the β_1 scaling factor

$$\left.\frac{\partial r_{4p}}{\partial \beta_1}\right|_{\beta_1=0} = \frac{2i(1-r_{01p}^2)/r_{12p}}{(1+r_{01p}r_{12p})^2} \times \frac{(1+r_{23p}e^{i2\beta_2})^2 + r_{23p}e^{i2\beta_2}\left(\dfrac{1+r_{12p}^2}{r_{12p}} - 2\right)}{(1+r_{02p}r_{23p}e^{i2\beta_2})^2}. \tag{10}$$

For $\theta_0 \approx \theta_B$, r_{02p} and r_{23p} are small compared to 1 [≈0.083 and 0.072, respectively, for $\theta_0 = 70°$ and $\epsilon_{\mathrm{Si}} = 15.07 + i0.15$) [see Ref. 13]. Neglecting products of these terms yields

$$\left.\frac{\partial r_{4p}}{\partial \beta_1}\right|_{\beta_1=0} \approx \frac{2i(1-r_{01p}^2)}{(1+r_{01p}r_{12p})^2}\,[r_{12p} + r_{23p}e^{i2\beta_2}(1+r_{12p}^2)]. \tag{11}$$

The term in the brackets can be given a simple geometric interpretation. In the transparent region of GaP, β_2 is real. Therefore, with increasing layer thickness $e^{i2\beta_2}$ describes a circle of unit radius in the complex plane. If the surface reaction layer is also transparent, r_{12p} is also real and the term in the brackets is largest/smallest for $e^{i2\beta_2} = \pm 1$ depending on the sign of r_{12p} and r_{23p}. Since for $e^{i2\beta_2} = \pm 1$ the interference term shows extrema the fine structure will be largest/smallest at the minima/maxima of the PRS signal (see Fig. 1). This effect is most pronounced when $r_{12p} \approx r_{23p}(1+r_{12p}^2)$, a condition that is fulfilled for reasonable values of ϵ_1 (for example, for $\epsilon_1 = 8$ the two terms are nearly equal). For small surface absorption we can write $r_{12p} = |r_{12p}|(1 + i\delta)$, where $\delta \ll 1$. Then

$$[r_{12p} + r_{23p}e^{i2\beta_2}(1+r_{12p}^2)] \sim [|r_{12p}| + i|r_{12p}|\delta + r_{23p}e^{i2\beta_2}(1+|r_{12p}|^2)], \tag{12}$$

where the term $\delta r_{23p}|r_{12p}|^2$ can be neglected. Now the center of the circle with radius $r_{23p}(1+|r_{12p}|^2)$ is located a distance $|r_{12p}|$ on the real axis and $\delta|r_{12p}|$ above the real axis. The amplitude of the fine structure reaches its maximum value at $\tan 2\beta_2 = |r_{12p}|\delta/(r_{23p}[1+|r_{12p}|^2])$. This no longer coincides with the extrema of the interference term. Therefore, from the relative phases of the interference and fine-structure signals we can obtain information about $\mathrm{Im}(r_{12p})$, and thus the surface absorption.

If $\mathrm{Im}(r_{12p}) \gg \mathrm{Re}(r_{12p})$, which is the case for a metallic surface layer with $\mathrm{Re}(\epsilon_1) < 0$, $\mathrm{Im}(\epsilon_1) > 0$, no phase matching of r_{12p} and $r_{23p}[1+r_{12p}^2]$ can occur. However, if $\mathrm{Re}(r_{12p}) \ll 1$ then we again obtain the equation of a circle, but one with its center near the imaginary axis.

Combining the previous results the general expression for the PRS signal, to first order in β_1, is

$$\mathrm{PRS} \sim R_{4p} \approx R_{4p0} + 2\,\mathrm{Re}\left[r^*_{4p0}\beta_1 \left.\frac{\partial r_{4p}}{\partial \beta_1}\right|_{\beta_1=0}\right]. \tag{13}$$

Therefore, the amplitude of the fine structure that is caused by a variation of β_1, i.e., variations in ϵ_1 and/or d_1, depends also on r^*_{4p0} and $\partial r_{4p}/\partial\beta_1|_{\beta_1=0}$. Neglecting for simplicity the squared terms of the Fresnel coefficients we obtain

$$\mathrm{PRS} \sim R_{4p} \approx R_{03p} + 2[R_{23p} - \mathrm{Re}(r^*_{03p}r_{23p})][1-\cos 2\beta_2] \times \mathrm{Re}\{r_{02p}(1-e^{-i2\beta_2})\beta_1 2i[r_{12p} + r_{23p}e^{i2\beta_2} \times (1+r_{12p}^2)]\}. \tag{14}$$

Therefore, the amplitude of the fine structure is also periodic in $2\beta_2$. In general, the situation is complicated because the fine-structure term is multiplied by real and imaginary parts of r^*_{4p0}, which results in terms $(1-\cos 2\beta_2)$ and $\sin 2\beta_2$ that are not in phase. For one extremum $2\beta_2 = \pi$, in which case

$$2\,\mathrm{Re}\left[r^*_{4p}(\beta_1=0)\beta_1 \left.\frac{\partial r_{4p}}{\partial \beta_1}\right|_{\beta_1=0}\right] = -r_{02p}\frac{16\pi d}{\lambda}\,[\mathrm{Re}(\sqrt{\epsilon_1 - \sin^2\theta_0}) \times \mathrm{Im}(r_{12p} - r_{23p}[1+r_{12p}^2]) + \mathrm{Im}(\sqrt{\epsilon_1 - \sin^2\theta_0}) \times \mathrm{Re}(r_{12p} - r_{23p}[1+r_{12p}^2])]. \tag{15}$$

This vanishes for real ϵ_1. The same occurs for $2\beta_2 = 0$. In contrast, we often observe that the fine structure is largest near the maxima of PRS signals. Consequently, in these cases $\mathrm{Im}(\epsilon_1)$ must be significant. This also follows directly from Eq. (1). At the extrema, $\exp(i2\beta_2) = \pm 1$, and for GaP on Si at 632.8 nm all Fresnel coefficients are real. Then the absolute value of r_{4p} contains only terms in $\cos 2\beta_1$, which is quadratic in β_1. We note also that for both parts of the PRS signal the zero-order term due to interference, $|r_{4p}(\beta_1=0)|^2$, and the fine structure term, $2\,\mathrm{Re}[r^*_{4p}(\beta_1 = 0)\beta_1(\partial r_{4p}/\partial\beta_1)|_{\beta_1=0}]$, are both of the order of the square of the Fresnel coefficient. Therefore, the large descrepancy in value between the two terms can only be explained by β_1 being small as suggested above. The ratio in the maxima is approximately

$$2\,\mathrm{Re}\left[r^*_{4p}(\beta_1=0)\beta_1 \left.\frac{\partial r_{4p}}{\partial \beta_1}\right|_{\beta_1=0}\right] \Big/ |r_{4p}(\beta_1=0)|^2 = -1/r_{02p}\frac{4\pi d}{\lambda}\,[\mathrm{Re}(\sqrt{\epsilon_1 - \sin^2\theta_0}) \times \mathrm{Im}(r_{12p} - r_{23p}[1+r_{12p}^2]) + \mathrm{Im}(\sqrt{\epsilon_1 - \sin^2\theta_0}) \times \mathrm{Re}(r_{12p} - r_{23p}[1+r_{12p}^2])]. \tag{16}$$

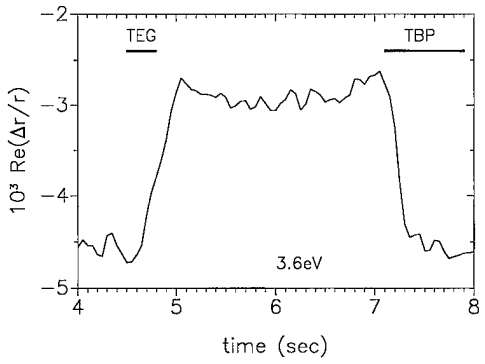


FIG. 2. RDS transient at 3.6 eV for one exposure cycle during heteroepitaxial growth of GaP on a vicinal wafer orientated 6° off (001) toward [110]. The response to the TEG pulse is slower than that to the TBP pulse. The upper level is not flat but has overlaid structure.

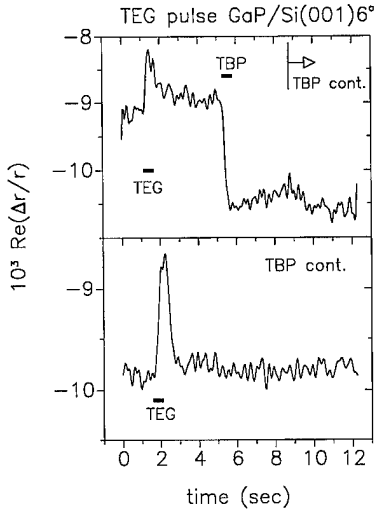


FIG. 3. Comparison between RDS responses to a single pulse of TEG on a surface oriented 6° off (001) toward [110], where the TBP exposure is either continuous (lower panel) or pulsed (upper panel). All pulse durations were 0.5 s, as indicated by the solid bars. With continuous TBP exposure the baseline is essentially invariant, indicating that the surface recovers quickly after the TEG pulse.

From Fig. 1 this ratio is 0.02. If we assume that the real and imaginary parts of $(r_{12p}-r_{23p}[1+r_{12p}^2])$ are of the order of $1/r_{02p}$, which is a crude approximation, and neglect $\sin^2\theta_0$ the ratio becomes

$$\frac{4\pi d}{\lambda}[\mathrm{Re}(\sqrt{\epsilon_1})+\mathrm{Im}(\sqrt{\epsilon_1})]. \tag{17}$$

If we further assume that d is about 0.5 nm, then $4\pi d/\lambda$ is 0.01 and $[\mathrm{Re}(\sqrt{\epsilon_1})+\mathrm{Im}(\sqrt{\epsilon_1})]$ is of the order of 2. This sets upper limits of 10 to the imaginary and absolute values of the real part of ϵ_1.

Although our motivation for this analysis is the understanding of the PRS data for GaP on Si, this analysis holds in general for any heterostructure where the epitaxial layer has a high refractive index and is not optically absorbing.

IV. RESULTS AND DISCUSSION

In the following we discuss the results of three different types of experiments: continuous cyclic growth, single pulses of TEG during continuous TBP exposure, and single pulses of both TEG and TBP. These experiments were performed for (113) surfaces and surfaces cut 6° off (001) toward [110] as indicated either in the figures or in the captions. Similar results were obtained for both orientations. So far, we have found no evidence of a growth dependence on step density.

A. Continuous cyclic growth

Figure 2 shows a typical RDS transient obtained at 3.6 eV and 25 ms averaging for a single 4 s cycle of alternating TEG and TBP exposures during the growth of GaP on a Si surface oriented 6° off (001) toward [110]. At 3.6 eV this GaP layer is optically thick. The onsets of the rising and falling parts of the transient correlate with the onsets of TEG and TBP exposures, respectively, but the rising and falling durations do not coincide with precursor exposures. Although TEG exposure was 0.3 s the rise time is broadened to about 0.5 s, whereas the decay time is less than 0.25 s for a TBP pulse of 0.8 s. An overshoot also appears on both rising and falling edges. The difference between response and exposure times are likely caused by surface processes.

B. Single precursor pulses

To investigate this behavior further we applied single 0.3 and 0.5 s TEG pulses to freshly grown GaP annealed at the growth temperature by a continuous or pulsed supply of TBP. The RDS response to a 0.5 s TEG pulse during TBP exposure for a surface annealed in TBP for 54 s following growth is shown in the lower panel of Fig. 3. At this point TBP exposure was terminated but the molecular hydrogen flow remained. The upper panel shows the result obtained when this surface was exposed to TEG followed 4 s later by a TBP pulse and 2.5 s later by continuous exposure to TBP. A repeat of this sequence yielded the same results except that the starting level was lower than that shown in the upper panel of Fig. 3. This level shift indicates that the surface becomes Ga rich when TBP exposure is terminated.

The RDS response to TEG clearly depends on whether TBP exposure is interrupted or continuous. With continuous TBP exposure the RDS signal increases linearly during the TEG pulse and afterward decays exponentially to its starting value. With interrupted TBP exposure a TEG pulse generates a fast (rise time <0.25 s) response followed by an exponential decay of time constant of about 0.7 s to a level higher than the original baseline. The original baseline is recovered only after further TBP exposure. We expect that the change of the RDS signal is related to the total amount of Ga present on the surface. If we take the RDS signal of the phosphorous-rich, TBP-annealed surface as the reference, the change in the RDS signal after TBP exposure in the upper transient ($t\approx 5$ s) is within 20% of the change induced by

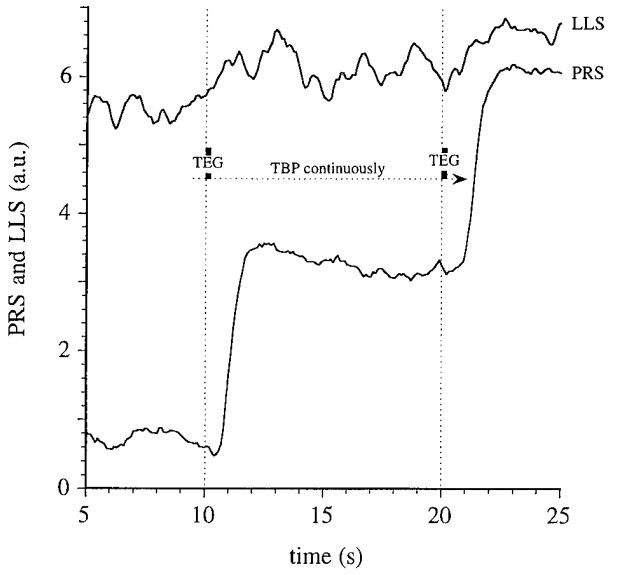


FIG. 4. Typical PRS and LLS transients for TEG pulses for a (113) surface under continuous exposure to TBP.

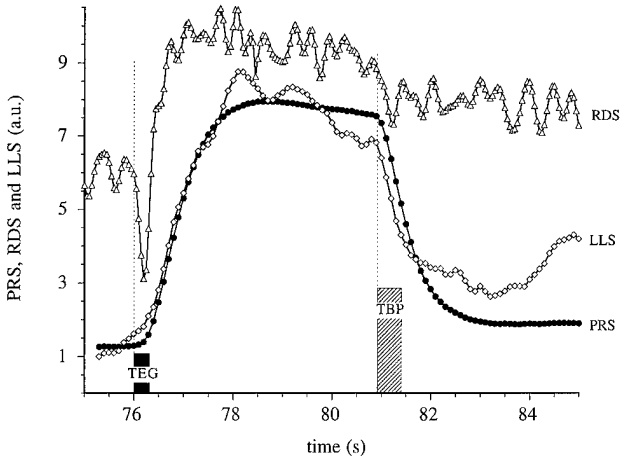


FIG. 5. Comparison of PRS, LLS, and 3.6 eV RDS responses upon exposure of a (113) surface to a TEG pulse followed by a TBP pulse.

TEG for continuous TBP exposure in the lower transient. It is also comparable to the change in signal for continuous cyclic growth when TBP is pulsed (see Fig. 2).

The PRS and LLS data provide another perspective of these processes, as illustrated in Fig. 4. For continuous TBP exposure the scattered light intensity increases only slightly if at all following a TEG pulse, while the PRS transient exhibits steplike behavior. From Eq. (7) we know that during steady-state growth, the increasing film thickness yields a sinusoidal PRS response, which is a much weaker edge than observed. We can therefore conclude that more than one process is active on the surface. This is further supported by the surprising fact that Ga does not begin to be incorporated into the GaP layer until after the TEG pulse is complete. This is not in contradiction to the faster RDS response of Fig. 2, because at 3.6 eV RDS is sensitive only to changes of surface anisotropy and not to layer thickness.

More insight into these processes can be obtained by examining responses to separate pulses of TEG and TBP. In Fig. 5 we compare LLS, PRS, and RDS transients for separate TEG and TBP pulses for GaP layers grown on (113) Si surfaces. As with (001) surfaces TEG exposure results in a broad feature in all three sets of data. However, the onsets for the LLS and PRS responses are delayed substantially with respect to that of the TEG pulse. However, at substrate temperatures well below those used for normal growth, an immediate PRS and LLS response was obtained. Under present conditions a TBP pulse generates an immediate exponential decay in the PRS and LLS transients while the RDS signal recovers essentially linearly within 0.4 s. As shown in Fig. 6 the PRS and LLS decays are well described by exponentials with time constants of 0.54 and 0.70 s, respectively. The RDS transient is more complicated and is discussed below.

Since the TEG pulse length was 0.5 s and the PRS and LLS responses occurred after the TEG pulse had terminated, the PRS and LLS delay times are at least 0.5 s. Since this delay time is much longer than the specified response time of the valves and since the response to TBP exposure is effectively immediate, the results show that rate-limiting intermediate surface processes are involved. Light scattering is caused by structures on the surface with a characteristic length, i.e., a feature size or separation, of order λ (here 632.8 nm), which are removed or smoothed by exposure to TBP. The reflectance of the material in these structures is probably rather high because it is very unlikely that these structures are large but instead should have diameters of the order of a few nm. Since the refractive index of TEG is low—smaller than 1.002 for the atmospheric-pressure gas phase[14]—and TEG is unstable at the temperatures used here, the delayed LLS response indicates that the LLS signal is

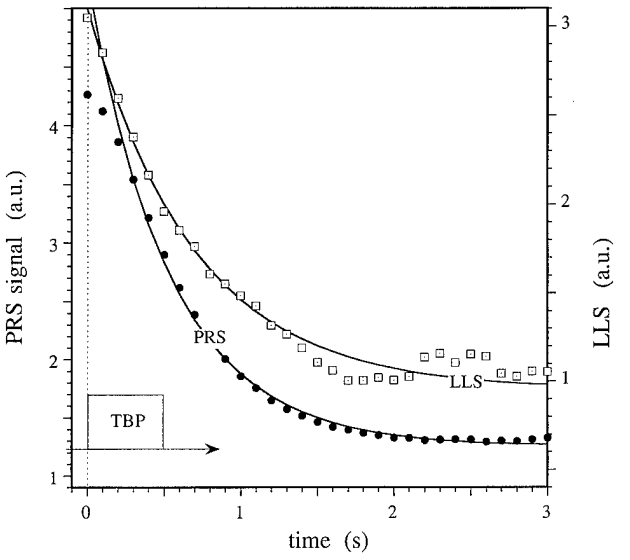


FIG. 6. Exponential fit to the PRS and LLS decay transients after TBP exposure for a (113) surface with a pulse sequence similar to that used in Fig. 5.

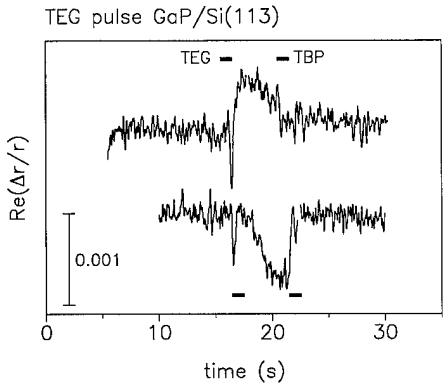


FIG. 7. Two RDS transients for a (113) surface annealed for 30 s in TBP followed by the same pulse sequence as in Fig. 5.

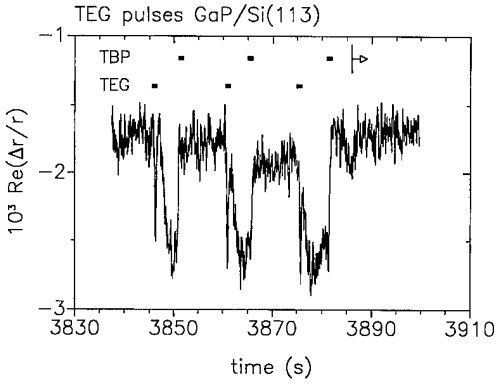


FIG. 8. RDS transients for a (113) surface for various combinations of TEG and TBP pulses followed by continuous TBP exposure.

due to the formation of Ga-containing clusters. The most likely constituent is metallic Ga, since neither monoethylgallium (MEG) nor diethylgallium (DEG) are likely to aggregate. Furthermore, even in liquid-phase densities their refractive indices should not exceed 1.5, which is too small to explain the LLS signal. However, the species that diffuse to form the clusters may be MEG, DEG, or even TEG, with complete dealkylization occurring only after the molecule arrives at the cluster. If MEG is the diffusing species, the fact that all its outer electrons are in filled orbitals suggests that it is relatively stable and hence there exists some probability that it could also desorb.

In this picture the delayed onset of the LLS and PRS responses would be due to complete and partial TEG dealkylization, which at the low pressures (low 10^{-5} mbar) and substrate temperatures (typically around 360 °C) used here is expected to occur on the surface instead of in the gas phase. The Ga-rich clusters remain stable because the incorporation of Ga as GaP requires the presence of P.

Although the LLS and PRS signals are related their origins are different. For sufficiently low TEG exposures the LLS signal vanishes but a strong PRS response remains. Moreover, depending on the thickness of the GaP layer and therefore the phase shift of the electromagnetic wave within the layer, the peaklike feature in the PRS transient due to the TEG pulse can be inverted as discussed above in the context of Eq. (11). Consequently, it is likely that the PRS signal is caused by a thin surface layer whose dielectric function differs strongly from that of the underlying film.

Figure 7 shows two RDS responses for a TEG pulse obtained on a GaP layer grown on a (113) Si surface. The lower transient is taken at a later time but under otherwise identical conditions. Clearly, each transient has at least two components. One is a brief negative spike representing a fast process that lasts less than 0.5 s. The other is a broad feature between the negative spike and the TBP pulse that is similar to that observed in PRS and LLS. The shape of the broad feature is history dependent, and unlike the negative spike it can appear either above or below the TBP-established base line. This variation of level cannot be due to interference from a backreflected wave within the GaP layer, as is the case for PRS, since at 3.6 eV the 20 nm penetration depth is much smaller than the layer thickness. Also, this behavior cannot be explained by surface chemistry alone, because if the surface only changes between P-rich and Ga-rich conditions the final state should be similar for the two transients. However, it is well known that the presence of anisotropic structures on the surface, for example ellipsoidal clusters, cause large RDS responses,[15] with the RDS signal depending on the shape and the optical properties of the clusters. For randomly distributed isotropic clusters the RDS signal must vanish by symmetry. A broad RDS response could be explained by supposing that initially isotropic clusters become anisotropic. The cluster material itself is not expected to be anisotropic since bulk Ga is liquid at the growth temperature. We also observe a change in the corresponding PRS transient.

Curiously, the sharp negative spike does not always appear at the onset of TEG exposure but, as shown in Fig. 8, is further delayed with successive TEG/TBP cycles. The TBP-established baseline also decreases with successive cycles, recovering only upon continuous exposure to TBP, as also shown in Fig. 8. This indicates that the surface retains some morphology remnant or fraction of fully or partially dealkylized TEG. The recovery effect of continuous TBP exposure is most pronounced in Fig. 9, where a steplike transient is observed.

C. Possible surface processes

It is puzzling that surfaces being continuously exposed to TBP react much more slowly to TEG than when TBP exposure is interrupted. No such variation is observed for TBP: the PRS, LLS, and RDS data all show that TBP appears to react immediately upon reaching the surface, regardless of its initial state. In particular, no LLS response would be expected if TBP were only physisorbed on the surface, since the refractive index of TBP is small. One possible explanation for the slower TEG response with TBP coexposure is that the butyl group requires time to desorb. The butyl group probably does not desorb directly as a radical or as fragments. Because the butyl group is also very bulky, it is also

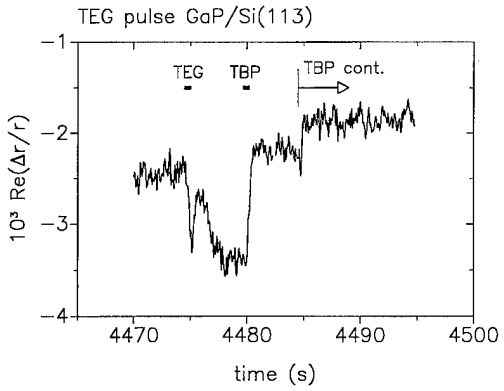


FIG. 9. As in Fig. 8 but after further deposition of GaP. A steplike response occurs when TBP exposure becomes continuous.

unlikely that two butyl groups will react or that butyl groups are incorporated into the growing layer. Most likely the butyl groups will react with hydrogen that originates from either (i) PH_2, (ii) the butyl group itself by conversion to *t*-butene (C_4H_8), or (iii) β-exchange within the ethyl groups to form C_2H_4. Hence, the likely possibilities are that (1) butyl reacts with atomic hydrogen at the surface, desorbing as *t*-butane or (2) *t*-butene is formed directly. In the former case we can expect that the attachment of TEG or its fragments to the P sites will be delayed. A similar argument should hold for TBP.

For surfaces terminated by P, one P bond is unsaturated. Two neighboring P atoms may form a dimer, or the dangling bond can be saturated by hydrogen. It is not clear at present whether hydrogen desorbs from the surface at these growth temperatures. However, the behavior of Si surfaces may give some hints about the nature of the P dangling bond. The Si dimer bond is strong and the desorption temperature of H from the dihydride phase is lower (about 300 °C) than that of H from the monohydride phase (around 520 °C). Since P has a lone pair orbital, a P-rich surface terminated by H should more closely resemble the Si dihydride than the Si monohydride. Growth temperatures near 360 °C are therefore above the temperature for H desorption. Hydrogen desorption may proceed in three ways: (1) as H_2; (ii) as ethane formed by β-exchange within the ethyl group; and (iii) as *t*-butane (C_4H_{10}) by saturation within the butyl group. However, in all cases all surface dangling bonds are saturated and the surface is passivated in a manner similar to that of As-terminated Si.[16,17]

V. SUMMARY AND CONCLUSIONS

Comparison of RDS, PRS, and LLS data for exposure of GaP layers grown on vicinal (001) and (113) Si surfaces to single pulses of TEG and TBP reveals that RDS, PRS, and LLS deliver complementary information, a necessary prerequisite for an unambigious interpretation of growth mechanisms. However, the observed chemistry is very complex. Further investigations using a quadrupole mass spectrometer in addition to the optical probes and an optical multichannel analyzer (OMA) detection system to obtain PRS spectra are in progress.

We have derived linearized expressions for the fine structure in the PRS data, which shows that a reaction layer including all not-yet-incorporated surface species is present on the surface of the growing GaP layer. This reaction layer is optically absorbing for the wavelengths used here. The time constants of the surface reaction processes preceding Ga and P incorporation are of the order of 1 s under present conditions, and appear to be substantially faster for TBP than for TEG. At least two processes occur when the surface is exposed to TEG. From LLS data we find evidence for Ga clustering, which is consistent with the observation that the surface reaction layer is optically absorbing. Unreacted species that are not directly incorporated into the growing layer also remain on the surface possibly for tens of seconds. For this system we find no evidence for atomic layer epitaxy, i.e., that growth is self-limiting. Our results are not specific to GaP on Si and are equally valid for GaP homoepitaxy as well. We expect that the ternary system $Ga_xIn_{1-x}P$ will behave similarly.

ACKNOWLEDGMENTS

This work was supported by the Alexander von Humboldt Foundation, the Office of Naval Research under Contract No. N-00014-93-1-0255, ARPA/AFOSR Grant No. F49620-95-1-0437, and DOD/AFSOR MURI Grant No. F49620-95-1-0447.

[1]D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).
[2]D. E. Aspnes, J. P. Harbison, A. A. Studna, L. T. Florez, and M. K. Kelly, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[3]W. Richter, Philos. Trans. R. Soc. London Ser. A **344**, 453 (1993).
[4]N. Dietz, A. Miller, J. T. Kelliher, D. Venables, and K. J. Bachmann, J. Cryst. Growth **150**, 691 (1995).
[5]N. Dietz, U. Rossow, D. E. Aspnes, and K. J. Bachmann, J. Electron. Mater. **24**, 1569 (1995).
[6]D. J. Robbins, A. J. Pidduck, C. Pickering, I. M. Young, and J. L. Glasper, Proc. SPIE **1002**, 25 (1988).
[7]K. J. Bachmann, U. Rossow, N. Sukidi, H. Castleberry, and N. Dietz, J. Vac. Sci. Technol. **14**, 3019 (1996).
[8]K. J. Bachmann, U. Rossow, and N. Dietz, Mater. Sci. Eng. B **35**, 472 (1995).
[9]N. Dietz, U. Rossow, D. E. Aspnes, and K. J. Bachmann, J. Cryst. Growth (in press).
[10]N. Dietz, U. Rossow, D. E. Aspnes, and K. J. Bachmann, Appl. Surf. Sci. (in press).
[11]N. Dietz and K. J. Bachmann, Vacuum **47**, 133 (1996).
[12]R. M. A. Azzam and N. M. Bashara: *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1977).
[13]D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 983 (1983).
[14]Refractive indices of gases with similar attached groups are well below 1.002. See *Handbook of Chemistry and Physics*, 64th ed. (Chemical Rubber, Boca Raton, FL, 1983), p. E-365.
[15]E. Steimetz, T. Zettler, W. Richter, D. I. Westwood, D. A. Woolf, and Z. Sobiesierski, J. Vac. Sci. Technol. **14**, 3058 (1996).
[16]R. D. Bringans, Crit. Rev. Solid State Mater. Sci. **17**, 353 (1992).
[17]U. Rossow, U. Frotscher, W. Richter, and D. R. T. Zahn, Surf. Sci. **287/288**, 718 (1993).

Implications of excess strain in As compound/P compound III–V multilayer superlattices grown by metal-organic vapor-phase epitaxy

A. R. Clawson[a)]
ECE Dept-0407, University of California San Diego, La Jolla, California 92093-0407

C. M. Hanson
NCCOSC RDTE DIV 8505, San Diego, California 92152-5070

(Received 22 January 1996; accepted 2 April 1996)

The As/P exchange behavior at interfaces of As compound/P compound heterojunction superlattices has been assessed by observing changes in average lattice spacing when very thin but different thickness As compound layers are inserted into InP, or P compound layers inserted into GaAs. Compositional changes are inferred from strain behavior measured by x-ray diffraction. Implications of the data are: (1) As rapidly displaces P and P slowly displaces As; (2) As atoms displace P atoms and reconstruct as a discrete single layer above 600 °C and as a double layer below 600 °C when InP surfaces are exposed to AsH_3; and (3) interfaces and layers of different P compounds on GaAs show the same temperature-dependent inclusion of As. © *1996 American Vacuum Society.*

I. INTRODUCTION

InGaAs/InP multiple quantum wells are often unintentionally compressively strained relative to the InP substrate. This strain is variously attributed to interaction of AsH_3 with the InP surface as InGaAs growth is started,[1] to carry over of excess As from InGaAs into InP,[2] or to different bond lengths at heterojunction interfaces.[3]

The purpose of this study was to better understand the formation of As/P heterojunction interfaces grown by metal-organic vapor-phase epitaxy (MOVPE) and the implications for growth of both InGaAs/InP and InGaP/GaAs materials systems. Superlattices were formed by inserting very thin strained layers (equivalent to a few angstroms) of a P compound III–V semiconductor into GaAs, or alternatively an As compound into InP. The strain of the superlattices was analyzed from x-ray-diffraction rocking curves. Three implications of As–P exchange have been made and are discussed. These are that

(1) As rapidly displaces P and P slowly displaces As,

(2) As atoms displace P atoms and reconstruct as a discrete single layer above 600 °C and as a double layer below 600 °C when InP surfaces are exposed to AsH_3, and

(3) interfaces and layers of different P compounds on GaAs show the same temperature-dependent (In)$GaAs_{1-x}P_x$ composition.

II. EXPERIMENTAL DETAILS

The superlattices were grown in a 50-mm-diam horizontal quartz chamber using a lamp-heated graphite susceptor. Use of 20 Torr chamber pressure and 1500 sccm total gas flow provided gas velocity of ~1 m/s to allow very rapid change of source gases and controllable growth periods of less than 1 s. The growth parameters are summarized in Table I. The thin strained P (or As) compound III–V semiconductors were inserted into GaAs (or InP) at intervals of 180 Å for 30 repetitions.

The multilayer superlattice strain was analyzed with x-ray-diffraction rocking curves using superlattice peak structure about the Cu $K\alpha_1$ (004) peak from (001)-oriented substrates. Changes in the average lattice parameter are measured from the strain in the superlattice.

III. METHOD OF INTERPRETATION

The x-ray diffraction from the (004) peak directly measures the superlattice strain as

$$(\Delta a_\perp/a)_{SL} d_{SL} = (\Delta a_\perp/a)_L d_L + (\text{unknown}), \quad (1)$$

where $(\Delta a_\perp/a)_{SL}$ is the superlattice strain, d_{SL} is the period thickness, $(\Delta a_\perp/a)_L$ is the inserted layer strain, d_L is the layer thickness, and the unknown is the interface strains multiplied by the interface thicknesses. Typically in these superlattices there is strain in excess to what can be accounted from the inserted layer alone. By using measured and known data we can extract the unknown excess strain contribution. If Vegard's law is then used to relate lattice parameter to alloy composition, inferences can be made about the As–P intermixing at the heterojunctions.

The analysis on a sample series is done by looking at the strain as a function of inserted-layer growth time such as shown in Fig. 1. Two key parameters used in the interpretation of the data are the excess strain extrapolated to zero time $\Delta\sigma(0)$, and the excess strain as a function of time $\Delta\sigma(t)$. $\Delta\sigma(0)$ is thought to represent two strain contributions, one from the inserted-layer interface as the inserted-layer growth is initiated, and one from the interface as the inserted-layer growth is terminated. $\Delta\sigma(t)$ is thought in most cases to be excess strain within the inserted layer.

a)Electronic mail: clawson@ece.ucsd.edu

TABLE I. Epilayer growth conditions.

Source partial pressures		Growth rates	
PH_3	1.3×10^{-2}	InP	2.2 (Å/s)
AsH_3	8.0×10^{-3}	InAs	2.0
TMIn	$\sim5\times10^{-5}$	GaAs	2.2
TMGa	$\sim5\times10^{-5}$	GaP	2.5

The signs of $\Delta\sigma(0)$ and $\Delta\sigma(t)$ are useful to give a quick idea of the composition which could cause the excess strain. In InP, only alloys with InAs can give a positive (compressive) strain. Although the precise $InAs_{1-x}P_x$ composition cannot be known, we assume for simplicity that the composition is pure InAs and determine what the thickness would be. Vegard's law assures that the strain for the As is the same whether it is an InAs layer or is distributed as an InAsP alloy. Similarly, for GaAs, only alloys with GaP can give a negative (tensile) strain.

Several assumptions must be made for this strain analysis, as follows.

(1) Growth rates for thin layers must be the same as has been measured for thick layer structures. An accurate equivalent thickness is needed to calculate the expected strain contribution introduced by the inserted layer. Prior published work has shown this to be a good assumption.[4]

(2) Vegard's law relates lattice parameters to compositions of very thin layers. This has also been shown to be a good assumption, supported by correct correlation of strain to thickness and mismatch for fractions of a monolayer (ML).

(3) Excess strain at interfaces can only occur from exchange of atoms on the group-V sublattice, i.e., only from As–P. This again is thought to be valid because at the high V/III ratios being used, essentially all available group III atoms are incorporated. Even if the group-III atoms were intermixing across the heterojunction, there is no excess strain because there are no excess group-III atoms incorporated.

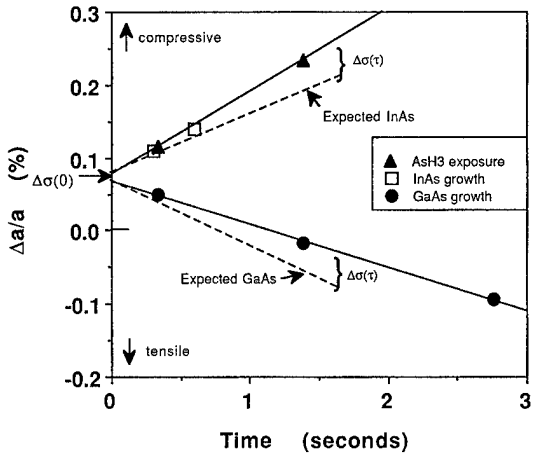


FIG. 1. InP superlattices (30 periods of 180 Å) formed at 650 °C by (1) AsH_3 exposure of InP, (2) InAs thin layer growth, and (3) GaAs thin layer growth. The data demonstrate an excess InAs-like strain $\Delta\sigma(0)$ at $t=0$, and an excess InAs-like time-dependent strain $\Delta\sigma(t)$.

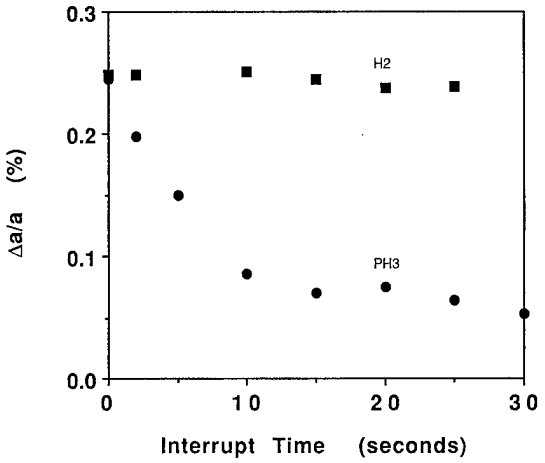


FIG. 2. Strain in InP superlattices (30 periods of 180 Å) formed at 650 °C by two steps: (1) 1.4 s exposure of InP to AsH_3 to form an As-terminated surface; (2) growth interrupt under (H_2+PH_3) or H_2 only to desorb the surface As.

(4) The strained regions are coherent with the substrate lattice, and tetragonal distortion perpendicular to the interface accommodates difference in the lattice spacing. The perpendicular lattice parameter $a_\perp$ is related to the relaxed lattice parameter a by the elastic modulii as[4]

$$\Delta a=\Delta a_\perp[C_{11}/(C_{11}+2C_{12})]\approx0.5\Delta a_\perp . \quad (2)$$

A. Implication 1: As rapidly displaces P, and P slowly displaces As

InAs-like excess strain components, $\Delta\sigma(0)$ and $\Delta\sigma(t)$, form when the InP surface is exposed to AsH_3 as shown in Fig. 1, and this occurs for both simple AsH_3 exposure of the InP surface and when layers of InGaAs, InAs, or GaAs are grown.[5] The As for $\Delta\sigma(0)$ is attributed to As rapidly displacing P surface atoms when InP is exposed to AsH_3, but P not quickly displacing As surface atoms exposed to PH_3.

Deliberate exposure of an As-terminated InP surface to PH_3 at 650 °C as shown in Fig. 2 results in rather slow displacement of the As, confirming that As atoms are slowly displaced by P. Removal of As atoms becomes very slow when the strain is equivalent to a surface coverage less than a single layer. The As-terminated surface is stable when exposed to interrupt in a hydrogen atmosphere for up to 25 s.

In Fig. 3 for GaAs exposed to PH_3 a partial displacement of As surface atoms by P to form $\Delta\sigma(0)$ occurs within less than 1 s, but P incorporation to form $\Delta\sigma(t)$ is very slow. Rapid displacement of P by As is demonstrated in this case by establishing P surface atoms with exposure of the GaAs surface to PH_3 for 20 s at 650 °C, then exposing the P surface to AsH_3 for 1 s. The strain dropped to zero and x-ray satellite peaks disappeared. The P atoms were desorbed under AsH_3. It is evident without the AsH_3 interrupt that P atoms remain

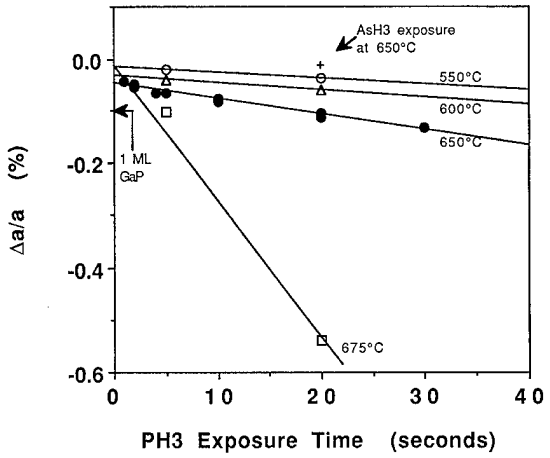


FIG. 3. Temperature dependence of strain in superlattices (30 periods of 180 Å) formed by periodically exposing the MOVPE-grown GaAs growing surface to PH_3. $\Delta\sigma(0)$ is attributed to GaP formed by P adsorbed on the GaAs surface. An additional excess time-dependent GaP-like strain component $\Delta\sigma(t)$ is slowly formed. The (+) data point results when a 20 s exposure to PH_3 is followed by a 1 s exposure to AsH_3.

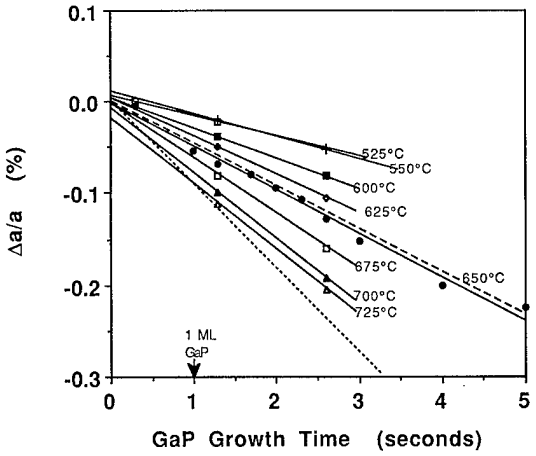


FIG. 5. Temperature dependence of strain in superlattices (30 periods of 180 Å) formed by inserting thin GaP layers periodically into GaAs. The dotted line is the calculated strain $\sigma(t)$ for GaP, and the dashed line is $\sigma(t)$ calculated for $GaAs_{0.5}P_{0.5}$. Unintended incorporation of residual As results in layers of $GaAs_{1-x}P_x$ in which x increases with temperature.

when a GaAs overlayer is grown using AsH_3 in combination with a trimethylgallium.

B. Implication 2: As atoms displace P atoms and reconstruct as a discrete single layer above 600 °C and as a double layer below 600 °C when InP surfaces are exposed to AsH_3

The InAs-like strain $\Delta\sigma(0)$ from brief AsH_3 exposure of InP is shown in Fig. 4 to be equivalent to ~1 monolayer InAs for growth temperatures above 600 °C and to be equivalent to ~2 monolayers InAs for growth temperatures below 600 °C. The implied As surface reconstruction on the InP is ~1 layer of As atoms above 600 °C and ~2 layers of As atoms below 600 °C, with virtually all the surface As incorporated as InP growth resumes.

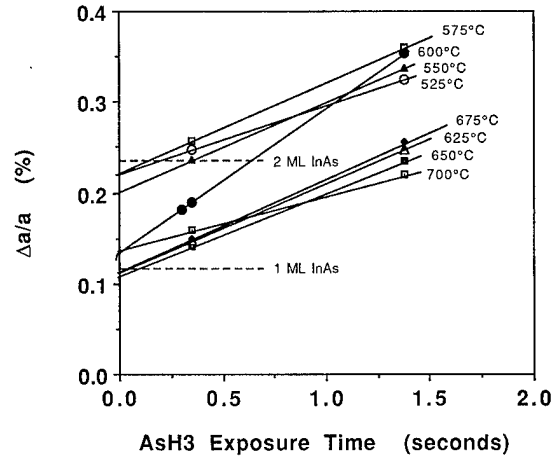


FIG. 4. Temperature dependence of strain in superlattices (30 periods of 180 Å) formed by periodic exposure of InP surface to AsH_3.

C. Implication 3: Interfaces and layers of different P compounds on GaAs show the same temperature-dependent (In)$GaAs_{1-x}P_x$ composition

Temperature dependence of strain from exposure of the GaAs surface to PH_3 is shown in Fig. 3. The negative $\Delta\sigma(0)$ and $\Delta\sigma(t)$ indicate that there is replacement of As atoms in the GaAs with P atoms. $\Delta\sigma(0)$ for 650 °C growth is equivalent to that of either 0.5 ML of GaP or its equivalent of 1 ML of $GaAs_{0.5}P_{0.5}$. As the growth temperature is reduced, $\Delta\sigma(0)$ decreases, i.e., becomes more GaAs-like. From $\Delta\sigma(t)$ behavior, further replacement of As by P is a slow process for temperatures of 650 °C or less, but is much faster as the temperature is raised to 675 °C. The fact that the 650 °C excess strain for long PH_3 exposures increases beyond that equivalent to a single ML of GaP (strain of 1 ML GaP in 180 Å GaAs is ~0.10%) suggests that the time-dependent P substitution for As is not limited to the surface but penetrates below the top-most layer.

In Fig. 5 is shown the strain for deliberately grown thin GaP layers. $\Delta\sigma(0)=0$, at least for $T=650$ °C. The $\Delta\sigma(t)$ behavior does not correspond to that calculated for the GaP thickness, but is consistent with alloys of $GaAs_{1-x}P_x$ where x varies with growth temperature. Considerable unintended As is incorporated during growth of the initial monolayers of GaP, and the amount of As increases at lower temperatures. At 650 °C the layer composition is nominally $GaAs_{0.5}P_{0.5}$. The existence of a large reservoir of residual As is evident from the 650 °C growths which ranged in layer thicknesses up to 5 ML without significant As depletion. Whether the residual As is from excess adsorbed on the GaAs surface or from surrounding reactor deposits is unresolved.

The strain for thin InP layer superlattices in GaAs is shown in Fig. 6. The InAsP$\Delta\sigma(0)$ is seen to be temperature dependent, however, the $\Delta\sigma(t)$ over the growth temperature range 550–650 °C has the same slope as calculated for InP

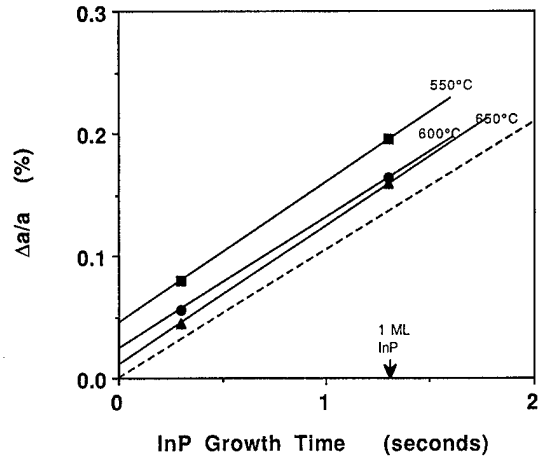


FIG. 6. Temperature dependence of strain in superlattices (30 periods of 180 Å) formed by inserting thin InP layers periodically into GaAs. The dotted line is the calculated strain $\sigma(t)$ for InP.

layer growth. There is no evidence of excess strain due to residual As incorporating into the InP.

The results for thin $In_{0.5}Ga_{0.5}P$ layer superlattices in GaAs are shown in Fig. 7. Rather than the expected $\Delta\sigma(t)=0$ for nominally lattice-matched $In_{0.5}Ga_{0.5}P$, a compressive strain exists suggestive of an excess $In_{0.5}Ga_{0.5}As$ component. The $\Delta\sigma(0)$ are of like magnitude to those of InP in Fig. 6. $\Delta\sigma(t)$ for the $In_{0.5}Ga_{0.5}P$ layer at 650 °C best fits a layer composition of $In_{0.5}Ga_{0.5}As_{0.5}P_{0.5}$, showing the same $x\approx 0.5$ residual As incorporation as seen for thin GaP layer growth at 650 °C in Fig. 5. Again, the amount of residual As available for incorporation is surprising. Layers as thick as 10 ML continue to show the same composition without evidence of depleting the available As.

All of the above GaAs superlattices show evidence of unintended excess As incorporating at the interface boundaries and/or in the initial monolayers of P compound growth. The concentration of As and its temperature dependence can be analyzed, at least qualitatively, for each case as shown in Fig. 8. For GaAs exposed to PH_3 or with inserted layers of InP or $In_{0.5}Ga_{0.5}P$, the As vs P composition can be estimated from $\Delta\sigma(0)$. With PH_3 exposure it is assumed that $\Delta\sigma(0)$ is due to formation of a single $GaAs_{1-x}P_x$ monolayer sandwiched in GaAs. The layer composition of the embedded mixed As/P layer is calculated from the strained lattice parameter of this monolayer using Vegard's law. The InAsP-like $\Delta\sigma(0)$ of both InP and $In_{0.5}Ga_{0.5}P$ consists of the combined strains of the initial and terminal interfaces bridging the P compound layer to the GaAs. The interface compositions are not obvious, but the temperature dependence of $\Delta\sigma(0)$ shows that the As/P composition is similar to that for PH_3 exposure. To estimate the interface As/P composition it was assumed that for an $In_{0.5}Ga_{0.5}P$ layer the combined interfaces are equivalent to 2 ML $In_{0.25}Ga_{0.75}As_{1-x}P_x$, where each interface consists of a monolayer with bonds of $Ga-(As_{1-x}P_x)-(In_{0.5}Ga_{0.5})$. This same interface composition was also applied to the InP data to estimate x.

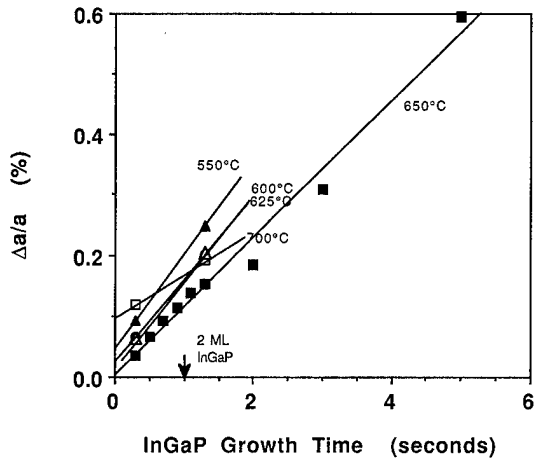


FIG. 7. Temperature dependence of strain in superlattices (30 periods of 180 Å) formed by inserting thin $In_{0.5}Ga_{0.5}P$ layers periodically into GaAs. There should be no $\sigma(t)$ for the nominally lattice-matched layer. Unintended incorporation of residual As results in layers of $In_{0.5}Ga_{0.5}As_{1-x}P_x$.

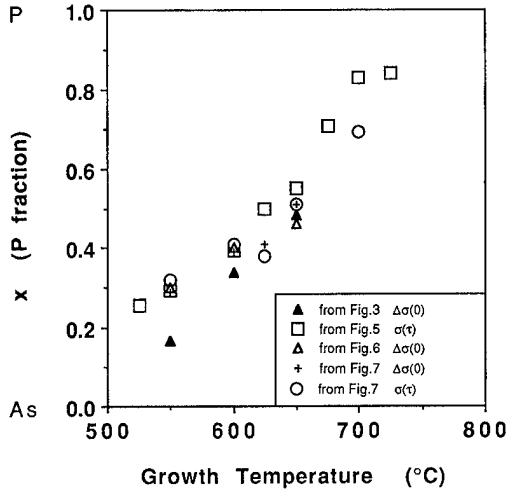


FIG. 8. Temperature dependence of the group-V sublattice composition x using the strain data: (a) from $\Delta\sigma(0)$ of PH_3-formed GaP layers in Fig. 3; (b) from the $\sigma(t)$ slope of GaP-grown layer in Fig. 5; (c) from $\Delta\sigma(0)$ of InP layers in Fig. 6; (d) from $\Delta\sigma(0)$ of $In_{0.5}Ga_{0.5}P$ layers in Fig. 7; (e) from the $\sigma(t)$ slope of $In_{0.5}Ga_{0.5}P$-grown layer in Fig. 7.

The concentrations of residual As in the GaP and $In_{0.5}Ga_{0.5}P$ layer growth were determined from $\Delta\sigma(t)$ using the growth rates and compositions of $GaAs_{1-x}P_x$ and $In_{0.5}Ga_{0.5}As_{1-x}P_x$, respectively. The As concentrations determined from $\Delta\sigma(0)$ and $\Delta\sigma(t)$ for the above cases are shown in Fig. 8 to be remarkably similar with virtually identical temperature dependence. The mechanism for apportioning the As/P seems to be in common, suggesting that arsenic and phosphorus concentrations on the growing surface are in equilibrium.

IV. SUMMARY

The measured strain in multilayer superlattices formed both by inserting thin As compound layers periodically into

MOVPE-grown InP and by inserting thin P compound layers periodically into GaAs has been used to derive information about As–P intermixing at the heterojunction interfaces. Implications of the strain behavior are as follows.

(1) As rapidly displaces P, and P slowly displaces As. This results in the dominance of As replacing P at the heterojunctions.

(2) As atoms displace P atoms and reconstruct as a discrete single layer above 600 °C and as a double layer below 600 °C when InP surfaces are exposed to AsH_3. More surface adsorbed As is available to carry over into InP overgrowth at the lower temperatures.

(3) Interfaces and layers of different P compounds on GaAs show the same temperature-dependent (In)$GaAs_{1-x}P_x$ composition. The As concentration in the solid appears to be in equilibrium with residual As that is present following GaAs growth and P from PH_3.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research, Program Element 62234N, Microelectronics Project RS34M40.

[1]R. Bhat, M. A. Koza, C. E. Zah, C. Caneau, C. C. Chang, S. A. Swartz, A. S. Gozdz, P. S. D. Lin, and A. Yi-Yan, J. Cryst. Growth **107**, 871 (1991).
[2]A. R. Clawson, X. S. Jiang, P. K. L. Yu, C. M. Hanson, and T. T. Vu, J. Electron. Mater. **22**, 423 (1993).
[3]M. S. Hybertson, J. Vac. Sci. Technol. B **8**, 773 (1990).
[4]A. R. Clawson and C. M. Hanson, J. Electron. Mater. **24**, 781 (1995).
[5]A. R. Clawson and C. M. Hanson, J. Electron. Mater. **25**, 739 (1996).

In situ characterization of ZnSe/GaAs(100) interfaces by reflectance difference spectroscopy

T. Yasuda,[a] L. H. Kuo,[b] K. Kimura,[b] S. Miwa,[b] C. G. Jin,[c] K. Tanaka,[a),c] and T. Yao[a),d]
Joint Research Center for Atom Technology (JRCAT), 1-1-4 Higashi, Tsukuba 305, Japan

(Received 23 January 1996; accepted 24 March 1996)

We discuss the structure and formation mechanism of ZnSe/GaAs(100) interfaces as probed by reflectance difference spectroscopy (RDS). First we describe a simple analytic procedure that separates the surface and interface contributions in the RD spectra obtained for a heterostructure. The procedure opens up the new possibilities of RDS as an *in situ* interface probe. We have applied this technique to characterize the ZnSe/GaAs interfaces prepared on various GaAs(100) surfaces. The interface-anisotropy spectra thus obtained and the cross-sectional transmission electron microscopy, show clearly that the chemical composition of the interfacial layer can be changed by controlling the reconstruction and the pregrowth Zn and Se treatments of the initial GaAs surface.
© *1996 American Vacuum Society.*

I. INTRODUCTION

It is now widely recognized that *in situ* diagnostics is one of the key technologies for advanced semiconductor processing. While there are a number of surface probes that can be used in processing environments, there are few spectroscopies or microscopies that are practically applicable for *in situ* interface diagnostics. As such, a versatile *in situ* interface probe is strongly desired to facilitate our understanding and control of the structure and formation process of the interfaces. Reflectance-difference spectroscopy (RDS), which measures in-plane reflectance anisotropy and requires only a compact optical system, has been increasingly used as a surface probe.[1,2] Although RDS is also a potential optical probe for *interface* studies, its application in this regard has been limited to the simple cases where the surface contribution was negligible.[3,4]

In this article we first describe an analytic procedure that allows surface-induced and interface-induced anisotropies to be determined simultaneously when the heterostructure of interest shows both anisotropies.[5] We have applied the *in situ* RDS technique to the ZnSe/GaAs heterostructure. Combined with *ex situ* cross-section transmission electron microscopy (TEM), the structure of the ZnSe/GaAs(100) interfaces prepared on various GaAs surfaces were studied. The results prove clearly that RDS is indeed an excellent tool for interface diagnostics, and illustrate how the ZnSe/GaAs interface can be controlled at atomic level by controlling the reconstruction and pregrowth treatment of the initial GaAs surface.

The fundamental interest in the ZnSe/GaAs consists in its heterovalent nature, i.e., II–VI on III–V. Since the lattice mismatch of this system is as small as 0.27%, heterovalency is expected to play a dominant role in the process of interface formation. Specifically, rearrangements and/or reconstructions possibly take place at this interface to satisfy the electron-counting requirements.[6] On the other hand, control of the ZnSe/GaAs interface is also technologically important. We have found that the microscopic structure and formation chemistry of this interface is strongly related to the formation of the various grown-in defects.[7] Reduction of such defect is strongly desired to improve the lifetime of ZnSe-based light-emitting devices.[8] It was also reported that the valence-band offset depends on the interface structure.[9]

In the following, we first describe the mathematical procedure we have developed to analyze the RDS data from the heterostructures. After discussing the effects of the Zn and Se treatments on variously reconstructed GaAs surfaces, we show the interface-induced anisotropy spectra and the cross-sectional TEM images. On the basis of these results, we discuss the structure and formation process of the ZnSe/GaAs interface, with an emphasis on the effects of the pregrowth treatments.

II. EXPERIMENT

Our homemade RDS system is similar to the one reported by Aspnes *et al.*,[10] except for the use of the calcite Glan–Taylor prism as the incidence polarizer. In RDS, one measures the difference between the normal-incidence reflectances for two orthogonal polarization directions in the surface plane. The results are displayed in terms of $\Delta\tilde{r}/\tilde{r}=\Delta r/r+i\Delta\theta$, where $\tilde{r}=r\exp(i\theta)$ is the complex reflectance. We define $\Delta\tilde{r}$ as $\Delta\tilde{r}=\tilde{r}_{\bar{1}10}-\tilde{r}_{110}$, where the suffices denote the incident polarization vectors. By minimizing the parasitic anisotropies in the optical system, we were able to collect reliable $\Delta\theta$ data, which is a prerequisite to obtain accurate interface-induced anisotropies by the procedure described below. In this article we will show only the $\Delta r/r$ spectra. The spectra cover an energy range from 1.5 to 5.8 eV, with a resolution better than 1 nm.

The heteroepitaxial growth of ZnSe on GaAs(100) was carried out in a dual-chamber molecular beam epitaxy (MBE) system. First, an undoped homoepitaxial GaAs layer

[a]Also with National Institute for Advanced Interdisciplinary Research (NAIR), Tsukuba 305, Japan; Electronic mail: tyasuda@jrcat.or.jp
[b]Also with Angstrom Technology Partnership (ATP), Tsukuba 305, Japan.
[c]Also with University of Tsukuba, Tsukuba 305, Japan.
[d]Also with Institute for Materials Research, Tohoku University, Sendai 980, Japan.

was grown in the III–V chamber to prepare an atomically flat surface. The high quality of the surface was confirmed in a separate experiment using a scanning tunneling microscope connected to the MBE system.[11] The sample was then transferred *in vacuo* to the II–VI MBE chamber for the ZnSe growth and for *in situ* characterization by RDS and reflection high-energy electron diffraction (RHEED). We were able to control the surface reconstruction (or, in other word, the As coverage) by controlling both the flux, or the background pressure, of As and the sample temperature during the sample transfer out of the III–V MBE chamber, and by a proper anneal prior to the growth in the II–VI chamber. During the ramp-up of the sample temperature and the pregrowth treatment, great care was taken to avoid uncontrolled reaction of the GaAs surface with the background Se in the II–VI chamber.

In the present study, the initial GaAs surfaces were prepared by the following procedures: (1) A Zn exposure on an As-terminated $c(4\times4)$ GaAs surface, which is intended to form a Zn–As layer at the interface. (2) A Zn exposure on an As-terminated (2×4) GaAs surface, which is known to be the optimal treatment for reducing the defect density in the final ZnSe/GaAs structure.[7] (3) A Se exposure on a Ga-terminated (4×6) surface followed by 798 K annealing for approximately 2 min. This processing forms a well-ordered Se-terminated (2×1) surface.[12]

These conditions will be called Zn/$c(4\times4)$, Zn/(2×4), and Se/(4×6), respectively. In either case, the exposure was continued until the RD signal at 2.3 eV (for the Zn exposure) or 3.0 eV (for the Se exposure) was saturated, the total exposure time being typically 60 s. For the surface treatment and the ZnSe growth, the beam-equivalent pressures of Zn and Se were $(6.7\pm1.0)\times10^{-5}$ Pa and $(1.1\pm0.1)\times10^{-4}$ Pa, respectively. The sample temperature was 533 K throughout the Zn or Se treatment and the ZnSe growth. During the growth, the incident light was blocked to avoid any possible photoassisted effects.

III. OPTICAL MODEL

In a heterostructure, there are a number of possible origins of optical anisotropies, such as anisotropic structures (reconstructions and steps) of the surface[13] and buried interfaces,[3,4] bulk anisotropies induced by the linear electro-optic effect (LEOE),[14,15] anisotropic relaxation of the mismatched lattice, development of anisotropic morphology,[2,16] and so on. Thus, we need to use a proper model to sort out the origins of the observed anisotropies.

In the following, we describe a simple data-analysis procedure using a three phase model.[5] As shown later, the RD spectra from our ZnSe/GaAs samples are significantly modulated by the interference between the surface and interface reflections. The procedure we describe below is useful for us to eliminate the interference artifacts from the data, and to identify which processes give rise to the observed anisotropy: the reflection at the surface or at the interface, or the propagation through the overlayer. While a better simulation of the observed spectra may be done using an optical model with more constituent layers, the primary advantage of the following approach is its simplicity.

In the three-phase model that consists of a substrate, an overlayer, and an ambient, the overall reflectance $\tilde{r}$ is given as:[17]

$$\tilde{r}=\frac{Z\tilde{r}_{so}+\tilde{r}_{oa}}{1+Z\tilde{r}_{so}\tilde{r}_{oa}}, \tag{1}$$

where $Z=\exp(4\pi i\tilde{n}_o L/\lambda)$, $\tilde{n}_o$ is the complex refractive index of the overlayer, L is the overlayer thickness, λ is the wavelength of light, $\tilde{r}_{oa}$ and $\tilde{r}_{so}$ are the complex reflectances at the overlayer-ambient and substrate-overlayer boundaries, respectively. A total derivative, $d\tilde{r}$, can be expressed using the partial differentials of $\tilde{r}$ with respect to the three variables that are potentially anisotropic: $\tilde{r}_{oa}$, $\tilde{r}_{so}$, and $\tilde{n}_o$. By performing the partial-differential calculations, we obtain the expression for $\Delta\tilde{r}/\tilde{r}$:

$$\frac{\Delta\tilde{r}}{\tilde{r}}=A^{-1}\left(\frac{\Delta\tilde{r}_{oa}}{\tilde{r}_{oa}}B+\frac{\Delta\tilde{r}_{so}}{\tilde{r}_{so}}C+\Delta\tilde{n}_o C\,\frac{4\pi iL}{\lambda}\right), \tag{2}$$

with

$$A=(1+Z\tilde{r}_{so}\tilde{r}_{oa})(Z\tilde{r}_{so}+\tilde{r}_{oa}), \tag{3a}$$

$$B=\tilde{r}_{oa}(1-Z\tilde{r}_{so})(1+Z\tilde{r}_{so}), \tag{3b}$$

$$C=Z\tilde{r}_{so}(1-\tilde{r}_{oa})(1+\tilde{r}_{oa}). \tag{3c}$$

In the following, we call $\Delta\tilde{r}_{oa}/\tilde{r}_{oa}$ and $\Delta\tilde{r}_{so}/\tilde{r}_{so}$ the surface and the interface anisotropies, respectively. If we neglect the $\Delta\tilde{n}_o$ term in Eq. (2), which is an adequate assumption in the ZnSe/GaAs case with $L<100$ nm, the surface and interface anisotropies can be determined using two RD measurements at different thicknesses, L_1 and L_2. Otherwise, if the surface anisotropy is known from separate experiments on homoepitaxial samples or thick-films samples, we can estimate the interface contribution by subtracting the surface contribution from the measured RD signal using Eq. (2). We emphasize that this procedure involves no parameter fitting, making itself suitable for the real-time diagnostics of the surface and interface.

IV. RESULTS AND DISCUSSION

A. Pregrowth surface treatments

Figure 1 compares the RD spectra of the GaAs surfaces prepared by the three different procedures as identified in Sec. II. All the beam exposures and the measurements were carried out at 533 K. The dotted curves in Figs. 1(a), 1(b), and 1(c) were obtained for the GaAs surfaces reconstructed into $c(4\times4)$, (2×4), and (4×6), respectively. Each of these curves is in good agreement with the corresponding spectrum reported by Kamiya *et al.*[18]

When the $c(4\times4)$ surface, which has double-layer As, is exposed to the Zn beam for 60 s, the spectrum showed a significant change below 4 eV, indicating the Zn atoms reacted with the surface As. The negative peak at 2.8 eV for the $c(4\times4)$ surface has been attributed to the top-layer As

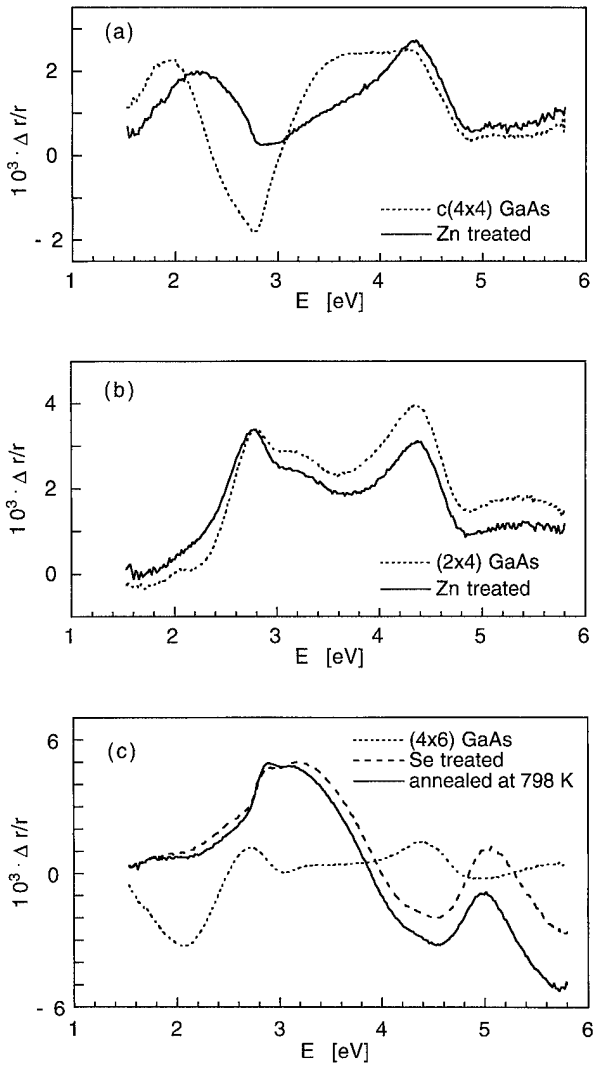


FIG. 1. RD spectra of the GaAs surfaces exposed to the Zn or Se beam. (a) Zn treatment on a $c(4\times4)$ surface; (b) Zn treatment on a (2×4) surface; (c) Se treatment on a (4×6) surface, and the effect of annealing at 798 K.

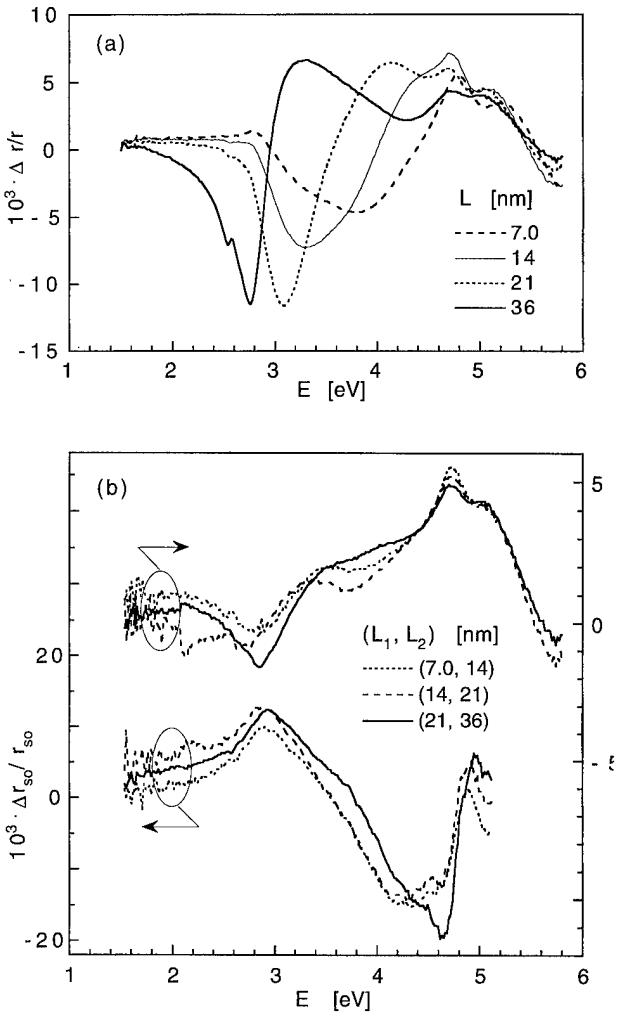


FIG. 2. (a) Evolution of the RD spectra of the ZnSe/GaAs heterostructure with the ZnSe thickness. The data taken under the Se beam are shown. The initial GaAs surface was the Zn/(2×4) surface, see Fig. 1(b). (b) The surface (top) and interface (bottom) contributions in the spectra shown in (a).

dimers.[18] Following this peak assignment, the reduction of this peak upon the Zn treatment suggests disappearance of the top-layer As dimers. This interpretation is consistent with the RHEED result that the Zn exposure transformed the surface reconstruction from $c(4\times4)$ to (1×4).

In contrast to the case of $c(4\times4)$, the (2×4) surface, which is also terminated with As but does not have the double-layer structure, shows a relatively small change upon the Zn treatment, as shown in Fig. 1(b). The RHEED pattern remained (2×4) after the treatment.[11] These results indicate that Zn did not interact with the majority of the surface As dimers.

In the Fig. 1(c) run, the broken curve was measured immediately after the Ga-terminated (4×6) surface was exposed to the Se beam. The sample was then annealed at 798 K, and the solid curve was measured at 533 K. The RHEED pattern of the as-treated surface showed (2×1) reconstruction with weak half-order streaks, and the following anneal at 798 K transformed the surface to a well-ordered (2×1) structure, as reported in the literature.[12] The RD spectrum showed a dramatic change upon the Se exposure, while the annealing at 798 K induced only minor changes, specifically sharpening of the spectral features. Our STM observation on a similarly prepared sample showed that the surface after the annealing is atomically flat with well-ordered dimer rows.[11]

B. Separation of surface and interface contributions

Figure 2(a) shows an example of apparent evolution of the RD spectra with the ZnSe thickness. For this experiment, the initial GaAs surface was the one characterized in Fig. 1(b), namely, the Zn/(2×4) surface. The growth was interrupted at the thicknesses noted in the figure legend, and at each thickness two spectra were measured, first under the Se beam and then under the Zn beam. The surface reconstructions under the Se and the Zn beams were, respectively, (2×1) and $c(2\times2)$, as commonly observed for MBE-grown ZnSe.[19] In Fig. 2, we only show the spectra measured under

the Se beam. The comparison between the spectra under the Se and Zn beams was reported in our previous article.[5] The roughness-induced signal, which usually appears as a broad baseline,[2,16] was negligible in all the experiments discussed in this article.

As clearly seen in Fig. 2(a), the apparent RD spectra show prominent changes with the ZnSe thickness. This is due to the interference between the surface and interface reflections. By applying the analytic procedure described in Sec. III, the interface and surface contributions were separated, and the results are displayed in Fig. 2(b). In this figure, we obtained each spectrum from the corresponding two source spectra in Fig. 2(a), as noted in the legend. The spectral lineshapes are essentially independent of epilayer thickness, which confirms that the interference artifacts have been successfully eliminated. The dielectric functions used in this calculation were the 500 K data for GaAs[20] and the 473 K data for ZnSe.[21] Because the 473 K ZnSe data terminate at 5.1 eV, the calculated spectra also terminate at the same energy. The exceptions are the surface-anisotropy spectra for (12, 21 nm) and (21, 26 nm), where the source spectra had a negligible interface contribution above 4.3 eV due to strong absorption in the ZnSe epilayer. The surface-anisotropy spectra [upper half of Fig. 2(b)] could be fairly simulated by energy scaling of the ZnSe dielectric-functions spectrum, E Im $\epsilon(E)$, suggesting that many-body screening is the dominant mechanism for the anisotropy of the Se-terminated (2×1) ZnSe surface.[5,16,22]

There are two observations that support the interface origin of the spectra shown in the lower half of Fig. 2(b). First, switching the surface termination from Se to Zn hardly changes the spectral lineshape.[5] Second, the spectral lineshape changes as the structure and chemical composition of the initial GaAs surface is changed, as shown later in this article.

Before discussing the interface structure in detail, we address the characteristic notch observed near 2.6 eV in the 36 nm spectrum in Fig. 2(a). A similar feature was reported by Yang *et al.*,[4] and was interpreted as the crossover transition from the valence band of ZnSe to the quantum-well state at the interface. We have found that, for all the samples prepared under different conditions, this feature becomes noticeable only when the ZnSe thickness exceeds 20 nm. This observation can be well understood if this feature is related to the quantum-well state at the interface. For the quantum well to be formed at the interface, the band bending needs to be large enough, which requires the ZnSe thickness to be more than a certain value.

C. ZnSe/GaAs interfaces

In Fig. 3(a), we compare the RD spectra of the ZnSe/GaAs interfaces formed on the three different GaAs surfaces characterized in Fig. 1. The source data from which these spectra were calculated were measured under the Se beam, and in a similar way as in Fig. 2. The ZnSe thicknesses for these data were in the range from 7 to 18 nm.

The broadband centered at 2.9 eV changes its peak intensity with different surface preparations, the Se/(4×6) sample showing the highest peak. The sharp feature above 4.6 eV is tentatively interpreted as due to the bulk anisotropy induced by the linear electro-optic effect at the E_1 critical point of ZnSe.[5] This bulk-induced feature can be separated from the interface-induced signals by modifying both the optical model and the dielectric functions of the constituent materials to take the linear electro-optic effect into account. In the following, we focus the discussion on the interface-induced 2.9 eV band.

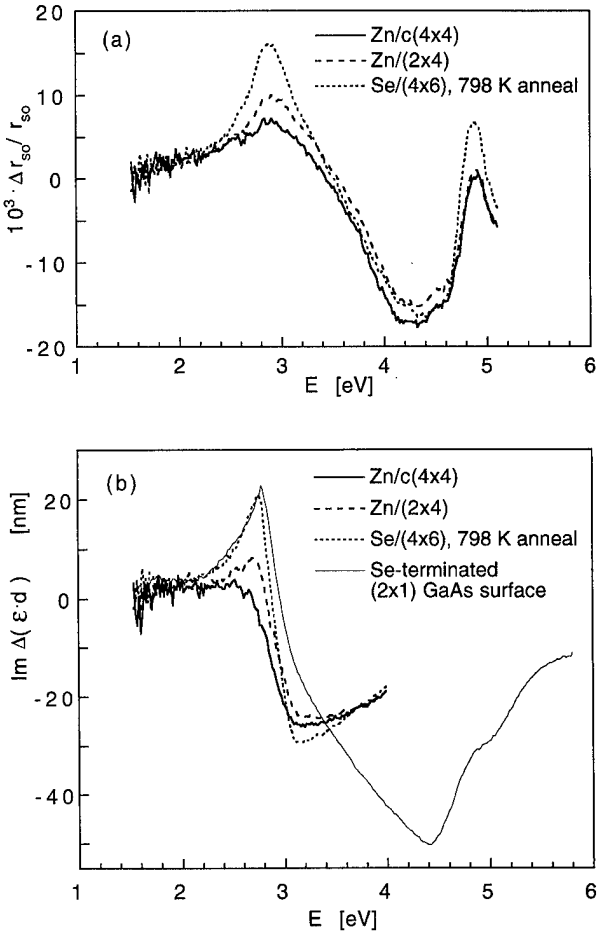


FIG. 3. (a) RD spectra of the ZnSe/GaAs interfaces prepared on the three different GaAs surfaces (see Fig. 1 for the RD spectra of these surfaces). (b) Imaginary part of $\Delta(\epsilon d)$. The curves were obtained by converting the spectra in (a). Also shown is the surface anisotropy of the Se/(4×6) sample obtained by converting the solid curve in Fig. 1(c).

The differences between the three interfaces are more clearly illustrated in Fig. 3(b), where the RD spectra in Fig. 3(a) were converted into anisotropy spectra of the dielectric response of the thin interface layer, $\Delta(\epsilon d)$. Shown in Fig. 3(b) are the imaginary part of $\Delta(\epsilon d)$. This conversion was done by applying the expression for the surface-induced optical anisotropy,[16] replacing the dielectric constant of the ambient with the dielectric functions of the overlayer. In brief, the model assumes an optically anisotropic thin layer at the boundary between two isotropic media, whose thickness d satisfies $d \ll \lambda$. $\Delta(\epsilon d)$ of this thin layer is defined, in the present case, as $\Delta(\epsilon d) = (\epsilon_{\bar{1}10} - \epsilon_{110})\, d$, where $\epsilon_{\bar{1}10}$ and ϵ_{110}

are the complex dielectric functions for polarization vectors parallel to the $[\bar{1}10]$ and [110] directions, respectively.

In Fig. 3(b), the Se/(4×6) sample shows a remarkable sharp peak at 2.8 eV. Also shown in this figure is the $\mathrm{Im}\,\Delta(\epsilon d)$ spectrum of the Se/(4×6) surface prior to the growth, namely, the conversion of the solid curve in Fig. 1(c). The agreement between the two curves below 3 eV is striking. This observation indicates that the anisotropic surface layer formed by the Se treatment is so stable and inert[23] that it maintained its structure even after the ZnSe epilayer was grown on it.

In Fig. 4(c), a cross section TEM image of this interface is shown. We can clearly observe an interfacial layer with zincblende structure. A high-resolution TEM observation showed the thickness of this layer is about 4 monolayers. Li *et al.*[24] reported similar TEM results, and proposed that this layer is a Ga–Se compound with zincblende structure. A known example of such a compound is Ga_2Se_3. Our scattering-factor analysis on 002 and 004 dark-field images also showed that this layer is a vacancy-containing form of gallium selenide, although unambiguous determination of its chemical stoichiometry has not been successful so far.[7]

We note that although bulk Ga_2Se_3 is known to show optical anisotropy,[25] the $\Delta(\epsilon d)$ in Fig. 3 do not exhibit any features at the critical-point energies of Ga_2Se_3. For example, E_0 and E_1 transitions of zincblende Ga_2Se_3 are at 2.1 and 3.9 eV, respectively.[26] We point out that the peak position in Fig. 3(b) is instead close to the E_1 critical point of GaAs at the growth temperature.[20]

The Zn/c(4×4) spectrum shows nearly zero anisotropy below 2.8 eV, which indicates that the formation of the Ga–Se layer was suppressed in this case. The cross-sectional TEM image of this interface, shown in Fig. 4(a), reveals an interfacial layer with an opposite contrast to Fig. 4(c). A high-resolution observation showed that this layer is 2–4 monolayers thick. Our analysis of the 002 and 004 dark-field images suggested that this layer contains Zn and As, which is consistent with the RDS result. The suppression of the Ga–Se formation and formation of the Zn–As layer can be reasonably understood if we recall that the c(4×4) surface has a double layer of As which presumably acts as a barrier for the Ga–Se reaction, and that the Zn atoms react with As on the c(4×4) surface, as shown in Fig. 1(a).

The $\Delta(\epsilon d)$ spectrum of the Zn/(2×4) sample falls between the Se/(4×6) and Zn/c(4×4) spectra. This suggests that the Ga–Se bonds account for, if at all, only a part of the interface bonds. On the other hand, in the TEM picture shown in Fig. 4(b) we observe an interfacial layer of the same contrast as in Fig. 4(a), but with less thickness. Thus, similarly to the case of Fig. 4(b), the interfacial layer seen in Fig. 4(b) is suggested to contain Zn and As. Thus, while the RDS and TEM data suggest a rather complex interface structure for the Zn/(2×4) treatment, a quantitative discussion of the interface structure is difficult at this stage. It is now widely recognized that the Zn/(2×4) treatment is the most effective in reducing the density of the stacking faults originating at the interface.[7] Therefore, the interface on the Zn/(2×4) surface is of technological interest. We are continuing the effort to understand the structure of this interface in more detail, for example by analyzing the chemical stoichiometry of this interface by x-ray photoemission spectroscopy.

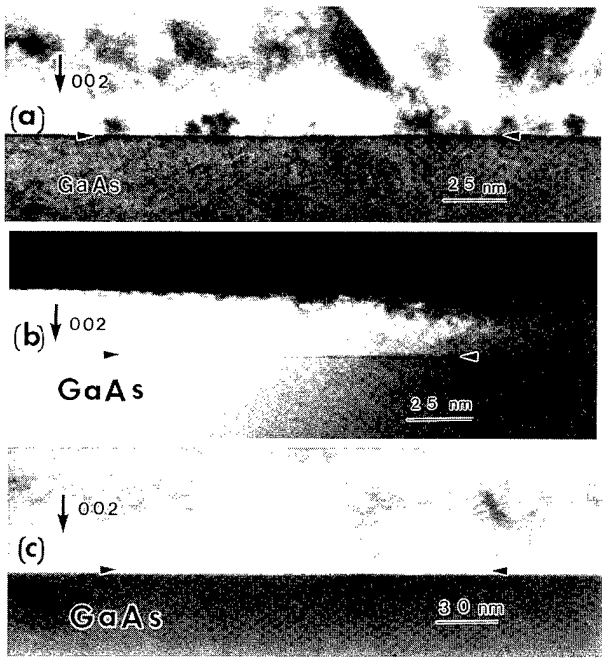


FIG. 4. 002 dark-field TEM images of the ZnSe/GaAs cross sections. The samples were prepared on the three different surfaces: (a) Zn/c(4×4); (b) Zn/(2×4); (c) Se/(4×6). See Fig. 1 for the RD spectra of the initial GaAs surfaces, and Fig. 3 for the corresponding interface-anisotropy spectra.

V. CONCLUSIONS

We have described the analytic procedure that separates the surface and interface contributions in the RD signal from a heterostructure. With this procedure, RDS is now a unique tool that can characterize buried interfaces *in situ* during processing. Taking advantage of the *in situ* characterization by RDS and RHEED, we discussed the structure and formation mechanisms of ZnSe/GaAs(100) interfaces on three different GaAs surfaces. The results are summarized below.

(1) When the (4×6) GaAs surface was exposed to Se, a stable Se-terminated GaAs surface layer was formed. Annealing of the Se-treated GaAs surface induced a well-ordered (2×1) reconstruction. The surface layer thus prepared remained intact, as long as the optical anisotropy is concerned, even after the ZnSe layer was grown on it. The interfacial layer, which is about 4 (ML) thick, is composed of Ga and Se, and also contains vacancies.

(2) When the c(4×4) GaAs surface was exposed to the Zn beam, both RDS and RHEED indicated that the reaction of Zn with the GaAs surface induced disappearance of the top-layer As dimers. The RDS showed that when the ZnSe was grown on this surface the formation of the Ga–Se layer was suppressed, being consistent with the TEM results suggesting that the interfacial layer contains Zn and As.

(3) When the (2×4) GaAs surface was exposed to Zn, both RDS and RHEED showed that Zn did not react with majority of the surface As. Similarly to the case (2) above, the TEM results suggested that the interface layer formed on this surface contains Zn and As. On the other hand, the interface RD spectrum showed a sign of Ga–Se bond formation. These TEM and RDS results can be consistent with each other if the Ga–Se bonds are formed only partially in coverage at this interface.

ACKNOWLEDGMENTS

The authors would like to thank Professor R. W. Collins for allowing us to use their optical data. This study, partly supported by New Energy and Industrial Technology Development Organization (NEDO), was performed at JRCAT under the research agreement between NAIR and ATP.

[1]D. E. Aspnes, Thin Solid Films **233**, 1 (1993).
[2]B. Drevvillon and V. Yakovlev, *Physics of Thin Films*, edited by M. H. Francombe and J. L. Vossen (Academic, San Diego, 1994), Vol. 19, p. 2.
[3]T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, J. Vac. Sci. Technol. A **12**, 1152 (1994).
[4]Z. Yang, G. K. Wong, I. K. Sou, and Y. H. Yeung, Appl. Phys. Lett. **66**, 2235 (1995).
[5]T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, Phys. Rev. Lett. (to be published).
[6]K. Oda and T. Nakayama, Jpn. J. Appl. Phys. **31**, 2359 (1992).
[7]L. H. Kuo, K. Kimura, T. Yasuda, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, Appl. Phys. Lett. **68**, 2413 (1996).
[8]L. H. Kuo, L. Salamanca-Riba, B. J. Wu, G. Hofler, J. M. DePuydt, and H. Cheng, Appl. Phys. Lett. **67**, 3298 (1995).
[9]R. Nicolini, L. Vanzetti, G. Mula, G. Bratina, L. Sorba, A. Franciosi, M. Peressi, S. Baroni, R. Resta, A. Baldereschi, J. E. Angelo, and W. W. Gerberich, Phys. Rev. Lett. **72**, 294 (1994).
[10]D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[11]S. Miwa, K. Kimura, T. Yasuda, L. H. Kuo, C. G. Jin, K. Tanakan and T. Yao, Appl. Surf. Sci. (to be published).
[12]S. Takatani, T. Kikawa, and M. Nakazawa, Phys. Rev. B **45**, 8498 (1992).
[13]For example, T. Yasuda, L. Mantese, U. Rossow, and D. E. Aspnes, Phys. Rev. Lett. **74**, 3431 (1995), and references therein.
[14]V. L. Berkovits and D. Paget, Thin Solid Films **233**, 9 (1993).
[15]H. Tanaka, E. Colas, I. Kamiya, D. E. Aspnes, and R. Bhat, Appl. Phys. Lett. **59**, 3443 (1991).
[16]D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).
[17]D. E. Aspnes, *Optical Properties of Solids: New Developments*, edited by B. O. Seraphin (North-Holland, Amsterdam, 1976), p. 799.
[18]I. Kamiya, D. E. Aspnes, L. T. Florez, and J. P. Harbison, Phys. Rev. B **46**, 15894 (1992).
[19]T. Yao and T. Takeda, Appl. Phys. Lett. **48**, 160 (1986).
[20]P. Lautenschlager, M. Garriga, S. Logothetidis, and M. Cardona, Phys. Rev. B **35**, 9174 (1987).
[21]R. W. Collins and J. Lee (private communication).
[22]W. L. Mochan and R. G. Barrera, Phys. Rev. Lett. **55**, 1192 (1985).
[23]D. Li and M. D. Pashley, J. Vac. Sci. Technol. B **12**, 2547 (1994).
[24]D. Li, M. Gonsalves, N. Otsuka, J. Qiu, M. Kobayashi, and R. L. Gunshor, Appl. Phys. Lett. **57**, 449 (1990).
[25]T. Okamoto, N. Kojima, A. Yamada, M. Konagai, K. Takahashi, Y. Nakamura, and O. Nittono, Jpn. J. Appl. Phys. **31**, L143 (1992).
[26]G. Guizzetti, F. Meloni, and Baldereschi, J. Phys. Soc. Jpn. **49**, 93 (1980).

Optical monitoring of the development of InAs quantum dots on GaAs(001) by reflectance anisotropy spectroscopy

E. Steimetz, J.-T. Zettler, and W. Richter
Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstr. 36, D-10623 Berlin, Germany

D. I. Westwood, D. A. Woolf, and Z. Sobiesierski
Department of Physics and Astronomy, University of Wales, Cardiff, P.O. Box 913, Cardiff CF2 3YB, United Kingdom

(Received 24 January 1996; accepted 22 April 1996)

Reflectance anisotropy spectroscopy (RAS) in combination with reflection high-energy electron diffraction (RHEED) was used to study *in situ* the initial steps of molecular beam epitaxial growth of InAs on GaAs(001). Due to the large lattice mismatch InAs is known to grow in Stranski–Krastanov mode leading to the formation of quantum dots after the transition from two- to three-dimensional growth mode. In this article the precise determination of the growth mode transition and the subsequent development of the islands have been of particular interest. During the growth of the two-dimensional InAs layer, the RHEED-pattern changed from the $c(4\times4)$ of the clean GaAs to a (1×3) surface reconstruction. Accordingly, the RAS-spectra, taken every 0.2 ML, indicate changes of the As-dimer configuration. At 1.8 ML (spotty RHEED-pattern) a saturation of the intensity of the dimer related RAS-signal around 2.6 eV was found. The relaxation of the InAs layer and the formation of the quantum dots was followed by time-resolved RAS at 2.6 and 4 eV. It is shown here, that the time constant of this process, the thickness of the InAs wetting layer and the equilibrium morphology of the islands are strongly temperature dependent. The remaining equilibrium InAs wetting layer thickness at the surface was estimated to be about 1 ML (0.8 ML at 625 K and 1.2 ML at 725 K). © *1996 American Vacuum Society.*

I. INTRODUCTION

The preparation of InAs islands, by relaxation of the first few monolayers (MLs) grown on GaAs substrates, show great promise to be used for growth of high quality quantum dots in optoelectronic devices. Therefore, there is presently a great interest in understanding the mechanism of their formation as part of an attempt to find ways to improve their uniformity, the main stumbling block to their application in practical devices.[1]

Targeting a one step deposition of quantum dots without any pre-patterning, many investigations have been performed in nearly all highly lattice mismatched semiconductor systems such as InAs/GaAs[2] or InP/GaInP.[3] The successful deposition of coherent InAs islands embedded in AlGaAs on high index substrates, especially on the (311B)-GaAs surface[4] proves the high potential of this growth technique. However the reproducibility of the dot formation which is essential for the growth of dot-stacks, still causes many problems. A search for greater control of the whole process, including the growth mode transition and the island development seems to be necessary for establishing this process as a reliable technique. The first step toward control is to monitor in real-time the growth mode transition. Reflection high-energy electron diffraction (RHEED) is capable to monitor this by the change from streaky diffraction patterns for two-dimensional to spotty ones for three-dimensional growth.[5] However, the subsequent development of the three-dimensional islands cannot be studied by this technique. In addition, RHEED is only applicable in ultra high vacuum systems such as molecular beam epitaxy (MBE) and metal organic molecular beam epitaxy (MOMBE), but not in gas phase environments as used in metal organic vapor phase epitaxy (MOVPE). Therefore, reflectance anisotropy spectroscopy (RAS) has already shown its capability for *in situ* analysis. The growth rate and the composition of ternary compounds, for example, were determined by RAS in both MOVPE and MBE systems.[6] In a former MBE-study[7] RAS measurements of InAs grown on GaAs(001) up to a total coverage of 100 ML were reported. The distinctly different spectra for clean and slightly InAs covered GaAs surfaces yielded evidence for different surface reconstructions. Growth beyond a total coverage of 4 ML caused large anisotropies which were explained by roughness effects due to three-dimensional growth. This hypothesis was confirmed by effective medium calculations.

Building on this study, simultaneously performed time resolved RAS and spectroscopic ellipsometry (SE) measurements at certain photon energies were successfully used in MOVPE to monitor the Stranski–Krastanov growth mode transition.[8] Complete spectra taken over a range from 1.5 to 5.5 eV revealed characteristic changes due to reconstruction and growth mode transition. For continuous deposition of 4 ML InAs on GaAs(001) a large increase in anisotropy with increasing deposition temperature appeared. It was explained by the formation of large clusters elongated in [110] direction. This result, derived from effective medium theories was proven by the observation of elongated islands in high resolution scanning electron microscopy. Thus RAS might also allow estimates on island size.

The intention of the investigations presented here was to

study in detail the relaxation of the InAs layer and to monitor the island formation until the surface equilibrium is reached. This was done by performing simultaneous RAS and RHEED measurements. In order to gain a better understanding of the processes governing the dot formation different deposition parameters were chosen.

II. EXPERIMENT

The experiments were carried out in a MBE-system[7] allowing simultaneous RHEED and RAS measurements. Details concerning the flux measurement have been described elsewhere.[9] The RAS[10] was coupled to the MBE chamber via a low strain silica window. The temperature calibration was made using the (2×4) to (4×2) reconstruction change in the RHEED pattern. Before InAs deposition a 1200 nm thick GaAs buffer layer was grown at 820 K (growth rate $r=1\mu$/h).

Three different types of experiments were performed:

(i) Stepwise deposition: InAs was deposited stepwise (0.2 ML steps) at different deposition temperatures (625 K, 725 K, and 775 K) and RAS spectra were taken in a range from 1.5 to 5.5 eV immediately after every deposition. These experiments were carried out up to an effective coverage of 4 ML with a growth rate of 0.05 ML/s. The RHEED pattern were observed simultaneously.

(ii) Continuous deposition monitored in real time: during the continuous deposition of 2 ML and 3 ML of InAs time resolved RAS measurements were performed at photon energies of 2.6 eV and 4.0 eV. The measurements were extended after growth and were finished in order to monitor the formation of QDs and the approach of the system to equilibrium state in real time. The growth rate for this experiment was 0.025 ML/s and it was performed at different growth temperatures (625 K and 725 K).

(iii) Postgrowth evolution of the surface towards equilibrium: For 2 ML InAs deposited at different growth temperatures the post-growth development of the surface was also investigated by taking RAS spectra until no changes in the spectra could be distinguished.

The spectra obtained for the surface equilibrium were compared to those achieved for stepwise deposition in order to determine the thickness of the remaining two-dimensional InAs wetting layer.

III. RESULTS AND DISCUSSION

A. RAS spectra for stepwise deposition of InAs on GaAs

Characteristic RAS spectra gained for stepwise deposition at 625 K up to an overall coverage of 3 ML are given in Fig. 1(a). Starting from the well known $c(4\times4)$ As-rich GaAs reconstruction already after the deposition of 0.2 ML significant changes in the As-dimer related part of the spectrum, i.e., a minimum at 2.6 eV occurred. The RHEED pattern changed to an asymmetric (1×3) surface reconstruction. For additional InAs deposition the RHEED-pattern became clearer and more symmetric. In the RAS spectra a change in

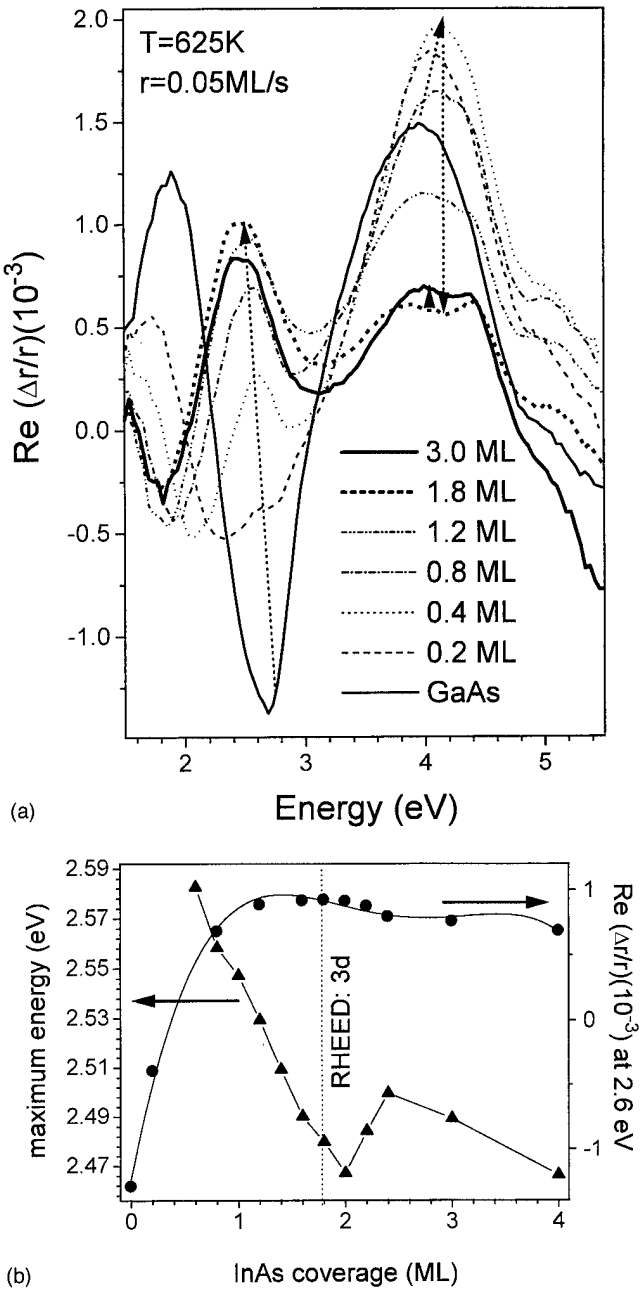


FIG. 1. RAS data taken immediately after the deposition of well defined amounts of InAs on GaAs(001). (a) RAS spectra: the strong signal change at 2.6 eV reflects the reconstruction change from an arsenic rich $c(4\times4)$-GaAs (001) surface to a (1×3)-reconstructed one. The signal change (increase) at photon energies around 4 eV for coverages beyond the critical thickness (1.8 ML) originates from the formation of anisotropic dots or clusters of InAs. (b) Spectral position of the As-dimer related maximum (left axis) and the RAS signal at 2.6 eV (right axis) [data gained from (a)].

the sign around 2.6 eV takes place leading now to a maximum at this energy. The occurrence of the change in the sign indicates the rotation of the As-dimers from the [110] to the [$\bar{1}$10]-direction. The intensity of the As-dimer related maximum in dependence of InAs coverage is shown in Fig. 1(b).

Besides the changes in signal intensity also a shift in the

energy position of the maximum appears. Figure 1(b) reveals a decrease of its spectral position with increasing InAs coverage. The lowest energy position is found for a total coverage of 2 ML. This corresponds to a movement of the As-dimer related maximum from the typical GaAs related position at 2.6 eV to a more InAs like position. RAS studies on InAs bulk material have shown a maximum at about 2.3 eV.[6] For InAs coverages above 2.2 ML the spectral position of the maximum shifted back to energies found before for lower total coverages. At a InAs coverage of 1.8 ML the appearance of spotty RHEED pattern indicates the change from two dimensional layer-by-layer to three-dimensional island growth.

Knowing this, the changes in the RAS spectra can be interpreted as follows: During two-dimensional growth the intensity of the As-dimer related signal increases because of reconstruction changes, but stays constant at the growth mode transition, indicating a constant As-dimer coverage and orientation at the surface. When the island formation starts (2 ML to 2.2 ML), and the nucleation of three-dimensional islands takes place (as previously reported in Ref. 5), material from the two-dimensional wetting layer is transfered to the islands resulting in a decrease of the maximum signal intensity and a shift of the maximum in energy position, that should be related to the layer thickness of the two-dimensional InAs-layer below the dots. The thicker the InAs layer the closer is the spectral position of the signal to that of InAs. The opposite shift of the maximum position for coverages above the critical thickness is characteristic for InAs coverages below the critical layer thickness and indicates a mass transfer out of the two-dimensional layer towards the InAs islands. This effect can be used to determine the residual wetting layer thickness.

Further InAs deposition lead to an apparent downshift of the RAS peak-position and the RAS-amplitude. The latter can be understood by comparing the whole spectra for coverages beyond the critical thickness (3 ML) to those of the smooth two-dimensional case (1.8 ML) (Fig. 1). A decreased RAS-signal is evident for the whole spectral range (except of the 4 eV peak), a phenomenon that can be attributed to an increased surface roughness, caused by the islands. A correlation between the size and morphology of the anisotropic islands and the RAS spectra was investigated in Ref. 8. Performing effective medium calculations (similar to those described in Ref. 8), we found that the increased RAS-signal at 4 eV may be caused by an increased island density, while the decrease in the RAS-signal may be induced by an increased island anisotropy (elongation in $[\bar{1}10]$).

Performing the same experiment at a substrate temperature of 725 K resulted in a similar behavior as seen at 625 K. The critical coverage was found to be approximately the same as for the lower temperature (1.8 ML). Taking an RAS spectrum for 4 ML deposited at 725 K half an hour after the last deposition, we observed a significant increase of the anisotropy at high photon energies (above 4.0 eV). This was a first hint that the island formation even at this relatively high temperature is a long lasting process. The equilibrium properties for different temperatures are studied in detail later (Figs. 3(a) and 3(b)).

For depositions at even higher temperatures (T =775 K) the changes from one 0.2 ML step to the next one were found to be much smaller than for lower temperatures and both a saturation of the dimer related anisotropy and a spotty RHEED-pattern (indicating three-dimer growth) occurred at significantly higher nominal coverages (about 3 ML). Thus it is concluded that at this temperature reevaporation of indium from the surface already occurs and that the effective growth rate is lower as expected.

B. Time resolved measurements during continuous InAs deposition

In order to follow the dynamic changes of reconstruction and morphology on line, time resolved RAS measurements at the characteristic energies 2.6 eV (the As-dimer related energy) and 4.0 eV (sensitive to roughness) were performed for continuous deposition of 2 or 3 ML InAs at two different temperatures (625 K and 725 K). After deposition the RAS transients were continued for additional 270 s.

In Fig. 2(a) the dimer related change of the RAS signal at 2.6 eV from a negative to a positive one, as already found for the stepwise deposition (Fig. 1(b)) is shown. Its sudden increase when InAs is deposited proves the immediate change of the surface reconstruction. For an InAs coverage around the critical thickness the RAS signal at 2.6 eV saturates indicating that the As-dimer coverage now stays constant. When the growth mode changes, because it becomes energetically favorable for the surface to form islands, the InAs layer thickness is reduced, as measured by a slight decrease of the RAS signal at 2.6 eV. This decrease is found for both coverages, at both temperatures, but it is most evident for deposition of 3 ML InAs at 725 K. The preceding increase of the signal (only observed for 3 ML at 725 K (bottom of Fig. 2(a)) that already starts during growth and continued when the deposition was stopped, can be interpreted by the superposition of a roughness related contribution since after growth mode transition island formation leads to an increasing surface roughness. This behavior was not found in the spectra taken for stepwise deposition, possibly because growth interruptions allowed a smoothening of the surface between two deposition steps. However, the following decrease in anisotropy, found in all experiments, should not only be related to the smoothening of the surface, but also to a transport of material to the islands (i.e., a decrease of the As-dimer coverage of the surface) as already discussed in the previous section.

RAS transients at 4 eV are given in (Fig. 2(b)). They have been taken to gain additional information about surface roughness introduced by island formation. At high photon energies, because of the shorter wavelengths light scattering due to surface roughness is stronger than at 2.6 eV thus causing larger changes in the RAS signal.

For all experiments (2, 3 ML and 625, 725 K) during two-dimensional growth (i.e., below 1.8 ML) the evolution

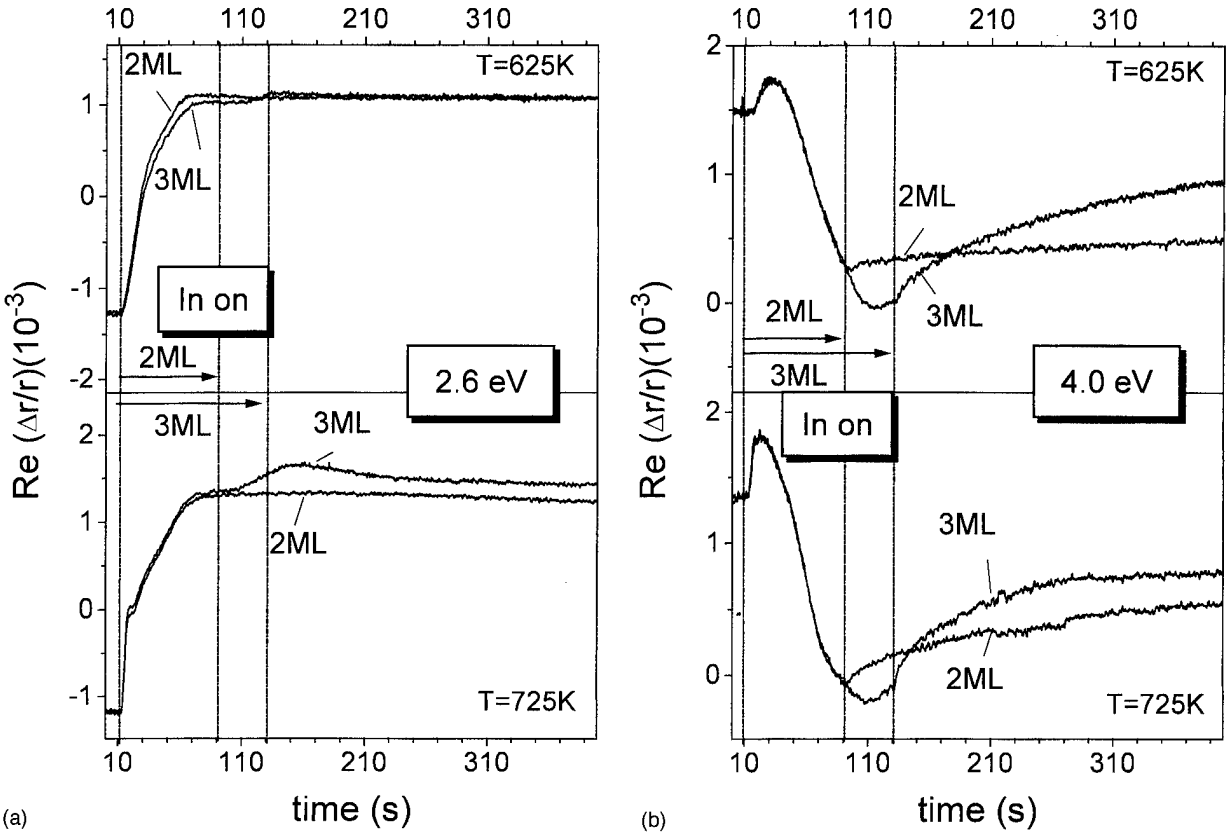


FIG. 2. RAS transients during InAs deposition for two total coverages (2 ML, 3 ML) and two temperatures (625 K, 725 K) measured at a photon energy of (a) 2.6 eV and (b) 4.0 eV. The horizontal arrows starting at t=10 s indicate the deposition time. The initial changes (up to 110 s) describe the surface modification from $c(4\times4)$ GaAs to a (1×3) InAs. The further signal increase in (b) is caused by surface roughening.

of the transients is governed by changes of the reconstruction as shown in Fig. 1(a).

After that, for deposition of 2 ML InAs at 625 K only a weak signal change occurs, even when RHEED already showed three-dimensional growth. Thus surface roughness should be low, i.e., only small islands had formed on the surface. However, the slight increase that occurred after deposition was halted, shows that the equilibrium surface state is not yet reached.

At the higher deposition temperature (bottom of Fig. 2(b)) the roughness related increase is enhanced and changes of the surface toward equilibrium seem to be much faster indicating that kinetic processes are responsible for the island formation.

For the 3 ML deposition of InAs, the response of the RAS to three-dimensional islands becomes evident by a sign change in the RAS signal that occurs already during the deposition. Stopping the deposition additionally increases this effect. Again for the higher deposition temperature the evolution of the surface towards the surface equilibrium is accelerated.

Depositing 3 ML InAs at 725 K results in a larger anisotropy as compared to that occurring for low temperature deposition, and can be understood from the surface diffusion properties. Increasing the temperature means increasing the diffusion length. Enhanced diffusion leads to an enlargement of the individual islands and hereby to a lowering of the island density, because big islands (beyond the critical island size[11] found in several TEM investigations) grow on the extense of smaller ones (Oswald ripening). Thus taking RAS transients at 4 eV we can decide *in situ* whether large islands have formed or not by looking at the anisotropy increase after the deposition is halted.

C. Post growth evolution of the islands

Since the time resolved measurements demonstrated, that the evolution of the islands continues after the deposition is halted, additional investigations of the post growth behavior were performed.

Therefore InAs was deposited to an amount slightly above the critical layer thickness and RAS-spectra were taken until no significant changes were observed anymore and the equilibrium of the islanded surface was reached. These spectra are given in Fig. 3(a). A significant drop of the signal at 2.6 eV (dimer related) and a shift of the maximum position, both toward values as found for lower InAs coverages (see Fig. 1) were measured. This behavior again gives evidence that the two-dimensional InAs-layer is reduced in favor of the islands. For the lowest deposition temperature under investigation (625 K) anisotropy changes originating from this rearrangement of the surface were found to take up to 70 min

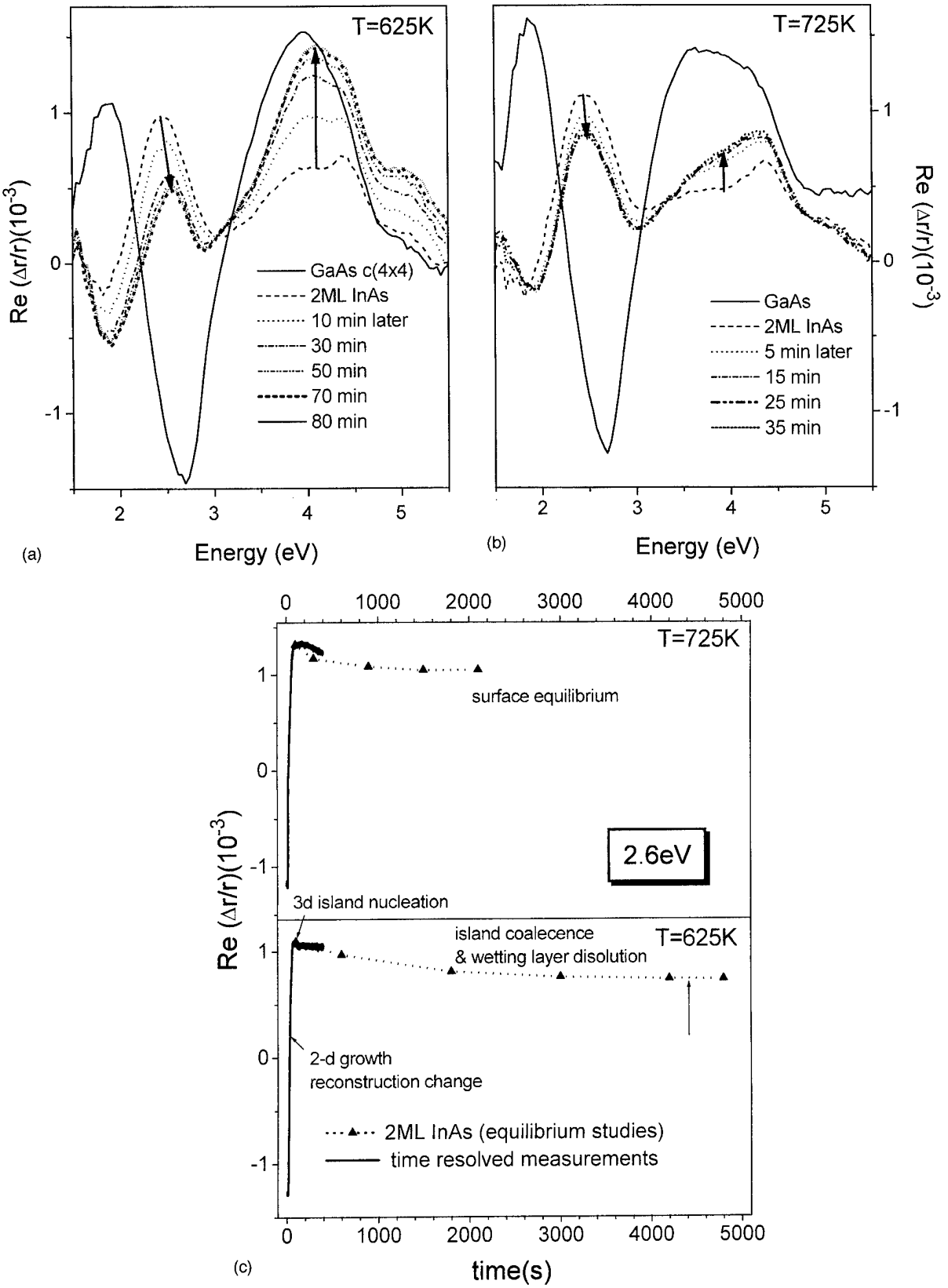


FIG. 3. Development of the RAS signal for extended times (up to 80 min) after deposition of 2 ML InAs on GaAs (001) for two temperatures (625 K, 725 K). (a) The RAS spectra indicate, especially at the lower temperature that still changes, e.g., surface modifications occur for times compared long with the deposition time. (b) Time resolved RAS signal at 2.6 eV taken from (a).

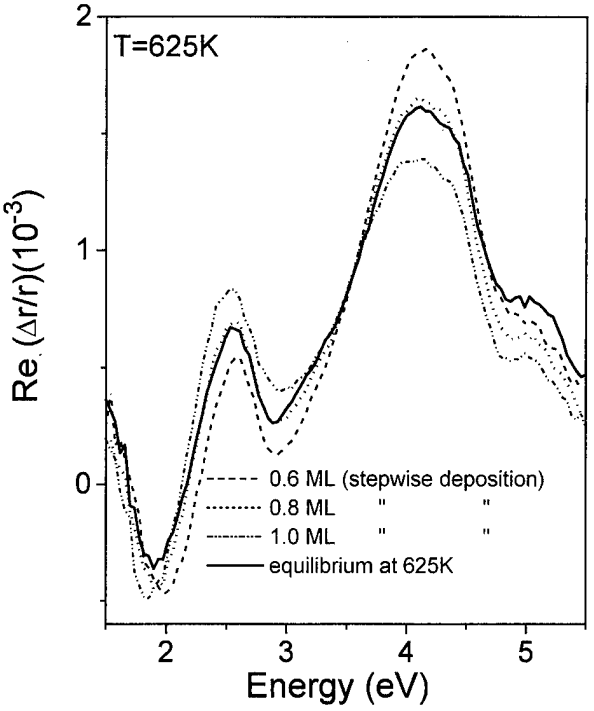


FIG. 4. Comparison of a RAS spectrum obtained 70 min after a one step deposition of 2 ML InAs/GaAs ("equilibrium") with different submonolayer deposition spectra (0.6, 0.8, 1.0 ML) obtained under the same conditions: The "equilibrium"-wetting layer thickness was determined to be 0.8 ML at 625 K.

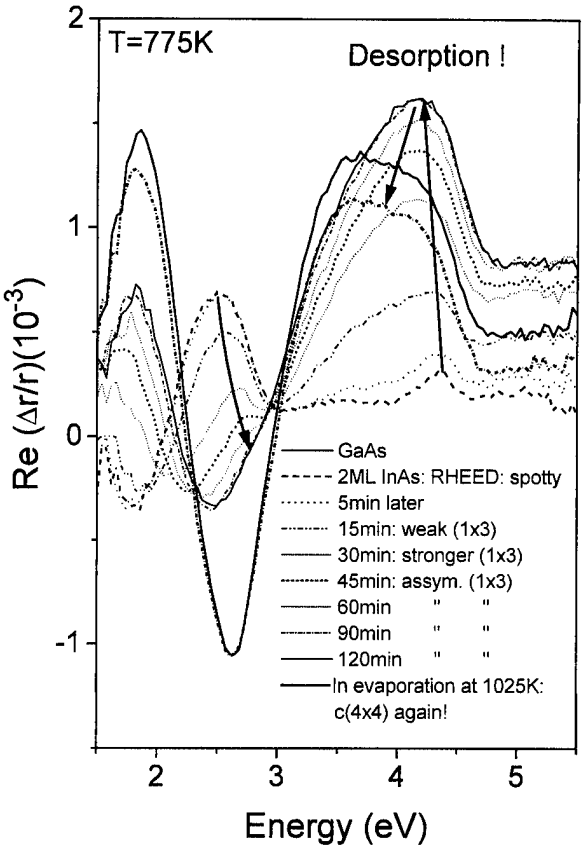


FIG. 5. "Equilibrium" studies at elevated temperature (775 K) for 2 ML InAs deposited on GaAs (001). Due to In desorption at this temperature the spectra return to spectra measured for a total coverage of only 0.2 ML. The remaining InAs is found to be very stable: only annealing at a even higher temperature (1025 K) removes the last InAs and delivers the starting surface, the $c(4\times4)$-GaAs, again.

after deposition. The process is enhanced for higher deposition temperatures and the equilibrium state is thus reached earlier (35 min at 675 K and 25 min at 725 K).

Figure 3(b) summarizes the whole island formation process by combining the time resolved measurements (Fig. 2(a)) with the 2.6 eV RAS response gained from the equilibrium studies (Fig. 3(a)) at defined times after deposition of 2 ML. The whole process is shown, from two-dimensional wetting layer growth over island nucleation and coalescence until surface equilibrium. This shows the good agreement of time resolved and spectral measurements.

An estimate of the final wetting layer thickness is obtained by a comparison of the equilibrium spectra (Fig. 3(a)) with the spectra for stepwise depositions (Fig. 1(a)) at the same temperatures. The comparison is given in (Fig. 4).

For low temperature deposition (625 K) the best agreement was found for a residual InAs layer thickness of 0.8 ML. For deposition at 725 K 1.2 ML InAs remained as two-dimensional wetting-layer. This difference in layer thickness may be caused by different growth rates for the submonolayer depositions (due to desorption of Indium at 725 K the growth rate may be lower), but could also be caused by different equilibrium surface states for different growth conditions (including different island shapes and sizes for different temperatures).

This behavior is very different from MOVPE-investigations where the final quantum-dot state was reached soon after the deposition was halted. This difference may be due to the higher surface mobility in MOVPE[12] leading to a faster transport of material to the islands. Besides chemical radicals at the surface may cause a faster reduction of the wetting layer.

Performing the same studies at a even higher temperature (Fig. 5) we found, that the equilibrium was reached only hours after growth. Here two effects are superimposed: desorption of In from the surface and the islands and additionally transformation of material from the wetting layer and the smaller QDs to the larger islands. However, the final state shows clearly that reevaporation is the dominating process at this temperature. Comparing the spectra taken two hours after deposition with those taken for submonolayer deposition at this temperature we found excellent agreement with the 0.2 ML-spectra. The RHEED pattern also changes back from a spotty to a streaky, first weak, later more pronounced asymmetric 1×3-reconstruction. This indicates that all islands have gone and only small, two-dimensional InAs regions are left on the surface. Continuation of this experiment for about 2 h gave no further changes, showing that the last 0.2 ML are very stable at this temperature. Annealing the sample to 790 K and cooling it down to 775 K again, yields a clear

$c(4\times4)$ reconstruction in both, RAS and RHEED, that was comparable to that of the clean GaAs surface. Thus annealing to 790 K leads to the desorption of the complete InAs layer.

IV. SUMMARY

In this work RAS and RHEED investigations for different coverages of InAs on GaAs(001) were performed. The growth mode transition was observed and the main interest was on the development of the islanded surface towards equilibrium. For this purpose time resolved measurements during growth and afterward were correlated to the reconstructions observed in RHEED and detailed information presenting the initial steps of dot evolution were performed. For two-dimentional wetting layer growth the intial change from the GaAs $c(4\times4)$-reconstruction to a (1×3)-reconstructed InAs surface was confirmed. Analyzing the shifts of the As-dimer related RAS maximum towards lower energies and the corresponding RHEED pattern, the critical thickness of the InAs layer was determined precisely to be 1.8 ML for 625 K deposition. Time resolved RAS studies with the intention to monitor not only the transition from two-dimensional to three-dimensional growth but additionally to study in detail the further development of the islands were performed. With increasing deposition beyond the critical thickness and for higher growth temperatures (775 K) an increase of the anisotropy due to roughening of the surface by large islands was found. A strong temperature dependence of the equilibrium surface state was found. For a coverage slightly above the critical thickness the equilibrium for the 625 K deposition was reached 70 min after deposition and at 725K it took only 25 min. At 775 K reevaporation of In from the surface was observed and a residual coverage of only 0.2 ML after about 2 h was approximated.

ACKNOWLEDGMENTS

This work was supported by the DAAD, the U.K. EPSRC Low Dimensional Structures and Device Initiative and the European H.C.M. "GOODS" (Growth of One Dimensional Structures) program. Helpful discussions with Ursula Resch-Esser are gratefully acknowledged.

[1]P. Petroff and S. DenBaars, Superlatt. Microstruct. **15**, 15 (1994).
[2]D. Leonard, S. Fafard, K. Pond, Y. Zhang, J. Merz, and P. Petroff, J. Vac. Sci. Technol. B **12**, 2516 (1994).
[3]N. Carlsson, W. Seifert, A. Petersson, P. Castrillo, M. Pistol, and L. Samuelson, Appl. Phys. Lett. **65**, 3093 (1994).
[4]R. Noetzel, J. Temmyo, and T. Tamamura, Jpn. J. Appl. Phys. **33** (1994).
[5]J. Moison, F. Houzay, F. Barthe, L. Leprince, E. Andre, and O. Vatel, Appl. Phys. Lett. **64**, 196 (1994).
[6]W. Richter and J.-T. Zettler, Appl. Surf. Sci. (to be published).
[7]S. M. Scholz, A. B. Müller, W. Richter, D. R. T. Zahn, D. I. Westwood, D. A. Woolf, and R. H. Williams, J. Vac. Sci. Technol. B **10**, 1710 (1992).
[8]E. Steimetz, J.-T. Zettler, F. Schienle, and W. Richter, Appl. Surf. Sci. (to be published).
[9]D. I. Westwood, D. A. Woolf, and R. H. Williams, J. Cryst. Growth **98**, 782 (1989).
[10]D. E. Aspnes, J. P. Harbison, A. A. Studna, L. T. Florez, and M. K. Kelly, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[11]K. Tillmann, D. Gerthsen, P. Pfundstein, A. Forster, and K. Urban, J. Appl. Phys. **78**, 3824 (1995).
[12]K. Ploska, M. Pristovsek, W. Richter, J. Jönsson, I. Kamiya, and J.-T. Zettler, Phys. Status Solidi A **152**, 49 (1995).

Reflectance anisotropy spectroscopy study of GaAs overlayer growth on submonolayer coverages of Si on the GaAs(001)-$c(4\times4)$ surface

Z. Sobiesierski,[a] D. I. Westwood, and D. A. Woolf
Department of Physics, University of Wales College of Cardiff, Cardiff CF2 3YB, Wales

(Received 22 January 1996; accepted 24 March 1996)

Reflectance anisotropy spectroscopy (RAS) has been employed to characterize the overlayer growth of GaAs onto sub to one monolayer coverages of Si δ layers deposited on the GaAs(001)-$c(4\times4)$ surface. The low growth temperature (400 °C), required to avoid spreading of the dopant away from the δ plane, has meant that the study of a RAS feature related to the linear electro-optic (LEO) effect is complicated by disordering at the GaAs surface. This disordering is induced not only by the growth temperature, but also by the presence of the Si δ layer itself. Variable thickness studies indicate that the LEO-induced signal is dependent on the field profile in the surface layer. It has been observed that the intensity of the LEO feature, as a function of Si coverage, reaches a maximum at ~0.01 ML ($\sim6.4\times10^{12}$ atoms cm^{-2}) in agreement with previous studies of the site occupancy of Si δ layers. © *1996 American Vacuum Society.*

I. INTRODUCTION

Considerable effort has been employed, over many years now, to characterize fully the nature of the reconstructions obtained from GaAs surfaces prepared under different growth conditions. Such studies have formed a key starting point when trying to predict both the structural and electronic characteristics of interfaces formed by GaAs with other materials. More recently, similar attention has become focused on the way in which these reconstructions become altered after planes of dopant atoms, such as Si or Be, are deposited on the GaAs surface. The majority of these studies have employed either reflection high-energy electron diffraction (RHEED), or a combination of RHEED and scanning tunneling microscopy (STM) as tools to probe the atomic bonding that gives rise to the observed surface reconstructions.[1–3] While these investigations have explored a wide manifold of possible reconstructions, using a broad range of growth conditions, only a small number of studies have been aimed specifically at probably the most technologically relevant Si/GaAs interfaces, i.e., those prepared at low temperature (400 °C and below).[4–6] At these growth temperatures, the spread of Si atoms away from the δ plane remains small enough for practical δ-doping applications.

We have recently used a combination of reflectance anisotropy spectroscopy (RAS) and RHEED measurements to study submonolayer coverages of Si on the GaAs(001)-$c(4\times4)$ surface, at a substrate temperature of 400 °C.[7,8] One of the observations to come out of this and other RAS studies[9,10] of the GaAs(001) surface is that there are characteristic spectral features that vary in a systematic manner with the reconstruction of the GaAs surface. The most well known of these features occurs at an energy of 2.65–2.7 eV and changes in both shape and sign from a pronounced minimum for the $c(4\times4)$ reconstruction, to a maximum for the (2×4) reconstructed GaAs(001) surface.[7,10] This change reflects an alteration in surface bond orientation from along [110] to $[\bar{1}10]$, and makes RAS measurements extremely sensitive to the degree of dopant coverage. In our case, we have been able to discern overlayer coverages of as little as 0.005 ML of either Si or Be.[8] In addition, there are a number of characteristic energies, linked to specific surface reconstructions, that act as nodal (isoplethtic) points through which the RAS spectra evolve with increasing Si coverage.

In this article, RAS data are presented for GaAs overlayer growth on top of different submonolayer coverages of Si on the GaAs(001)-$c(4\times4)$ surface. This is a natural progression of our earlier Si/GaAs work and is of direct relevance to the inclusion of δ layers within a III–V semiconductor device. The interpretation of the RAS spectra becomes somewhat more complicated, in this case, since the measured anisotropy reflects not only changes in the surface order but is also influenced by the surface field and the thickness of the overlayer. The dependence on surface field has been studied previously for uniformly doped GaAs[11,12] and has been shown to give rise to a feature in the RAS spectrum at around 3 eV. This feature has been attributed, in turn, to the linear electro-optic (LEO) effect, associated with the E_1 and $E_1+\Delta_1$ interband transitions. The strength of the electric field in the near-surface region can be altered either by changing the concentration of Si atoms in the δ layer, or by varying the position of the δ layer with respect to the surface. Hence, it has been our aim in this study to probe the contribution of the LEO effect to the RAS signal and to increase our understanding of the surface ordering that takes place when GaAs is deposited on top of Si/GaAs. Once again, all growth has been restricted to 400 °C in order to limit the diffusion of Si atoms away from the δ plane.

II. EXPERIMENT

All growth took place on GaAs(001) on-axis substrates within a VG Semicon V80H MBE reactor, fitted with a VG LEG 110 RHEED system. Experimental procedures such as substrate preparation, growth temperature, and MBE flux calibration details can be found elsewhere.[13] The in-house

[a]Electronic mail: spxzs1@cf.ac.uk

constructed RAS system was positioned at a strain-free pyrometer viewport situated on the MBE reactor growth chamber. A description of the integrated MBE/RAS system can also be found elsewhere, the design being based upon that due to Scholz *et al.* and Aspnes *et al.*[14,15] The RAS system, which has a working spectral range from 1.5 to 5.5 eV, measures the difference (Δr) between the anisotropic complex reflectance (r) along the [$\bar{1}$10] and [110] optical eigenaxes within the (001) surface crystallographic plane, normalized to the mean reflectance ($\bar{r}$):

$$\frac{\Delta r}{\bar{r}} = 2\frac{r[\bar{1}10]-r[110]}{r[\bar{1}10]+r[110]}.$$

Only the real part of the RAS signal was investigated, since even small residual strain effects, associated with the pyrometer viewport, significantly affect the imaginary component of the RAS signature.[14]

Following the complete thermal desorption of the GaAs surface oxides at ~600 °C under an As_4 flux of $F_{As}\sim 5\times10^{14}$ molecules $cm^{-2}\,s^{-1}$, a sharp, clear GaAs(001)-(2×4)-β reconstruction was observed by RHEED. A 1 μm, undoped GaAs buffer layer was then grown at a rate of 1 $\mu m\,h^{-1}$, with an As_4/Ga flux ratio of ~0.8, at a temperature of 580 °C. The sample was cooled to 400 °C for both the deposition of Si and subsequent overgrowth with GaAs. The Si was deposited from a standard 40 cm^3 VG Mk.II effusion cell, which had been rigorously calibrated from a large series of Hall effect and electrochemical $C-V$ profiling measurements. The effusion cell temperature was set to $T_{Si}=1120$ °C, which yielded a flux of $F_{Si}\sim 2.8\times10^{11}$ atoms $cm^{-2}\,s^{-1}$, such that a single monolayer would be deposited in ~0.64 h. To replicate the growth conditions encountered during δ doping, the As_4 flux was incident onto the sample surface at all stages throughout the experiment. Si coverages of 0.001, 0.0025, 0.005, 0.01, 0.1, and 1.0 ML were investigated in this study. GaAs overlayer coverages started at 1 ML and doubled with each successive deposition, up to a maximum of 512 ML. RAS spectra were recorded for the clean GaAs(001)-c(4×4) surface, at 400 °C, and after each deposition. The RAS intensities presented here differ from those in our previous publications,[7,8] as noted in Ref. 16.

III. RESULTS AND DISCUSSION

Figure 1(a) shows a typical RAS spectrum for a clean GaAs(001)-c(4×4) surface, i.e., one where the GaAs buffer layer has been grown at 580 °C and the sample cooled to the measurement temperature of 400 °C. In our experiments, both the Si deposition and subsequent GaAs depositions have been carried out at 400 °C; hence, we need to ascertain the similarity between GaAs surfaces grown at 400 °C and those grown at 580 °C, then cooled to 400 °C. Our results indicate that, under the growth conditions employed here, low temperature (400 °C) growth results in a disordered surface that, when growth is terminated, slowly recovers to a well-ordered surface once more. To demonstrate this, Fig. 1(b) displays a time-resolved RAS measurement at an energy of 2.65 eV, the energy characteristic of direct excitation of

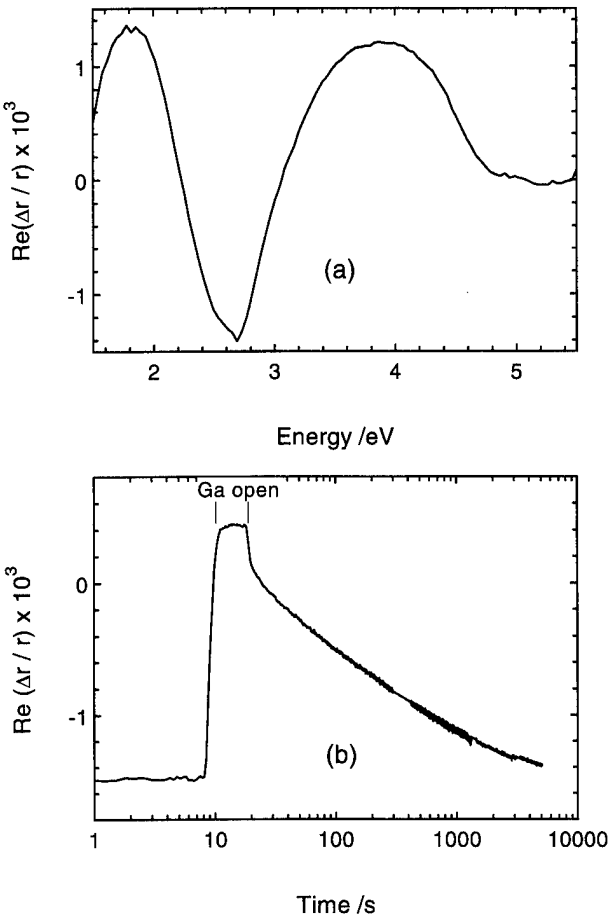


FIG. 1. (a) Typical RAS spectrum obtained at 400 °C from a GaAs(001)-c(4×4) surface, and (b) time-resolved RAS measurement at 2.65 eV before, during, and after 10 s GaAs deposition on GaAs(001), at 400 °C.

surface dimer species.[7] The initial equilibrium signal, the different equilibrium reached during the 10 s growth, and the return towards the initial equilibrium following growth are evident. As can be seen, a timescale of >1 h is required to recover the intensity of the 2.65 eV minimum at this growth temperature.

This is a somewhat unrealistically long time to use when studying GaAs overgrowth on Si/GaAs by a cycle of sequential growth and RAS measurements, under UHV conditions. Hence, we have chosen to systematically record RAS spectra 5 min (300 s) after each deposition. Since each scan takes 3 min (180 s) to complete, it is important to determine the degree by which the RAS intensity varies between 300 and 480 s after each deposition. Reference to Fig. 1(b) reveals that the intensity of the 2.65 eV minimum changes by 7.5% during this time interval, an amount that we have taken to be an acceptable systematic uncertainty in our study of GaAs overgrowth on Si/GaAs.

A. 0.01 ML Si

Figure 2(a) shows a series of RAS spectra that correspond to the initially clean GaAs(001)-c(4×4) surface, its modification with the addition of 0.01 ML Si, and the first stages of overgrowth with GaAs. The RAS spectra for GaAs overlayer

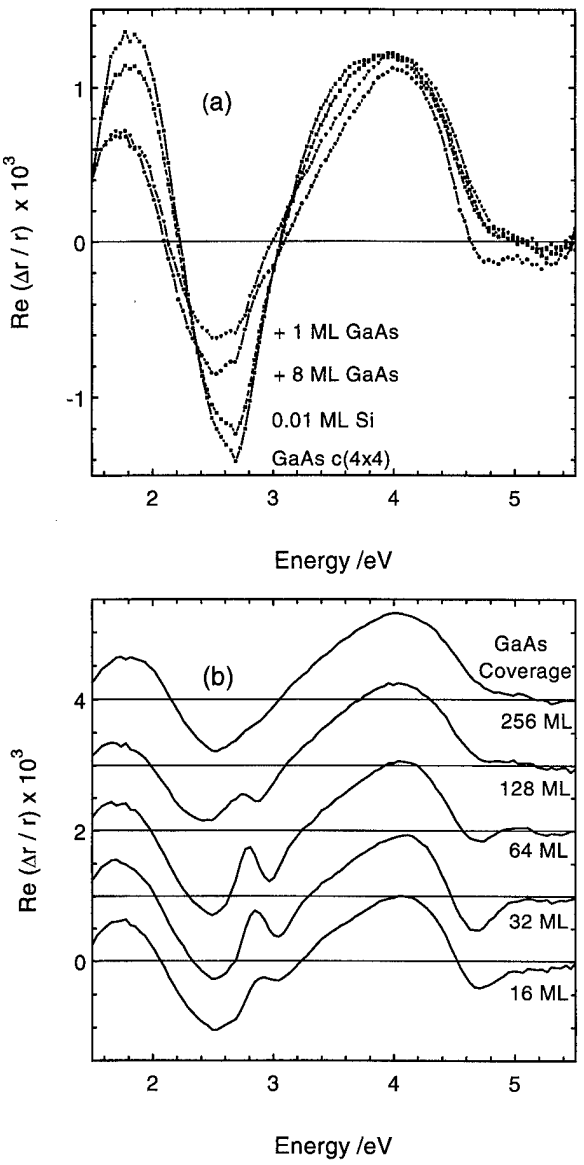


FIG. 2. RAS spectra for (a) initial GaAs(001)-$c(4\times4)$ surface, following addition of 0.01 ML Si and the first stages of overgrowth with GaAs, and (b) increasing GaAs coverage on top of 0.01 ML Si/GaAs.

thicknesses of 2 and 4 ML have been omitted for clarity. Most noticeably, deposition of 0.01 ML Si is seen to give rise to an upward shift of the 2.65 eV minimum, which becomes enhanced greatly on overgrowth with 1 ML GaAs. This is contrary to what one would expect if the deposited GaAs was simply restoring a well-ordered $c(4\times4)$ surface. Increasing GaAs coverage systematically reverses the shift obtained for 1 ML GaAs (the omitted RAS spectra for depositions of 2 and 4 ML lie between those shown for 1 and 8 ML).

While it is tempting to attribute the shift of the 2.65 eV minimum solely to a growth temperature induced decrease in surface order, the influence of the LEO effect and the possible variation of RAS signal with GaAs coverage above the Si δ layer cannot be totally excluded. The RAS spectrum for 8 ML GaAs on top of 0.01 ML Si/GaAS shows an inflection at ~2.9 eV, the signature of an LEO-related feature. Previous studies of the LEO effect in n-GaAS,[12] at 300 K, have shown the LEO-related feature to be centered around 3.05 eV in energy and only affect the lineshape of the RAS spectrum between 2.85 and 3.25 eV. If we apply this finding to our own RAS spectra, measured at 400 °C, then we might reasonably expect the contribution of the LEO feature to be restricted to the spectral range from 2.6 to 3.2 eV. This could distort the shape of the 2.65 eV minimum and might cause a displacement of this minimum to lower energy. However, our observation that there is no further displacement of the minimum between 1 and 8 ML GaAs coverage, added to the fact that the LEO feature only provides a weak contribution to the RAS spectrum for 8 ML, leads us to believe that the spectral changes are related, in some way, to the overgrowth with GaAs.

Figure 2(b) reveals the effect of increasing GaAs overlayer coverage. Each spectrum is plotted using the same absolute scale, but displaced vertically for clarity. The position of the zero line has been included in each case. The intensity of the LEO feature, identified for 8 ML GaAs/0.01 ML Si/GaAs, is found to reach a maximum at an overlayer thickness of 64 ML GaAs and then decrease with further GaAs coverage. At the same time, a minimum appears in the RAS spectrum at ~4.7 eV. Such a minimum has been observed previously for n-GaAs,[17] where the authors concluded that its behavior was independent of both the degree and type of doping. From Fig. 2(b) it is evident that the 4.7 eV minimum follows the behavior of the LEO feature, indicating that it is also related to the surface field, although its origin remains unclear.

Previous studies of the LEO effect in n-GaAs[11,12] have shown that it is possible to establish a direct correspondance between the integrated "LEO area" and the degree of bulk doping, since the LEO intensity is proportional to the field that arises due to pinning of the Fermi level at the GaAs surface. An accurate determination of the integrated LEO area was facilitated by being able to subtract the RAS spectrum for an undoped GaAs(001) sample from the spectra obtained for GaAs layers with different degrees of bulk doping, but similar surface structure. In the present case, the RAS spectra for GaAs overgrowth on 0.01 ML Si/GaAs are not identical to that for the clean GaAs(001)-$c(4\times4)$ surface, hence; the integrated LEO area cannot be obtained by a simple process of spectral subtraction.

By solving Poisson's equation self-consistently, we have calculated that 0.01 ML Si, a fully electrically active δ layer, separated by 64 ML GaAs from the sample surface corresponds to a surface field of $\sim5\times10^5$ V cm^{-1}. Moreover, the strength of the surface field increases, as expected, as we bring the δ layer closer to the surface of the sample. However, the intensity of the LEO feature is observed to decrease for coverages both above and *below* 64 ML GaAs. The decrease for thicknesses >64 ML GaAs can be explained by a combination of (i) the surface field decreasing with increasing coverage, and (ii) the δ layer passing beyond the penetration depth of the incident radiation (~170 Å at 3 eV[18]). One possible explanation for the decrease in signal below 64 ML

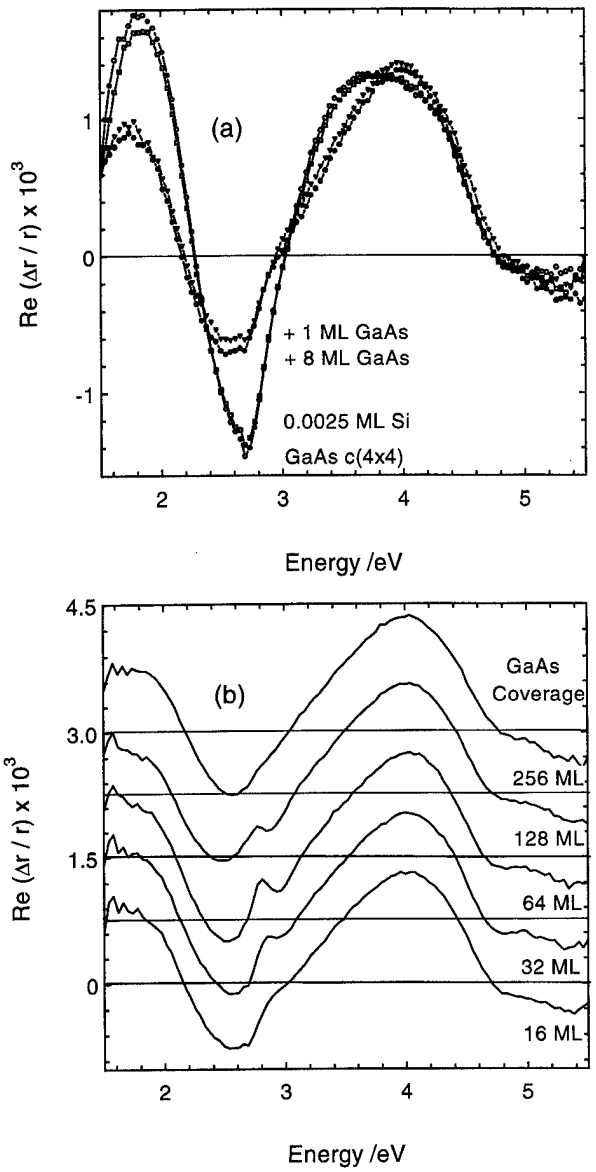


FIG. 3. RAS spectra for (a) initial GaAs(001)-$c(4\times4)$ surface, following addition of 0.0025 ML Si and the first stages of overgrowth with GaAs, and (b) increasing GaAs coverage on top of 0.0025 ML Si/GaAs.

is that the RAS technique averages over the near-surface region. This is beyond the scope of the present article and will be discussed elsewhere.[19]

B. 0.0025 ML Si

Figure 3(a) contains RAS spectra for GaAs overgrowth on 0.0025 ML Si/GaAs. By examining the minimum at 2.65 eV, it can be seen that it is virtually impossible to distinguish the clean surface $c(4\times4)$ spectrum and that obtained following deposition of 0.0025 ML Si. The RAS spectrum obtained following overgrowth with 1 ML GaAs is very similar to that shown for the same coverage in Fig. 2(a). However, the additional vertical shift of the 2.65 eV minimum, for a GaAs overlayer thickness of 8 ML, is reduced by comparison with Fig. 2(a). In fact, it is a general observation that the degree of vertical displacement of the 2.65 eV minimum, for overgrowth by 1–8 ML GaAs, decreases with the Si submonolayer coverage. The implication of this is that there are three effects that act to modify our RAS spectra when GaAs overgrowth is performed. These are the LEO effect, the disordering induced by the low growth temperature and, finally, the fact that overgrowth is performed on top of a Si δ layer. Indeed, the results are consistent with the presence of a disordering that increases with Si δ-layer coverage and acts in addition to the growth temperature induced disordering.

Figure 3(b) follows the evolution of the RAS spectrum for GaAs thicknesses in excess of 8 ML. Once again, the LEO feature is observed to develop initially with increasing GaAs coverage, up to 64 ML GaAs, before starting to decrease in intensity and disappearing completely by 256 ML GaAs. The sensitivity of RAS to the presence of dopant atoms, such as Si or Be, on the GaAs surface has previously been shown to be limited to 0.005 ML,[8] the minimum coverage that provides a resolvable shift of the 2.65 eV minimum. The fact that the LEO-related feature is still clearly observable for GaAs overgrowth on 0.0025 ML Si/GaAs means that the RAS technique shows increased sensitivity (1.6×10^{12} atoms cm^{-2}), for submonolayer coverages of planar dopant atoms, under these conditions.

C. 0.001–1.0 ML Si

Up until this point we have looked only at GaAs overgrowth on Si coverages of <0.01 ML Si/GaAs, where the RHEED reconstruction is exclusively $c(4\times4)$. It is to be noted, from our previous RAS measurements for Si on GaAs(001),[7] that a Si coverage of 0.1 ML corresponds to a crossover in the behavior of the 2.65 eV feature. At coverages of <0.1 ML Si, both $c(4\times4)$ and (2×1) reconstructions appear to coexist and so the overall RAS signal contains contributions from both surface phases. For Si coverages >0.1 ML, the (1×2) reconstruction dominates and the 2.65 eV feature shows a positive rather than a negative peak. In Fig. 4 we display RAS spectra for 64 ML GaAs overgrowth, the thickness at which the LEO feature is fully developed, on all the Si submonolayer coverages we have studied. Once again, all spectra have been plotted with the same absolute scale, but displaced vertically for clarity. The overall shapes of the RAS spectra are remarkably similar, excluding the contribution of the LEO-related feature, considering that the Si coverages span three orders of magnitude from 0.001 to 1.0 ML. It can be seen that the intensity of the LEO-related feature increases with Si submonolayer coverage up to 0.01 ML (equal to $\sim6.4\times10^{12}$ atoms cm^{-2}), and then decreases slightly with further coverage up to 1.0 ML Si. A similar correspondence has been reported between the density of Si_{Ga} (i.e., Si on donor sites) and the total Si coverage, up to a coverage of $\sim10^{13}$ cm^{-2} (0.016 ML).[20] In that case, the measured density of Si_{Ga} then remained approximately constant up to a coverage of $\sim10^{14}$ cm^{-2}, in very good agreement with the results presented here.

It is worth noting that the LEO feature remains evident in the RAS spectrum for 64 ML GaAs on top of 0.001 ML

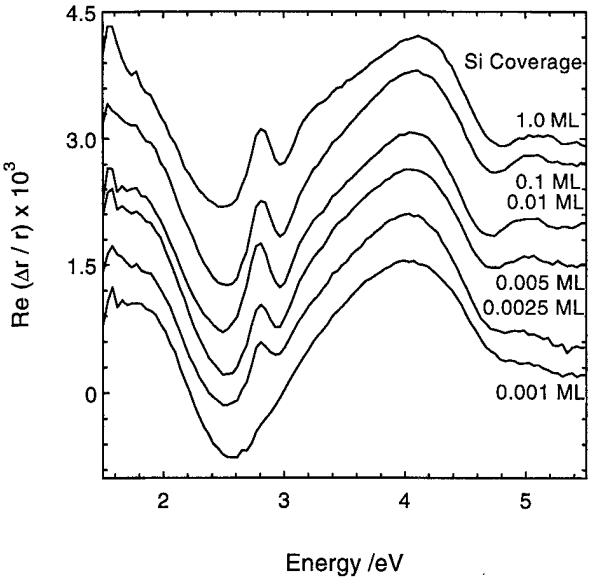


FIG. 4. RAS spectra for 64 ML GaAs deposited on Si/GaAs, where the Si coverages vary from 0.001–1.0 ML.

Si/GaAs, although it only appears as a kink around 2.85 eV. However, its presence has been confirmed by comparing RAS spectra that have been differentiated. In this case, the field at the surface has been calculated to be $\sim 9\times10^4$ V cm^{-1}.

IV. CONCLUSIONS

The RAS technique has been used to characterize the overlayer growth of GaAs onto sub to one monolayer coverages of Si δ layers deposited on the GaAs(001)-$c(4\times4)$ surface. The low growth temperature (400 °C) required to avoid spreading of the dopant away from the δ plane has meant that the study of an LEO-related signal is complicated by disordering at the GaAs surface. This disordering is induced not only by the growth temperature, but also by the presence of the Si δ layer itself. Variable thickness studies indicate that the LEO-induced signal is dependent on the field profile in the surface layer. Finally, the intensity of the LEO feature as a function of Si coverage reaches a maximum at ~0.01 ML ($\sim 6.4\times10^{12}$ atoms cm^{-2}) in agreement with previous studies of the site occupancy of Si δ layers.

ACKNOWLEDGMENTS

The authors are indebted to Martin Elliott for performing the electrical calculations and they would also like to acknowledge the financial support of the United Kingdom SERC Low Dimensional Structures and Devices initiative.

[1]M. R. Fahy, M. J. Ashwin, J. J. Harris, R. C. Newman, and B. A. Joyce, Appl. Phys. Lett. **61**, 1805 (1992).
[2]L. Daweritz, K. Hagenstein, and P. Schutzendube, J. Cryst. Growth **127**, 1051 (1993).
[3]M. Wassermeier, J. Behrend, L. Daweritz, and K. Ploog, Phys. Rev. B **52**, R2269 (1995).
[4]A. R. Avery, J. Sudijono, D. M. Holmes, T. S. Jones, and B. A. Joyce, Appl. Phys. Lett. **66**, 3200 (1995).
[5]A. R. Avery, D. M. Holmes, J. L. Sudijono, T. S. Jones, M. R. Fahy, and B. A. Joyce, J. Cryst. Growth **150**, 202 (1995).
[6]A. H. Levermann, D. A. Woolf, D. I. Westwood, and J. E. Macdonald, Surf. Sci. (in press).
[7]D. A. Woolf, K. C. Rose, J. Rumberg, D. I. Westwood, F. Reinhardt, S. J. Morris, W. Richter, and R. H. Williams, Phys. Rev. B **51**, 4691 (1995).
[8]D. A. Woolf, K. C. Rose, S. J. Morris, D. I. Westwood, J. Rumberg, F. Reinhardt, W. Richter, and R. H. Williams, J. Cryst. Growth **150**, 197 (1995).
[9]M. Wassermeier, I. Kamiya, D. E. Aspnes, L. T. Florez, J. P. Harbison, and P. M. Petroff, J. Vac. Sci. Technol. B **9**, 2263 (1991).
[10]I. Kamiya, D. E. Aspnes, L. T. Florez, and J. P. Harbison, Phys. Rev. B **46**, 15 894 (1992).
[11]H. Tanaka, E. Colas, I. Kamiya, D. E. Aspnes, and R. Bhat, Appl. Phys. Lett. **59**, 3443 (1991).
[12]S. E. Acosta-Ortiz and A. Lastras-Martinez, Phys. Rev. B **40**, 1426 (1989).
[13]D. I. Westwood, D. A. Woolf, and R. H. Williams, J. Cryst. Growth **98**, 782 (1989).
[14]S. Scholz, A. B. Muller, W. Richter, D. R. T. Zahn, D. I. Westwood, D. A. Woolf, and R. H. Williams, J. Vac. Sci. Technol. B **10**, 1710 (1992).
[15]D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[16]The RAS intensities are a factor of 3.7 lower in this paper than in our previous publications (Refs. 7 and 8) since the correction factor with which the ac/dc signal ratio is multiplied is now set to the theoretical value that comes out of Aspnes *et al.*'s (Ref. 15) calculation of the signal shape.
[17]S. E. Acosta-Ortiz and A. Lastras-Martinez, Solid State Commun. **64**, 809 (1987).
[18]D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).
[19]Z. Sobiesierski, D. I. Westwood, and M. Elliott (unpublished).
[20]M. J. Ashwin, M. Fahy, J. J. Harris, R. C. Newman, D. A. Sansom, R. Addinall, D. S. McPhail, and V. K. M. Sharma, J. Appl. Phys. **73**, 633 (1993).

Interpretation of surface-induced optical anisotropy of clean, hydrogenated, and oxidized vicinal silicon surfaces investigated by reflectance-difference spectroscopy

U. Rossow,[a] L. Mantese, and D. E. Aspnes
Physics Department, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 14 March 1996; accepted 26 April 1996)

Using reflectance-difference spectroscopy, we determine surface-induced optical anisotropy (SIOA) spectra of clean, hydrogenated, and oxidized (113) and vicinal (001) Si surfaces to obtain a better understanding of the origin of the optical response of surfaces and interfaces. Hydrogenation was performed either by etching in dilute HF or by exposing clean surfaces to atomic hydrogen. Hydrogenated and oxidized vicinal surfaces show energy-derivativelike spectra that roughly scale with offcut angle, indicating step-induced behavior, and exhibiting features near 3.4 and 4.2 eV, the threshold energies of the (E_1, E_0') and E_2 interband critical points of bulk Si, respectively. The appearance of derivativelike line shapes indicates that bulk threshold energies become dichroic near the surface due to the surface-induced modification of the potential, as supported by model calculations. However, direct integration yields dielectric functions somewhat different from bulk values, indicating that the surface affects the near-surface dielectric function. SIOA spectra for clean surfaces are qualitatively different, exhibiting dielectric-functionlike line shapes that appear to originate primarily from terraces, although indications of smaller derivativelike contributions are also found. © *1996 American Vacuum Society.*

I. INTRODUCTION

In this work, we use reflectance-difference spectroscopy (RDS) to investigate surface-induced optical anisotropy (SIOA) spectra of a broad range of clean, hydrogenated, and oxidized Si surfaces with the objective of obtaining a better understanding of the optical response of surfaces and interfaces. RDS measures the difference in reflectance between two principal axes in the surface plane and is, therefore, well suited for isolating surface optical responses. Since the bulk optical response of cubic materials is isotropic, it cancels upon subtraction leaving an RDS signal originating from the lower-symmetry surface.

Because the Si surface has been well studied and has been shown to be relatively structurally simple, Si is an ideal material for such investigations. It has been shown that clean room-temperature (RT) (001) surfaces form asymmetric dimers that are parallel to each other within a given terrace.[1] Adjacent terraces separated by single-atomic-height steps form dimer domains rotated 90° relative to each other. Because the two sublattices are chemically equivalent, in the absence of any driving mechanism for the formation of double-height steps, single-height steps will predominate, the areas associated with the two different domains will be approximately equal, and no macroscopic RDS response will be observed. However, for angles greater than approximately 4°, atomically clean, annealed surfaces are dominated by double-height steps so that dimers in adjacent terraces are parallel.[2,3] In this situation, a net RDS response is obtained that in general contains both terrace and step contributions.

One way to investigate step and terrace contributions is to chemically modify the surface and observe changes in the RDS signal. In this work, we used both hydrogen and oxygen to saturate dangling bonds and terminate the surface. Hydrogen termination is of interest since the H-terminated (001) Si surface exists in two stable configurations depending on coverage. At low coverages, a single hydrogen atom bonds to the single dangling bond of each Si surface atom. In this so-called monohydride phase, dimer bonding is retained but the initially asymmetric dimers become symmetric.[4] Upon further exposure, the surface converts to the dihydride phase where the dimer bonds are broken and both Si dangling bonds on each atom are saturated by hydrogen.[5] Other phases are also possible[6] with H exposure, however, they are not as stable as the mono- and dihydride surfaces. While the exact structure of the clean and chemically modified surfaces is still a matter of debate, it is clear that the backbonds of the Si dimer atoms are strained. This allows the effect of localized strain on optical spectra to be studied.

II. EXPERIMENT

SIOA spectra were obtained using a RDS spectrometer[7–9] that operates at near-normal incidence from 1.5 to 5.5 eV. The anisotropy of the complex reflectance $\tilde{r}$ is given by $\Delta\tilde{r}/\tilde{r}=2(\tilde{r}_\alpha-\tilde{r}_\beta)/(\tilde{r}_\alpha+\tilde{r}_\beta)=\Delta r/r+i\Delta\theta$, where $\tilde{r}_\alpha$ and $\tilde{r}_\beta$ are the complex reflectances for light linearly polarized along the α and β principal axes, respectively. Here we report data for $\Delta r/r$ with α defined to be along $[\bar{1}10]$, and β to be orthogonal to the $[\bar{1}10]$ direction and the surface normal. Wafer orientations were verified using low energy electron diffraction (LEED). Several chemically treated samples were measured on a rotational stage where spectra were taken at

[a]Electronic mail: rossow@unity.ncsu.edu

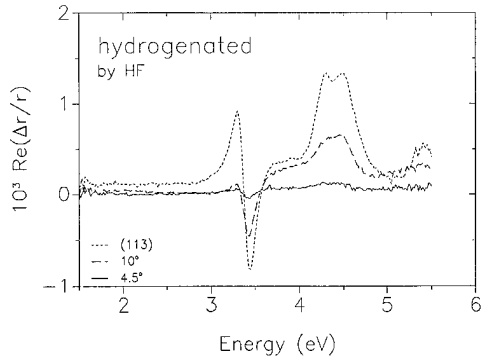


FIG. 1. RDS spectra of various vicinal Si surfaces hydrogenated by HF.

two orthogonal orientations and subtracted to minimize artifacts in $\Delta r/r$ due to nonidealities of the optical components.

We investigated a broad range of Si surfaces, specifically: (a) vicinal (001) surfaces cut 4°, 6°, and 10° off (001) toward [110]; and (b) the (113) surface which can be considered as an (001) surface miscut 25° off (001) toward [110]. The wafers were lightly p- or n -doped with resistivities from 1 to 10 Ω cm and were chemically prepared by an RCA clean.[10] Further treatment consisted of a 30 s dip in dilute HF followed by a de-ionized-water rinse to obtain a hydrogenated surface or a 30 s dip in H_2O_2 to obtain an oxidized surface. Data were taken directly after treatment with the samples in air. To establish that the surfaces did not degrade during measurement, we also obtained data on samples immersed in dilute HF or H_2O_2 contained in a quartz beaker.

For comparison, we measured the evolution of the SIOA response of single-domain surfaces exposed to atomic H. These samples were prepared by an RCA clean with no final HF dip, then transferred to a UHV system with a low 10^{-10} Torr base pressure where the residual oxide layer was desorbed and the surface reconstructed by annealing for 2 min at 920 °C. LEED patterns exhibited sharp, well-defined spots with splittings in the direction of the steps, along the β direction. We measured the evolution of the RDS signal as the surface was exposed to atomic H, which was generated by dissociating H_2 in the 10^{-6} mbar range with a hot W filament.

III. RESULTS AND DISCUSSION

Figure 1 shows RDS spectra of (113) and two vicinal (001) Si surfaces hydrogenated in dilute HF.[11] RCA cleaned surfaces with no final HF dip show similar RDS signals. It is important to note that the magnitude of the signal increases with offcut angle. Since the total area of the terraces decreases slightly with offcut angle and the areal density of steps increases, it follows that these RDS signals are dominated by step-induced contributions, namely (001) double-height steps. This has been addressed previously in Ref. 11.

These line shapes all exhibit three main features: The first is an energy-derivativelike structure near 3.4 eV; the second and third form a double structure near 4.3 eV. These energies are near those of the (E_0', E_1), and E_2 interband critical points of bulk silicon, respectively. Previous work[12–14] has shown that RDS spectra can exhibit energy-derivativelike features and can even be dominated by such features[11] which can be described as follows. For a thin anisotropic overlayer on an isotropic bulk, the RDS signal is given to first order in d/λ by

$$\frac{\Delta\tilde{r}}{\tilde{r}} = -\frac{4\pi i E d}{hc}\,\frac{\epsilon_\alpha - \epsilon_\beta}{\epsilon_b - 1}, \tag{1}$$

where E is the photon energy, d is the thickness of the anisotropic overlayer, λ is the wavelength of light, and hc is the product of Planck's constant and the speed of light. ϵ_α and ϵ_β are the components of the surface dielectric tensor along α and β, respectively, and ϵ_b is the bulk dielectric function. If we assume that ϵ_α and ϵ_β are both functions of ϵ_b, but with slightly different threshold energies and broadening parameters, i.e., that the associated critical points are dichroic, then in the vicinity of a critical point, we can write approximately

$$\epsilon_\alpha = \epsilon_b(E - E_g - \Delta E_{g\alpha} + i\Gamma + i\Delta\Gamma_\alpha),$$

$$\epsilon_\beta = \epsilon_b(E - E_g - \Delta E_{g\beta} + i\Gamma + i\Delta\Gamma_\beta),$$

whence

$$\frac{\Delta\tilde{r}}{\tilde{r}} = \frac{4\pi i E d}{hc}\,\frac{(-\Delta E_g + i\Delta\Gamma)}{\epsilon_b - 1}\,\frac{\partial\epsilon_b}{\partial E}, \tag{2}$$

where $\Delta E_g = \Delta E_{g\alpha} - \Delta E_{g\beta}$ and $\Delta\Gamma = \Delta\Gamma_\alpha - \Delta\Gamma_\beta$.

Although the assumption is typically made, it is not clear *a priori* that the energy derivative in Eq. (2) should be with respect to ϵ_b or whether a surface-modified bulk dielectric function may be more appropriate. Since extra scattering channels are available at the surface, we would expect at least an increase in broadening of the interband critical point structures near the surface. To examine this further, we determined whether we could recover ϵ_b from RDS data by integrating Eq. (2). We define the result as a near-surface dielectric constant, ϵ_s, given by

$$\epsilon_s(E)\cdot d = \epsilon_s(E_0)\cdot d - i\,\frac{hc}{4\pi}\,\frac{1}{(-\Delta E_g + i\Delta\Gamma)} \times \int_{E_0}^{E} dE'\,\frac{(\epsilon_b - 1)}{E'}\,\frac{\Delta\tilde{r}}{\tilde{r}}, \tag{3}$$

where E_0 is a lower limit of integration, here equal to 1.5 eV. For Im($\epsilon_s \cdot d$) we assume an integration constant of 0 which is reasonable since Im($\epsilon_s \cdot d$) is observed to be nearly constant below 3 eV and a constant nonzero absorption is unlikely. We also assume that $\Delta\Gamma$ is small compared to ΔE_g, which was found to be 80 meV Å in the case of Si(113).[11] This point will be discussed later. The resulting calculated spectra Im($\epsilon_s \cdot d$) for hydrogenated or oxidized Si samples cut 10° off (001) toward (110) are shown in Fig. 2. For scaling purposes, we divided Im($\epsilon_s \cdot d$) by 0.3 nm, a reasonable thickness for an anisotropic surface layer.

The calculated spectra show a feature near 3.33 eV, which is very near the energies of the (E_0',E_1) interband critical points of bulk silicon, and a second broad feature near 4.2

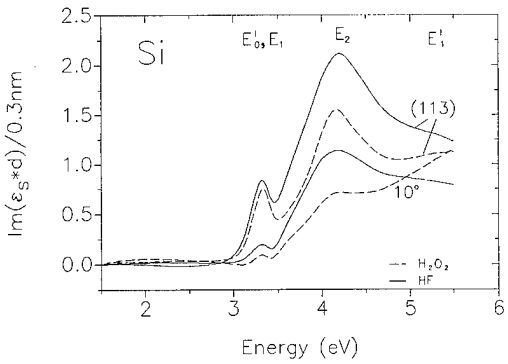


FIG. 2. $\mathrm{Im}(\epsilon_s \cdot d)$ spectra obtained by applying Eq. (3) to RDS data for hydrogenated and oxidized (113) and 10° vicinal (001) Si surfaces. The hydrogenated data are from Fig. 1. The energy locations of the bulk interband critical points (E_0', E_1), E_2, and E_1' are indicated.

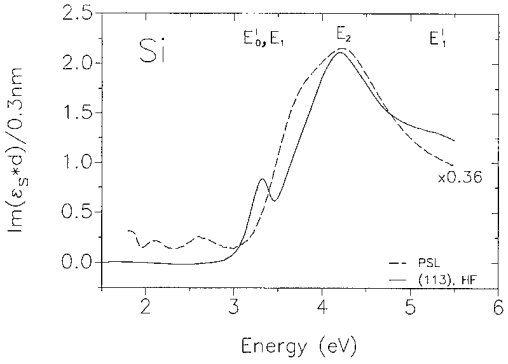


FIG. 4. Comparison between the hydrogenated (113) $\mathrm{Im}(\epsilon_s \cdot d)$ spectrum of Fig. 2 (solid line) and $\mathrm{Im}(\epsilon_b)$ of porous Si layers (dashed line).

eV, which is near the E_2 critical point. Figure 3 shows a comparison between the bulk crystalline spectrum $\mathrm{Im}(\epsilon_b)$ and the calculated spectrum $\mathrm{Im}(\epsilon_b \cdot d)$ from Fig. 2 for the (113) surface. There is qualitative agreement between the two spectra, however, the (E_0', E_1) structure for the (113) surface is weaker, and the features are somewhat broader and shifted. The different amplitudes for the (E_0', E_1) and E_2 structures may not be unreasonable since it is not apparent *a priori* why the scaling factor ΔE_g should be the same for the two transitions. In addition, with respect to scattering from the (113) surface with respect to the bulk, which would explain the enhanced broadening of the integrated spectrum. We can gain further insight into the effect of the surface on the dielectric response by comparing the same (113) spectrum from Fig. 2 to that of bulk porous Si,[15–17] which is strongly influenced by the very large internal surface area of its pores. The result is shown in Fig. 4. In this case, there is better agreement near E_2, but the (113) spectrum retains some structure associated with the (E_0', E_1) structure although it is completely absent in the porous Si spectrum. These results suggest that the correct representation for ϵ_b in the derivative of Eq. (2) appears to lie somewhere between the two limits of crystalline and porous Si.

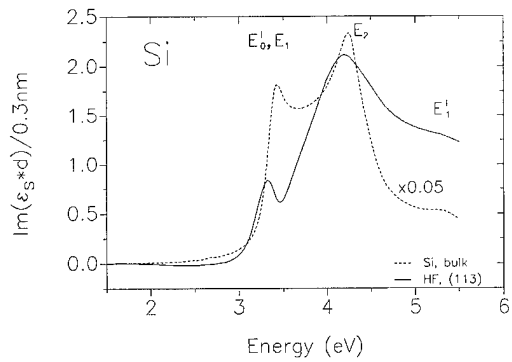


FIG. 3. Comparison between the hydrogenated (113) $\mathrm{Im}(\epsilon_s \cdot d)$ spectrum of Fig. 2 (solid line) and $\mathrm{Im}(\epsilon_b)$ of bulk crystalline Si (dashed line).

In calculating the spectra of Fig. 2, we assumed that $(-\Delta E_g + i\Delta\Gamma)$ is essentially real, which means that $\Delta E_g \gg \Delta\Gamma$. If this were not the case, $\epsilon_s(E) \cdot d$ would be multiplied by a complex phase factor $\exp(i\theta)$, which would mix real and imaginary parts of $d\epsilon/dE$. Although no strong justification can be given at present for our assumption, we observed no improvement of the fit to the spectra when nonzero values of $\Delta\Gamma$ were used. In addition, the spectra calculated for nonzero values of $\Delta\Gamma$ showed features shifted unreasonably to lower energies. Note that the statement $\Delta\Gamma \ll \Delta E_g$ implies only that $\Delta\Gamma_\alpha \approx \Delta\Gamma_\beta$ on the scale of ΔE_g, which still allows a significant increase of Γ with respect to bulk values. We also assumed that both ΔE_g and $\Delta\Gamma$ are essentially independent of photon energy, which is a reasonable assumption near a critical point. However, the complete spectrum contains contributions from several critical points and thus the use of constant values for ΔE_g and $\Delta\Gamma$ for all critical points is a simplification. The question as to whether ϵ_b should also be replaced in the denominator of Eq. (2) depends on whether the three-phase (ambient/anisotropic surface layer/isotropic bulk) model is sufficient to describe the optical response. The question has not been resolved, however, it will be discussed further in Ref. 18.

We now discuss the origin of the threshold energy shifts of the bulk critical points that lead to the derivativelike features. The obvious causes are surface strain and chemical induction, which is a change in the subsurface electron distribution caused by chemisorption. Both cases result in near-surface Si atomic positions that are different from those of the bulk. To estimate the possible effect of strain on critical-point energies near the surface, we consider clean (001) Si using atomic positions reported in Ref. 19 for the $c(2\times4)$ reconstruction. Although a (2×1) reconstruction is observed at RT, there is evidence that the $c(2\times4)$ unit is still the building block even at high temperatures.[20] From the atomic positions given in Ref. 19, the length of the asymmetric dimer bond is 2.30 Å, which is slightly less than the length of the bulk Si–Si bond (2.352 Å at RT). Due to dimer formation, the Si–Si back bonds are rotated and strained. The back bonds of the lower atom of the asymmetric dimer have a length 2.315 Å and thus are in 1.6% compressive strain, whereas, the back bonds for the higher atom have a length

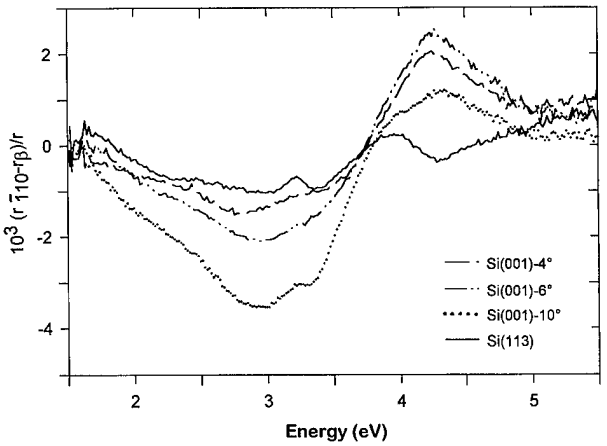


FIG. 5. RDS spectra for clean vicinal (001) and (113) Si surfaces. The spectra are offset to yield the value zero at the low energy limit of the figure.

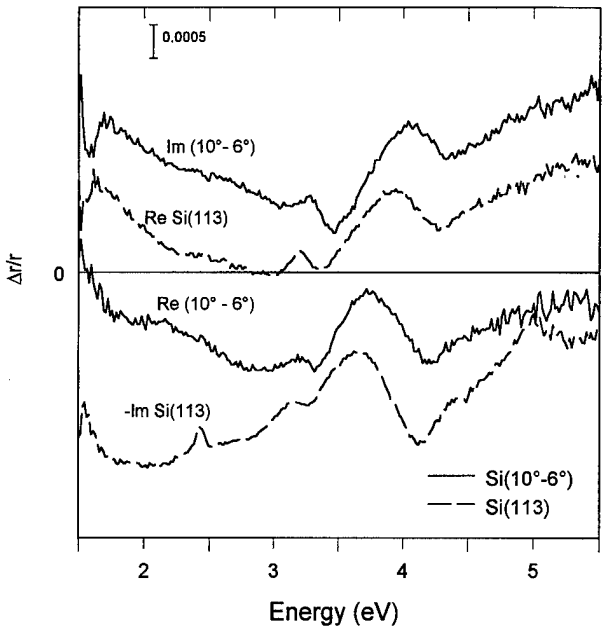


FIG. 6. Comparison between the difference between the 10° and 6° vicinal Si(001) RDS spectra (solid line) and the Si(113) spectra (dashed line). The (113) spectra are offset for clarity and the imaginary component has been multiplied by (−1).

2.360 Å and thus are in 0.4% tensile strain. Therefore, the net effect is a compressive strain of the Si–Si back bonds. From the elastic stiffness constants $(c_{11}+2c_{12})=2.93$ Mbar,[21] hydrostatic strain shifts the (E_0',E_1) critical point of bulk Si by 5 meV/kbar.[22] We calculate a gap shift of about 0.2 eV to higher energies. Most likely the apparent shift would be averaged over several layers, in which case shifts would be of the order of 0.1 eV Å, which was observed.

The monohydride surface is more difficult to explain using strain arguments. Here the dimer is symmetric, so all back bonds will be strained equally. However, measured and calculated values for the dimer length range from 2.40 Å (Ref. 4) to 3.0 Å (Ref. 23) compared to a bulk separation of 3.84 Å. The strain induced by the back bonds strongly depends on this dimer length. Taking a value of 2.34 Å for the dihydride phase as a lower limit,[4] we estimate a strain of 0.4% and a shift of about 47 meV.

In addition to strain, surface electric fields can create first-derivativelike features for crystals *without* inversion symmetry. This is the well-known linear electro-optic effect,[24] which is usually related to the electric field of the space charge region generated from Fermi level pinning by surface states.[25] This mechanism does not contribute to the optical response of bulk Si because of its inversion symmetry. However, this symmetry is broken at the surface and therefore a local electric field may contribute to the RDS signal. Such local electric fields might arise from the charge transfer due to the chemisorbed species. The Pauling electronegativity for H (2.1) is higher than that of Si (1.8), which results in a net charge transfer from Si to H. The charge transfer for a single Si–H bond would be $0.02e_0$, which is in good agreement with the calculated value of $0.029e_0$ in Ref. 26. The related electric field, therefore, would be strong enough to create a derivativelike feature.

A change of the threshold energies from their bulk values, as indicated by the observed derivativelike line shapes, can only occur if the states involved are near-bulklike and localized. We discuss this issue in a forthcoming article.[18]

Unlike the hydrogenated and oxidized Si surfaces, the clean surfaces exhibit dielectric-functionlike line shapes as shown in Fig. 5. A broad feature is apparent near 4.2 eV, similar to that seen for porous Si. The main difference is the finite value between 2 and 3 eV, which is most likely related to the dangling-bond states of the Si dimers.[11] No significant change in signal amplitude is observed for increasing offcut angle for the clean Si surface,[11] which indicates that step contributions are minor. The fact that the smallest signal is observed for the (113) surface, which can be considered to be all steps, is consistent with this interpretation.

To isolate possible step contributions to the RDS spectra of the clean surfaces, we calculate the difference between the (001) spectra for the 10° and 6° miscut angles. The resulting real and imaginary parts are shown in Fig. 6. The result indicates that

$$\frac{\Delta\tilde{r}}{\tilde{r}}(10°)-\frac{\Delta\tilde{r}}{\tilde{r}}(6°)\approx i\,\frac{\Delta\tilde{r}}{\tilde{r}}(113). \tag{4}$$

The simplest interpretation is that the RDS signal arises from a linear combination of step and terrace contributions:

$$\frac{\Delta\tilde{r}}{\tilde{r}}=A_t\left(\frac{\Delta\tilde{r}}{\tilde{r}}\right)_{\text{terr}}+A_s\left(\frac{\Delta\tilde{r}}{\tilde{r}}\right)_{\text{step}}, \tag{5}$$

where A_t and A_s are weighting functions, $A_t+A_s=1$, and A_t and A_s are proportional to the areas of terrace and step. However, this model cannot explain the result expressed as Eq. (4), especially if we assume that the RDS signal for the (113) surface is dominated by steps. Even if the difference between the 10° and 6° spectra has only step contributions, it is not clear why its amplitude should be similar to that of the (113)

spectrum. The phase shift indicated by i could be explained if we assume that the RDS spectrum for the clean (113) surface has a derivativelike line shape. This would require that the character of the RDS spectra changes with offcut angle. The appearance of i then indicates that the system evolves from being dominated by a dichroic band gap ΔE_g to being dominated by a dichroic broadening parameter $\Delta\Gamma$. However, the small feature near 3 eV in the (113) spectrum resembles that shown in Fig. 2 for $\mathrm{Im}(\epsilon_s \cdot d)$ although slightly shifted to lower energies. This point needs to be clarified by further experiments.

IV. SUMMARY AND CONCLUSIONS

We have investigated terrace- and step-induced contributions to the RDS spectra of clean and chemically modified Si surfaces. Hydrogenated and oxidized surfaces show derivativelike line shapes arising from bulk critical points that have become dichroic as a result of a potential that induces either chemical induction or strain at the surface. By integrating these spectra, we obtain near-surface dielectric functions which resemble those of porous Si, a material that is dominated by surface contributions and exhibits features that can be related to significantly broadened bulk interband critical points. In contrast, RDS line shapes of clean surfaces are dielectric-functionlike and exhibit only minor contributions from steps.

ACKNOWLEDGMENTS

This work is supported by the Alexander von Humboldt Foundation and the Office of Naval Research under Contract No. N-00014-93-1-0255.

[1]D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979).
[2]E. Schroeder-Bergen and W. Ranke, Surf. Sci. **259**, 323 (1991).
[3]D. J. Chadi, Phys. Rev. Lett. **59**, 1691 (1987).
[4]J. E. Northrup, Phys. Rev. B **44**, 1419 (1991).
[5]T. Sakurai and H. D. Hagstrum, Phys. Rev. B **14**, 1593 (1976).
[6]Y. J. Chabal and K. Raghavachari, Phys. Rev. Lett. **54**, 1055 (1985).
[7]D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).
[8]D. E. Aspnes, J. P. Harbison, A. A. Studna, L. T. Florez, and M. K. Kelly, J. Vac. Sci. Technol. A **6**, 1327 (1988).
[9]W. Richter, Philos. Trans. R. Soc. London Sect. A **344**, 453 (1993).
[10]W. Kern and D. A. Puotinen, RCA Rev. **31**, 187 (1970).
[11]L. Mantese, U. Rossow, and D. E. Aspnes, Appl. Surf. Sci. (in press).
[12]D. E. Aspnes and A. A. Studna, J. Vac. Sci. Technol. A **5**, 546 (1987).
[13]T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, J. Vac. Sci. Technol. A **12**, 1152 (1994).
[14]T. Yasuda, L. Mantese, U. Rossow, and D. E. Aspnes, Phys. Rev. Lett. **74**, 3431 (1995).
[15]U. Rossow, H. Münder, M. Thönissen, and W. Theiss, J. Luminescence **57**, 205 (1993).
[16]U. Rossow, U. Frotscher, C. Pietryga, W. Richter, and D. E. Aspnes, Appl. Surf. Sci. (in press).
[17]U. Rossow, L. Mantese, U. Frotscher, D. E. Aspnes, and W. Richter, Mater. Res. Soc. Symp. Proc. **406**, 371 (1996).
[18]U. Rossow, L. Mantese, and D. E. Aspnes (to be published).
[19]J. E. Northrup, Phys. Rev. B **47**, 10 032 (1993).
[20]A. Shkrebtii and R. Del Sole, Appl. Surf. Sci. (in press).
[21]*Properties of Silicon* (INSPEC, London, 1988), p. 14.
[22]R. Zallen and W. Paul, Phys. Rev. **155**, 703 (1967).
[23]Y. Wang, M. Shi, and J. W. Rabalais, Phys. Rev. B **48**, 1678 (1993).
[24]S. E. Acosta-Ortiz and A. Lastras-Martinez, Solid State Commun. **64**, 809 (1987).
[25]U. Resch, S. M. Scholz, U. Rossow, A. B. Müller, and W. Richter, Appl. Surf. Sci. **63**, 106 (1993).
[26]B. I. Craig and P. V. Smith, Surf. Sci. Lett. **226**, L55 (1990).

Ab initio calculations of the reflectance anisotropy spectrum of the GaAs(001) $c(4\times4)$ surface

J. M. Bass[a] and C. C. Matthai
Department of Physics and Astronomy, University of Wales, College of Cardiff, P.O. Box 913, Cardiff CF2 3YB, United Kingdom

(Received 21 January 1996; accepted 5 April 1996)

Using an *ab initio,* nonlocal pseudopotential method, we have calculated the reflectance anisotropy spectrum for a particular model of the GaAs(001) $c(4\times4)$ reconstructed surface. Excellent agreement with experiment was obtained, supporting the chosen model. The reflectance anisotropy was found to originate from optical transitions between bulk like valence band states and surface states in the conduction band. The nature and distribution of the electronic states involved is discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

When light is reflected at normal incidence from the surface of a cubic crystal, the intensity of the reflected light should not be dependent on its polarization. In practice, a small difference in the reflectance of light polarized along particular symmetry directions is observed. As the complex dielectric function, which describes the optical response of a solid, is, in a centro-symmetric crystal, reduced to a complex scalar, any anisotropy in the normal incidence reflectivity of cubic crystals must arise from the surface region. Reflection anisotropy spectroscopy (RAS), which measures the difference in the reflection for light of two orthogonal polarizations normally incident on a surface as a function of the photon energy, provides useful information about the surface electronic structure. However, to interpret the features observed in the spectrum one has to resort to quite detailed calculations.

Theoretical work has progressed along two different approaches. Mochán and Barrera[1] identified surface local field effects as being responsible for the observed anisotropy. By contrast, the calculations of Selloni *et al.*[2] were performed within a single-particle picture in which the electronic eigenstates of the crystal were calculated explicitly and the optical functions derived by considering the electronic transitions. Subsequent work by Manghi *et al.*[3] and Chang *et al.*[4] based on this latter approach resulted in calculated spectra showing reasonable agreement with experiment, suggesting this to be appropriate when dealing with *semiconductor* surfaces.

The GaAs(001) surface has been the subject of much study in recent years because of its perceived technological importance. It displays a wide variety of surface reconstructions, the most important being the (4×2), the (2×4), and the $c(4\times4)$ reconstructions. The (2×4) reconstruction is As terminated and occurs in three different phases, known as the α, β, and γ,[5] depending on the growth conditions. In all three phases the surface is comprised of As dimers directed along the $[\bar{1}10]$ direction. For coverages of As of more than 0.75 of a monolayer extra As atoms adsorb on the top As layer breaking the dimer bonds and forming new dimers along the [110] direction. At sufficiently high As coverages the $c(4\times4)$ reconstruction forms. The widely accepted model for this reconstruction, as deduced from scanning tunneling microscopy (STM) images[6,7] and supported by x-ray diffraction experiments,[8] is shown in Fig. 1. Kamiya *et al.*[9] have shown that the RAS spectral line shapes for the three primary reconstructions of the GaAs(001) surface are quite different from each other making RAS an effective technique in distinguishing between the different reconstructions.

In our earlier study[10] of the GaAs(001) (2×4) reconstructed surfaces, we demonstrated that *ab initio* calculations of the electronic structure and optical properties based on the one-electron picture gave RAS spectra that are directly comparable with experimental results. We also showed that a comparison of the calculated and experimental spectra can be used to yield information about the atomic structure of the surface. In particular, we found that the low temperature experimental spectrum of the (2×4) surface could be interpreted using a combination of two structural models for this surface, viz. the $\alpha(2\times4)$ and the $\beta2(2\times4)$ structures as denoted by Northrup and Froyen.[11] The quality of the fit between the calculated and the experimental spectra is illustrated by Fig. 2.

In growth studies of GaAs(001) surfaces, the oscillations in the RAS intensity at a photon energy of 2.7 eV are generally ascribed to the As dimers on the surface alternating between the $[\bar{1}10]$ and [110] directions as the surface undergoes transitions between the (2×4) and $c(4\times4)$ reconstructions. It is therefore important that any theoretical study aimed at arriving at an understanding of the RAS spectrum must give a good description of both these reconstructions. Following from our study of the (2×4) reconstructed surface, the aim of this paper is to calculate the RAS spectrum of the $c(4\times4)$ reconstructed surface and to interpret its features in terms of electronic structure.

II. CALCULATION METHOD

The calculations were performed in a manner similar to that used in our previous work,[10] using *ab initio* nonlocal pseudopotentials, the local density approximation and the conjugate-gradient method to minimize the total energy

[a]Electronic mail: bass@cf.ac.uk

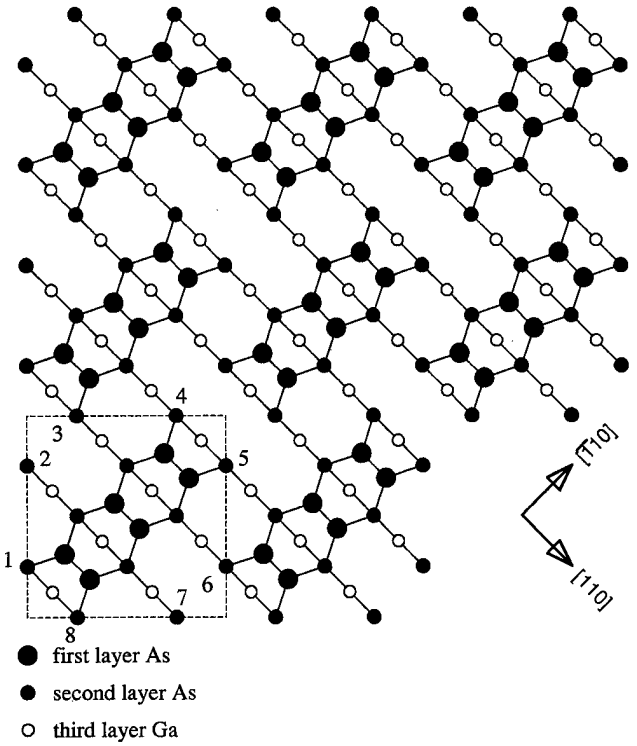


FIG. 1. The model for the GaAs(001) $c(4\times4)$ reconstructed surface.

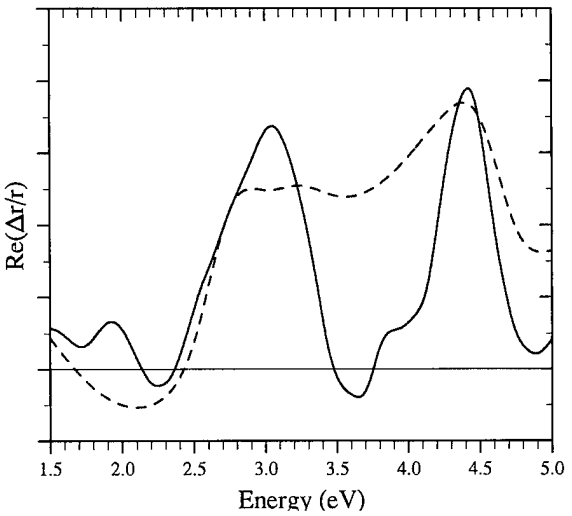


FIG. 2. The calculated RAS spectrum (solid line), as obtained from the combination of two structural models, and the low-temperature experimental spectrum (dashed line), for the GaAs(001) (2×4) reconstructed surface.

through the relaxation of the electronic and atomic degrees of freedom.[12] The surface was represented using a repeated slab geometry, with a slab containing four layers of GaAs, one layer of hydrogen atoms terminating the "bottom surface" and six extra As atoms on the "top surface" forming the three dimers oriented in the [110] direction as shown in Fig. 1. Six layers of "vacuum" separated the two surfaces, making the total slab thickness equivalent to eight layers of GaAs. Due to the symmetry properties of GaAs, any slab terminated with As on both surfaces has the dimers on each surface perpendicular to the other. Such a construction would be counterproductive as the anisotropies of the perpendicular dimers on either surface would cancel each other out and lead to a null result. As the optical matrix elements are calculated in reciprocal space it is hard to separate the contributions of the two surfaces, so the most practical way to find the anisotropy of the top surface is to ensure that the bottom surface is isotropic. This is the purpose of the hydrogen atoms on the bottom surface which are positioned in such a way that each hydrogen atom is *directly below* the Ga atom in the next layer. This has the effect of forcing the dangling bonds to hang vertically downwards, thus eliminating their directionality. Although this arrangement is not physical and does not give the lowest total energy or satisfy electron counting heuristics, it was found, when compared to other possible bottom surface terminations, to give the best results.[10] The concern that the surplus of charge on one of the two surfaces will set up an electric field which will affect the optical, electronic, and structural properties of the other surface was tested by considering another bottom surface termination which had hydrogen atoms positioned below alternative Ga atoms such that half the Ga dangling bonds are occupied, half are empty and that electron counting is satisfied. Although there was slight charge transfer across the slab in both cases, nearly identical optical properties were obtained. As long as the slab was thick enough an imbalance of charge was found not to be a problem. It was further found that the electronic states which were strongly located on the hydrogen atoms were close to the Fermi energy and that transitions involving these states were largely in the region below 1 eV, which is not of interest. A plane wave basis set corresponding to plane waves with an energy of up to 136 eV was used. Although this relatively small cutoff does not provide an accurate description of the electronic states, in particular of the higher unoccupied electronic states, realistic values of the optical matrix elements and more pertinently the difference in the reflectivity, in the energy range of interest, might be expected. Furthermore test calculations on the optical properties of bulk GaAs agreed well with experiment for energies up to $\sim$6 eV. Only at higher energies was it found that the matrix elements became unreliable.

It was noted, however, that one of the consequences of using a relatively small basis set is that the unoccupied bands are shifted up in energy. This upwards shift in energy of the conduction band states is not to be confused with the error due to the use of density functional theory which results in an underestimate of the band gap. When the cutoff energy is increased it was found that the energies of the conduction band states move downwards until convergence is reached and only then is there an underestimate of the band gap. The energy cutoff which we used in these calculations has the novel property that these two opposing errors cancel so that the band gap of bulk GaAs is in very close agreement with the experimental value. Thus there was no need to apply any kind of "scissors operator" and gave us much more confidence in our results. It should also be noted that by minimiz-

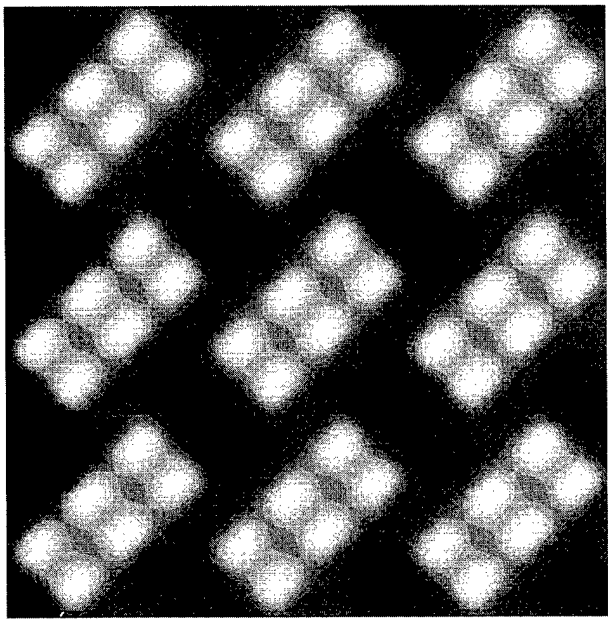


FIG. 3. Simulated scanning tunneling microscopy image for the GaAs(001) $c(4\times4)$ reconstructed surface.

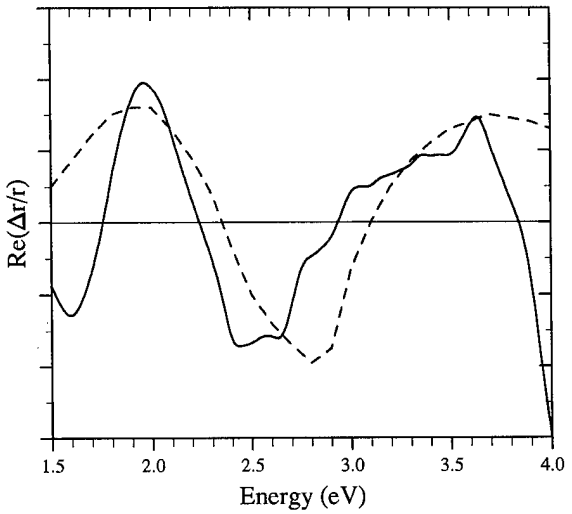


FIG. 4. The calculated RAS spectrum (solid line), and the experimental spectrum (dashed line), for the GaAs(001) $c(4\times4)$ reconstructed surface.

ing the total energy only the wave functions of the occupied states could be obtained. In order to get the conduction band energies and wave functions a separate procedure was used. Using the self-consistent charge density found in the first part of the calculation, an alternative procedure was used under which the band wave functions and energies were found by minimization of the eigenvalues (rather than the minimization of the total energy) under the constraint that each band must be orthogonal to all the bands below it. Optical matrix elements were then calculated for each polarization direction, using the momentum operator, and the imaginary part of the dielectric function calculated according to

$$\varepsilon_2(\omega)=\frac{e^2\hbar^2}{\pi m^2\omega^2}\sum_{v,c}\int_{\mathrm{BZ}}|\langle\psi_{c,\mathbf{k}}|-i\nabla_{\mathbf{r}}|\psi_{v,\mathbf{k}}\rangle|^2 \times\delta(E_{vc,\mathbf{k}}-\hbar\omega)d\mathbf{k}.$$

The real part was then calculated from this equation via a Kramers–Kronig transform. From the dielectric functions it is very straightforward to calculate the complex refractive indices and the reflection coefficients from which the reflectance anisotropy is obtained.[13] It was possible to ignore the nonlocal corrections as these were found to be negligible in the energy range of interest. Care was also taken to ensure that integration of the slab Brillouin zone did not introduce artificial anisotropies. The special point scheme of Chadi and Cohen,[14] with the criterion

$$\sum_{i=1}^{n}\alpha_i A_m(\mathbf{k}_i)=0,\quad m=1,2,\ldots,N$$

was applied to generate the four special points used in the Brillouin zone integration.

III. RESULTS

Prior to calculating the RAS spectrum we were, with very little extra effort, able to generate some theoretical scanning tunneling microscopy (STM) images of the GaAs(001) $c(4\times4)$ reconstructed surface one of which is shown in Fig. 3. This is a realistic simulation involving a tip consisting of four Al atoms, a fixed tunneling gap of 2 Å, a bias voltage of -3 V and a tip–surface interaction given by the Bardeen formalism.[15,16] This image shows the characteristic "brickwork" structure as seen in experiment.[6,7] Furthermore each top layer As atom is individually resolved, indicating tunneling from occupied lone pair orbitals on the As atoms. It was reported by Avery *et al.*[7] that, in their images taken at a bias voltage of -3 V, the central As dimer appears less bright than the outer As dimers. Our simulated image at -3 V does not show this but our images at lower bias voltages do, such that at -1 V the central As dimer is barely visible. In summary the simulated STM images can be considered to add extra credibility to the model used for the $c(4\times4)$ reconstruction.

The calculated RAS spectrum for the GaAs(001) $c(4\times4)$ reconstructed surface is shown in Fig. 4 together with the "low" temperature (189 °C) experimental spectrum of Kamiya *et al.*[9] Note that the magnitude of the calculated spectrum has been scaled to give a similar magnitude to the experimental spectrum. Such a scaling is necessary because the penetration depth of visible light is much greater than the thickness of the slab used in the calculation resulting in the total calculated reflectance being too small, while the difference in reflectances is largely unaffected, thus causing the reflectance anisotropy to be too large. Test calculations on slabs of different thicknesses have supported this explanation. The agreement between calculation and experiment in the energy range of 1.5–3.5 eV is remarkably good. Beyond that, the calculated spectrum falls to zero and negative values fall much more rapidly than is experimentally observed. Part

of the reason may be attributable to the various approximations used in the calculation as a similar departure between calculated and observed RAS spectra was also found for the (2×4) surface (Fig. 2). As it is the low energy structure of the spectra that is used to characterize the two surfaces, the excellent agreement between the calculated and observed spectra in this energy range allows the calculation to be useful in interpreting the spectrum.

In analyzing the spectrum it is useful to note that the anisotropy in the surface dielectric function, $\mathrm{Im}(\varepsilon_{[\bar{1}10]} - \varepsilon_{[110]})$, resembles the RAS spectrum. This has been observed in these calculations and has also been reported in experiment.[17] Thus, features in the reflection anisotropy may be related to specific optical transitions. However, rather than look at specific transitions, $\langle \psi_{c,\mathbf{k}} | \nabla_{\mathbf{r}} | \psi_{v,\mathbf{k}} \rangle$, it is more instructive to examine which valence bands, ψ_v, and conduction bands, ψ_c, contribute most to the dielectric function anisotropy in a specific energy range. The two characteristic features in the theoretical RAS spectrum are the dip at 2.55 eV and the peak at 3.65 eV so we therefore choose an energy range of 0.5 eV centered at each of these energies. The first point to note is that in both energy ranges the contribution of any given valence or conduction band to the dielectric function anisotropy is either positive, zero or negative. The peak (dip) in the dielectric function anisotropy occurs when the positive (negative) contributions are dominant. We then observe that the set of ψ_v which are responsible for the largest contributions to the dielectric function anisotropy (either positive or negative) at the lower energy is very similar to that which are responsible for the largest contributions to the dielectric function anisotropy at the higher energy. The two sets of ψ_c, however, are quite different with those contributing to the dip at 3.65 eV generally being higher in energy as would be expected. An examination of the charge density of these various groups of valence and conduction bands reveals that, for both features, the ψ_v which are responsible for the dielectric function anisotropy are only weakly localized in the vicinity of the surface whereas the ψ_c are strongly localized in the vicinity of the surface. The other valence and conduction bands show no preferential localization. These points would indicate that the dielectric function anisotropy and hence the RAS spectrum has its origins in transitions between bulk like valence band states and surface states in the conduction band. Furthermore there seems to be no fundamental difference in the origin of the dip and peak in the RAS spectrum except for the relative magnitude of the transitions involved. A final point to note regarding the spatial distribution of the electronic states, in the plane of atoms parallel to the surface, is that there seems to be a slight localization of the ψ_v which are responsible for the dielectric function anisotropy on the second layer As atoms labeled 1–8 in Fig. 1 whereas the ψ_c appear to be located around the first layer As atoms.

To further clarify the issue of surface states we have calculated the local density of states (LDOSs) associated with the first layer of As atoms and the GaAs layer below it, the results of which are shown in Fig. 5. There are a number of surface states common to both layers suggesting that they represent bonding states between the top two As layers. However, at both valence and conduction band edges, there are surface states associated with the first layer of As alone. Detailed analysis of the LDOS and optical transition matrix elements show that some of the wave functions that are responsible for the most prominent features in the valence band of the surface LDOS are not those which contribute most significantly to the dielectric function anisotropy. However the distinct features in the conduction band surface LDOS at ~1 eV and ~1.7 eV above the Fermi level are attributable to electronic states which do contribute significantly to the dielectric function anisotropy.

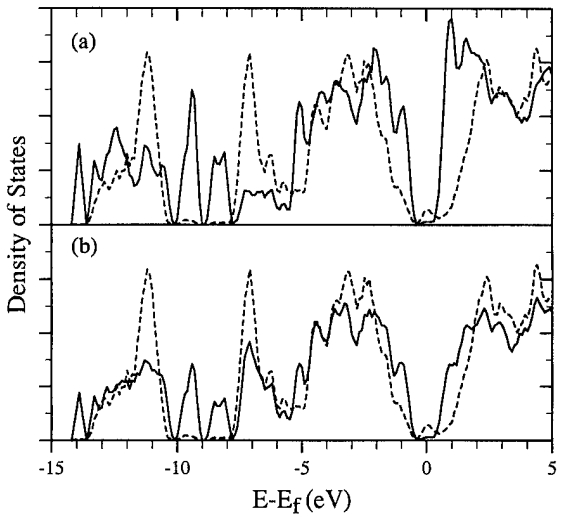


FIG. 5. The local density of states (solid lines) for (a) the first layer of As atoms and (b) the GaAs layer below it. The density of states of bulk GaAs is given by the dashed lines.

IV. CONCLUSIONS

In conclusion, we have shown that RAS spectra for the GaAs(001) reconstructed surfaces calculated using an *ab initio* pseudopotential method within the one-electron framework can produce results which are directly comparable with observed spectra at low energies. Such results can be used to differentiate between different model structures and to yield information about the electronic wave functions responsible for, and the origins of, the observed reflectance anisotropy.

ACKNOWLEDGMENTS

This work was supported by the EPSRC, UK Grant No. GR/J80320. The EPSRC is also acknowledged for the provision of resources on the Intel iPSC/860a supercomputer at the Daresbury Laboratory.

[1]W. L. Mochán and R. G. Barrera, Phys. Rev. Lett. **55**, 1192 (1985).
[2]A. Selloni, P. Marsella, and R. Del Sole, Phys. Rev. B **33**, 8885 (1986).
[3]F. Manghi, R. Del Sole, E. Molinari, and A. Selloni, Surf. Sci. **211/212**, 518 (1989); F. Manghi, R. Del Sole, A. Selloni, and E. Molinari, Phys. Rev. B **41**, 9935 (1990).

[4]Y.-C. Chang, S.-F. Ren, and D. E. Aspnes, J. Vac. Sci. Technol. A **10**, 1856 (1992); S.-F. Ren and Y.-C. Chang, Phys. Rev. B **44**, 13 573 (1991).
[5]H. H. Farrell and C. J. Palmstrøm, J. Vac. Sci. Technol. B **8**, 903 (1990).
[6]D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L.-E. Swartz, Phys. Rev. B **41**, 5701 (1990).
[7]A. R. Avery, D. M. Holmes, J. Sudijono, T. S. Jones, and B. A. Joyce, Surf. Sci. **323**, 91 (1995).
[8]M. Sauvage-Simkin, R. Pinchaux, J. Massies, P. Calverie, N. Jedrecy, J. Bonnet, and I. K. Robinson, Phys. Rev. Lett. **62**, 563 (1989).
[9]I. Kamiya, D. E. Aspnes, L. T. Florez, and J. P. Harbison, Phys. Rev. B **46**, 15 894 (1992).
[10]S. J. Morris, J. M. Bass, and C. C. Matthai, Phys. Rev. B **52**, 16 739 (1995).
[11]J. E. Northrup and S. Froyen, Phys. Rev. B **50**, 2015 (1994).
[12]M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[13]J. D. McIntyre and D. E. Aspnes, Surf. Sci. **24**, 417 (1971).
[14]D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).
[15]J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).
[16]J. M. Bass and C. C. Matthai, Phys. Rev. B **50**, 11 212 (1994).
[17]D. E. Aspnes, Y. C. Chang, A. A. Studna, L. T. Florez, H. H. Farrell, and J. P. Harbison, Phys. Rev. Lett. **64**, 192 (1990); D. E. Aspnes, L. T. Florez, A. A. Studna, and J. P. Harbison, J. Vac. Sci. Technol. B **8**, 936 (1990).

Anisotropic optical reflection by stepped surfaces

P. L. de Boeij,[a] C. M. J. Wijers, and E. Zoethout
Faculty of Applied Physics, Twente University, P.O. Box 217, 7500 AE Enschede, The Netherlands

(Received 22 January 1996; accepted 20 April 1996)

Discrete dipole calculations of the double cell type have been used to study the anisotropic reflection at normal incidence of stepped Si(00$\bar{1}$) 2×1-type surfaces. Only D_B-type steps have been used. The maximum of the anisotropy turns out to be in a direction rotated with respect to the principal axes, and the anisotropy itself depends strongly on terrace width. Further the crossed polarizer configuration is interesting for experimental work since it has no offset. © *1996 American Vacuum Society.*

I. INTRODUCTION

Scanning tunneling microscopy (STM) has given a considerable new impetus to surface science. As a result many detailed features of the surface have become the subject of elaborate studies, one of those features being the step. It extends beyond the scope of this article to give an extensive overview of all research about stepped surfaces mainly because the optical properties of stepped surfaces are the prime interest of this article. There is hardly any literature in this field. The work coming closest is the study by Yasuda and Aspnes[1] about several Si surfaces. Very recent is the work of Mantese *et al.* about Si(001) with high miscut angles.[2] Further, there exists some theoretical work about the influence of missing row reconstructions on the surface optical response of Au (Ref. 4) and Si.[3] The aim of this article is to study the influence of steps on the optical response of stepped Si (00$\bar{1}$) (2×1) surfaces and to understand the extent steps can be ignored when measuring and interpreting the optical anisotropy of this type of surface. This question is relevant since samples used in practice always have some kind of miscut. The steps are well-defined structures on top of a surface. This opens unique possibilities to study the short range interactions by varying the density of the steps. Since steps are by their very nature microscopic structures on the surface, the best way to study them is by means of a microscopic model. Unfortunately, terraces with a width of more than 10 lattice spacings are beyond the capacity of most microscopic methods particularly electronic structure calculation methods. We will show in this article that discrete dipole types of calculations, using the double cell technique,[5,6] are able to handle such high Miller index surfaces. However the cellular polarizabilities, which are the basic input for this class of calculations, are not yet available from *ab initio* calculations. The present calculations have been performed using model polarizabilities.

II. EXTENDED DOUBLE CELL METHOD

How to calculate the optical response of the surface of a semi-infinite crystalline system using the discrete dipole model has been described by us in a number of articles.[5,6] The basic idea is that the entire system becomes represented by a number of discrete sources: the point dipoles $\mathbf{p}_l$, located at sites $\mathbf{r}_l$. Each dipole becomes linearly induced by the local field $\mathbf{E}_{\mathrm{Loc},l}$ at the dipole position, according to its polarizability α_l. In this way, a set of simultaneous linear equations is obtained for any system which must be solved numerically. Hence very large systems and especially semi-infinite systems cannot be treated in a straightforward way. Only if symmetry can be taken into account, being the case for crystalline systems, semi-infinite systems can be transformed to a problem of manageable proportions. The double cell method[5,6] is one of the most efficient methods to solve the discrete dipole equations for an arbitrary choice of crystalline surfaces.

Two kinds of symmetry are exploited by the double cell method. The translational symmetry remains intact for directions parallel to the surface. Deep inside the crystal, however, the influence of the surface can be ignored, and there effectively, three-dimensional translational symmetry exists. There the response of all bulk dipoles can be efficiently described by normal modes. For this reason, the system is split in the double cell method into a surface and a bulk part, described by two normal modes. Two complex numbers, governing the strength of these normal modes, are the only free parameters for the bulk. All dipole–dipole interactions of the system are taken into account if we consider the interactions between the characteristic dipoles $\mathbf{p}_i$ of the surface layer and the bulk normal modes. The characteristic dipole is a single dipole which can represent an entire lattice plane of dipoles as a result of parallel translation symmetry. Upon solution of the dipole–dipole (electrodynamic) interactions governing the system, the individual dipoles of the system are obtained. Now a rather straightforward procedure can be used to obtain the remote optical response (reflection).

To enable treatment of the the optical response of stepped surfaces, the double cell method needs adaptation to handle systems having a triclinic bulk unit cell with two or more dipoles. Until now this method could handle only systems having a simple monoclinic bulk unit cell (with one dipole)[5] spanned by $\mathbf{s}_1$,$\mathbf{s}_2$ and the perpendicular axis $\mathbf{d}_B$. If more than one dipole had to be taken into account, we reduced them to effectively one, by means of the commensurability theorem.[5] Most of the adaptation is shown in Fig. 1 and concerns only the true bulk (normal mode) part of the double

[a]Electronic mail: boeij@tn.utwente.nl

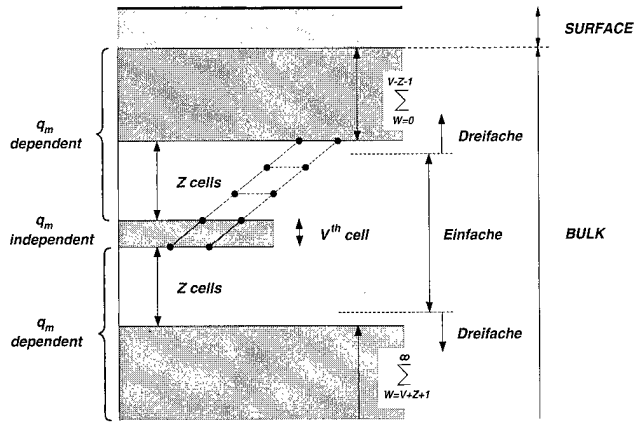


FIG. 1. Schematic view of the double cell geometry with thin triclinic bulk unit cells and the $\pm Z$ layers using the one-fold integral transform.

cell method. The key quantity for this type of discrete dipole calculations is the transfer tensor $\mathbf{f}$, governing the interaction between a dipole and a plane of dipoles. All bulk transfer tensors were evaluated by means of the Ewald threefold integral transform. This integral transform expands the transfer tensor into a series of partial waves controlled by the wave vectors $\mathbf{k}_{pq}$ and transverse projectors $\mathbf{d}_{pq}$.[6] With triclinic bulk unit cells the distance between neighboring cells in a direction perpendicular to the surface can become very small, turning down the efficiency of the threefold integral transform. Therefore, for a certain number Z, we calculate the interactions due to planes being at most Z bulk unit cells below or above the bulk cell of interest V by means of Ewald's one-fold integral transform[7] (Fig. 1). This requires a rederivation of the secular matrix/determinant and the interaction equations. Technically the derivation is almost exactly as given in,[6] but we get a slightly different expression for the secular matrix (determinant):

$$\begin{aligned}\vec{\mathscr{M}}_{vw} &= \vec{\alpha}_v^{B-1}\delta_{vw} - \alpha_o^{-1}\sum_{z=-Z}^{+Z}\vec{\mathbf{f}}_{uVw(V+z)}\exp(iz\mathbf{k}_m\mathbf{d}_B) \\ &+ \alpha_0^{-1}\sum_{pq}\left[\frac{\exp\{[i\mathbf{k}_{pq}(\mathbf{r}_v^B-\mathbf{r}_w^B+Z\mathbf{d}_B)-\mathbf{k}_mZ\mathbf{d}_B]\}}{1-\exp[i(\mathbf{k}_m-\mathbf{k}_{pq})\mathbf{d}_B]}\vec{\mathbf{d}}_{pq}\right. \\ &\left.+\frac{\exp\{[i\underline{\mathbf{k}}_{pq}(\mathbf{r}_v^B-\mathbf{r}_w^B-Z\mathbf{d}_B)+\mathbf{k}_m\mathbf{d}]\}}{1-\exp[i(\underline{\mathbf{k}}_{pq}-\mathbf{k}_m)\mathbf{d}_B]}\right]\vec{\underline{\mathbf{d}}}_{pq}, \qquad (1)\end{aligned}$$

where we have used the following definition for the normal modes:

$$\mathbf{p}_{vV} = \sum_{m=1}^{M}\nu_m\mathbf{u}_{mv}e^{iV\mathbf{k}_m\mathbf{d}_B}, \qquad (2)$$

where $\mathbf{k}_m$ is the wave vector, $\mathbf{u}_{mv}$ the eigenvector, and ν_m the strength of the normal modes. Further, we use for the bulk sites of the dipoles $\mathbf{r}_{vV}$ the definition $\mathbf{r}_{vV}=\mathbf{r}_v^B+\mathbf{d}_S+V\mathbf{d}_B$, where $\mathbf{r}_v^B$ are the sites of the dipoles within the bulk unit cell, and $\mathbf{d}_S$ is the vector connecting the origins of bulk and surface unit cell. Equation (1) reduces to the result given in Ref. 6, by making $Z=0$. The crucial difference is in the first summation. All transfer tensors $f_{vVw}(V+z)$ are calculated now, using the one-fold integral transform. In a similar way, we get the expression for the four different types of matrix elements of the double cell interaction matrix $\mathscr{M}$:

$$\mathscr{M}_{ij}^{SS} = \vec{\alpha}_i^{-1}\delta_{ij} - \alpha_0^{-1}\vec{\mathbf{f}}_{ij}, \qquad (3)$$

$$\mathscr{M}_{im}^{SB} = \alpha_0^{-1}\sum_{w=1}^{N_B}\left[\sum_{pq}^{\infty}\left(\frac{\exp\{i[\underline{\mathbf{k}}_{pq}(\mathbf{r}_i-\mathbf{r}_w^B-Z\mathbf{d}_B-\mathbf{d}_S)+Z\mathbf{k}_m\mathbf{d}_B]\}}{1-\exp[i(\mathbf{k}_m-\underline{\mathbf{k}}_{pq})\mathbf{d}_B]}\right)\vec{\underline{\mathbf{d}}}_{pq}+\sum_{W=0}^{Z-1}\exp(iW\mathbf{k}_m\mathbf{d}_B)\vec{\mathbf{f}}_{iwW}\right]\mathbf{u}_{mw}, \qquad (4)$$

$$\mathscr{M}_{nj}^{BS} = -\frac{ik^2}{2\epsilon_0|\mathbf{s}_1\times\mathbf{s}_2||k_z|}\exp[i\mathbf{k}(\mathbf{d}_s-\mathbf{r}_j)]\hat{\mathbf{t}}_n^T, \qquad (5)$$

$$\mathscr{M}_{nm}^{BB} = -\frac{ik^2}{2\epsilon_0|\mathbf{s}_1\times\mathbf{s}_2||k_z|}\hat{\mathbf{t}}_n^T\mathbf{P}_m^B, \qquad (6)$$

where $\mathbf{r}_i$ is a surface dipole site, $\hat{t}_n$ the direction of the analyzer and P_m^B is defined as

$$\mathbf{P}_m^B = \mathbf{P}_m^B(\mathbf{k}) = \frac{\Sigma_{w=1}^{N_B}\exp(-i\mathbf{k}\mathbf{r}_w^B\mathbf{u}_{mw})}{1-\exp(i[\mathbf{k}_m-\mathbf{k}]\mathbf{d}_B)}. \qquad (7)$$

The superscript labels S and B will always refer to surface and bulk, respectively. The only place where we see explicit changes due to the one-fold integral transform, is for the SB type of matrix elements. We will use frequently the anisotropic azimuth Ω, for this article, the angle between the electric field and $\hat{\mathbf{x}}$. The inhomogeneous vector of the interaction equations remains unchanged.[6] Equations (1)–(7) comprise the bulk of the modifications necessary to let our software handle oblique bulk unit cells with any number of dipoles.

When more than (effectively) one dipole occupies the bulk unit cell, the calculation of the normal mode parameters slows down substantially. Therefore, the performance of the root searching procedure, used in the normal mode part of our discrete dipole package, has been improved. Historically, a general purpose Broyden routine has been used to find the roots from the modulus of the determinant, but this turned out to be far too slow. In a second version, we used a quadratic (complex) extrapolation of the secular determinant while searching for zeros. In the present version, we have used a method which does not use the value of the determinant as such. In this method, we bring the secular matrix to triangular shape, using full pivoting of the matrix elements.

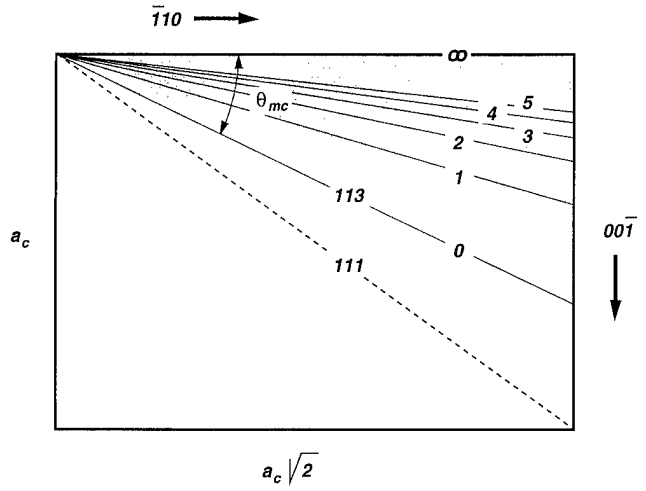


FIG. 2. Intersection of the $\bar{1}10$-$00\bar{1}$ plane and the conventional cube showing miscut angle θ_{mc} (the integer numbers refer to N_D).

For the triangular secular matrix we know we can write the determinant as:

$$\|\mathscr{A}\| = \prod_{i=1}^{N_B} A_{ii}. \tag{8}$$

This product will become zero if and only if one of the diagonal elements becomes zero. We are looking for zeros of the diagonal elements: since we have done full pivoting, the zeros will always be found at the lower right corner of the matrix. The point to notice is the matrix element which becomes zero also approaches zero in a linear way even for two-fold degenerate matrices. Using a Regula–Falsi-like technique for the last two diagonal matrix elements, the relevant zeros were found very fast. The advantages of this procedure are clear: It handles matrices of any size, uses linear extrapolation, and creates highly integrated code, since we can use the triangular matrix at the root immediately to produce the normal mode eigenvectors.

III. STEPPED SURFACES AND GEOMETRY

The basic geometry of stepped surfaces on Si($00\bar{1}$), has been given by Chadi,[8] and we follow his treatment and arguments here. We use in this study only double layer stepped surfaces. Of the two types of double layer steps, only the D_B-type was found to be stable,[8,9] so we use this type of step to create terraces. Terraces enclosed between D_B-type steps have all surface dimers oriented in the same direction, parallel to the step edge. For this type of stepped surfaces, the origin of a change in anisotropy can only be in the amount of step edges. If we would have used single step surfaces, there would have been two types of terraces with dimer orientation perpendicular to each other—then the change in anisotropy would depend on both the amount of step edges and the ratio of the surface areas with different orientation.

Chadi gives only isolated steps in his article. A topview of the D_B-type step is shown in the right panel of Fig. 5. This topview shows the atomic lattice. The atomic lattice as such cannot be used directly for discrete dipole calculations. Ideally, each atom should be represented by one dipole; however, such correspondence does not work.[10] For semiconductors, the discrete dipole model overestimates the short range interactions. Partially this effect can be overcome by assigning one dipole to each couple of atoms, as Mochán and Barrera did. In this study, we do the same, but this forces us to devise a suitable dipole lattice. The dipole geometry—used by us—is shown in Figs. 3 and 4. The dipole site containing the rebonded atom is, to a certain extent, questionable. From the position of the rebonded atom itself, we can choose the middle of either of two downward directed bonds, both choices being arbitrary. The reconstruction of the Si($00\bar{1}$) 2×1 surface is commonly called *rows of dimers* reconstruction, those rows being perpendicular to the D_B steps. In addition, we use *lines of dimers*, being all dimers on the same line parallel to the steps. Rows of dimers belonging to successive terraces have to shift over a $a/4$ distance parallel to the step edge. On the next terrace, this shift can be the same

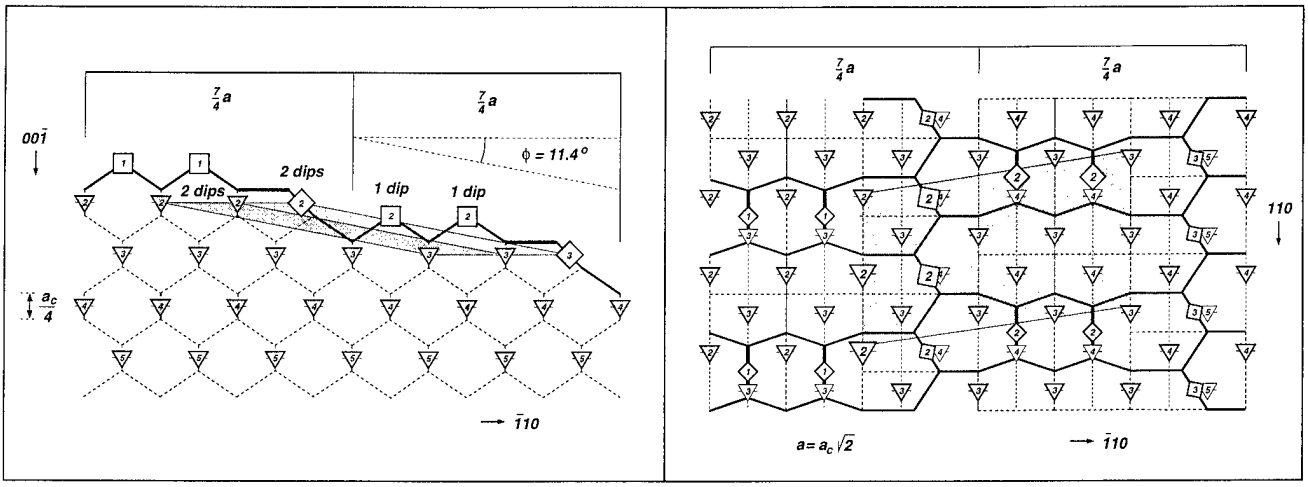


FIG. 3. Location of dipole sites used in the calculations. Left: sideview Si ($00\bar{1}$), 2 dimer lines seen from (110). Right: topview Si ($00\bar{1}$), 2 dimer lines seen from (001).

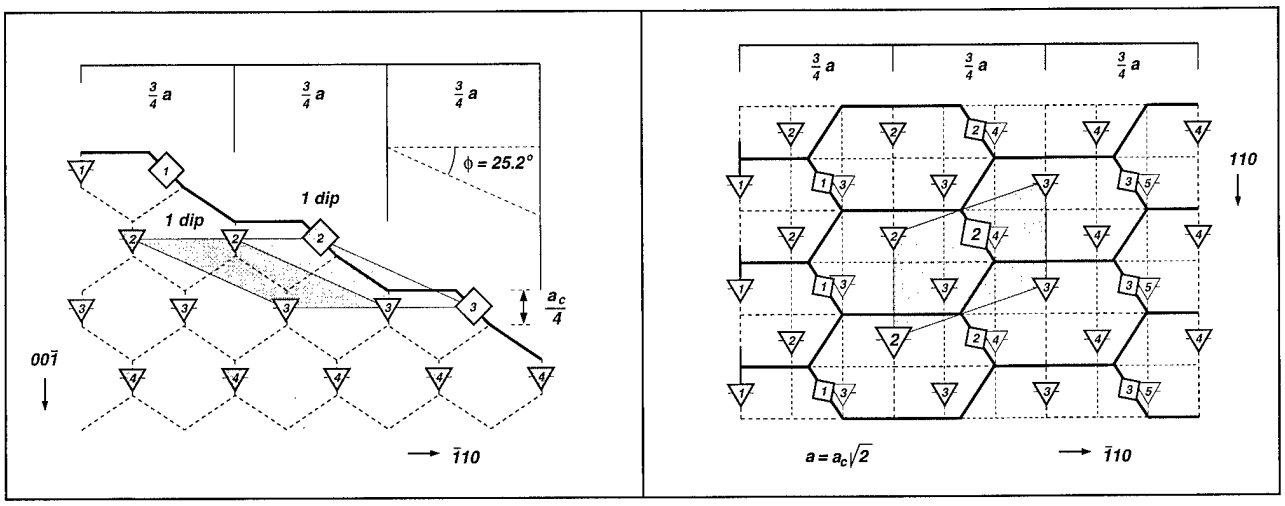


FIG. 4. Location of dipole sites used in the calculations. Left: sideview Si (113) seen from (110). Right: topview Si (113) seen from (00$\bar{1}$).

or reversed, resulting in two different types of surfaces. We chose a continued shift, yielding an oblique surface unit cell.

We use D_B-type stepped surfaces, having the miscut angle θ_{mc} in the ($\bar{1}$10) (00$\bar{1}$)-plane of the original 2×1 reconstructed (00$\bar{1}$)-surface, as shown in Fig. 2. Only planes in the shaded area are considered. The limiting cases are the (00$\bar{1}$)- and (11$\bar{3}$)-surface—the latter plane having only D_B-type steps and no dimers. The terraces themselves were supposed to have the (00$\bar{1}$) 2×1 reconstruction. The type of surface morphology obtained in this way is valid for orientations as steep as (117). Beyond that point, nonrebonded step edges are introduced and surface periodicity becomes less well defined [at least until (114), $\theta_{mc} = 19.5°$].[11] So our modeling of terraces, having less than 2 dimer lines of width, is not very realistic, but we have added them, nevertheless, to get a better feeling for possible trends.

To obtain a stepped surface having a terrace width containing N_D dimer lines requires a miscut angle θ_{mc} given by

$$\tan(\theta_{mc}) = \left(\frac{\sqrt{2}}{3+2N_D}\right). \tag{9}$$

For the cases studied here, we list values of θ_{mc} in Table I.

The coordinates of the sites will be given first in a base frame spanned by the (110), ($\bar{1}$10), (00$\bar{1}$) directions. For all surfaces with $N_D \neq 0$, there are two dipoles in the bulk unit cell and $2+N_D$ dipoles in the free surface unit cell. The two bulk dipoles are located at (0,0,0) and at $a(\frac{1}{2},0,0)$ with respect to the origin of the bulk unit cell, where $a=\sqrt{2}a_c$, a_c being the conventional lattice parameter (5.43 Å for Si). Those sites will not be affected by a rotation to the surface oriented frame. The two dipoles marking the step have sites $\boldsymbol{\tau}_S^{1,2}$, given in the base frame by

$$\boldsymbol{\tau}_S^1 = a[0,\tfrac{1}{2}(1-\tfrac{1}{8}\sqrt{2}),0],$$

$$\boldsymbol{\tau}_S^2 = a[\tfrac{1}{2},\tfrac{1}{2}(1-\tfrac{1}{8}\sqrt{2}),0]. \tag{10}$$

The N_D terrace dipoles have sites $\boldsymbol{\tau}_S^i$ given in the base frame by

$$\boldsymbol{\tau}_S^{2+i} = a[\tfrac{1}{8},1+\tfrac{1}{2}(i-1),\tfrac{1}{16}\sqrt{2}]. \tag{11}$$

Those sites are with respect to the origin of the free surface unit cell. The origins of bulk and free surface unit cell are connected by $\Delta\boldsymbol{\tau} = -a\hat{\mathbf{y}}$ and the third vector $\mathbf{d}_B$ spanning the bulk unit cell is in the base frame $-a\hat{\mathbf{y}}$. From these data, the proper coordinates in the (stepped) surface oriented frame are obtained by a coordinate transformation over θ_{mc}, as described. In the surface frame the parallel translational symmetry is governed by the spanning vectors $\mathbf{s}_1$ and $\mathbf{s}_2$:

$$\mathbf{s}_1 = a(1,0,0),$$

$$\mathbf{s}_2 = a[\tfrac{1}{4},\tfrac{1}{4}\sqrt{2+(3+2N_D)^2},0]. \tag{12}$$

This description of the geometry does not include the two limiting cases (00$\bar{1}$) and (11$\bar{3}$). The first should be obvious and the (11$\bar{3}$) case follows basically the treatment given here but is *halved* in the (110) direction.

The last part of the description of the configuration used concerns the polarizabilities. All polarizabilities will be derived from bulk optical data using Clausius–Mossotti in the following way:

$$\left(\frac{\alpha}{\alpha_0}\right) = \left(\frac{a_c}{a}\right)^3 \frac{3}{16\pi}\left(\frac{\epsilon-1}{\epsilon+2}\right), \tag{13}$$

TABLE I. θ_{mc} as a function of # of dimer lines N_D.

N_D	Miscut angle θ_{mc} (°)
0	25.239 402
1	15.793 169
2	11.421 754
3	8.930 142
4	7.326 037
5	6.208 545
∞	0.000 000
(111)	35.264 39

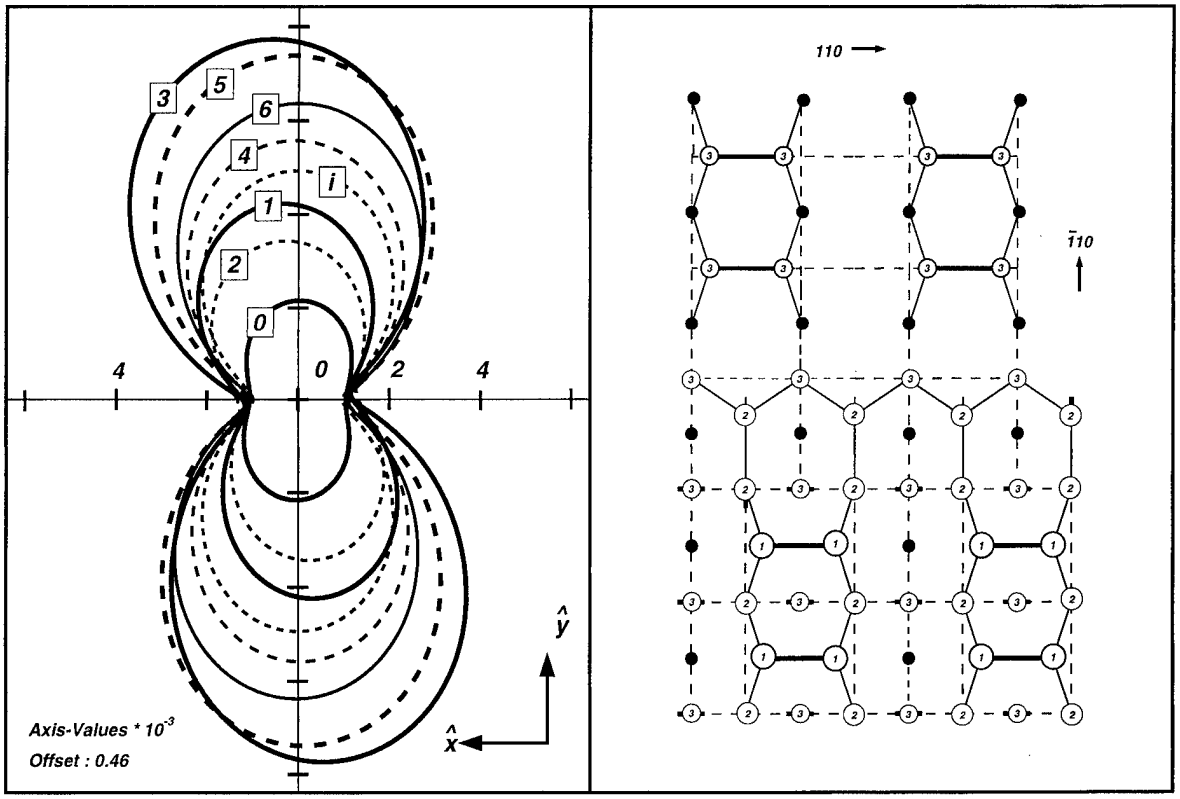


FIG. 5. Left: Azimuthal plot pp reflectance as a function of Ω. Right: topview separate D_B step (atomic sites).

where ϵ is the relative dielectric constant, a_c/a depends on the surface [$(11\bar{3})$ is different], and $\alpha_0=4\pi\epsilon_0 a^3$. The numerical value for α_0 equals $5.039\,857\,3\times10^{-38}$ F m^2 for all cases except for (113). There the value is $6.299\,821\times10^{-39}$ F m^2. Calculations are isotropic if only isotropic polarizabilities are used and show the structural contribution to the anisotropy. To investigate a possible electronic structure influence, anisotropic polarizabilities are required. We have restricted this anisotropy to the surface dipoles, the edge and terrace ones. Taking the z direction parallel to the dimer axis, we introduce anisotropy as

$$r_{\mathrm{Ani}}=\alpha_{zz}/\alpha_{xx,yy},$$
$$\alpha_{yy}=\alpha_{xx}. \tag{14}$$

We have kept in our calculations this intrinsic anisotropy modest, effectively only 2%. Hence, we use only $r_{\mathrm{Ani}}=1.0$ (isotropic), and $r_{\mathrm{Ani}}=1.02$ (anisotropic). The anisotropic polarizability has to be rotated to its proper position. This requires interchange of indices for the terrace dipoles. The edge dipoles have been rotated twice by means of an Euler rotation. First they have been given a rotation over the Euler angles $\Psi=-45°$, $\theta=-45°$ to get the correct values for the base frame:[12]

$$\alpha^L_{pq}(\theta_M)=\sum_{i,j} S(\phi,\theta,\Psi)_{pi}\alpha^M_{ij}[S(\phi,\theta,\Psi)^{-1}]_{jq}, \tag{15}$$

where $S(\phi,\theta,\Psi)$ is the Euler transformation matrix as given by Goldstein.[13] The second rotation from the base frame to the surface frame is over the angle θ_{mc} and is applied to the entire system. After that, the polarizability tensors no longer contain zero valued components. Numerical values for the anisotropic components of the polarizability prior to rotation are $\alpha_{zz}=1.914\,052\times10^{-2}+i2.085\,819\times10^{-4}$ and $\alpha_{xx,yy}=1.876\,522\times10^{-2}+i2.044\,921\times10^{-4}$ in units of α_0 [except (113)].

IV. RESULTS

Using the previously mentioned geometry and optical input data, model calculations have been performed for perpendicular incidence. We have used p type of polarization for the incoming light, letting the direction of polarization coincide with anisotropic azimuth Ω. The frequency for all calculations has been $\hbar\omega=3.0$ eV unless stated otherwise. The polarizabilities were also calculated at this frequency. The first results of this type (pp combinations) are shown in Fig. 5, left panel. For those calculations, all surface dipoles have been given anisotropic polarizabilities. The numbers shown in boxes in the curves themselves indicate the number of dimer lines in the reconstruction. The two limiting cases are given by the label (i) for $(00\bar{1})$, and 0 for $(11\bar{3})$. Since the anisotropy is weak as compared to the average signal, an offset of 0.46 has been subtracted to emphasize the results. The right panel shows the atomic configuration of a single D_B step, shown in the same orientation as the left panel. It is

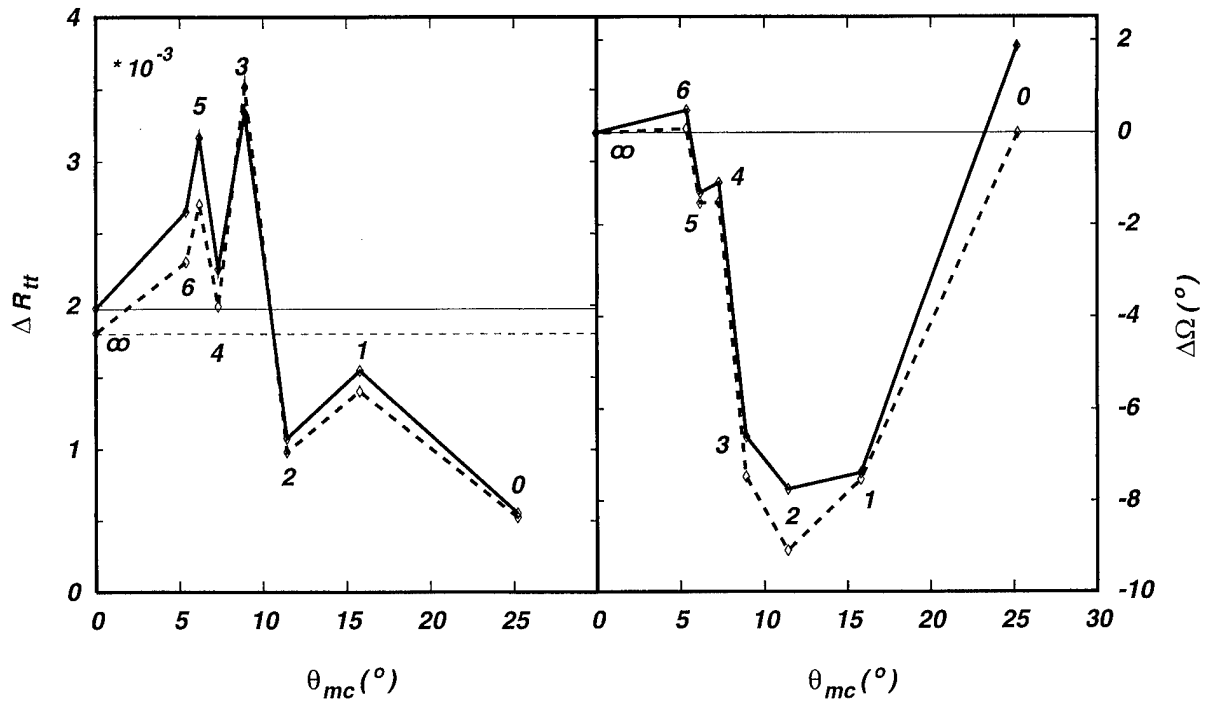


FIG. 6. Left: Top–top modulation of R_{pp} (reference anisotropies indicated by horizontal lines). Right: anisotropic azimuth offset $\Delta\Omega$. Solid line: anisotropic results. Dashed line: isotropic results.

immediately clear that the pp plots are not exactly lattice oriented. Neither the bulk lattice, nor the step seems to control the orientation of the plots.

The maximum value of the pp reflectance is obtained for the polarization direction approximately perpendicular to the steps. We know from experience that single linear chains of dipoles have increased emission when the polarization direction is parallel to the chains. If we assume the dimer rows will act like separate chains of dipoles, we can expect that for these stepped surfaces, the maximum reflectance is obtained for polarization parallel to the rows. We will use the anisotropy of the original $(00\bar{1})$ surface as a reference in the remainder of this article. It is surprising to see that this reference anisotropy is not the maximum anisotropy. By intuition, we would expect the interruption of the dimer rows by the steps to decrease the reference anisotropy of the unstepped surface. Apparently this is not the case. Why the steps sometimes happen to increase, or to decrease the reference anisotropy has no simple explanation. It is clear, nevertheless, that steps have a strong influence on this anisotropy. They can increase or decrease it by a factor of two. Finally, we notice that the plots are tilted into the same direction as the shift of the dimer rows, belonging to consequent terraces. This is also expected.

The maximum terrace width shown here is for $N_D=6$. Technically we can go well beyond $N_D=10$, but at the expense of highly increased runtimes. The reason for the increment is the number of dipoles required by the matching layer between the free surface cell and the bulk normal mode regime. This number increases from 60 to 90, going from $N_D=1$ to $N_D=6$. Even those large numbers are not enough to keep the convergence rate at the level observed for the $(00\bar{1})$ and $(11\bar{3})$ limiting cases. We have always used such a number of matching dipoles that convergence had no influence on our results.

To have more direct access to the relevant parameters, we have made a fit through the calculated data. We have used as a fit function:

$$R_{pp}(\Omega)=A+B\cos(N_\Omega\Omega)+C\sin(N_\Omega\Omega), \tag{16}$$

where we have taken $N_\Omega=2$. After performing the least squares procedure, we arrived at a very good fit. The fit parameters have subsequently been transformed to more directly understandable data as follows:

$$R_{pp}(\Omega)=A+\Delta R_{tt}\sin[N_\Omega(\Omega-\Delta\Omega)],$$

$$\Delta R_{tt}=\sqrt{B^2+C^2},$$

$$\Delta\Omega=\arctan(B/C)/N_\Omega. \tag{17}$$

We will focus upon the top–top modulation ΔR_{tt} and the azimuthal shift $\Delta\Omega$. The solid curves in Fig. 6 belong to the anisotropic cases, shown as azimuthal plots in the left panel of Fig. 5. From the top–top modulation, we learn that the anisotropy is not a smooth function of the miscut angle or terrace width. There is something like an even–odd behavior superimposed on top of a simple curve having a single maximum. This maximum shows up for about 3 dimer lines of terrace width. It turns out that for pp reflectance the top–top modulation cannot be used experimentally as an indicator for the terrace width. The same holds for the azimuthal shift. The influence of the azimuthal shift $\Delta\Omega$ has been explicitly

omitted by Yasuda and Aspnes in their article.[1] The only misoriented (001) surface in their paper has a $\theta_{mc}=4°$ corresponding to a terrace width of 9 dimer lines, just beyond the region where we find the strongest shifts. This justifies their assumptions. The minimum of the azimuthal shift is at $N_D=2$ and these shifts arrive faster at the reference level (≈ 0) for increasing terrace width, than the corresponding top–top modulation.

To figure out how much of the effects found is structural (geometry, dipole sites), and how much is electronic (polarizability), we have repeated the calculations with isotropic polarizabilities throughout. Those isotropic results are also shown in Fig. 6 as the dashed curves. Although the results clearly change, the overall pattern remains unaffected. Only for the azimuthal shift $\Delta\Omega$, we see that for the $(11\bar{3})$ case the tilt disappears. This also happens for terraces with $N_D \geq 6$, but there the shift is already very low. For the values of r_{Ani} used in this article (about 1.02), the anisotropy of stepped surfaces is predominantly structural. For a larger value of r_{Ani} stronger electronic contributions will be expected. From these calculations we get a first impression how a larger r_{Ani} will affect the anisotropy. Better polarizability data have to be obtained from other calculations.

Since most of the observed behavior has to be understood as structural, we have also tried to figure out whether wiggling plays a role. Those rapid oscillations of the dipole strength have been observed by us before[5] and we associate them with structural effects. The new phenomenon, which might play a role here is lateral wiggling, wiggling along the surface. Previous observations of wiggling were always perpendicular to the surface. The even–odd behavior of the top–top modulation might be caused by it. Since wiggling depends strongly on frequency, we have repeated the anisotropic calculations with $\hbar\omega=6.0$ eV. The results indicate that the influence of wiggling is weak. A more direct observation is given by the individual dipole strengths (in Fig. 7). All dipole strengths of the free surface layer are shown as a function of the y-coordinate. Indeed lateral wiggling with a two-fold period can be observed.

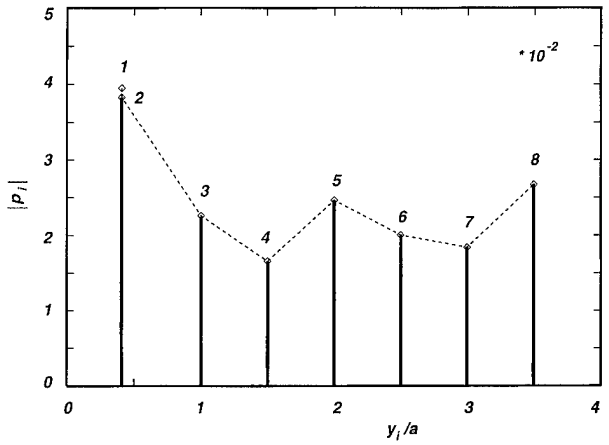


FIG. 7. Modulus of dipole strength $|\mathbf{p}|_i$ as a function of y. $|\mathbf{p}|_i$-axis values to be multiplied by 10^{-2}.

Discrete dipole calculations depend critically on the choice of the dipole geometry. To verify how sensitive our present calculations are as to this point, we have repeated the calculations with the edge dipole moved to the position of

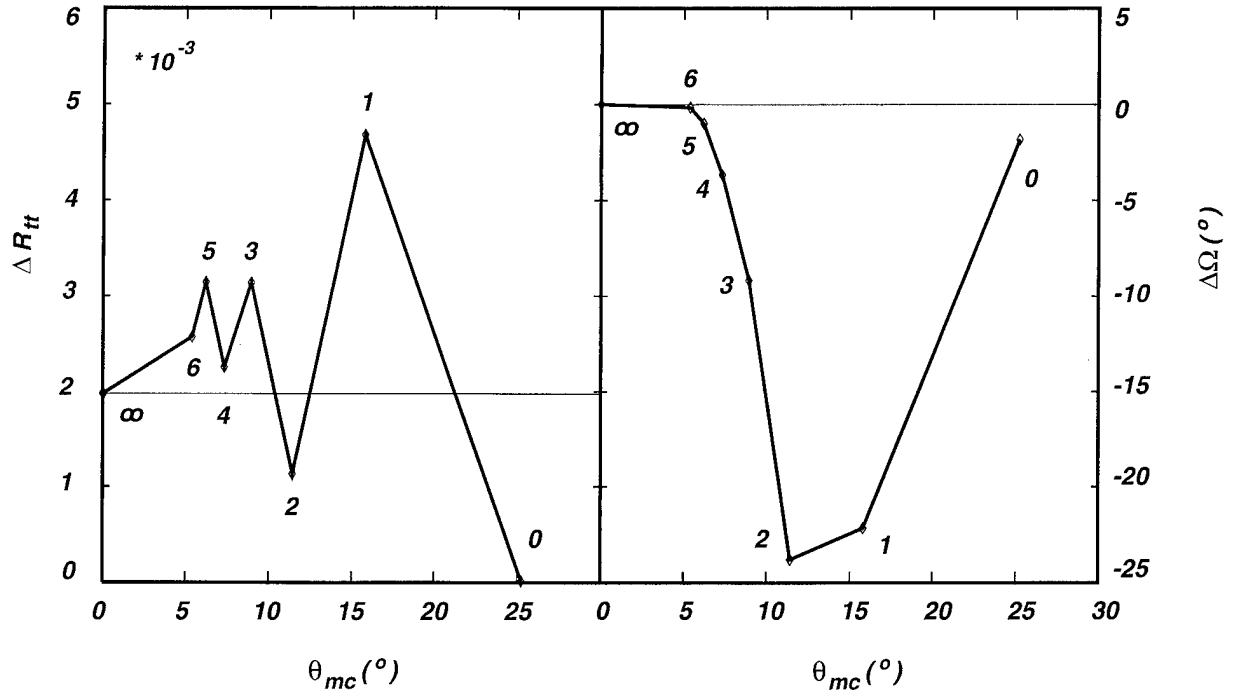


FIG. 8. Left: top–top modulation of R_{pp} (horizontal line: reference). Right: anisotropic azimuth offset $\Delta\Omega$. Anisotropic results, $\hbar\omega=3.0$ eV edge dipole at rebonded atom position.

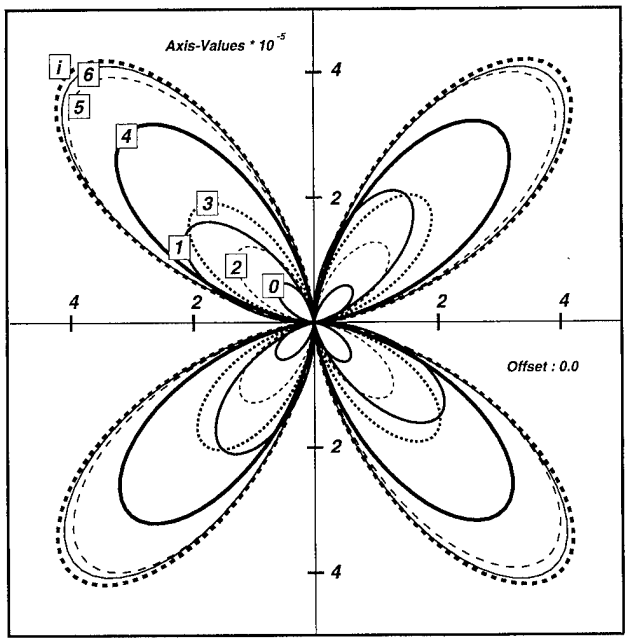


FIG. 9. Azimuthal plot ps reflectance as a function of Ω.

the rebonded atom. To remove asymmetry due to the tilting of the polarizability tensor of this site, we have given it the same polarizability tensor as the dimers. The results are shown in Fig. 8. There are a few major differences. The top–top modulation for the N_D=1-case has increased by a factor of 3, the anisotropy for the N_D=0-case has virtually disappeared, and the azimuthal shifts have increased systematically by a factor of 2. Those changes are large, but the overall behavior has not been affected by the shift of the edge dipole. When $N_D \geqslant 3$, the changes are not significant.

For the stepped surfaces treated here, another outspoken new feature of the calculation is that the polarization no longer remains in the plane of incoming polarization. For this reason, we have investigated also this ps-type (crossed polarizer) configuration. The azimuthal results for the anisotropic case are shown in Fig. 9. The result is a series of four-lobed patterns. Those patterns are obtained without using an offset. At first glance, those patterns look fundamentally different from the ones shown in Fig. 5, but this is only seemingly. The major point is that there is really no offset. The complex ps-reflection coefficient has the same two period behavior as the pp component, but since it is symmetric around zero, the modulus-procedure $R=r^*r$ turns it into a four period phenomenon.

The ps-type of reflectances have again been fitted by us by means of Eq. (16), using now N_Ω=4. The results obtained this way, are shown in Fig. 10. Again the solid lines represent the anisotropic cases. A major result shows up already after inspection of the top–top modulation. For N_D>2 this modulation is a monotonous function of N_D. The even–odd behavior is absent. As compared to their pp-counterparts, the modulation has dropped by two orders of magnitude. Further the $N_D=\infty$ [original (001)] results turn out to be the highest now. Only in the results for the azimuthal shift, still the even–odd behavior is visible but not as pronounced as for the pp cases. The differences between the azimuthal shift data comparing pp and the ps-cases are smaller than between the corresponding top–top modulation data. We have also

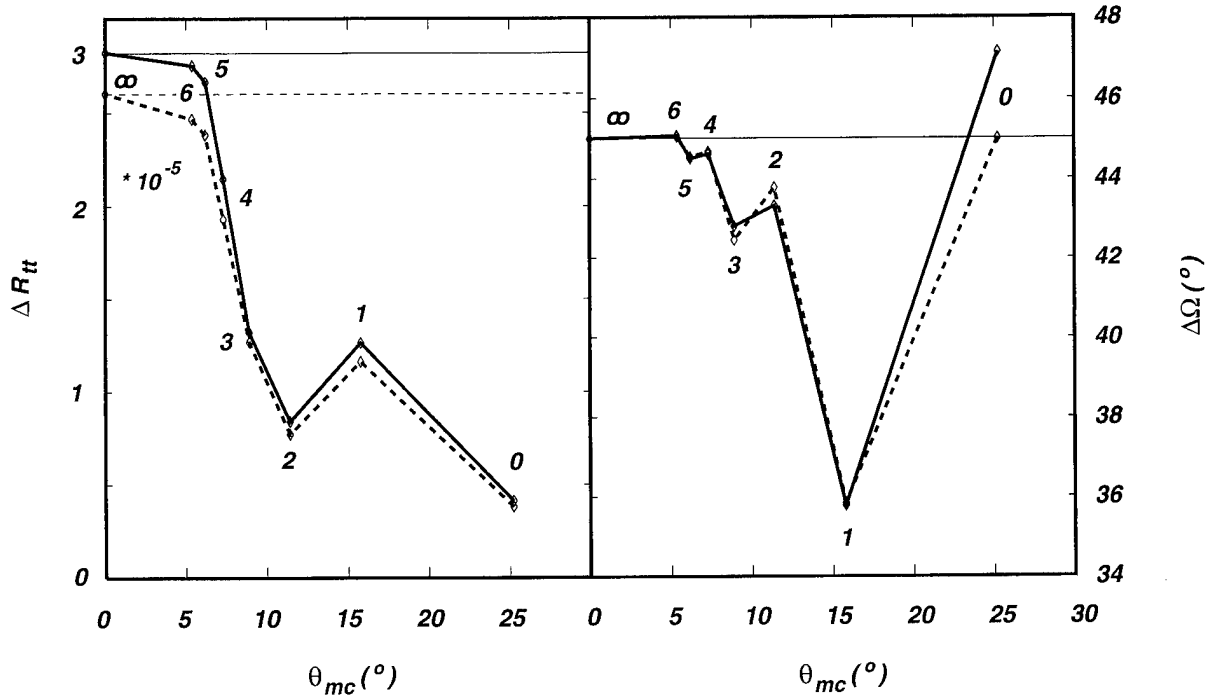


FIG. 10. Left: top-top modulation of R_{ps} (horizontal lines: reference). Right: anisotropic azimuth offset ΔΩ. Solid line: anisotropic results. Dashed line: isotropic results.

repeated the calculations using isotropic polarizabilities. Again the influence is small.

The major conclusion from the ps type of reflection coefficients is in the way experiments can be done. Despite the fact that the ps-type results are in strength two orders of magnitude below the pp-type results, those data can be obtained without offset. This means that the full sensitivity of the detection system can be used. If our calculations turn out to be reliable, it means that the ps results can be easier correlated with the terrace width, since the ps-response curve is monotonic.

V. CONCLUSION

In this article, we have shown a number of model calculations based on the discrete dipole model for the optical response of stepped surfaces. We have shown that it is possible to perform those calculations for terrace widths, having up to 10 dimer lines. Most of the step influence is gone for terraces being more than 6 dimer lines wide. Measurements on Si $(00\bar{1})$, with up to 4° miscut, need no correction for those effects. For higher miscut angles, the influence of steps becomes too strong to be ignored. The step contribution to the anisotropy can be as large as the anisotropy of the original Si $(00\bar{1})$ 2×1 surface. That is the regime where interesting fundamental work can be done. The point where our calculations and the experimental data will possibly deviate will give an indication for the effective range of the electrodynamic interaction. This asks for further experimental work into this direction. For perpendicular incidence, as studied here, the main axes of the polarization plots do not follow the crystallography of the surfaces. From our calculations, follows that the easier way to measure the influence of steps on optics is by measuring in crossed polarizer (ps) mode. The expected advantages are in a better sensitivity and in easier interpretation of results.

[1]T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, J. Vac. Sci. Technol. A **12**, 1152, (1994).
[2]L. Mantese, U. Rossow, and D. E. Aspnes, *Proceedings of the ACSI-3 Conference*, 1995 (to be published).
[3]G. P. M. Poppe, H. Wormeester, A. Molenbroek, C. M. J. Wijers, and A. van Silfhout, Phys. Rev. B **43**, 12 122 (1991).
[4]S. Wang, W. L. Mochán, and R. G. Barrera, Phys. Rev. B **42**, 9155 (1990).
[5]G. P. M. Poppe and C. M. J. Wijers, Physica B **167**, 221 (1990).
[6]C. M. J. Wijers and G. P. M. Poppe, Phys. Rev. B **46**, 7605 (1992).
[7]G. P. M. Poppe, C. M. J. Wijers, and A. van Silfhout, Phys. Rev. B **44**, 7917 (1991).
[8]D. J. Chadi, Phys. Rev. Lett. **59**, 1691 (1987).
[9]H. J. W. Zandvliet (private communication).
[10]W. L. Mochán and R. G. Barrera, Phys. Rev. Lett. **55**, 1192 (1985).
[11]A. A. Baski, Naval Research Laboratory, Washington DC (to be published).
[12]R. Bersohn, Y. Pao, and H. L. Frish, J. Chem. Phys. **45**, 3184 (1966).
[13]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA 1974).

Reflection anisotropy spectroscopy, surface photovoltage spectroscopy, and contactless electroreflectance investigation of the InP/$In_{0.53}Ga_{0.47}As$(001) heterojunction system

M. Leibovitch, P. Ram, L. Malikova, and Fred H. Pollak[a)]
Physics Department and New York State Center for Advanced Technology in Ultrafast Photonic Materials and Applications, Brooklyn College of the City University of New York, Brooklyn, New York 11210

J. L. Freeouf
Interface Studies, Inc., Katonah, New York 10536-0982

L. Kronik, B. Mishori, and Yoram Shapira
Department of Electrical Engineering, Tel Aviv University, Ramat-Aviv 69978, Israel

A. R. Clawson
ECE Department-0407, University of California San Diego, La Jolla, California 92093

C. M. Hanson
NCCOCS RDTE Div 555, San Diego, California 92152-5070

(Received 24 January 1996; accepted 20 April 1996)

Using the optical methods of reflection anisotropy spectroscopy, surface photovoltage spectroscopy, and contactless electroreflectance, we have conducted an *ex situ* investigation of (a) the InP/$In_{0.53}Ga_{0.47}As$(001) heterojunction interface as a function of InP overlayer thickness (50–1000 nm) and (b) the surfaces of *n*- and *p*-doped $In_{0.53}Ga_{0.47}As$(001). All samples were fabricated by organometallic vapor phase epitaxy. The results from these optical probes make it possible to form a comprehensive quantitative picture of the InP/InGaAs heterojunction, including conduction and valence band offsets of 275 and 325 meV, respectively, as well as the (001) surface of InGaAs (surface Fermi level=200 mV from the conduction band edge). © *1996 American Vacuum Society.*

I. INTRODUCTION

The technological importance of InP/$In_{0.53}Ga_{0.47}As$ heterojunctions is due to their integral part in InP-based microelectronic and photonic devices. Two microelectronic devices in particular, InP/InGaAs high electron mobility transistors (HEMTs) and heterojunction bipolar transistors (HBTs), have impressive microwave properties including high frequency gain, noise characteristics, and power handling capacity. InP-based HEMTs are being pursued in integrated circuits (ICs) such as 60 GHz low noise amplifiers,[1] and likewise HBTs are of interest in ICs such as high resolution, high speed analog to digital converters.[2] InP-based photonic devices can uniquely meet today's needs for optical transport and access networks. Lasers have been pursued commercially for some time and photonic integrated circuits (PICs) are also being developed. Examples of PICs that have been recently reported are distributed feedback laser integrated with an electroabsorption modulator for long haul transmission at 2.5 GHz, and electro-optic integrated circuits for receivers capable of wide bandwidths of 6 GHz.

The device properties of InP/InGaAs heterojunction devices are strongly influenced by the electronic and optical properties of the interface.[3] Better understanding of this heterojunction can potentially improve the performance of devices even further. Perhaps more importantly it can aid in the commercial pursuit of InP-based devices where the emphasis is on reproducibility, reliability, circuit yield, and increased IC complexity at the lowest possible cost.

Using the optical methods of reflection anisotropy spectroscopy (RAS),[4–10] surface photovoltage spectroscopy (SPS),[11–14] and contactless electroreflectance (CER),[15] we have conducted an *ex situ* investigation of the InP/$In_{0.53}Ga_{0.47}As$(001) heterojunction interface as a function of InP overlayer thickness (50, 100, 150, 200, 250, 300, and 1000 nm). The results from these optical probes make it possible to form a comprehensive quantitative picture of the InP/InGaAs heterojunction. To gain further information about this heterojunction we also have studied the surfaces of *n*- and *p*-doped $In_{0.53}Ga_{0.47}As$ using RAS and CER. In the former system we have (a) determined the conduction (ΔE_C) and valence (ΔE_V) band offsets to be 275 and 325 meV, respectively, and (b) observed the evolution of the interface electric field with InP overlayer dimension. The behavior of the interface electric field with InP coverage indicates a lack of significant interface charge and has been explained by a self-consistent Poisson's calculation of the properties of the heterojunction. The CER signal from the direct band gap of the InP confirmed the low doping level in this material. Also, the observed below band gap CER features may be related to the presence of some As in the InP, as suspected from the growth conditions.[16] For the InGaAs surface we have (a) observed that the RAS signal increases linearly with the square root of the doping level (and hence surface electric field) due to the linear electro-optic effect

[a)]Also at the Graduate School and University Center of the City University of New York, New York, NY 10036; Electronic mail: fhpbc @cunyum.cuny.edu

associated with the E_1, $E_1+\Delta_1$ optical features[5–8] and (b) evaluated the surface Fermi level to be 200 mV below the conduction band edge.

Reflection anisotropy spectroscopy measures the polarization anisotropy of light linearly polarized along the [110] and [1$\bar{1}$0] principal axes in the plane of the (001) surface of zincblende-type semiconductors. It has been demonstrated that RAS can be employed to determine the sign and magnitude of near surface electric fields from the anisotropy produced by the linear electro-optic effect in the vicinity of the spin-orbit split E_1, $E_1+\Delta_1$ optical features.[5–8] Our work demonstrates that RAS can be employed to gain information about the electric fields at buried interfaces[10] in addition to its previous applications to (a) the *in situ* surface chemistry of MBE and OMCVD growth[4] and (b) surface electric fields.[5–8] SPS employs a Kelvin probe for the measurement of the changes in surface potential as a function of heterojunction overlayer thickness. This method produces a unified picture about the potential profile in the entire structure including band offsets, interface and bulk defect states, as well as net carrier concentrations.[11–14] CER is a contactless form of electromodulation in which the modulating electric field is applied to the sample using a capacitorlike configuration. The observed spectra, which may display Franz–Keldysh oscillations (FKOs), yields information about band gaps and built-in electric fields (from the FKOs).[15]

II. EXPERIMENTAL DETAILS

The samples studied in this investigation, all fabricated by organometallic vapor phase epitaxy (OMVPE), consisted of two groups of material: (a) 1.3-μm-thick n (Si)- and p-(Zn)-doped $In_{0.53}Ga_{0.47}As$/InP(001) and (b) 1-μm-thick n-type Si-doped (1×10^{17} cm^{-3}) $In_{0.53}Ga_{0.47}As$ layers grown on n^+ (1.1×10^{18} cm^{-3}) Sn-doped (001) InP substrates, with undoped (expected unintentional n doping) layers of InP (50, 100, 150, 200, 250, 300, and 1000 nm) on top of the InGaAs. For the former set the doping concentrations were 4, 10, 30, 200, 1000, 5000, 8000 (in units of 1×10^{15} cm^{-3}) for the n-type samples and 10, 20, 300, 7000 (in units of 1×10^{15} cm^{-3}) for the p doping. The former were grown on n^+ (3×10^{18} cm^{-3}) S-doped InP substrates while the latter were fabricated on n (2×10^{16} cm^{-3}) InP. For the CER measurements, the modulating voltage was $\pm$500 V at 200 Hz.

The samples were grown in a 50-mm-diam horizontal, lamp-heated, quartz chamber reactor at 650 °C, 100 Torr. The sources were trimethyl indium, trimethyl gallium, arsine, and phosphine for the InP and InGaAs growth and silane diluted in H_2 and dimethylzinc for the n and p-doping, respectively.

III. EXPERIMENTAL RESULTS

A. n- and p-doped $In_{0.53}Ga_{0.47}As$ samples

1. Reflection anisotropy difference measurements

Reflection difference anisotropy experiments have been used to study the surface electric fields in n- and p-doped GaAs(001) and ZnSe(001) as well as InSb and GaAs(110).

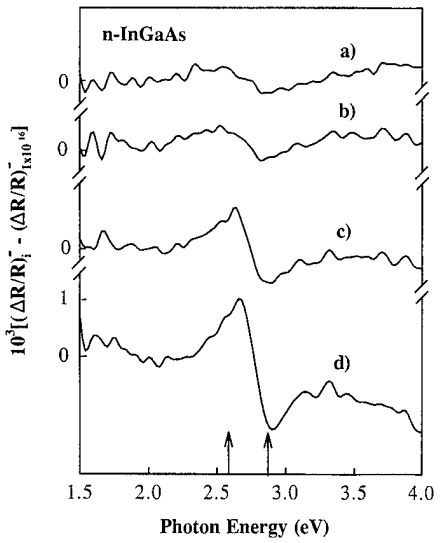


FIG. 1. $(\Delta R/R)_i^- - (\Delta R/R)_{1\times10^{16}}^-$ for a series of n-doped InGaAs samples: $n=3\times10^{16}$ cm^{-3} (curve a), 2×10^{17} cm^{-3} (curve b), 1×10^{18} cm^{-3} (curve c), and 8×10^{18} cm^{-3} (curve d). The energies of the E_1 (2.57 eV) and $E_1+\Delta_1$ (2.82 eV) optical transitions in $In_{0.53}Ga_{0.47}As$ are denoted by arrows at the bottom of the figure.

In zincblende-type semiconductors the RAS signal in the vicinity of the E_1, $E_1+\Delta_1$ optical features [interband transitions along the Λ lines ($\langle111\rangle$ directions) of the Brillouin zone] contains a linear electro-optic effect.[5–8] Also, there may be a background signal due to any improper alignment of the modulator and analyzer. Since it is the linear electro-optic effect that provides the information of interest, we would like to eliminate the influence of the background term.

We define $(\Delta R/R)_a$ as

$$(\Delta R/R)_a=(R_{[110]}-R_{[\bar{1}10]})/(R_{[110]}+R_{[\bar{1}10]}), \tag{1a}$$

where R_{110} and $R_{1\bar{1}0}$ are the reflectivities for light polarized along the [110] and [$\bar{1}$10] directions, respectively. On the other hand, the signal $(\Delta R/R)_b$ is

$$(\Delta R/R)_b=(R_{[\bar{1}10]}-R_{[\bar{1}\bar{1}0]})/(R_{[\bar{1}10]}+R_{[\bar{1}\bar{1}0]}). \tag{1b}$$

By taking the difference between these two signals, i.e., $(\Delta R/R)^-=(\Delta R/R)_a-(\Delta R/R)_b$, we can eliminate any background terms. We term $(\Delta R/R)^-$ as the RAS spectrum.

Plotted in Fig. 1 is $(\Delta R/R)_i^- - (\Delta R/R)_{1\times10^{16}}^-$ in the range 1.5–4 eV for a series of n-doped samples, where $(\Delta R/R)_{1\times10^{16}}^-$ is the RAS signal for the 1×10^{16} cm^{-3} material and $i=3\times10^{16}$ cm^{-3} (curve a), 2×10^{17} cm^{-3} (curve b), 1×10^{18} cm^{-3} (curve c), and 8×10^{18} cm^{-3} (curve d). The energies of the E_1 (2.57 eV) and $E_1+\Delta_1$ (2.82 eV) optical transitions in unstrained $In_{0.53}Ga_{0.47}As$[17] are denoted by arrows at the bottom of the figure. Note that the amplitude (peak-to-valley) of the signal increases as the doping concentration (and hence surface electric field) is raised, in a manner similar to that observed in GaAs(001).[5,6]

In Fig. 2 is displayed $(\Delta R/R)_i^- - (\Delta R/R)_{1\times10^{16}}^-$ for a set of p-type samples with $i=2\times10^{16}$ cm^{-3} (curve a), 3×10^{17}

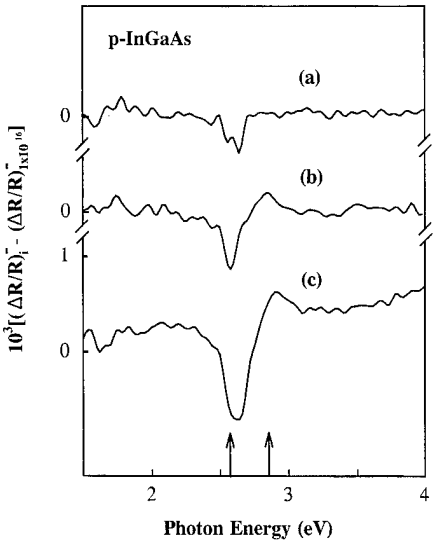


FIG. 2. $(\Delta R/R)_i^- - (\Delta R/R)_{1\times10^{16}}^-$ for a set of p-type InGaAs samples with $p=2\times10^{16}$ cm^{-3} (curve a), 3×10^{17} cm^{-3} (curve b), and 7×10^{18} cm^{-3} (curve c). The energies of the E_1 (2.57 eV) and $E_1+\Delta_1$ (2.82 eV) optical transitions in $In_{0.53}Ga_{0.47}As$ are denoted by arrows at the bottom of the figure.

cm^{-3} (curve b), 7×10^{18} cm^{-3} (curve c). The energies of the E_1 (2.57 eV) and $E_1+\Delta_1$ (2.82 eV) optical transitions in unstrained $In_{0.53}Ga_{0.47}As$ are denoted by arrows at the bottom of the figure. Again there is a shift in oscillator strength from the $E_1+\Delta_1$ to E_1 structure as the doping concentration, and hence surface electric field, is increased. However, note that the phase of the RAS signals is reversed in the n- and p-type samples, thus making it possible to determine the nature of the band-bending at the surface/interface.[5–8]

We have found that the amplitude of the RAS signal increases linearly as the square root of the carrier concentration (surface electric field) as reported for GaAs(001) in Refs. 5 and 6.

2. Contactless electroreflectance

In order to obtain more information about the surface electric fields in these samples we also have performed CER measurements in the vicinity of the direct gap (E_0) at 300 K. Shown in Fig. 3 is the CER spectrum from the 2×10^{17} cm^{-3} n-doped material. The data exhibit well-defined FKOs from which it is possible to evaluate the surface electric field F_s. The positions of the Nth extrema in the FKOs are given by[15]

$$N\pi=\tfrac{4}{3}(E_N-E_0/\hbar\Theta)^{3/2}+\chi, \tag{2a}$$

where the electro-optic energy $\hbar\Theta$ is

$$(\hbar\Theta)^3=q^2\hbar^2F^2/2\mu_{\|} \tag{2b}$$

and E_N is the photon energy of the Nth extrema, E_0 is the band gap, F is the electric field, $\mu_{\|}$ is the reduced interband effective mass (0.0376 in units of the free electron mass)[18] in the direction of $\vec{F}$ and χ is an arbitrary phase factor. It has been shown that for the nonuniform field in a space charge layer the field evaluated from the FKOs is F_s. This is the

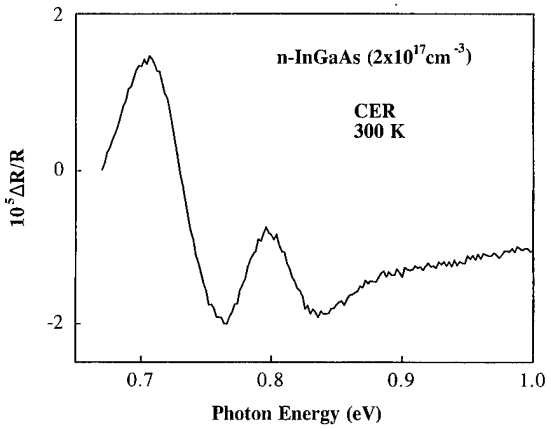


FIG. 3. Contactless electroreflectance spectrum of an n-doped (2×10^{17} cm^{-3}) InGaAs sample.

case when the ac modulating field $F_{ac}\leq0.15F_s$, independent of the penetration depth of the light.[15] The obtained value of $F_s=90\pm10$ kV/cm.

The surface Fermi level $V_{F,s}^C$ (relative to the conduction band edge) can be evaluated from the relation[19]

$$F_s=[(2qN_d/\epsilon_0\kappa)(V_{F,s}^C-V_{F,b}^C-kT/q)]^{1/2}, \tag{3}$$

where N_D is the donor concentration, ϵ_0 is the permittivity of free space, $\kappa(=14)$[18] is the static dielectric constant, and $V_{F,b}^C$ is the bulk Fermi position, relative to the conduction band edge. We have neglected any photovoltaic effects. Even though CER does not employ a pump beam (to produce the modulation) there still may be a photovoltaic effect (typically about 50–100 meV at 300 K) due to the probe light.[15,20] From Eq. (3) with $V_{F,b}^C$ = 15 mV[19] we find $V_{F,s}^C$ = 200 ± 40 mV (relative to the conduction band), in reasonable agreement with the barrier height reported by Tawari and Frank.[21]

B. InP/$In_{0.53}Ga_{0.47}As$ system

1. Surface photovoltage spectroscopy measurements

The SPS response of the above structure has been studied numerically by simultaneously solving the continuity and Poisson equations. The details of the simulation program have been elaborated elsewhere.[14,22] The calculated SPS response for several overlayer thicknesses is shown in Fig. 4, using values of the absorption coefficient and carrier lifetime taken from Ref. 18. The spectra of Fig. 4 display two distinct spectral regions according to the dominant absorption site: Below the band gap of the InP overlayer (≈1.35 eV), the photons are absorbed in the InGaAs layer. Above that energy, they are absorbed in both the InP overlayer and the InGaAs layer. With increased overlayer thickness both the shape of the spectra and their magnitude undergo very significant changes. It is of particular importance to note that the "knee" in the spectrum at about 1.35 eV, clearly seen in the spectra of samples with relatively thick overlayers, is completely absent from the spectra of samples with thinner ones.

Furthermore, in order to test the sensitivity of the spectral features to the values of the various parameters we have

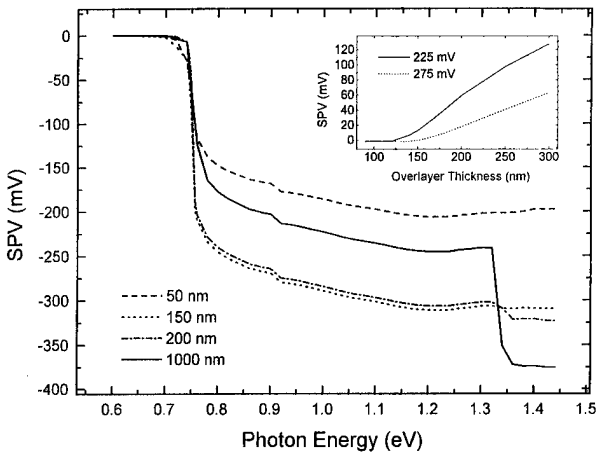


FIG. 4. Numerical simulation of SPS spectra of the InP/InGaAs(001) structure for various InP overlayer thicknesses (for clarity the curve for 200 nm is down shifted by 20 mV). Inset: dependence of the spectral "knee" at 1.35 eV vs InP overlayer thickness for $\Delta E_C = 275$ meV and $\Delta E_C = 225$ meV.

performed additional simulations using parameters (absorption coefficient and carrier lifetime) deviating from the standard ones by plus/minus an order of magnitude. We find that while these quantities may significantly influence the amplitude of the SPS spectrum, they have no appreciable effect on the critical thickness, W_c, at which the above mentioned "knee" appears. However, changing the value of the band offsets has a very large effect on W_c.

The magnitude of the above mentioned "knee" as a function of overlayer thickness for both $\Delta E_C = 275$ meV (used for the main portion of Fig. 4) and $\Delta E_C = 225$ meV is shown in the inset of Fig. 4. It is easily seen that indeed this "knee" appears quite abruptly at a thickness of W_c, and also that W_c is critically dependent on ΔE_C. Thus, once W_c has been determined experimentally, the band offsets may be found by numerical fitting.

The experimental SPS spectra for several overlayer thicknesses are shown in Fig. 5. Three regimes are apparent, i.e., photon energies below the 0.75 eV (E_0 of InGaAs), photon energies between the InGaAs and InP band gaps and photon

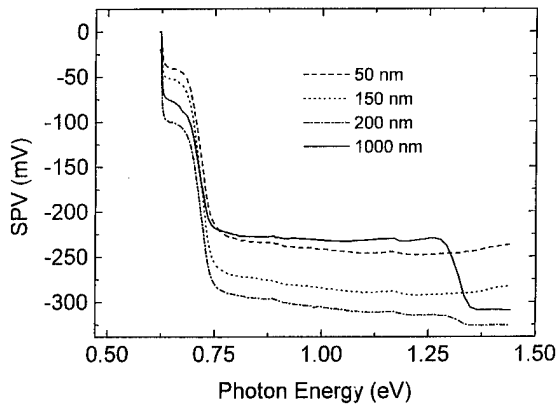


FIG. 5. Experimental SPS spectrum of the InP/InGaAs(001) structure for several InP overlayer thicknesses.

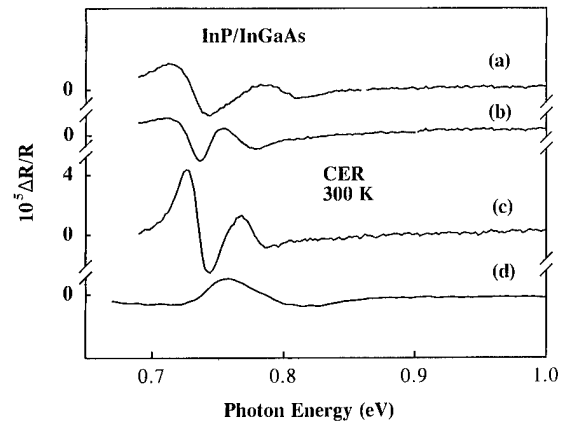


FIG. 6. The CER spectra in the vicinity of E_0 of InGaAs for (a) bare n-InGaAs ($n = 1\times10^{17}$ cm^{-3}), (b) InP(50 nm)/InGaAs ($n = 1\times10^{17}$ cm^{-3}), (c) InP(300 nm)/InGaAs ($n = 1\times10^{17}$ cm^{-3}), and (d) bare p-InGaAs ($p = 2\times10^{17}$ cm^{-3}).

energies above the InP band gap. The first regime clearly corresponds to absorption in a defect gap state (it does not appear in the numerical simulation since no specific defect gap states were included in it). The second and third regimes are very similar to those obtained in the numerical simulation. All features of the numerical simulation are also observed in the real experiment. These include a very large contribution from the InGaAs layer, which is not to be expected in the absence of charge transfer into the InP overlayer and a critical thickness ($W_c = 150$–200 nm) below which the InP-related "knee" does not appear in the spectrum. Further simulations which use ΔE_C and ΔE_V as fitting parameters show that the best fit for W_c is obtained for $\Delta E_C = 275 \pm 30$ meV and $\Delta E_V = 325 \pm 30$ meV. The error results mainly from the experimental uncertainty in the precise value of W_c. These values are in excellent agreement with values reported in the literature.[23–25] Note that since the dependence of W_c on the band offsets is very strong a rather large uncertainty in W_c leads to a rather small error in the band offsets—clearly an advantage of the proposed approach. The full details of such SPS measurements and analysis will be the subject of a future work.[26]

2. CER measurements

Shown in Figs. 6(a), 6(b), and 6(c) are the CER spectra in the vicinity of E_0 of InGaAs for bare n-InGaAs ($n = 1\times10^{17}$ cm^{-3}), InP(50 nm)/InGaAs ($n = 1\times10^{17}$ cm^{-3}), and InP(300 nm)/InGaAs ($n = 1\times10^{17}$ cm^{-3}) samples, respectively. For comparison purposes, we also have displayed in Fig. 6(d) the CER data for bare p-InGaAs ($p = 2\times10^{17}$ cm^{-3}). The sign of the CER signal for curves a, b, and c indicates that all the surface/interface fields are "n type," i.e., upward band bending towards the surface/interface. For the doping levels used in these samples the space charge region is expected to be fully depleted. The fields determined from the FKOs of curves a, b, and c are 50 ± 5 kV/cm, 27 ± 3 kV/cm, and 30 ± 3 kV/cm, respectively. Thus, with increasing InP overlayer thickness the interface electric field decreases, in agreement

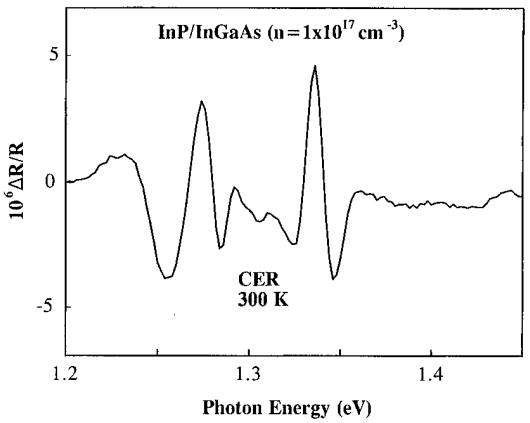


FIG. 7. The CER spectrum for the InP(300 nm)/InGaAs sample in the region of E_0 of InP.

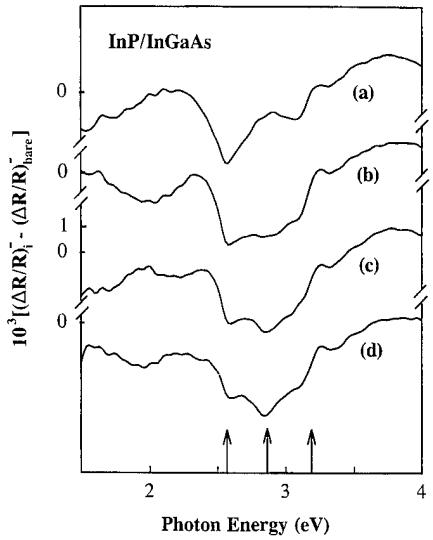


FIG. 8. $(\Delta R/R)_i^- - (\Delta R/R)_{\text{bare}}^-$ for a series of InP/InGaAs samples with InP(50 nm) [curve (a)], InP(100 nm) [curve (b)], InP(150 nm) [curve (c)], and InP(200 nm) [curve (d)]. The arrows at 2.57 and 2.82 eV are the E_1 and $E_1+\Delta_1$ optical transitions, respectively, of InGaAs, while the arrow at 3.2 eV denotes E_1, $E_1+\Delta_1$ of InP.

with a computer simulation. For the bare InGaAs sample, from Eq. (3) with $V_{F,b}^{C} = 30$ mV we find $V_{F,s}^{C} = 160 \pm 30$ mV, in reasonable agreement with the CER measurement of Sec. II A 2. The discrepancy is probably due to the photovoltaic effect, which will be smaller in the more heavily doped material. However, the trend with InP coverage is clear.

In Fig. 7 we display the CER spectrum for the InP (300 nm)/InGaAs sample in the region of E_0 of InP. The peak at 1.33 eV corresponds to the direct gap of the InP. The absence of FKOs confirms the low doping level in this material. The criteria for the appearance of FKOs is that the electro-optic energy $\hbar\Theta \geq \Gamma$,[15] the broadening parameter. With $\Gamma=6$ meV (from Fig. 7) and $\mu_{\parallel}=0.069$[18] (in units of the free electron mass) we find $F_s \leq 4$ kV/cm. The observed below band gap CER features may be related to the presence of some As in the InP, as suspected from the growth conditions.[16] These resonances are the subject of further investigation.

3. RAS measurements

Displayed in Fig. 8 are the results of the RAS investigation of the InP/InGaAs heterojunction as a function of InP layer thickness. In the spirit of Figs. 1 and 2 we have plotted the parameter $(\Delta R/R)_i^- - (\Delta R/R)_{\text{bare}}^-$, where $(\Delta R/R)_{\text{bare}}^-$ is the RAS signal from the bare InGaAs ($n=1\times10^{17}$ cm^{-3}) surface and i=InP(50 nm) [curve (a)], InP(100 nm) [curve (b)], InP(150 nm) [curve (c)], and InP(200 nm) [curve (d)]. The arrows at 2.57 and 2.82 eV are the E_1 and $E_1+\Delta_1$ optical transitions, respectively, of relaxed InGaAs while the arrow at 3.2 eV denotes E_1, $E_1+\Delta_1$ of InP (in this material the spin-orbit splitting Δ_1 is only about 100 meV).[27] These measurements are in agreement with the above CER results, i.e., increased InP coverage reduces the interface electric field. However, comparison of curve (a) with Figs. 1 and 2 seems to indicate that the interface field is "p type" (downward band bending towards the surface/interface), in contrast to our CER results. This may be due to the differing probe depths of the two techniques. One explanation of this difference is the possibility of a "camel-back diode" type of band bending within the InGaAs near the InP interface. If the maximum conduction band potential is not directly at the interface (denoted as $z=0$) but is reached a few hundred angstroms beyond the interface ($z=z_{\max}$) there will be two electric fields, i.e., a "p-type" field between $0<z<z_{\max}$ and an "n-type" field beyond this point. Thus, the much larger "n-type" region may be probed by CER and SPS while RAS, which has a penetration depth of only a few hundred angstroms, senses the former field.

IV. SUMMARY

We have conducted an *ex situ* investigation of (a) the surfaces of n- and p-doped $In_{0.53}Ga_{0.47}As(001)$ and (b) the $InP/In_{0.53}Ga_{0.47}As(001)$ heterojunction interface as a function of InP overlayer thickness (50–1000 nm) using the optical methods of RAS, SPS, and CER. The results from these optical probes make it possible to form a comprehensive quantitative picture of the InP/InGaAs heterojunction as well as the (001) surface of InGaAs. In the former system we have (a) determined $\Delta E_C=275$ meV and $\Delta E_V=325$ meV from SPS and (b) observed the evolution of the interface electric field with InP overlayer dimension using SPS, CER, and RAS. The behavior of the interface electric field with InP coverage indicates a lack of significant interface charge and has been explained by a self-consistent Poisson's calculation of the properties of the heterojunction. The RAS measurement indicated a different type of interface InGaAs electric field in relation to the SPS and CER studies. We have postulated one possible explanation for this apparent discrepancy. The CER signal from the direct band gap of the InP confirmed the low doping level in this material. Also, the observed below band gap CER features may be related to the presence of some As in the InP, as suspected from the growth conditions. For the InGaAs surface we have observed that the RAS signal increases linearly with the square root of the

doping level (and hence surface electric field) and that there is a change in phase between the n- and p-type material. These observations are due to the linear electro-optic effect associated with the E_1, $E_1+\Delta_1$ optical features. Also we have evaluated the surface Fermi level to be 200 mV below the conduction band edge using CER.

ACKNOWLEDGMENTS

The authors wish to thank D. E. Aspnes and J. M. Woodall for a number of very informative and useful discussions. M. L., P. R., L. M., and F. H. P. acknowledge the support of National Science Foundation Grant No. DMR-9414209, PSC/BHE Grant No. 666424, and the New York State Science and Technology Foundation through its Centers for Advanced Technology program. The authors A.R.C. and C.M.H. acknowledge the support of the Office of Naval Research, Program Element 62234N, Microelectronics Project RS34M40.

[1]M. Berg, J. Dickman, W. Bischoff, S. Koblowski, and P. Narozny, *IEEE 7th International Conference on InP and Related Materials, Sapporo, Hokkaido, Japan* (IEEE, New York, 1995), Catalog No. 95CH35720.

[2]J. F. Jensen, A. E. Consand, W. E. Stanchina, R. L. Walden, T. Lui, T. K. Brown, M. Montes, K. Elliot, and C. G. Kirkpatrick, *IEEE 16th Annual GaAs Integrated Circuit Symposium, Philadelphia, PA* (IEEE, New York, 1994), Catalog No. 94CH34488, p. 224.

[3]V. Somner, A. Kohl, K. Weigel, A. Mesquida Kusters, and K. Hieme, *IEEE 7th International Conference on InP and Related Materials, Sapporo, Hokkaido, Japan* (IEEE, New York, 1995), Catalog No. 95CH35720, p. 857.

[4]See, for example, D. E. Aspnes, Mater. Res. Soc. Symp. Proc. **324**, 3 (1994); W. Richter, Philos. Trans. R. Soc. London A **344**, 453 (1993); O. Acher, S. M. Koch, F. Omnes, M. Defour, M. Razeghi, and B. Drévillon, J. Appl. Phys. **68**, 3564 (1990).

[5]S. E. Acosta-Ortiz and A. Lastraz-Martinez, Phys. Rev. B **40**, 1426 (1989); also Proc. SPIE **1286**, 31 (1990).

[6]H. Tanaka, E. Colas, I. Kamiya, D. E. Aspnes, and R. Bhat, Appl. Phys. Lett. **59**, 3443 (1991).

[7]H. H. Farrell, M. C. Tamargo, T. J. Gmitter, A. L. Weaver, and D. E. Aspnes, J. Appl. Phys. **70**, 1033 (1991).

[8]C. C. Kim and S. Sivanathan, Mater. Res. Soc. Symp. Proc. **406**, 319 (1996).

[9]K. C. Rose, S. J. Morris, D. I. Westwood, D. A. Woolf, R. H. Williams, and W. Richter, Appl. Phys. Lett. **66**, 1930 (1995).

[10]Z. Yang, G. K. Wong, I. K. Sou, and Y. H. Yeung, Appl. Phys. Lett. **66**, 2235 (1995).

[11]M. Leibovitch, L. Kronik, E. Fefer, V. Korobov, and Y. Shapira, Appl. Phys. Lett. **66**, 457 (1995).

[12]L. Kronik, M. Leibovitch, E. Fefer, V. Korobov, and Y. Shapira, J. Electron. Mater. **24**, 893 (1995).

[13]L. Kronik, M. Leibovitch, E. Fefer, L. Burstein, and Y. Shapira, J. Electron. Mater. **24**, 379 (1995).

[14]M. Leibovitch, L. Kronik, E. Fefer, and Y. Shapira, Phys. Rev. B **50**, 1739 (1994).

[15]See, for example, F. H. Pollak and H. Shen, Mater. Sci. Eng. **R10**, 275 (1993) and references therein; F. H. Pollak, in *Handbook on Semiconductors*, edited by M. Balkanski (North Holland, Amsterdam, 1994), Vol. 2, p. 527 and references therein.

[16]A. R. Clawson, X. Jiang, P. K. L. Yu, C. M. Hanson, and T. T. Vu, J. Electron. Mater. **22**, 155 (1993).

[17]C. Pickering, R. T. Carline, M. T. Emeny, N. S. Garawal, and L. K. Howard, Appl. Phys. Lett. **60**, 2412 (1992).

[18]*Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung, M. Schulz, and H. Weiss (Springer, New York, 1982), Landolt-Börnstein, New Series, Group III, Vol. 17a.

[19]S. M. Sze, in *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981).

[20]F. H. Pollak, W. Krystek, M. Leibovitch, H. Qiang, D. C. Streit, and M. Wojtowicz, *Proceedings of the International Workshop on Semiconductor Characterization*, edited by W. M. Bullis, D. G. Seiler, and A. C. Diebold (American Institute of Physics, Woodbury, NY, 1995), p. 669.

[21]S. Tawari and D. J. Frank, Appl. Phys. Lett. **60**, 630 (1992).

[22]G. Ashkinazi, M. Leibovitch, and M. Nathan, IEEE Trans. Electron. Devices **40**, 285 (1993).

[23]W. L. Chen, J. P. Sun, G. I. Haddad, M. E. Sherwin, G. O. Munns, J. R. East, and R. K. Mains, Appl. Phys. Lett. **61**, 189 (1992).

[24]M. A. Haase, N. Pan, and G. E. Stillman, Appl. Phys. Lett. **54**, 1457 (1989).

[25]D. V. Lang, M. B. Panish, F. Capasso, J. Allam, R. A. Hamm, A. M. Sergent, and W. T. Tsang, Appl. Phys. Lett. **50**, 736 (1987).

[26]M. Leibovitch, L. Kronik, B. Mishori, Y. Shapira, C. M. Hanson, and A. R. Clawson (unpublished).

[27]C. M. Herzinger, P. G. Synder, B. Johs, and J. A. Woollam, J. Appl. Phys. **77**, 4677 (1995).

Physics and chemistry of silicon wafer bonding investigated by infrared absorption spectroscopy*

M. K. Weldon,[a] Y. J. Chabal, D. R. Hamann, S. B. Christman, E. E. Chaban, and L. C. Feldman[b]
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

(Received 22 January 1996; accepted 13 April 1996)

Silicon wafer bonding is achieved by joining two particle-free silicon wafers and annealing to elevated temperatures (~1100 °C). We have used multiple internal transmission infrared absorption spectroscopy to probe the interface between the wafers upon initial joining and also during subsequent annealing steps. For atomically flat hydrophobic wafers (H passivated), we observe a pronounced shift in the Si–H stretching frequency due to the physical interaction (van der Waals attraction) that occurs when the surfaces come into intimate contact. The hydrogen eventually disappears at high temperatures (1000 °C) and Si–Si bonds are formed between the two surfaces. For hydrophilic wafers (oxide passivated), we initially observe three to five monolayers of water at the interface (providing the initial attraction through H bonding), as well as the presence of hydroxyl groups that terminate the oxide at low temperature. Upon moderate heating (<400 °C), the water trapped at the interface dissociates and leads to the formation of additional oxide. Between 400 and 800 °C, the hydroxyl groups disappear, resulting in a corresponding increase in oxide *and* the formation of Si–O–Si bridging linkages across the two surfaces. © *1996 American Vacuum Society.*

I. INTRODUCTION

Silicon wafer bonding[1] is an important technological process and represents a significant challenge in interface science. Direct[2] silicon wafer bonding occurs when two particle-free wafers are brought into contact and subsequently annealed. Initial joining is performed at room temperature by applying gentle pressure at the center so that the two wafers can come into intimate contact by expelling any ambient between them. Permanent bonding is then achieved via interface chemical reactions that are initiated by increasing the temperature. This method is used in industry to produce silicon-on-insulator materials,[3] replace epitaxy, fabricate *p-i-n* diodes and power devices, as well as micromechanical structures.[4] Despite this, little is known about the microscopic interface processes that occur in this constrained environment, effectively hindering precise control or optimization of the bonding process.

Consequently, the scientific challenge is to probe and understand such interface phenomena in this novel environment. Initially, the weak physical interaction between the inner surfaces provides the force necessary to bring the surfaces together and expel ambient gas. Yet, the surfaces are not necessarily atomically smooth or macroscopically flat, so that mechanical strain must also play a critical role in determining the bonding quality. Upon annealing, the interface species and inner surfaces can undergo a chemical evolution that may be unique to this highly constrained environment. Since no ambient molecule can reach the interface region after joining, this system is equivalent to a vacuum vessel in which pressures can locally vary from several atmospheres to near perfect vacuum. The issues range, therefore, from identifying chemical species and their interaction at the interface to mapping and understanding inhomogeneities, such as intrinsic void formation.

Characterization of the interface is particularly challenging and may be best achieved using infrared radiation. Traditional surface science techniques cannot penetrate through a wafer (0.5 mm thick). In addition, techniques such as x-ray photoelectron spectroscopy (XPS) and secondary ion mass spectrometry (SIMS) require grinding and etching prior to analysis, which is destructive and only possible for well bonded wafers (i.e., after a high temperature anneal). In contrast, infrared (IR) spectroscopy is well suited to the study of both the physical and chemical aspects of the joining at room temperature, and subsequent bonding by higher temperature anneal. In general, the IR radiation can easily reach the buried interface since Si is transparent above 1500 cm^{-1}. As a result, the characteristic high frequency (>2000 cm^{-1}) stretching vibrations of hydrogen-containing species such as Si–H, H_2O, OH, NH_x, and CH_x can be routinely probed in all of the different environments inherent to these systems.[5] Furthermore, the lower frequency (<1500 cm^{-1}) modes of important atoms such as O, C, N, and F can also be studied, although this frequency region is quite challenging because of Si lattice absorption, diminishing IR source intensities, and the lower sensitivity of IR detectors below 1500 cm^{-1}. Therefore, the sample dimensions become critical if one wishes to probe the 700–1500 cm^{-1} spectral range, as shown in the present work.

In this article, a review of recent IR studies is presented for both Si–Si bonding, starting from hydrophobic (H-terminated) surfaces, and Si–SiO_2–Si bonding, starting from

*Published without author corrections.

[a]Electronic mail: marcus@physics.att.com
[b]Permanent address: Dept. of Physics, Vanderbilt University, P. O. Box 1807, Station B, Nashville, TN 37202-1807.

hydrophilic (oxidized) surfaces. In Sec. II, the process used for wafer bonding is first described with a brief outline of the technical difficulties of this otherwise extremely simple method. Next, the technique of multiple internal transmission (MIT) is summarized, with emphasis on its use for spectroscopy below 1500 cm^{-1} and on its relative sensitivity for different optical configurations and polarizations. Following this, in Sec. III, the bonding of hydrophobic and hydrophilic surfaces is discussed, with the initial bonding mechanism being quite different in the two cases. For hydrophobic surfaces, we focus on the joining of the particularly simple H/Si(111)-(1×1) surfaces characterized by a nearly perfect, atomically smooth, ideally H-terminated structure. We then contrast these results with those obtained for the joining of two hydrophobic, atomically rough, surfaces. For hydrophilic surfaces, we study the joining of two chemically oxidized surfaces, characterized by a thin (~5 Å) chemical oxide. Finally, Sec. IV emphasizes the physical processes and chemical reactions taking place at the interface for the different surface preparations.

II. EXPERIMENT

A. Wafer bonding

The method used for cleaning and initially joining the silicon wafers is remarkably simple. Standard wet chemical treatments, such as RCA cleans, are performed to obtain oxidized wafers with negligible metallic and carbon contamination.[6] The chemical oxide thus formed is approximately 5 Å thick, usually contains some dissolved hydrogen, and has a lower stoichiometry than high quality thermal (gate) oxides.[6,7] Hydrophilic wafers can be joined immediately following this treatment. In contrast, hydrophobic wafers require an additional HF or NH_4F etch, followed by thorough rinsing in de-ionized (DI) water. The choice of etchant depends critically on the crystallographic face, and controls the final surface morphology. For Si(100), concentrated HF solutions are used to minimize the facetting and roughness of the surfaces,[8] whereas highly buffered HF solutions (e.g., concentrated NH_4F solutions) are used to achieve atomically flat, ideally H-terminated Si(111) surfaces.[9] The two wafers are then placed on a spinner that maintains a 3 mm interwafer spacing, while highly filtered DI water is forced between them for several minutes as they slowly rotate, to wash away any remaining particulate contaminants. The wafers are then covered and spun at 3000 rpm for 5 min, while being mildly heated by an IR lamp (<80 °C), as shown in Fig. 1(a). Immediately following this step, the wafers are pushed together[10] with a gentle pressure applied at the center [Fig. 1(b)].

The process of joining is fascinating and can be followed in real time, as the area of intimate contact can be seen growing outward from the center using an IR imaging system.[11] With near monochromatic IR radiation (1.2 μm) transmitted through the Si pair, the contrast imaged by the camera depends critically on the distance between the wafers. Initially, the center portion of the wafers (where the pinching occurred) gives the highest brightness as the wafers

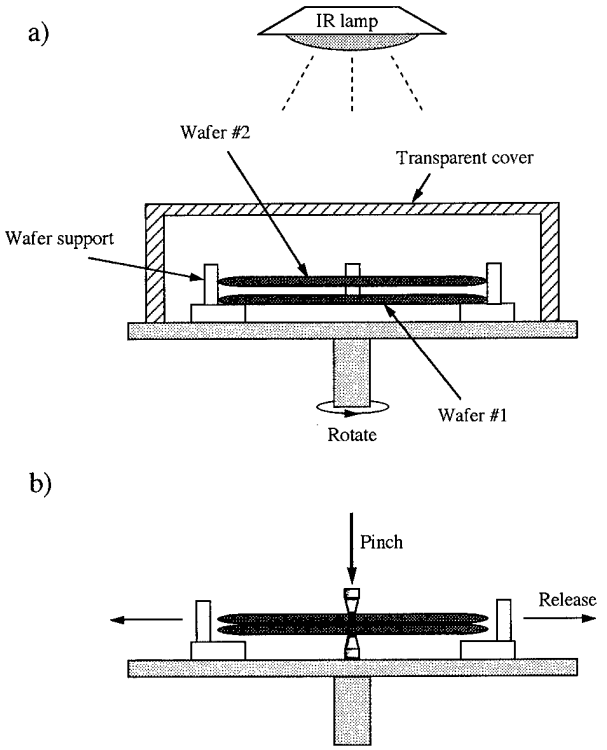


FIG. 1. Schematic representation of the apparatus used for wafer bonding: (a) the wafers are held in a "spinner" by teflon contacts and spun at 3000 rpm under a low intensity infrared lamp (T_{wafer}<80 °C); (b) after the requisite spin cycle, the wafers are contacted together and pinched at the center to initiate the "joining."

are much closer than the radiation wavelength ($d \ll 1$ μm). The rest of the area usually shows interference fringes and a darker outer portion. The central portion expands to encompass the whole wafer area (at which point joining is complete), with a "contact wave velocity" that is dependent on the particular conditions such as the ambient viscosity, wafer flatness, and density of physisorbed molecules.[11] At this stage, dust particles trapped at the interface generate highly visible, wide spots (1–10 mm diameter) with differing contrast. Clearly, the presence of voids or other contaminants will tend to weaken the bonding strength between the wafers and may compromise the electrical characteristics of such bonded substrates, so that higher resolution techniques such as acoustic microscopy or x-ray imaging are also widely used.[12]

The macroscopic bonding strength is an important indication of the quality of the bonding. It can be obtained in a straightforward manner by inserting a razorblade and measuring the crack propagation.[13] Such measurements indicate that wafers are only very weakly bonded after room temperature joining, with the bond strength of hydrophobic wafers being significantly less than that of hydrophilic wafers. Consequently, any handling of the wafers at this stage (particularly the cutting and edge polishing necessary for IR must be done with the utmost caution. However, the bonding strength is found to increase monotonically with annealing temperature so that, in general, wafer pairs are safe to handle after a

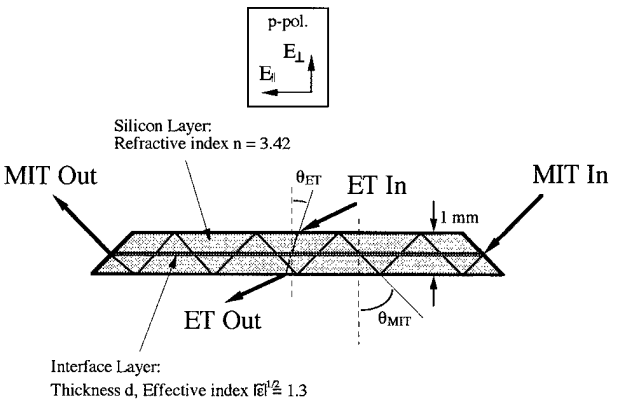


FIG. 2. Illustration of the multiple internal transmission (MIT) and external transmission (ET) optical configurations used.

500 °C anneal. Interestingly, hydrophobic wafers exhibit a higher bond strength than hydrophilic after a 1100 °C anneal, as the latter are limited by the properties of the oxide layer.[14] At intermediate temperatures (600–900 °C), the formation of intrinsic voids (not related to particles) often occurs, particularly for hydrophobic wafers.[15] A comprehensive microscopic picture of the relevant bonding mechanisms, intrinsic void formation processes, and ultimate interface structures after high temperature annealing has yet to emerge and is the main motivation for this work.

B. Infrared absorption spectroscopy

Probing the interface with sufficient sensitivity to observe submonolayer coverages requires the use of multiple internal *transmission* (MIT) spectroscopy.[5] The geometry is shown schematically in Fig. 2 and is termed MIT because the IR radiation is *transmitted* through the interface rather than reflected, even under the angular conditions where total internal reflection is observed for bulk media, because the interface is much thinner than the wavelength (tunneling). Importantly, this situation enhances the sensitivity to vibrational absorption *normal* to the interface, as conveniently summarized for p-polarized radiation using a three-layer model:[16]

$$\frac{\Delta I}{I_0}=\frac{2\pi d}{\lambda}\frac{1}{n\cos(\theta)}\left[\cos^2(\theta)\mathrm{Im}(\tilde{\epsilon})+n^4\sin^2(\theta)\mathrm{Im}\left(\frac{-1}{\tilde{\epsilon}}\right)\right]. \quad (1)$$

In this equation, the absorption ($\Delta I/I_0$) of IR radiation, incident with internal angle θ on a layer of thickness d and dielectric function $\tilde{\epsilon}$, is divided into two terms; one proportional to the imaginary part of $\tilde{\epsilon}$ and the other to the imaginary part of $1/\tilde{\epsilon}$, which describe the components of absorption parallel and normal to the interface, respectively. Noting that the silicon index of refraction ($n=3.42$ in the mid IR) is much larger than that of the interface layer ($|\tilde{\epsilon}|^{1/2}\approx 1.3$), two immediate conclusions can be drawn from this equation. The first is that the component normal to the interface is enhanced by

$$X_{\perp/\parallel(\mathrm{MIT})}=\frac{n^4\sin^2(\theta)}{|\tilde{\epsilon}|^2\cos^2(\theta)} \quad (2)$$

over the parallel component, which corresponds to a factor of 40–50 for a typical 45° incidence angle. The second consequence of this configuration is apparent on comparison of Eq. (1) to that obtained for internal reflection at the outer surfaces:[16]

$$\frac{\Delta R}{R_0}=\frac{2\pi d}{\lambda}\frac{1}{n\cos(\theta)}\left[I_x\,\mathrm{Im}(\tilde{\epsilon})+n_{\mathrm{vac}}^4 I_z\,\mathrm{Im}\left(\frac{-1}{\tilde{\epsilon}}\right)\right], \quad (3)$$

where n_{vac} is now the index of air or vacuum ($n_{\mathrm{vac}}=1$) and I_x ($=2.0$ for 45°) and I_z ($=2.4$ for 45°) are the normalized field intensities on the vacuum side of the outer surfaces.[16] The normal component of absorption is greater for MIT than for MIR by a factor

$$X_{\perp(\mathrm{MIT/MIR})}=\left(\frac{n}{n_{\mathrm{vac}}}\right)^4\frac{\sin^2(\theta)}{I_z}, \quad (4)$$

which is 28 for 45° and 42 for 60°!

The sensitivity of the MIT configuration is therefore 20 to 40 times that of the traditional MIR configuration. To illustrate the practical importance of this result, this statement can be rephrased as follows: using MIT, both the outer surface absorption (contamination) and the parallel component of the interface absorption are negligible (at least 20 times weaker) compared to the perpendicular component of the interface absorption. These two factors make the interpretation of the data more straightforward. In fact, the study of interfaces using the MIT configuration is nearly optimal since the experiment can be done outside a vacuum chamber (stable interface), the absorption from the outer surface contamination is negligible, and the sensitivity to the interface is over an order of magnitude better than any other surface configuration! The experimental challenge remains the low frequency region (<1500 cm^{-1}), where the sample dimensions become critical.

The problem with measurement below 1500 cm^{-1} is the silicon bulk lattice absorption which, although weak, dominates the spectrum for typical pathlengths. Just as in the MIR configuration, the probing IR radiation in the MIT configuration traverses several centimeters of silicon (typically 7–10 cm) in total. Consequently, the weak multiphonon absorption[17] in Si leads to orders of magnitude attenuation of the IR throughput, making it impractical to work in this frequency range, with such long path lengths. To study the interface, we would then be restricted to using a one pass external transmission configuration if it were not for realizing that the enhancement per pass of internal transmission over external transmission is substantial. Indeed, in external transmission, the refraction into the high index Si controls the internal incident angle to a value between 16.3° (for grazing external incidence) and 0° (for normal external incidence). In

contrast, using internal reflection, large angles can be achieved. The enhancement of the component perpendicular to the interface is then:

$$X_{\perp(\mathrm{MIT/ET})}=\frac{\sin^2(\theta_{\mathrm{int}})}{\sin^2(\theta_{\mathrm{ext}})}\frac{\cos(\theta_{\mathrm{ext}})}{\cos(\theta_{\mathrm{int}})} \tag{5}$$

per pass. Practically, the angle achieved with external reflection is close to 16° (Brewster condition) and the angle chosen for internal reflection is 60°. The enhancement per pass is therefore 19, suggesting that the transmission below 1500 cm^{-1} could be 19 times weaker for a similar S/N ratio. A convenient geometry is therefore the MIT configuration with 60° incidence in a 1-cm-wide sample. This gives an average of 5.8 passes through a total of 1.2 cm of Si, yielding an enhancement of 110 over a single external reflection. If intrinsic Si is used with a resistivity of 20 Ω cm or greater, signal levels roughly 10 times weaker than above 1500 cm^{-1} would be obtained in the 800–1500 cm^{-1} region, but the overall gain over a single external reflection is still over a factor of 10. The MIT configuration using a shorter sample is therefore theoretically preferable. In practice, precise control of the sample temperature and IR beam position is also critical since the overall signal is still dominated by the highly temperature dependent bulk phonon absorption. In addition, it is also important to use float-zone silicon instead of Czochralski-grown silicon to avoid the very strong interstitial oxygen absorption centered at 1100 cm^{-1}.

III. RESULTS AND DISCUSSION

A. Hydrophobic silicon wafer bonding

1. Experimental results for the H/Si(111)-(1×1) surface

The best surface to study in order to characterize the processes occurring during the bonding of hydrophobic wafers is the atomically flat, ideally H-terminated Si(111) surface prepared by wet chemical etching.[9,18] Such H/Si(111)-(1×1) surfaces are characterized by atomically flat domains extending over several thousand angstroms, separated by atomically straight double-layer steps.[19] Over small areas, therefore, these surfaces have the possibility to come within a few angstroms. However, over the macroscopic wafer (100 mm diameter), the flatness is no better than 5000 Å, making it impossible to be uniformly close without wafer deformation.

Figure 3(a) illustrates the differences between the MIR spectrum of a H/Si(111)-(1×1) surface and the MIT spectrum of two such surfaces after room temperature joining. The most striking observations are (1) the dramatically increased spectral intensity and (2) the significantly altered spectral shape after joining. Specifically, the sharp (<1 cm^{-1} wide) Si–H stretch absorption line at 2083.5 cm^{-1} of the free H/Si(111)-(1×1) surface (i) is dramatically broadened to lower frequency, with the appearance of a new feature centered at 2065 cm^{-1}, when the wafers are joined at room temperature (ii). In addition, the integrated Si–H stretching absorption of the joined wafers is 24.5 times larger than for the single H/Si(111)-(1×1) surface, in good agreement with the factor of 28 calculated for a 45° incident angle, using Eq.

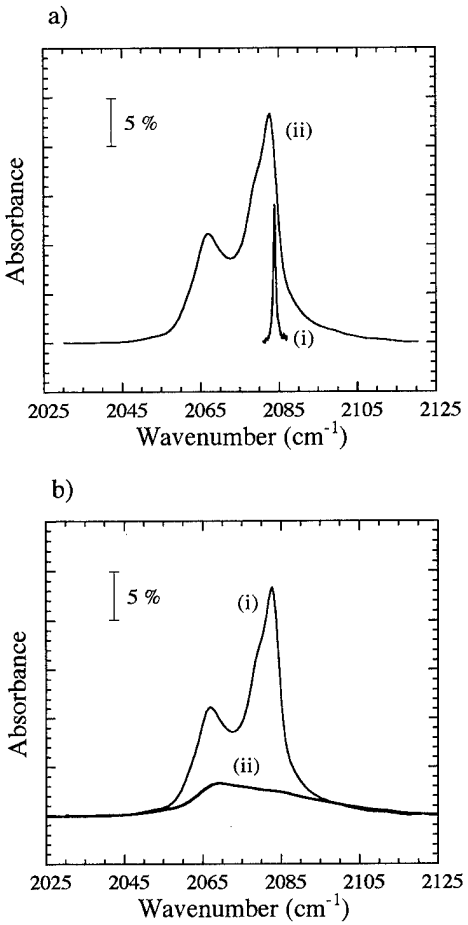


FIG. 3. (a) Demonstration of the profound changes apparent on joining hydrogen-terminated Si(111) wafers: (i) MIR spectrum of a single H/Si(111)-(1×1) sample, and (ii) MIT spectrum of two H/Si(111)-(1×1) wafers at room temperature (outer surfaces are oxide-passivated). Both spectra were obtained using the same optical setup (thickness=0.5 mm per Si, sample length=3.8 cm, internal angle of incidence=45°). (b) Comparison of the spectrum for (i) the 3.8 cm area shown in (a) and (ii) $a\sim1$ cm sample area.

(4). Interestingly, we find that the process is *completely* reversible, since the spectrum collapses back to a sharp line at 2083.5 cm^{-1}, if the joined wafers are separated, and broadens again as they are allowed to come back together.

We propose that the pronounced spectral broadening is due to the varying degrees of interaction between the macroscopic surfaces. The shift in frequency of the free-surface mode at 2083.5 cm^{-1} is then indicative of the interaction strength, with 2065 cm^{-1} reflecting the condition for maximum shift. To check this hypothesis, we have cut a Si(111) bonded wafer pair into many smaller pieces and probed each piece separately, as shown in Fig. 3b(ii). The predominance of the 2065 cm^{-1} mode and absence of the unperturbed free-surface mode at 2083.5 cm^{-1} confirms that the former frequency represents a stable configuration that can be achieved over areas as large as 1 cm *and* that the spectrum in Fig. 3b(i) is an average over the whole sample, which is comprised of regions of differing interaction strength. However, up to this point, we have not discussed the *precise* physical or chemical origin of this frequency shift.

Clearly, the spectral shift cannot be caused by a chemical reaction between the two surfaces, since the changes apparent on joining at room temperature are completely reversible. Furthermore, the possibility that a thin layer (~monolayer) of foreign molecules causes the shift by a hydrogen bonding or Van der Waals type interaction, can also be eliminated. We have investigated the former possibility by trapping water, HF, and ethanol between the two H/Si(111)-(1×1) surfaces prior to joining and found that, in all cases, a systematic broadening (50–75 cm^{-1}) of the Si–H spectrum occurs, but with an overall shift to *higher* frequencies of 10 to 20 cm^{-1}. Similarly, we find that the feature at 2065 cm^{-1} in the spectrum of joined Si(111) cannot be related to the presence of physisorbed contaminants (such as ambient C_xH_y or CO), since not only have such interactions been shown to increase the Si–H frequency on the isolated (111) surface,[20] but the 2065 cm^{-1} mode persists on annealing the joined samples to above 800 °C, by which point temperature decomposition of all molecular species has occurred. In summary, the spectrum shown in Fig. 3 cannot be reproduced by any of these (contaminant) species, so we conclude that the perturbation must arise from the direct interaction of the two clean surfaces. Specifically, we propose that the two surfaces can come sufficiently close under the influence of the van der Waals attraction, and that the opposing hydrogens interact, causing the vibrational frequency to shift.

The possibility of a direct physical interaction between the two surfaces is intriguing, as it necessarily requires that the separation be very small, and that most physisorbed contaminants are purged from the interface region upon joining. Clearly, the data suggest that such a clean and intimately contacted interface is indeed obtained. To provide a theoretical understanding of the spectral shifts, we have carried out a series of calculations on ideal H-terminated Si(111) surfaces in close proximity. These *ab initio* calculations are based on the local density approximation,[21] utilizing separable norm-conserving pseudopotentials to represent the atoms,[22,23] and a plane-wave basis set to expand the wave functions. We used the Ceperley–Alder exchange-correlation functional, as parametrized by Perdew and Zunger.[24] The Car–Parrinello method was used to simultaneously optimize the wave functions and atomic coordinates,[25] incorporating recently proposed modifications to accelerate the dynamics.[26]

The joined surfaces were simulated by two supercells consisting of 10 and 12 Si[111] layers plus terminating H layers on each surface. The surface spacing was varied by changing the lattice constant of the supercell in the surface normal direction. An in-plane lattice constant of 3.81 Å was determined by minimizing the energy of the bulk Si using the same pseudopotentials and cutoffs. For the 10 Si-layer supercell, the geometry is such that the H bonded to each surface approaches a three-fold hollow site of the facing H layer.[27] In contrast, the 12 Si-layer supercell yields "atop" registry, that is, the opposing H atoms approach head-on. Tests indicated that a plane wave cutoff of 15 Ry and a **k** sample of eight special points[28] in the irreducible wedge of the supercell Brillouin zone gave well-converged results. The damped

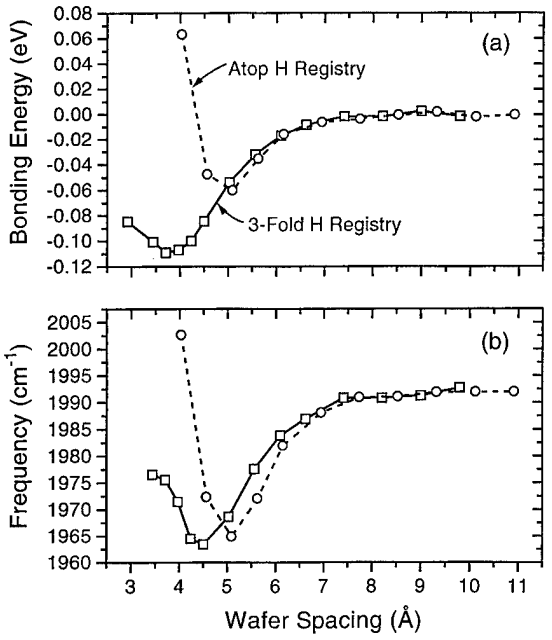


FIG. 4. Theoretical results for (a) the bonding energy and (b) the asymmetric Si–H stretch frequency for two joined, ideal H/Si(111)-(1×1) surfaces as a function of wafer spacing.

Car–Parrinello dynamics were run until residual forces on the atoms were less than 10^{-5} atomic units. After all the atoms were relaxed, the Si atoms were fixed, and the H atoms were moved adiabatically in steps of 0.05 Å in the (111) direction. The forces on the H atoms at five points were fit to the derivative of a Morse potential, and the force constants extracted from the fits. The vibrational frequencies were approximated using the reduced Si–H mass with these calculated force constants.

The results of our calculations are shown in Fig. 4. The abscissa is chosen to be "wafer spacing," so that the hollow- and atop-registry results can be compared in a physically reasonable way. We define this quantity as the spacing between the outermost Si layers of the facing surfaces, assuming the ideal bulk interlayer spacing. The actual surfaces display relaxation and slight elastic effects in the outermost layers. The bonding energy curves in Fig. 4(a) (which are the total supercell LDA energies referred to the large separation limit) show weak, but well-defined, attractive wells. The minima for the atop and three-fold registries correspond to H–H spacings of 2.11 and 2.38 Å, respectively. The potential displays a much steeper rise at small wafer spacings in the atop case. At distances greater than the atop minimum, the attraction is remarkably independent of registry. The existence of potential minima in each case may appear rather surprising, on first inspection, since the Si–H surface is a "closed shell" system, so that one might expect only repulsive interactions from a local theory. However, Lang has previously found that the LDA gives a reasonable account of rare gas adsorption, with energy minima roughly comparable in magnitude to those found here, and in rather good agreement with experiment.[29] The physical mechanism underlying

these results is discussed at some length in his article.

The Si–H vibrational frequency is calculated to be 1992 cm^{-1} at large spacing, which is 4% below the measured value (2083 cm^{-1}). This constitutes excellent agreement with experiment, with the residual discrepancy attributable to anharmonic effects, basis convergence, or the intrinsic limitations of the LDA. In Fig. 4(b), we see that there is a small, but systematic, softening of the optically active asymmetric stretch vibrational mode. Interestingly, the hollow and atop registry shifts are found to be very similar in magnitude, although the atop frequency versus spacing curve is shifted to greater wafer spacings by ~0.5 Å. The shifts have several sources. The Si–H bond lengths are calculated to be 0.0046 and 0.0066 Å longer at the spacing of maximum shift than at large spacings for the hollow and atop sites, respectively. Based on the anharmonicity of the Morse potential fit at large spacings, this "static" effect on the Si–H bond accounts for a large fraction of the total shift in each case. An additional source of the shifts is the dynamic effect due to the interactions of the vibrating H atoms on one surface with those on the other. These interactions also lead to significant splittings between the optically *in*active symmetric stretch mode and the asymmetric stretch at small wafer spacings.

In summary, we find that there is excellent agreement between the theoretically predicted, and experimentally observed, Si–H frequency shifts upon joining. We have not undertaken a full lineshape analysis since that would require many assumptions about the effects of steps and other defects on the spacing and registry of the wafers, as well as a detailed treatment of the concomitant elastic deformations. The principal consequence of these effects is that the maximum calculated shift may not be practically observable. Indeed, the (shifted) absorbance maximum is observed ~18 cm^{-1} lower in frequency than the constituent free-surfaces Si–H mode, compared to the 29 cm^{-1} shift, predicted above. However, a shift of 27 cm^{-1} has been observed for some samples after annealing to elevated temperatures, consistent with the attainment of an even more intimate interaction under these conditions. Clearly, the magnitude of the maximum observed shift is only consistent with the calculated model results if the interwafer spacing is at the atomic level, that is, within a few Å, over significant areas of the bonded interface formed upon room temperature joining of two Si(111) surfaces.

2. Results for atomically rough hydrophobic surfaces

The preceding discussion focused on the interaction of the atomically smooth Si(111) surfaces, where intimate contact be achieved over areas on the order of 1000 Å. The situation is clearly different if the hydrophobic surfaces are atomically rough. Such surfaces are prepared by etching Si(111) or Si(100) in dilute HF (1% to 10%) rather than in NH_4F.[8] As a result of this microscopic roughness, the interaction between the two surfaces is significantly decreased. In fact, we find that when these H-terminated surfaces are joined together the interface spectrum is *identical* to the clean surface spectrum (obtained prior to bonding), as shown in Fig. 5 for Si(100)

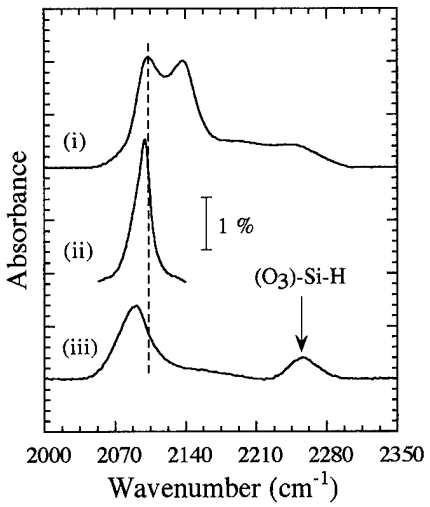


FIG. 5. MIT spectra of two joined Si(100)–H surfaces, as a function of annealing temperature: (i) room temperature; (ii) after annealing to 500 °C; and (iii) 600 °C. Note the presence of oxygen attached to–Si–H groups [ν(Si–H)>2150 cm^{-1}] at room temperature, which agglomerates to form predominantly (O_3)–Si–H, on annealing above 500 °C.

surfaces.[30] Recalling that only the components perpendicular to the surface are enhanced, the spectrum shows two main contributions at 2110 and 2140 cm^{-1}, corresponding to the perpendicular components of dihydride and trihydride vibrations of H-terminated Si(100) (see Fig. 6 of Ref. 8). There is, therefore, no evidence for physical interaction between the two surfaces over the vast majority of the wafer area, at room temperature (see Fig. 7).

The absence of a strong interaction between the two rough surfaces over the majority of the interface region is further supported by the observed behavior upon annealing to 500 °C. At this temperature, the Si–H stretch spectrum collapses into a single new feature at 2097 cm^{-1}, previously assigned to the symmetric stretch of the dimerized, monohydride-terminated H/Si(100)-(2×1) surface [Fig. 5(ii)].[31] Such an atomic rearrangement of the inner surfaces exactly mimics the behavior of a single atomically rough, H-terminated Si(100) surface upon annealing in UHV. The partial loss of hydrogen from the higher hydrides (SiH_2 and SiH_3) that occurs as the surface becomes smoother and assumes its most stable dimerized structure, also has important mechanistic implications. First, the occurrence of Si–H bond scission suggests that, although no interaction is apparent by analysis of the infrared spectrum alone, a minority of direct Si–Si linkages may in fact form between the two surfaces. Second, since the decomposition almost certainly proceeds by evolution of molecular hydrogen (as in UHV), gaseous H_2 is expected to be trapped at the interface (potentially leading to intrinsic voids or bubble formation), but with too low a concentration and pressure to be detectable using IR.[32] The former conclusion is confirmed by inspection of data obtained on annealing to 600 °C [Fig. 5(iii)], which show a distinct (−4 cm^{-1}) shift in the monohydride stretching frequency, consistent with perturbation of the isolated surface

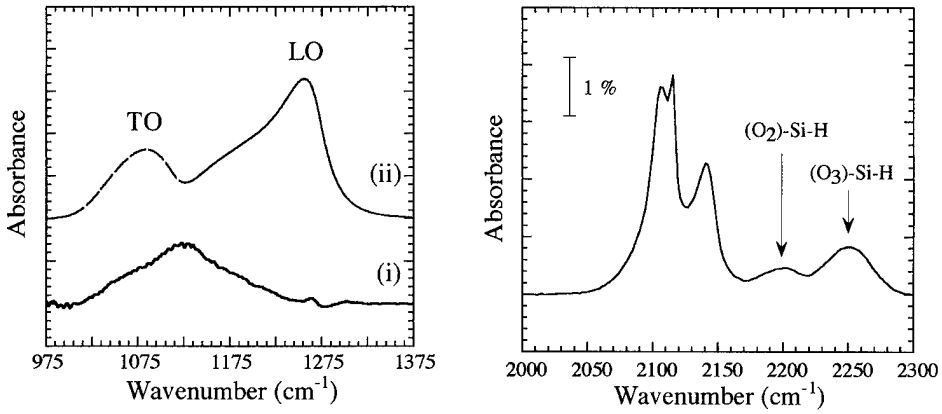


FIG. 6. Right panel: Si–H stretching modes of $(O_{2,3})$–Si–H species present at the interface of two joined hydrophobic Si(100) samples at room temperature. Left panel: (i) MIT spectrum of the low frequency (<1500 cm^{-1}) Si–O stretching modes of this surface oxygen; (ii) ET spectrum of 1600 Å of thermal SiO_2 (grown at 1100 °C) showing the TO and LO phonon absorption.

structure due to the proximity of the opposing surface at this temperature.

3. Oxygen contamination and oxidation at the Si/Si interface

It is clear from the spectra in Fig. 5 that some oxygen contamination of the silicon surface remains, even after careful HF etching and rinsing of the samples. This is well documented in the literature[33] [and is particularly prevalent for Si(100)] for isolated silicon surfaces, and is not therefore a unique property of joined wafers. However, an interesting difference exists for these highly constrained systems; the evaporation of SiO groups at elevated temperature that is observed for free Si surfaces is obviously suppressed in these environments. Consequently, although some dissolution of the oxygen into the bulk silicon may occur upon annealing to high temperature, there is a significant possibility that some oxygen will remained trapped at the interface, causing electrical degradation. We have investigated these possibilities by monitoring the thermal evolution of the (high frequency) Si–H modes that are shifted by the presence of oxygen in the Si backbone. With this approach, we assume implicitly that these Si–H modes are a sensitive probe of the oxygen environment,[34] i.e., that the oxygen remains in the Si backbone and the Si–H bond remains intact. As a result, no information can be obtained above ~900 °C, the temperature at which complete loss of interfacial hydrogen is observed. Consequently, we have also probed the SiO vibrations of the oxide directly, by modifying the experimental configuration, as described earlier.

Figure 6(i) shows the low frequency region of the infrared spectrum of a (typical) joined hydrophobic Si(100) sample that exhibits Si–H stretching modes between 2175–2275 cm^{-1}, indicative of oxygen in the silicon backbone (Fig. 6, right). Accordingly, we observe a prominent, broad absorption band centered at 1125 cm^{-1} that is assigned to the Si–O stretching modes of (O_x)–Si–H surface structures, by reference to previous (UHV) studies of the initial oxidation of Si,[35] for which bands at 1000–1200 cm^{-1} were assigned to local Si–O–Si modes. Although the precise interpretation awaits a detailed theoretical treatment, the width and frequency range of this absorption band is consistent with the existence of an inhomogeneous array of (predominantly) local Si–O modes. Furthermore, the absorption maximum is indicative of $(O_{2,3})$Si–H species, so that some aggregation or clustering of oxygen must occur, potentially resulting in the formation of small oxidized islands. Indeed, such islands would necessarily exhibit extended coulombic coupling, giving rise to the highest frequency component of the absorption band at 1230 cm^{-1}. To illustrate this, the spectrum of a thick (1600 Å) layer of thermal SiO_2 is also shown [Fig. 6(ii)].

For an SiO_2 network, the local Si–O modes are coupled together, giving rise to both transverse (TO) and longitudinal (LO) optical phonon modes.[36] In the thin film limit ($d \ll \lambda$), both the TO and LO phonons can be optically excited. Importantly, the LO (TO) phonon is polarized perpendicular (parallel) to the plane of the film and the LO phonon exhibits strong Coulombic coupling that shifts this mode to higher frequency than the corresponding TO phonon. In addition, as a result of this extended coupling, the LO mode frequency is effectively a (weighted) "average" over the extended network and is therefore not as sensitive to local inhomogeneities as the TO mode. On the other hand, the LO mode is sensitive to the layer thickness,[37] displaying a decrease in frequency with decreasing thickness (of ~1.5 cm^{-1}/Å), below 30 Å. The separation of the TO/LO modes is clearly apparent in the spectrum of thick SiO_2, for which the LO mode occurs at 1250 cm^{-1}, whereas the TO mode is observed at 1080 cm^{-1} [Fig. 6(ii)]. It is important to note that this spectrum was obtained for a single oxidized silicon wafer using the external transmission configuration, so that both the LO (perpendicular) and the TO (parallel) modes are observed with good sensitivity. In contrast, the LO phonon absorption would be 42 times more intense than the TO using MIT (at 60° incidence), so that it effectively dominates the observed spectrum. Given the above and the fact that the LO phonon of a continuous SiO_2 layer is predicted to be ~1200 cm^{-1} for vanishingly small thicknesses, the spectrum of the

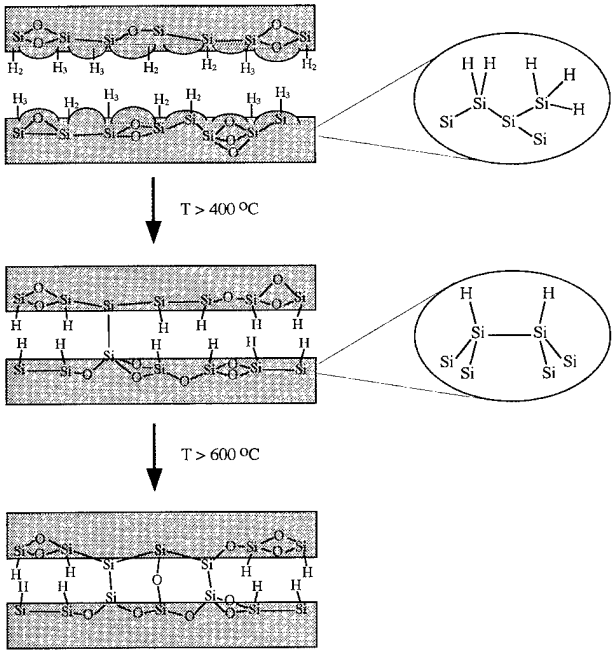


FIG. 7. Schematic of proposed mechanism for hydrophobic [Si(100)] wafer bonding.

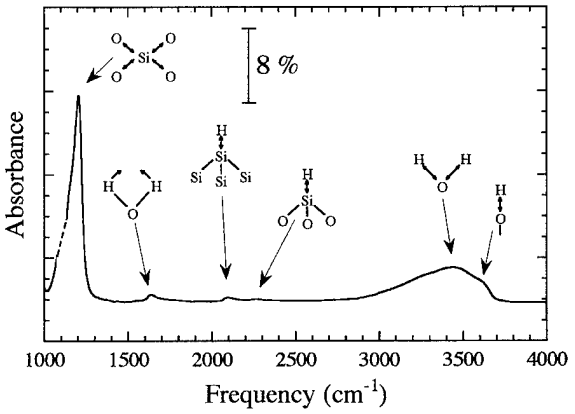


FIG. 8. Typical MIT infrared spectrum of two hydrophilic [Si(111)] wafers joined at room temperature. Schematic representations of the vibrational modes associated with each absorbance band are shown, for reference. Note the huge (~40%) absorbance due to the interfacial oxide initially present on these hydrophilic surfaces. The data were collected in a single acquisition and have not been corrected for baseline variations. The region at ~1100 cm^{-1} in the Si–O band is omitted since the interstitial oxygen present in these CZ wafers completely absorbs the IR radiation over this frequency range.

low frequency region of joined hydrophobic Si(100) is clearly more consistent with a local Si–O mode description. Finally, we see spectroscopic evidence for the formation of local Si–O–Si bridging bonds upon annealing these samples to elevated temperature, as illustrated in Fig. 7, which may (in part) explain the different electrical performance that is sometimes reported for direct-bonded Si substrates from that observed for bulk Si.

B. Hydrophilic silicon wafer bonding

1. Experimental results for chemically oxidized wafers

The growth of (thin) oxides is an important area of microelectronics.[38] It is also obviously a particularly pertinent issue in the formation of SOI substrates using wafer bonding. A central concern is the quality of the interface when two hydrophilic (oxidized) surfaces are bonded. We have addressed this problem by studying the bonding of two wet-chemically oxidized silicon surfaces.[39] The main advantage of these oxides is that they are very thin (~4–5 Å) so that the formation of new oxide can readily be identified. The extension of this work to thicker thermal oxides is necessary to describe the actual SOI systems, and is currently under way in our laboratory.

The spectrum observed after joining two hydrophilic Si(111) surfaces is shown in Fig. 8, along with the relevant vibrational assignments. The corresponding spectrum for hydrophilic Si(100) surfaces is similar[39] since most of the species (H_2O, OH, CH_x, and H) are at the interface or inside the oxide, and the nature of the chemical oxide is similar on both Si(111) and Si(100). A brief summary of the assignments is as follows:[30] the broad feature between 2800 and 3500 cm^{-1} along and the sharper feature at 1630 cm^{-1} are characteristic of H-bonded water molecules adsorbed on the internal surfaces; the strong shoulder at 3500–3650 cm^{-1} is due to H-bonded hydroxyl (–OH) groups that terminate the two oxide surfaces; the intense absorption band centered at 1205 cm^{-1} is assigned to the LO phonon of the thin chemical oxide network and the sharp feature at 2250 cm^{-1} to H dissolved in this oxide (chemically bonded to Si in the SiO_2 network).[40] Finally, the mode at 2100 cm^{-1} is best assigned to the stretching mode of hydrogen bonded to pure Si surfaces, which are therefore believed to exist at the Si/SiO_2 interface.[40] The interface region between the two oxide surfaces contains an estimated $3–4\times10^{15}$ H_2O/cm^2 and $1–2\times10^{15}$ OH/cm^2, all subject to H bonding. Clearly, in order to establish strong chemical bonds between the two oxide surfaces, these species must either react or dissolve into bulk material, thereby allowing the initial oxide surfaces to come into intimate contact.

In order to determine the precise fate of the interfacial OH and H_2O groups, we have spectroscopically monitored the evolution of both the molecular interface and the oxide layer as a function of annealing temperature (30 min in N_2 ambient for each temperature).[39] A few select spectra are shown in Figs. 9(a)–9(d) for the key annealing temperatures of 400, 600, 800, and 1000 °C, and the complete data are summarized in Figs. 10(a)–10(c). A substantial loss (~85%) of molecular water is observed after annealing to 400 °C, with complete loss evident at 600 °C. Interestingly, a small amount of reformation of H_2O occurs at 800 °C, indicating that the formal interface still persists at these temperatures, before complete closure occurs after annealing to 1000 °C. Correspondingly, there is a manifold (75%) increase in the absorbance of the oxide on annealing to 400 °C, corresponding to formation of additional ~7–8 Å of oxide ($3–4\times10^{15}$

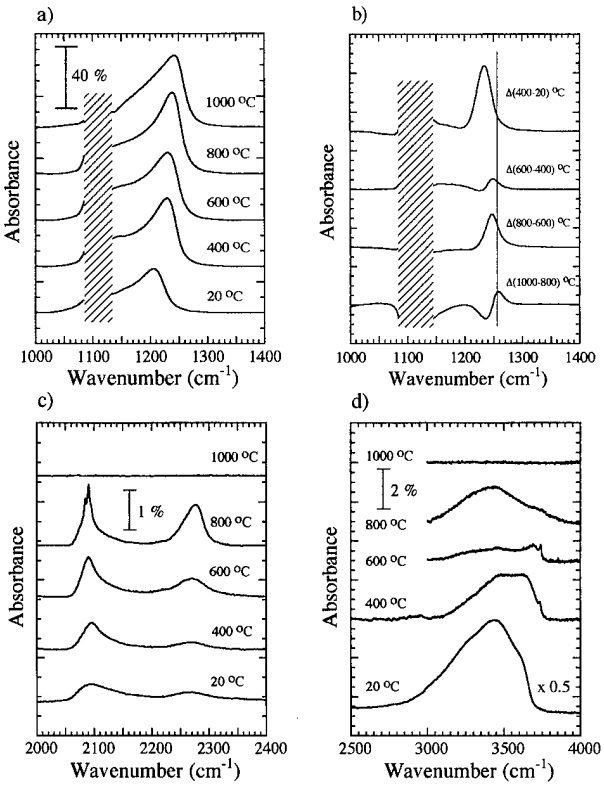


FIG. 9. MIT spectra of the ratio of p:s polarization for the different spectral regions of interest for hydrophilic Si(111) as a function of annealing temperature: (a) Si–O phonon modes; (b) difference spectra ($T_2 - T_1$) derived from the Si–O spectra shown in (a). The solid vertical line indicates the lower bound of the absorbance attributed to formation of bridging Si–O–Si groups; the hashed area indicates the spectral region for which no IR throughput is observed, due to absorption by the interstitial oxygen present in these CZ samples; (c) Si–H stretching modes; and (d) O–H stretching modes.

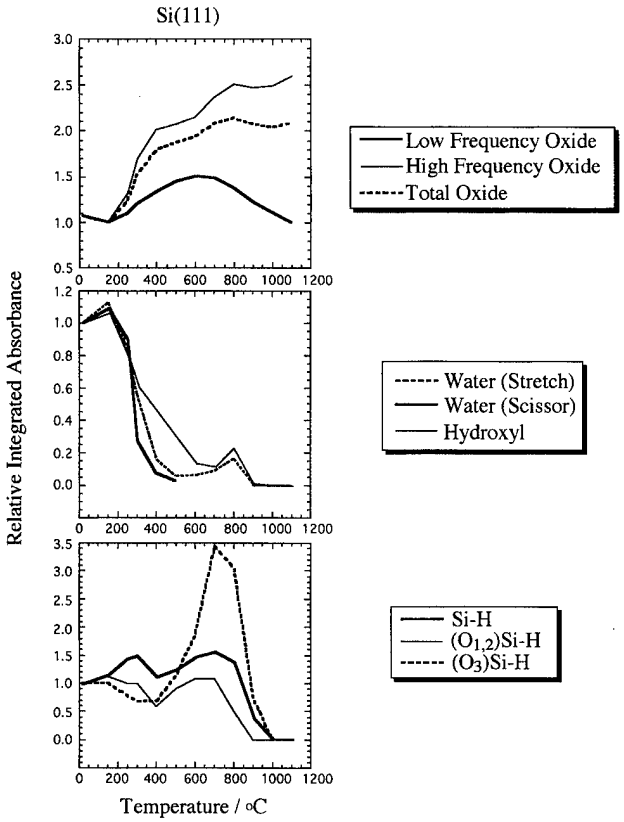


FIG. 10. Plots of the relative integrated absorbance of each of the modes shown in Fig. 8, as a function of temperature. Top panel: Si–O modes. "Low frequency oxide" is defined as absorbance between 1140 and 1180 cm^{-1}, whereas "high frequency oxide" is the remainder of the observed band (>1180 cm^{-1}); "total oxide" is the sum of the two. Middle panel: water and hydroxyl modes. Bottom panel: Si–H modes. The Si–H data are the integrated absorbance between 2050–2150 cm^{-1}; the (O$_{1,2}$)–Si–H and (O$_3$)–Si–H points are the integrated absorbances between 2175–2225 cm^{-1} and 2225–2300 cm^{-1}, respectively.

O atoms/cm^2). In addition, the peak (LO) frequency is shifted to 1235 cm^{-1}, consistent with growth of a more thermal-like (stoichiometric and less defective) oxide over this temperature range.

Significantly, the predominant molecular signature over the 400–800 °C temperature range is due to the interfacial hydroxyl species that terminate the oxide layers [Fig. 9(d)]. Specifically, the feature associated with these hydroxyl groups first sharpens and shifts to 3740 cm^{-1} (consistent with a loss of H bonding) on annealing to 600 °C, then subsequently broadens as the reformation of interface water occurs at 800 °C, before completely disappearing after a 900 °C anneal. Correspondingly, a high frequency oxide absorbance (1250–1260 cm^{-1}) appears over this temperature range, adding an extra 20%–30% to the original absorbance value [Fig. 8(b)]. This frequency is well above the highest LO frequency of the (thermal-like) oxide layer (1235 cm^{-1}), suggesting that it originates from a distorted Si–O–Si species associated with closure of the formal interface between the two surfaces.[41] Above 800 °C, the overall oxide absorbance remains constant, with a loss of the lower frequency component of the oxide band (1000–1150 cm^{-1}) and a sharpening of the LO phonon (1245 cm^{-1}), indicative of the formation of a good quality thermal oxide at these elevated temperatures.

The evolution of the Si–H spectra provides additional mechanistic information (Fig. 9(c)]. The most relevant fingerprint is the mode at 2250 cm^{-1}, associated with hydrogen inside the oxide matrix. The integrated absorbance of this mode initially decreases upon annealing to 400 °C (the range over which the molecular water signal is significantly attenuated), followed by a substantial further increase and blueshift at 700 °C. This increase suggests that a new source of hydrogen is available in the 400–700 °C range, with the most obvious candidate being the interface–OH groups, the thermal evolution of which was described above. Finally, there is another pronounced shift after a 900 °C anneal so that the absorbance maximum is now observed at 2285 cm^{-1}, suggestive of a change in the nature of the oxide, before complete loss of this feature occurs on annealing to 1000 °C. In contrast, the integrated absorbance of the 2100 cm^{-1} feature,

associated with hydrogen on internal Si surfaces at the Si/SiO_2 interface, is essentially constant up to 700 °C and decreases monotonically above this temperature. The fate of the hydrogen liberated between 800–1000 °C is not known: one possibility is that it simply migrates into the bulk (silicon) and diffuses away. However, a second intriguing possibility is that recombinant dihydrogen formation occurs, leading to bubble or intrinsic void formation in the interface region; indeed, we see preliminary spectroscopic evidence for such a reaction, although more work is required to substantiate these observations.

2. Model for oxidation

The primary (novel) observation in this work is the formation of a thicker oxide layer (and closure of the initial interface between the two thin oxide layers) upon annealing to 1100 °C, together with the disappearance of water, hydroxyl, and hydrogen. In addition, considerable insight into the pathways leading to oxidation can also be inferred from our data. First, there is a direct correlation between the dissociation of water initially trapped between the two joined hydrophilic surfaces and the increased "thermal-like" oxide below 400 °C. Indeed, the observed loss of ~3–4 monolayers of water ($\sim 3-4\times10^{15}$ oxygen/cm^2) over this temperature range is sufficient to form an additional 7–8 Å of SiO_2, in good agreement with experiment. Second, the thermal evolution of the "hydrogen-in-oxide" signature provides additional mechanistic information: the amount of hydrogen initially contained in the wet-chemical oxide decreases on heating to 400 °C, as the concentration of interface water decreases, but dramatically increases on annealing from 400 to 700 °C, as the disappearance of interface hydroxyl groups is observed. The former observation suggests that there is interaction between the oxidizing species and hydrogen in the oxide layer, i.e., there is migration of oxygen-containing species through the oxide layer. In contrast, the latter observation suggests that hydroxyl decomposition is the source of the additional hydrogen, over this temperature range. Indeed, the further oxidation observed between 400 and 800 °C is quantitatively consistent with the loss of the estimated $1-2\times10^{15}$ hydroxyl groups/cm^2. The high frequency component (1260 cm^{-1}) of the oxide band that becomes apparent over this temperature range, is also indicative of the formation of strained, Si–O–Si bridging bonds. The suggested model is shown schematically in Fig. 11 and summarized as follows.

(1) Below 400 °C, the water molecules trapped at the interface react and/or diffuse through the thin wet-chemical oxide to form additional oxide at the Si/SiO_2 interface. Using the tabulated values of $E_A=0.79$ eV and $D_0=10^{-6}$ cm^2 for diffusion of water through SiO_2,[42] we find that the transport of water through the thin oxide can take place at temperatures as low as 200 °C. However, the oxidation rate is predicted to be extremely slow at these temperatures, by extrapolation of data reported for conventional thermal oxidation processes.[43] The precise origin of this anomaly is not yet clear; it may be that the use of chemical oxides facilitates the oxidation process, by providing a reactive (highly strained or defective) Si/SiO_x interface. Alternatively, the constrained environment may promote the formation of an active oxidizing species that is unique to these systems.

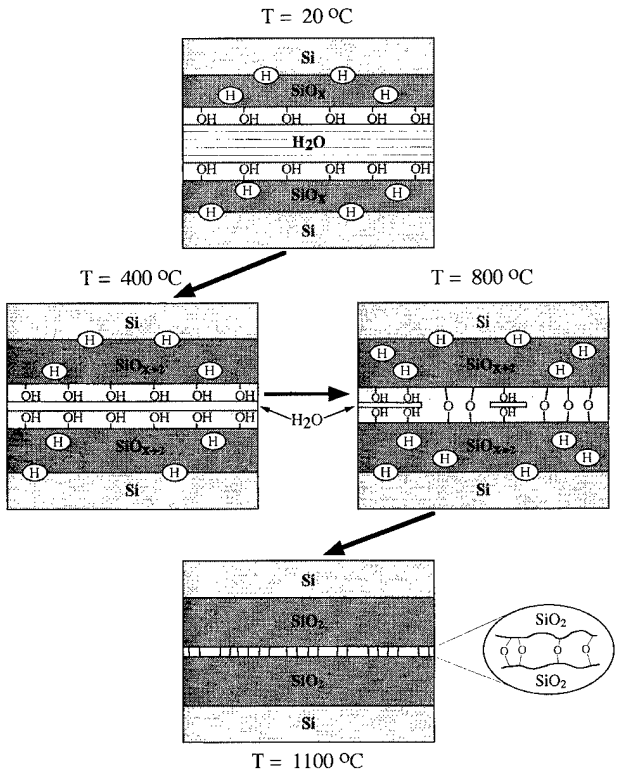


FIG. 11. Schematic of proposed mechanism for hydrophilic wafer bonding.

(2) Above 400 °C, additional Si–O–Si bond formation takes place, due to the decomposition of interface hydroxyl groups. The proposed mechanism is the elimination of molecular water by reaction of surface silanol groups:

$$-\mathrm{Si}-\mathrm{OH}+-\mathrm{Si}-\mathrm{OH}\Rightarrow-\mathrm{Si}-\mathrm{O}-\mathrm{Si}-+\mathrm{H_2O},$$

where some of the –Si–O–Si– linkages so-formed provide the chemical bridge between the two oxide surfaces, thereby closing the interface. The observation of molecular water reformation between 700–800 °C is clearly consistent with this reaction scheme. In addition, the appearance of a very high frequency band (1250–1270 cm^{-1}) between 400–800 °C, provides strong evidence for the formation of such bridging –Si–O–Si–, since this frequency range is most consistent with oxide growth remote from the Si/SiO interface (i.e., at the SiO/SiO interface).[41] Interestingly, this high frequency component of the oxide band displays the most inhomogeneity with lateral position along the interface, indicating that substantial variations exist in the formal closing of the interface. This inhomogeneity may be due, in part, to the intrinsic interface roughness or (possibly) the formation of voids or (H_2) bubbles. Studies are currently being undertaken using thin thermal oxide-terminated substrates, to further investigate these possibilities.

IV. CONCLUSIONS

In this article, we have presented an overview of the field of silicon wafer bonding. While the method is relatively simple, fundamental processes limit the quality of the bonding and ultimately its practical use. Experimentally, it is difficult to probe the interface spectroscopically since most of the traditional surface spectroscopies cannot penetrate through the silicon. Conversely, we have shown that not only can IR absorption spectroscopy be used effectively, but is in fact a very sensitive way to characterize the chemical and physical state of the interface. The results obtained so far provide the first glimpse at the relevant bonding mechanisms and raise additional important questions.

For hydrophobic wafers, the most exciting results are obtained on the flat H/Si(111)-(1×1) surfaces, where there is convincing evidence for strong physical interaction between the two contamination-free internal surfaces. The experimental challenge is to increase the level of interface perfection to the point where registry issues can be addressed (azimuthal rotation of one wafer with respect to the other). Furthermore, a detailed understanding of the thermal evolution of the hydrogen trapped at the interface, both as Si–H and possibly H_2, as well as the nature of the bonding between the two silicon surfaces upon annealing, is necessary to assess the validity of this method for low temperature epitaxy replacement. The actual bonding mechanism for atomically rough hydrophobic surfaces is difficult to establish, since it appears to involve only a small fraction of the surface, and therefore requires increased sensitivity to the atomic species at the interface. Such studies may be possible using the significantly higher intensity of synchrotron radiation sources which are optimal for the 300 to 800 cm^{-1} region,[44] where the majority of Si–X (X=O, C, N, F) stretching vibrations are known to occur.[45]

For hydrophilic wafers, the presence and evolution of molecular species such as H_2O, OH, CH_x and of dissolved hydrogen, have been well characterized using MIT. Based on our observations, detailed schematic models can be proposed. However, the challenge is clearly to probe and interpret the IR absorption associated with silicon oxide species. While submonolayer oxygen contamination experiments together with *ab initio* cluster calculations will help one to understand the local oxygen structures that are initially formed, the characterization of the oxide films will demand a more complex analysis. The complexity arises from the strong Coulombic coupling within the SiO_2 network, image charge effects for thin films, stoichiometric variations, different interface structures, and strain due to the mismatch between the silicon lattice and the SiO_2 film. However, meeting this challenge is critical since IR absorption measurements may be the only nondestructive way to monitor oxide formation in a wide variety of technologically relevant processes.

Finally, the formation of intrinsic voids at intermediate annealing temperatures has not been addressed in this article, although it remains a critical issue in silicon wafer bonding due to the potential effect on interface homogeneity. Indeed, while a number of studies of the size distribution and number density of interfacial voids have been reported in the literature,[15] a comprehensive understanding of the effect of surface cleaning and rinsing (i.e., surface passivation, residues, physisorbed species) on the concentration and nature of the resultant voids has not yet emerged. Therefore, there is clearly a need to devise ways to characterize both the interface homogeneity and the physical and chemical state of such voids as a function of annealing temperature.

[1] *Proceedings of the 1st, 2nd, and 3rd International Symposia on Semiconductor Wafer Bonding: Science, Technology and Applications*, The Electrochemical Society Proceedings Series (Electrochemical Society, Pennington, NJ, 1992–1995).
[2] J. B. Lasky, S. R. Stiffler, F. R. White, and J. B. Abernathey, IEDM Tech. Dig., 684 (1985); J. B. Lasky, Appl. Phys. Lett. **48**, 78 (1986).
[3] W. G. Maszara, J. Electrochem. Soc. **138**, 341 (1991).
[4] H. Baumann, S. Mak, and H. Münzel, Electrochem. Soc. Proc. **95–7**, 471 (1995).
[5] Y. J. Chabal, M. A. Hines, and D. Feijoo, J. Vac. Sci. Technol. A **13**, 1719 (1995).
[6] *Handbook of Semiconductor Wafer Cleaning Technology: Science, Technology and Applications*, edited by W. Kern (Noyes, Park Ridge, NJ, 1993), p. 433.
[7] B. E. Deal and D. B. Kao, in *Proceedings of the Tungsten and Other Refractory Metals for VLSI Applications II*, edited by E. K. Broadbent (Materials Research Society, Pittsburgh, PA, 1987), p. 27.
[8] Y. J. Chabal, G. S. Higashi, K. Raghavachari, and V. A. Burrows, J. Vac. Sci. Technol. A **7**, 2104 (1989).
[9] G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari, Appl. Phys. Lett. **56**, 656 (1990); G. S. Higashi, R. S. Becker, Y. J. Chabal, and A. J. Becker, *ibid.* **58**, 1656 (1991).
[10] R. Stengl, K.-Y. Ahn, and U. Gösele, Jpn. J. Appl. Phys. **27**, L2364 (1988).
[11] G. A. C. M. Spierings and J. Haisma, in *Proceedings of the First International Symposium on Semiconductor Wafer Bonding: Science, Technology and Applications*, edited by U. Gösele, J. Haisma, T. Abe, and M. A. Schmidt, The Electrochemical Society Proceedings Series (Electrochemical Society, Pennington, NJ, 1991), PV 92-97, p. 18.
[12] H. Baumgart, T. J. Letavic, and R. Egloff, Philips, J. Res. **49**, 91 (1995).
[13] W. Maszara, G. Goetz, A. Caviglia, and J. B. McKitterick, J. Appl. Phys. **64**, 2987 (1986).
[14] Y. Backlund, K. Ljungberg, and A. Soderbarg, J. Microeng. **2**, 158 (1992).
[15] T. Abe, A. Uchiyama, K. Yoshizawa, Y. Nakazato, M. Miyawaki, and T. Ohmi, Jpn. J. Appl. Phys. **29**, L2315 (1990); S. Bengtsson and O. Engstrom, J. Electrochem. Soc. **137**, 2297 (1990); Q.-Y. Tong and U. Gösele, J. Microelectromech. Syst. **3**, 29 (1994).
[16] Y. J. Chabal, Surf. Sci. Rep. **8**, 211 (1988).
[17] R. J. Collins and H. Y. Fan, Phys. Rev. **93**, 674 (1954).
[18] P. Jakob, Y. J. Chabal, and K. Raghavachari, Chem. Phys. Lett. **187**, 325 (1991); P. Jakob, Y. J. Chabal, K. Raghavachari, P. Dumas, and S. B. Christman, Surf. Sci. **285**, 251 (1993).
[19] P. Jakob and Y. J. Chabal, J. Chem. Phys. **95**, 2897 (1991).
[20] M. Suhren, Y. J. Chabal, K. Raghavachari, and S. B. Christman (unpublished).
[21] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965).
[22] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979); D. R. Hamann, Phys. Rev. B **40**, 2980 (1989).
[23] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[24] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980); J. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[25] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[26] F. Tassone, F. Mauri, and R. Car, Phys. Rev. B **50**, 10561 (1994).
[27] The three-fold site chosen was that above the empty site in the second Si layer. There was no indication that significantly different results would be obtained for the other threefold site.
[28] A. Baldereschi, Phys. Rev. B **7**, 5212 (1973); D. J. Chadi and M. L. Cohen, *ibid.* **8**, 5747 (1973); H. J. Monkhorst and J. D. Pack, *ibid.* **13**, 5188 (1976).
[29] N. D. Lang, Phys. Rev. Lett. **46**, 842 (1981).

[30]D. Feijoo, Y. J. Chabal, and S. B. Christman, Appl. Phys. Lett. **65**, 2548 (1994).
[31]Y. J. Chabal, Surf. Sci. **168**, 594 (1986).
[32]Y. J. Chabal and C. K. N. Patel, Phys. Rev. Lett. **53**, 210 (1984).
[33]See for example: L. Ling, S. Kuwabara, T. Abe, and F. Shimura, J. Appl. Phys. **73**, 3018 (1993); D. Graf, M. Grundner, and R. Schulz, J. Vac. Sci. Technol. A **7**, 808 (1989).
[34]J. A. Schaefer, D. Frankel, F. Stuki, W. Gopel, and G. J. Lapeyre, Surf. Sci. **139**, L209 (1984).
[35]H. Ibach, H. D. Bruchmann, and H. Wagner, Appl. Phys. A **29**, 113 (1982).
[36]D. W. Berreman, Phys. Rev. **130**, 2193 (1963).
[37]T. Yamazaki, C. H. Bjorkman, S. Miyazaki, and M. Hirose, Proc. 22nd ICPS/1994 **3**, 2653 (1994).
[38]*The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1988).
[39]M. K. Weldon, Y. J. Chabal, S. B. Christman, E. E. Chaban, L. C. Feldman, C. M. Hsieh, and C. A. Goodwin, Proceedings of the Electrochemical Society, Pennington, NJ (to be published).
[40]H. Ogawa and T. Hattori, Appl. Phys. Lett. **61**, 577 (1992).
[41]G. Lucovsky, J. T. Fitch, E. Kobeda, and E. A. Irene, *Proceedings of the Electrochemical Society, Atlanta* (Electrochemical Society, Pennington, NJ, 1988), p. 139.
[42]*VLSI Technology*, edited by S. M. Sze (McGraw-Hill, New York, 1983), Chap. 5, Sec. 7, Table 4.
[43]L. E. Katz, in *VLSI Technology*, 2nd ed., edited by S. M. Sze (McGraw-Hill, New York, 1988), p. 99.
[44]G. P. Williams, Rev. Sci. Instrum. **63**, 1535 (1992); C. Hirschmugl, G. P. Williams, Y. J. Chabal, and F. M. Hoffmann, J. Vac. Sci. Technol. A **12**, 2229 (1994).
[45]See P. Dumas and Y. J. Chabal, Chem. Phys. Lett. **181**, 537 (1991), and references therein.

Nonlinear optical spectroscopy of Si–heterostructure interfaces

C. Meyer,[a] G. Lüpke, Z. G. Lü,[b] A. Gölz, and H. Kurz
Institute of Semiconductor Electronics II, Rheinisch-Westfälische Technische Hochschule, D-52056 Aachen, Germany

G. Lucovsky
Departments of Physics, Material Science and Engineering, and Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 21 January 1996; accepted 20 April 1996)

Strain, dislocations, and electrically active defects at and near the interface of Si/SiO_2 and Si/GaP heterostructures are analyzed by optical second-harmonic spectroscopy. For plasma oxides deposited on Si(001) and Si(111), time-dependent second-harmonic experiments reveal that near-interface oxide defects trap charge by the tunneling of photoexcited electrons from the Si conduction band. The space-charge field-induced second-harmonic transients are resonantly enhanced by two-photon E_1 transitions in silicon. In GaP epilayers grown on Si(001) the bulk dipole-allowed electro-optical effect is suppressed by the formation of antiphase domains. In contrast, in GaP films grown on Si(111) and vicinal Si(001) the density of antiphase domains is considerably reduced yielding an enhancement of the second-order nonlinear optical response by two orders of magnitude. © *1996 American Vacuum Society.*

I. INTRODUCTION

Silicon-based heterostructures are of outstanding importance for micro- and optoelectronic device technology and of considerable scientific interest. Device limiting structural imperfections and process-induced strain at the Si/SiO_2 interface are a main obstacle on the way towards smaller and faster device structures, which require new approaches to oxide growth technology such as low temperature plasma deposition. Additionally, due to the exceptional perfection, purity, and low cost of silicon, great interest exists in developing Si-based optoelectronic devices which allow the integration of optical components on conventional electronical integrated circuits. The main problem encountered in heteroepitaxial growth of III–V compound semiconductors on silicon is the high density of dislocations in the epitaxial layer resulting from the formation of antiphase domains (APD), which can be overcome by careful choice and special treatment of the substrate material.[1–3] The development of new improved heterostructure devices requires powerful non-invasive analytical techniques for the characterization of the buried solid/solid interfaces, which are difficult to probe with traditional optical and electron-based surface diagnostics.

In recent years, optical second-harmonic generation (SHG) has proven to be an effective and sensitive tool for studies of a large variety of interface properties.[4,5] An electric dipole second-harmonic (SH) response is forbidden in centrosymmetric media such as Si or SiO_2, making SHG intrinsically sensitive to the interfacial region where the inversion symmetry is broken. Several studies on Si/SiO_2 interfaces have employed the tensorial characteristic of SHG to examine the chemical bonding arrangement at atomic-scale step[6,7] and kink sites.[8] These results could be correlated with electrical measurements of the midgap density of interface traps on the same samples. The spectral resolution of optical techniques has been used to identify a strain and disorder-induced redshift and broadening of the Si bulk direct band gap E_1 and E_0' transitions at the Si/SiO_2 interface.[9–11] Most recently, electric-field induced SHG has been used to study electroinduced effects at Si based MOS structures.[12–14] SH-transients on a time scale of seconds have been observed on native oxide interfaces by several groups.[15,16] However, the interpretation of this effect attributed to "charge transfer effects" by Mihaychuk *et al.* or explained entirely by a thermal-induced effect by Dadap *et al.* remains controversial.

In this article we report on dispersion and time-resolved SHG measurements for the characterization of electrically active defects at Si/SiO_2 interfaces and antiphase domains in GaP-epilayers grown on Si(111) and Si(001). For Si/SiO_2 interfaces, a space-charge field-induced SH transient reveals that a high density of near interface oxide traps occurs on plasma deposited oxides, which can be reduced significantly by a rapid thermal annealing of the oxide. A tunneling model is introduced to explain the observed SHG time dependence. Furthermore, SHG has been used to study the influence of APDs formed during the growth of Si/GaP heterostructures on the electrooptical response of the GaP-epilayer. The nonlinear optical response of the GaP-epilayers exhibits a high density of APDs on a flat Si(001) substrate due to its double domain structure, whereas on Si(111) and vicinal Si(001) the SH response is enhanced by about two orders of magnitude.

The article is organized as follows: Section II introduces the experimental technique and the data analysis procedure. In Sec. III the SH transients and spectra are presented for (a) the Si/SiO_2 interfaces and (b) the Si/GaP epilayers. The conclusions are given in Sec. IV.

[a]Electronic mail: meyer@basl.rwth-aachen.de
[b]Permanent address: Department of Applied Physics, South China University of Technology, Guangzhou, People's Republic of China.

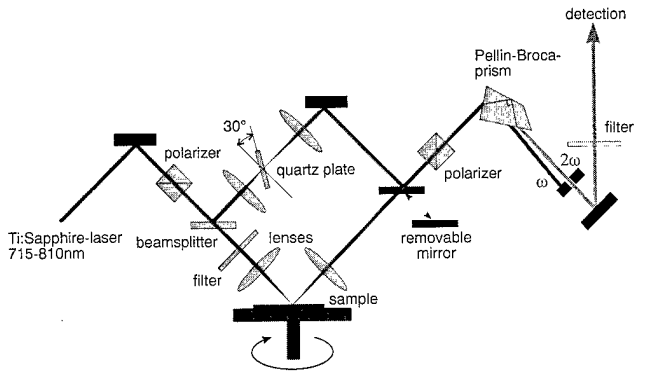


FIG. 1. Schematic view of the experimental setup used for the SHG measurements.

II. EXPERIMENTAL CONDITIONS AND SAMPLE PREPARATION

The optical measurements were performed using 710–815 nm radiation of a passively mode-locked titanium:sapphire laser generating 100 fs pulses at a repetition rate of 76.5 MHz. The output was attenuated to 30 mW average power and focused on the sample to a 15-μm-diam spot at a 45° angle of incidence. The experimental setup is shown in Fig. 1. The reflected light is spectrally filtered by a prism and glass filters and the SH signal is detected using a single photon counting device. All measurements were performed in air.

The electric field at the SH frequency, $E(2\omega)$, is generated by a nonlinear polarization $P(2\omega)$ determined by the nonlinear susceptibility tensor $\chi^{(2)}$, and can be expanded in terms of the electric field:

$$\begin{aligned} E(2\omega) &\propto P(2\omega) \\ &\propto \chi^{(2)}_{d,B} E(\omega)E(\omega) + \chi^{(2)}_{d,S} E(\omega)E(\omega) \\ &\quad + \chi^{(2)}_{q,B} E(\omega)\nabla E(\omega) + \chi^{(3)}_{d,B} \mathscr{E}(\omega_{\mathscr{E}} \sim 0) \\ &\quad \times E(\omega)E(\omega) + \cdots . \end{aligned} \tag{1}$$

The first term in this expansion, the dipole-allowed bulk nonlinearity $\chi^{(2)}_{d,B}$, is only allowed in non-centrosymmetric media such as III–V compound semiconductors. In silicon or SiO_2 this contribution is symmetry-forbidden, allowing the next terms to dominate the SH response: the dipolar surface term $\chi^{(2)}_{d,S}$, the bulk quadrupolar contribution $\chi^{(2)}_{q,B}$, and the dipolar bulk term $\chi^{(3)}_{d,B}$. The surface term $\chi^{(2)}_{d,S}$ appears because of the broken inversion symmetry at surfaces and interfaces, and is comparable with the bulk allowed $\chi^{(2)}_{q,B} E(\omega)\nabla E(\omega)$. The last term in Eq. (1) gives rise to an electric field-induced SH response induced by a (quasi) static electric field $\mathscr{E}$ via the bulk dipole $\chi^{(3)}_{d,B}$ which is also present in centrosymmetric media.[17,18] The different $\chi^{(2)}$ tensor elements reflect the sample's surface and bulk electronic structure and can be reduced to a set of phenomenological constants describing the SH response as a function of the sample orientation.

For SH dispersion measurements, we detected the p-polarized SH response $I^{(2\omega)}_{p,p}$ produced by a p-polarized fundamental field as a function of the sample azimuth ϕ. For a surface of C_{1v} and higher symmetry, $I^{(2\omega)}_{p,p}$ can be written phenomenologically:

$$I^{(2\omega)}_{p,p}(\phi) = \left| \sum_{m=0}^{4} a^{p,p}_m \cos(m\phi) \right|^2 , \tag{2}$$

where $a^{p,p}_m$ are the complex Fourier coefficients of the SH field and ϕ is defined with respect to a crystal mirror plane.[19] Since the same polarization combination is used throughout this work, the index p,p is omitted in the following. Using Eq. (2) to fit the azimuthal dependence of the SH intensity measured at different wavelengths, the dispersion of $|a_m|$ is deduced. These coefficients can be attributed to certain structural elements of the interface, e.g., a_1 is partially caused by the bonding arrangement at atomic steps on the substrate exhibiting C_{1v} symmetry.

Optically flat Si(001) and Si(111) standard wafers cut with misorientations up to 6° in the (110) or (11$\bar{2}$) direction, respectively, were used as substrates and submitted to RCA cleaning before growing the heterostructures. Si/SiO_2 interfaces were formed in two different ways, leading to oxide layers in the range of $\approx$30 nm. For deposited oxides, the wafers were subjected first to a preoxidation step at 300 °C in an oxygen plasma followed by a remote plasma enhanced chemical vapor deposition (RPECVD) step.[20] Additionally, a second set of samples was thermally oxidized in dry oxygen at 850 °C. The samples were partially subjected to a rapid thermal annealing (RTA) step (100 s at 1050 °C in argon) after oxidation. The GaP films are grown by pulsed chemical beam epitaxy (PCBE) to a thickness of 6–9 nm monitored *in situ* by p-polarized reflectance (PRS) and reflectance difference spectroscopy (RDS).[21]

III. RESULTS AND DISCUSSION

A. Nonlinear electroreflectance from Si/SiO_2 interfaces

Electric field-induced SHG allows the direct observation of charge trapping at near-interface defects, which alters the reflected SH signal due to a change of the static electric field $\mathscr{E}$ in the space-charge region at the interface. For this experiment, the samples were rotated to a position where only a_0 is detected [$\phi=90°$ for Si(111), $\phi=45°$ for Si(001)]. In this case, $\chi^{(3)}_{d,B}\mathscr{E}(\omega_{\mathscr{E}}\sim 0)E(\omega)E(\omega)$ contributes to $I^{(2\omega)}$, and a_0 can be split up into a complex interface term $a^{(int)}_0$ and the complex electric field–induced term $a^{(fi)}_0$. Thus, using Eqs. (1) and (2), the SH intensity can be written as

$$I^{(2\omega)}(\mathscr{E}) = |a^{(int)}_0 + a^{(fi)}_0 \mathscr{E}|^2 . \tag{3}$$

Time-dependent studies of the SH response as shown in Fig. 2 were performed by moving the sample laterally, then stopping the motor at $t=0$, and recording the SH signal as a function of time. Specific for the SH transient is the fast decay at the beginning, which slows down significantly at longer times. The time-dependent effect is observed on

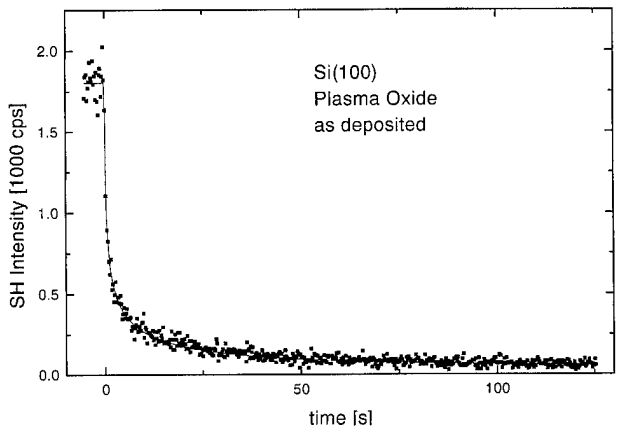


FIG. 2. Temporal decay of the SH intensity of a plasma oxide grown on Si(001).

plasma-deposited oxides, but after annealing of the interface, it is reduced by at least two orders of magnitude. A similar situation prevails for thermal oxides, where there is no detectable time dependence of the SHG response. Hence we can exclude laser-induced surface heating as a reason for the observed SH-transient as proposed earlier.[16] The temporal signal change in the unannealed plasma oxides cannot be described by a simple exponential decay, but can be well fitted by a tunneling model (solid line in Fig. 2), in which photoexcited electrons tunnel directly from the Si conduction band-edge. The tunneling process will be described later.

To identify the origin of this effect we performed time-dependent measurements at different input wavelengths. Figure 3(a) shows the spectral dependence of $I(\mathscr{E}(t=0))-I(\mathscr{E}(t=\infty))$ for a Si(111)–plasma–oxide interface. For comparison, and to avoid the influence of a possible wavelength dependent oxide trap-induced electric field, the same spectral regime was measured on a MOS structure, to which an externally controlled electric field has been applied. By using a thermal oxide the MOS sample shows no time-dependent effect. The metallization was made semitransparent to allow for optical measurements at the Si/SiO_2 interface. We used a metal contact made of 3 nm Cr and 6 nm Au. The spectral dependence of $\Delta^{(2\omega)}(U)=I^{(2\omega)}(U)-I^{(2\omega)}(0\mathrm{V})$ for $U=12$ V is shown in Fig. 3(a). Both spectra exhibit a strong resonance at the direct band gap energy of 3.4 eV, revealing that bulk electronic transitions resonantly enhance $\chi^{(3)}_{d,B}$ and thus showing that an electric field–induced effect caused by trapped oxide charges is responsible for the SH time-dependence observed on plasma-deposited oxides. The spectra are not corrected for the dispersion of the linear optical parameters, which cause only a small reduction (reabsorption of the SH radiation) and redshift (<10 meV) of the resonance features, as calculated from the dispersion data given in Ref. 22. The trap-induced alteration in the interfacial space charge field $\Delta\mathscr{E}$ is given by

$$\epsilon_{\mathrm{Si}}\Delta\mathscr{E}=-\Delta Q_{\mathrm{Si}}=q\int_0^d n_{ot}dz, \tag{4}$$

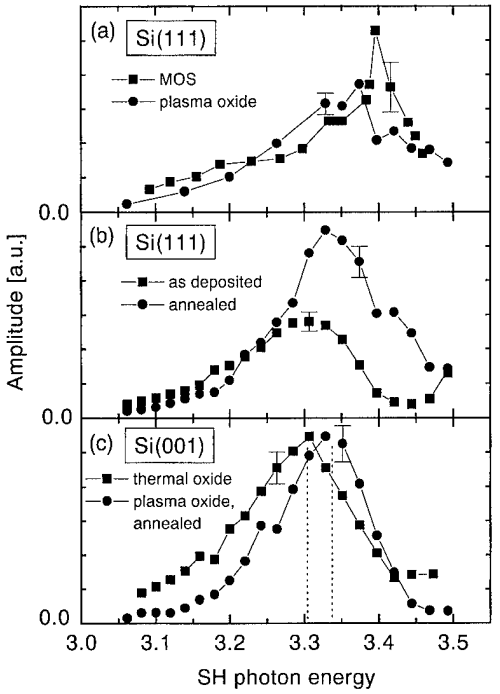


FIG. 3. Spectral dependence of (a) $|a_{fi}\Delta\mathscr{E}_z|^2$ of a plasma oxide grown on Si(111) and $\Delta^{(2\omega)}(U)$ of a Si(111) MOS structure biased at $U=-12$ V; (b) $I^{(2\omega)}(\mathscr{E})$ of a plasma oxide grown on Si(111) and after RTA; and (c) $I^{(2\omega)}$ of a thermal oxide and an annealed plasma oxide, both grown on Si(001).

where Q_{Si} and ϵ_{Si} are the space-charge density per unit area and the static dielectric constant of Si, q is the carrier (electron or hole) charge, d is the thickness of the interfacial oxide trap region, and n_{ot} is the volume density of trapped charges.

To determine the polarity of $\Delta\mathscr{E}$, and thus the signum of the trapped charge, we measured the rotational SH anisotropy from the MOS structure described above for a distinct applied external voltage U. Figure 4 shows the rotational anisotropy of both, the SH signal from the thermally grown $Si(111)/SiO_2$ interface of the MOS structure (●), i.e., without the metal contact, given by

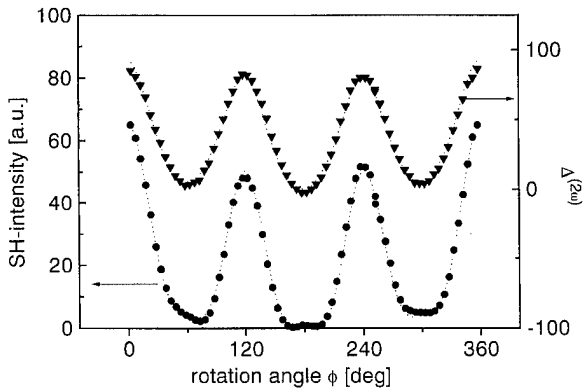


FIG. 4. Rotational anisotropy of (●) the SH intensity of a thermally grown oxide/silicon interface (▼) for the difference $\Delta^{(2\omega)}(U)=I^{(2\omega)}(U)-I^{(2\omega)}\times(0)$ for $U=3.0$ V measured on a MOS structure on Si(111).

$$I^{(2\omega)}(\phi)=|a_0^{(int)}|^2+|a_3|^2\cos^2(3\phi)+2|a_0^{(int)}a_3|\cos[\Delta\psi(a_0^{(int)},a_3)]\cos(3\phi), \quad (5)$$

and second the difference $\Delta^{(2\omega)}(U=4\text{V})$ measured from the MOS structure (▼) showing a rotational dependence of the form

$$\Delta^{(2\omega)}(\phi)=2\mathscr{E}|a_0^{(fi)}a_0^{(int)}|\cos[\Delta\psi(a_0^{(fi)},a_0^{(int)})]+2\mathscr{E}|a_0^{(fi)}a_3|\cos[\Delta\psi(a_0^{(fi)},a_3)]\cos(3\phi)+\mathscr{E}^2|a_0^{(fi)}|^2, \quad (6)$$

where $\Delta\psi(a,b)$ denotes the phase difference between the complex numbers a and b. Fourier analysis of both curves yields the relative phases between $a_0^{(int)}$ and a_3, and $a_0^{(fi)}$ and a_3, respectively. Figure 2 shows that $a_0^{(int)}$ and $a_0^{(fi)}$ are either in phase or anti-phase dependent on the polarity of $\Delta\mathscr{E}$. The rotational anisotropies shown in Fig. 4 clearly reveal that both isotropic contributions are in phase, leading to the conclusion that the decrease of the SH signal is caused by a negative $\Delta\mathscr{E}$, i.e., caused by electron capture at near-interface oxide traps.

The next question is concerned with the charge state of the observed oxide traps. To clarify whether the defects are initially positively charged or neutral, SH spectra have been taken from plasma oxides annealed or as grown on Si(111). Figure 3(b) shows the SH signal from both samples after illuminating the samples for at least 10 minutes, thereby saturating the electric field-induced SH contribution (see Fig. 2). Compared to the annealed plasma oxide, the as-grown sample reveals a lower SH intensity in the regime of the $\chi^{(3)}$ resonance at 3.4 eV, while the annealed sample which exhibits no SH transient shows the resonance feature of the interfacial $\chi_{d,S}^{(2)}$.[9–11] This indicates that the near-interface oxide traps are initially neutral.

Figure 3(c) shows a comparison between the dispersion of a thermal and a plasma oxide both grown on Si(001). The plasma deposited oxide was subjected to RTA, so that both samples exhibit no electric field induced effect by charging of near-interface oxide traps. However, the peak in the SH response is significantly shifted for the thermally oxidized sample. The observed resonance peaks at 3.35 eV for the deposited oxide, while the thermal oxide peaks at 3.3 eV. In previous SH studies this redshift from the bulk value (3.37 eV) has been attributed to electronic transitions in a few monolayers of strained Si–Si bonds in the interfacial region. The larger redshift in the SH response from the thermal oxide indicates a greater interfacial strain induced by the higher thermal budget and thus a different interface structure. This structure remains different even after annealing at 1050 °C, which has also been observed by infrared Si–O stretching frequency measurements.[23]

As mentioned before, the time dependence of the SH signal cannot be explained by a simple exponential decay. Furthermore, excitation of electrons by reabsorption of bulk generated third-harmonic light could be ruled out as a reason for the observed SH-transient, since the decrease in signal is visible after illumination with linearly and circularly polarized light, the latter showing a third-harmonic efficiency which is three orders of magnitude smaller. Since the photon energy of our laser radiation (~1.5 eV) exceeds the silicon indirect band gap energy of 1.11 eV, one possible mechanism is tunneling of electrons from the conduction band to the trap states.[24] In this model, the tunneling time τ depends exponentially on the tunneling distance z: $\tau(z)=\exp(\alpha_T z)$, where α_T denotes the oxide specific tunneling constant. Thus the trapping of photoexcited electrons in the oxide can be described by the following rate equation:

$$\frac{\partial n_{ot}}{\partial t}=-\frac{n_{ot}}{\tau_d(z)}+\frac{c_e}{\tau(z)}(N_{ot}-n_{ot}), \quad (7)$$

where $\tau_d(z)=\tau_d\exp(\alpha_T z)$ is the detrapping time, and $c_e=n_e/N_{ot}$ with N_{ot} and n_e the volume densities of oxide traps and photoexcited electrons in the interface region. This equation and the initial condition $n_{ot}(t_0)=0$ for the density of charged traps lead to the following expression for the time and z-dependent trapped-electron density:

$$n_{ot}(t,z)=N_{\max}(1-\exp\{-[\tau_d(z)^{-1}+c_e\tau(z)^{-1}]t\}), \quad (8)$$

with $N_{\max}=n_eN_{ot}/[N_{ot}(\tau/\tau_d)-n_e]$. The change in the electric field at the interface $\Delta\mathscr{E}$ can be calculated from Eq. (4) by integrating for $d\to\infty$, which is reasonable as typical tunneling distances are far below the oxide thickness and thus the density n_{ot} strongly decays away from the interface. The integration leads to

$$\Delta\mathscr{E}=\frac{qN_{\max}}{\alpha_T\epsilon_{Si}}\left[\gamma+\ln\left(\frac{t}{t_c}\right)+\mathrm{E}_1\left(\frac{t}{t_c}\right)\right], \quad (9)$$

where γ is Euler's constant, E_1 is the first order exponential integral, and $\tau_c=c_e\tau^{-1}+\tau_d^{-1}$. Inserting $\mathscr{E}=\mathscr{E}_0+\Delta\mathscr{E}$ in Eq. (3) gives $I^{(2\omega)}(\mathscr{E})$ as a function of time with the parameters τ_c, $|a_0^{(\text{int})}|$, $|a_0^{(fi)}|$, and $\Delta\psi(a_0^{(\text{int})},a_0^{(fi)})$, where $a_0^{(fi)}\mathscr{E}_0$ is absorbed in $|a_0^{(\text{int})}|$. The solid line in Fig. 2 shows a best fit using this function.

The detrapping time τ_d has been measured by performing several charging and discharging cycles yielding a $\tau_d\approx10^3$ s. For this experiment the sample was illuminated for 5 minutes and then kept in the dark for times ranging from 10 s to 2 days. To investigate the intensity dependence of τ_c, we repeated the measurement shown in Fig. 2 at different laser intensities ranging from 5 to 20 kW/cm^{-2}. Analysis of the data showed that $N_{\max}$ is nearly independent of the laser fluence, indicating that a saturation of charge trapping has been reached. Furthermore the values of τ_c^{-1} shown in Fig. 5 depend linearly on the laser intensity, indicating that the observed carrier trapping is rather dominated by tunneling of photoexcited electrons from the Si conduction band than by multiphoton electron transitions into the SiO_2 conduction band. From Fig. 5 a trapping time $\tau\approx10^{-1}$ s is deduced, which is four orders of magnitude shorter than τ_d, indicating that the band-to-trap tunneling process is followed by atomic relaxation at the defect.

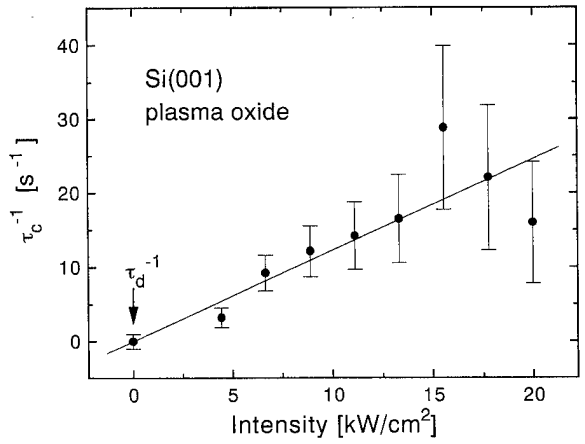


FIG. 5. Charging rate τ_c^{-1} as a function of $I^{(\omega)}$. The solid line is the best fit to the data.

B. Si/GaP

The integration of silicon and compound semiconductor devices is of tremendous interest in the context of future advanced microelectronic circuits, e.g., vertically integrated circuits, optically interconnected memory, and integrated sensor circuits. The main problem in growing III–V compounds on Si substrates is the high density of dislocations in the epitaxial layer. This has stimulated attempts to grow first a buffer layer of a lattice-matched compound, e.g., GaP, that seals the silicon surface, which is followed by the deposition of an appropriately graded layer of the desired lattice-mismatched compound.[25] However, despite many investigations mechanisms for the formation and annihilation of antiphase domains (APD), which are generated at single atomic steps on Si, have not been clarified yet.[1,26] This is due to the lack of experimentation and the difficulty of the evaluation of APD.

SHG rotational analysis is a promising technique to reveal the symmetry characteristics and to investigate the possibility of APD formation in III–V compound epilayers on silicon.[27] In contrast to Si, these compound semiconductors are non-centrosymmetric, and assuming a defect free epilayer, should possess an electrooptical coefficient or nonlinear susceptibility, $\chi^{(2)}$, which is orders of magnitude larger than in Si. For a (001) orientation of a cubic epilayer, e.g., GaX (X=As, P), this would lead to a rotational SH anisotropy containing a dominating a_2 contribution, while for a (111) surface a_0 and a_3 are dominating the SH signal.[27]

The dominant SH Fourier coefficients of the SHG rotational anisotropy of Si(111)/GaP and Si(001)/GaP epilayers obtained by fitting Eq. (2) to the data are summarized in Fig. 6 as a function of SH photon energy. For the Si(001) substrate, the a_2 coefficient is below the detection limit of our apparatus. The SH response is dominated by the isotropic component a_0 shown in Fig. 6(a), and is comparable in magnitude to the a_0 component of the Si/SiO_2 interface. However, the observed resonance at 3.3 eV is missing here, which can be explained by a broadening and damping of the E_1 resonance due to inhomogeneous strain present at the interface. The vanishing of a_2 can be explained by the crystalline structure of the Si(001) substrate. Such wafers are typically oriented to $<0.5°$, leading to ≈ 1000 monoatomic steps across our laser spot size. The local C_{2v} symmetry of the terraces rotates by 90° when a monoatomic step is encountered, leading to a macroscopic C_{4v} symmetry of the sample. The growth of a III–V epilayer on flat Si(001) may lead to the formation of APDs in the epilayer[1,26] with domain boundaries located at the monoatomic substrate steps. The averaging over many APDs cancels the net bulk dipole-allowed $\chi^{(2)}$ in the GaP film leading to fatal consequences for the application of III–V compound semiconductors grown on Si(001) as electrooptic materials.

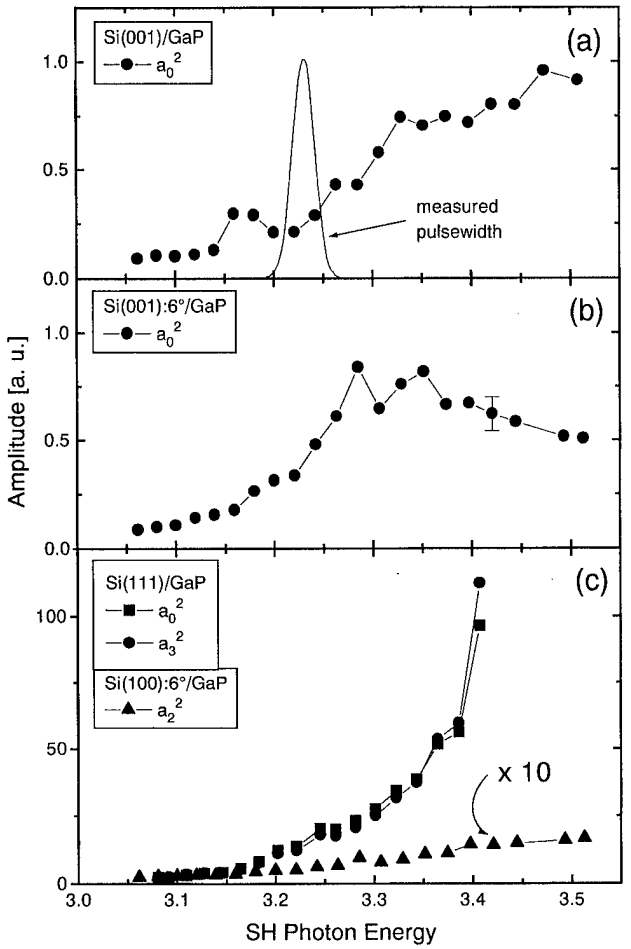


FIG. 6. Dispersion of (a) the isotropic contribution $|a_0|^2$ of a GaP-epilayer grown on Si(001) and (b) on vicinal Si(001), cut 6° towards (111); (c) the isotropic and threefold component of a GaP-film grown on Si(111) and the twofold component of a film grown on vicinal Si(001).

For an effective use of silicon-based heterostructures in optoelectronic devices, this problem has to be overcome. One possibility is to use vicinal Si(001), which possesses a dominant part of diatomic steps, assuring that contributions of different domains remain in phase. This has two effects on the measured SH spectra: as shown in Fig. 6(b) the isotropic component a_0 exhibits the redshifted 3.3 eV resonance, which is still significantly broader than those from Si/SiO_2 interfaces (see Fig. 3), indicating that the Si/GaP interface is

more ordered on vicinal than on flat Si(001) substrates. Additionally, as shown in Fig. 6(c) a significant twofold contribution a_2 is obtained from the vicinal Si(001)/GaP heterostructure, demonstrating that the density of APDs is considerably reduced.

Also shown in Fig. 6(c) are the data measured from GaP-films grown on Si(111) substrates. The results reveal a nonresonant dispersion curve with a signal lying two orders of magnitude above the Si/SiO_2 interface response. Note that the relative magnitude of the nonlinear optical response is still almost two orders of magnitude higher than for vicinal Si(001) substrates, indicating that a single-domain GaP film is not yet achieved on vicinal Si(001). Although it has been shown that an annihilation of APDs is expected for thick films,[1] our film thicknesses are well below the required value. Surface steps on Si(111) are naturally occurring as diatomic steps solving the problem of APDs in the GaP film and thus allowing a large electrooptic effect in GaP grown on Si(111).

IV. CONCLUSIONS

In this article we demonstrated that SHG is an effective tool to study strain, disorder, and defect densities at the interface of silicon heterostructures. For two different systems SHG was used to evaluate aspects of interface formation and transformation induced by the growth process and post-growth treatment. A detailed analysis of our SH spectra revealed structural differences between plasma and thermal oxides, which strongly depend on the thermal budget imposed by the oxidation process. On Si/SiO_2 interfaces we have directly observed the trapping of photoexcited carriers by near-interface defect states in the plasma oxide. A strong electric field-induced SH contribution is caused by electron capture in the oxide layer, which is resonantly enhanced at a SH photon energy close to the E_1 transitions in silicon. This effect provides a metallization-free method for characterizing near-interface trap states. For Si/GaP heterostructures we assume that the formation of anti-phase domains on Si(001) substrates significantly reduces the nonlinear optical response of the GaP films. The formation of interface defects and anti phase domains might be reduced by using vicinal Si(001) or Si(111) substrates.

ACKNOWLEDGMENTS

The authors would like to thank N. Dietz (North Carolina State University) for growing the GaP films and E. Stein von Kamienski (RWTH Aachen) for fruitful discussions. The research presented in this article was supported by the Deutsche Forschungsgemeinschaft, the National Science Foundation, the Office of Naval Research, and by a joint collaboration between the states of Nordrhein–Westfalen and North Carolina.

[1] T. Soga, H. Nishikawa, T. Jimbo, and M. Umeno, Jpn. J. Appl. Phys. **32**, 4912 (1993).
[2] S. Sakai, T. Soga, M. Takeyasu, and M. Umeno, Appl. Phys. Lett. **48**, 413 (1986).
[3] H. Shimizu, T. Egawa, T. Soga, T. Jimbo, and M. Umeno, Jpn. J. Appl. Phys. **31**, L1150 (1992).
[4] Y. R. Shen, Nature **337**, 519 (1989).
[5] T. F. Heinz, in *Nonlinear Surface Electromagnetic Phenomena*, edited by H.-E. Ponath and G. Stegeman (Elsevier, Amsterdam, 1991), Chap. 5, pp. 353–416.
[6] G. Lüpke, D. J. Bottomley, and H. M. van Driel, Phys. Rev. B **47**, 10389 (1993).
[7] U. Emmerichs *et al.*, Phys. Rev. B **50**, 5506 (1994).
[8] G. Lüpke, C. Meyer, U. Emmerichs, F. Wolter, and H. Kurz, Phys. Rev. B **50**, 17292 (1994).
[9] W. Daum, H.-J. Krause, U. Reichel, and H. Ibach, Phys. Rev. Lett. **71**, 1234 (1993).
[10] C. Meyer, G. Lüpke, U. Emmerichs, F. Wolter, H. Kurz, C. H. Bjorkman, and G. Lucovsky, Phys. Rev. Lett. **74**, 3001 (1995).
[11] J. R. Power, J. D. O'Mahony, S. Chandola, and J. F. McGilp, Phys. Rev. Lett. **75**, 1138 (1995).
[12] O. A. Aktsipetrov, A. A. Fedyanin, V. N. Golovkina, and T. V. Murzina, Opt. Lett. **19**, 1450 (1994).
[13] J. I. Dadap, X. F. Hu, M. H. Anderson, M. C. Downer, J. K. Lowell, and O. A. Aktsipetrov, Phys. Rev. B **53**, R7607 (1996).
[14] P. Godefroy, W. de Jong, C. W. van Hasselt, M. A. C. Devillers, and T. Rasing, Appl. Phys. Lett. **68**, 1981 (1996).
[15] J. G. Mihaychuk, J. Bloch, Y. Liu, and H. M. van Driel, Opt. Lett. **20**, 2063 (1995).
[16] J. I. Dadap, X. F. Hu, N. M. Russell, J. G. Ekherdt, J. K. Lowell, and M. C. Downer, IEEE J. Selected Topics Quant. Electron. **1**, 1145 (1995).
[17] G. Lüpke, C. Meyer, C. Ohlhoff, H. Kurz, S. Lehmann, and G. Marowsky, Opt. Lett. **20**, 1997 (1995).
[18] C. Ohlhoff, C. Meyer, G. Lüpke, T. Löffler, T. Pfeifer, H. G. Roskos, and H. Kurz, Appl. Phys. Lett. **68**, 1699.
[19] G. Lüpke, D. J. Bottomley, and H. M. van Driel, J. Opt. Soc. Am. B **11**, 35 (1994).
[20] T. Yasuda, Y. Ma, S. Habermehl, and G. Lucovsky, Appl. Phys. Lett. **60**, 434 (1992).
[21] N. Dietz, A. Miller, J. T. Kelliher, D. Venables, and K. J. Bachmann, J. Crystal Growth **132**, 134 (1995).
[22] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic, New York, 1985).
[23] J. T. Fitch, Ph.D. thesis, North Carolina State University (1990).
[24] P. Balk, editor, *The Si–SiO_2 System*, Vol. 32 of *Materials Science Monographs* (Elsevier, New York, 1988).
[25] K. J. Bachmann, N. Dietz, A. E. Miller, D. Venables, and J. T. Kelliher, J. Vac. Sci. Technol. A **13**, 696 (1995).
[26] B. N. Sverdlov, G. A. Martin, H. Morkoç, and D. J. Smith, Appl. Phys. Lett. **67**, 2063 (1995).
[27] D. J. Bottomley, G. Lüpke, J. G. Mihaychuk, and H. M. van Driel, J. Appl. Phys. **74**, 6072 (1993).

Spatial mapping of ordered and disordered domains of GaInP by near-field scanning optical microscopy and scanning capacitance microscopy

J-K. Leong, J. McMurray, and C. C. Williams[a)]
Department of Physics, University of Utah, Salt Lake City, Utah 84112

G. B. Stringfellow
Department of Materials Science and Engineering, University of Utah, Salt Lake City, Utah 84112

(Received 5 February 1996; accepted 13 April 1996)

Imaging of topography, locally induced photoluminescence and Fermi-level pinning in adjacent ordered and disordered domains on a cleaved GaInP sample is performed using a near-field scanning optical microscope and scanning capacitance microscope at room temperature in air. Highly localized photoluminescence spectra obtained by the near-field scanning optical microscope on these domains show spectral peaks at 680 nm (ordered) and 648 nm (disordered) GaInP. The near-field scanning optical microscope and scanning capacitance microscope data confirm previously published data, indicating that the electronic surface structure of ordered GaInP is significantly different from that of disordered GaInP. Both approaches indicate that the Fermi-level at the surface of ordered GaInP is pinned, while the Fermi-level at the surface of disordered GaInP is not pinned. The size, structure, and position of the ordered and disordered domains observed by the near-field scanning optical microscope and scanning capacitance microscope agree with those obtained by cathodoluminescence and Kelvin probe force microscopy. © *1996 American Vacuum Society.*

In the past few years, atomic ordering has been observed for a wide variety of III–V semiconductor alloys.[1] The ordering in GaInP affects the optical and electrical properties of the material and is therefore interesting and important for optoelectronic devices. It is found that under certain growth conditions, $Ga_xIn_{1-x}P$ lattice-matched to GaAs spontaneously orders on a group III sublattice with gallium and indium atoms preferentially occupying alternate {111} planes.[2] The ordering is identified by the appearance of superspots in transmission electron diffraction (TED) patterns. Recently, atomic force microscopy[3,4] (AFM) and Kelvin probe force microscopy[5] (KPFM) have been used to study the surface morphology and surface electronic properties of GaInP. The ordering of GaInP reduces the energy band gap, as observed by cathodoluminescence[6] (CL) and photoluminescence[7] (PL). The measurement of band gap by PL is macroscopic, while the measurement of CL provides approximately 1 μm spatial resolution. Recently, partially ordered GaInP is studied by a low temperature NSOM with a spatial resolution of 270 nm.[8] In this work, we report the results of direct imaging of PL from adjacent ordered and disordered domains in GaInP by NSOM (Ref. 9) at room temperature and correlate these measurements with scanning capacitance microscopy[10,11] (SCM) with a resolution of 150 nm.

Figure 1 shows a schematic of the NSOM used in this work. The NSOM consists of a shear force imaging (topography) system[12] and a near-field optical imaging system.[9] The shear force imaging system is used to control the distance between the tip and sample. The tip is vibrated at its resonance frequency (15–30 kHz) with an amplitude of about 10–20 nm. As the tip approaches the sample surface, the oscillation of the tip is damped because of the tip-sample interaction via shear force. Keeping a constant force results in a constant tip-sample distance. At the heart of the near-field imaging system is the tip, which is a tapered single mode fiber made opaque everywhere except at its very end by deposition of a layer of aluminum (100 nm).[9] The output of an Ar^+ laser (488 nm) is launched into the fiber. This light exits through the 150 nm fiber aperture exciting a small region of the sample and inducing PL from this region. The PL signal is collected by a 0.25 NA objective lens from the side. The pump light (488 nm) is filtered out by a notch filter. The PL signal is then sent to a monochromator and detected by an avalanche photodiode (APD) with a dark count of 8 counts per second.

Figure 2 shows a combination AFM and SCM.[10,11] A metal-coated silicon tip is brought to the surface of the sample. A contact mode AFM is used to position and scan the tip over the surface of the sample. A high sensitivity capacitance sensor (operating at 915 MHz) is connected to the tip to measure the capacitance between the tip and the sample. The tip is grounded by an inductor (at low frequencies) and a 10–100 kHz ac bias voltage is applied to the sample. The tip–sample capacitance change associated with the ac bias voltage is measured by the capacitance sensor. When the tip is scanned across the sample, a topography and a capacitance change image are obtained simultaneously.

NSOM and SCM measurements have been performed on a GaInP sample (AO-98). The ordering in this sample is controlled by substrate misorientation. The GaInP layer is grown by organometallic vapor phase epitaxy (OMVPE) on a grooved GaAs (001) substrate misoriented by 9° toward [$\bar{1}$10]. The period of the groove is 40 μm, and the depth is 4 μm. The grooves run along the [110] direction. The growth

a)Electronic mail: clayton@physics.utah.ed

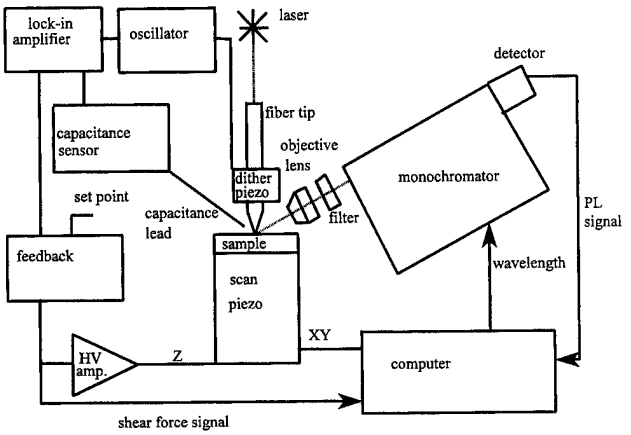


FIG. 1. Schematic diagram of the near-field scanning optical microscope.

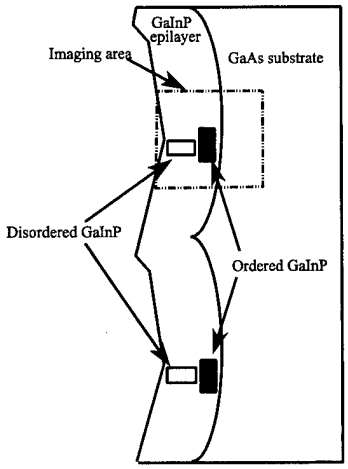


FIG. 3. Schematic drawing of the groove structure of sample AO-98 (in cross section).

temperature is 670°C and growth rate is 2 μm/hr. The thickness of the GaInP epilayer is approximately 5 μm. The sample is cleaved perpendicular to the surface and grooves. CL measurements indicate the existence of a highly ordered domain of several micrometer length and approximately 2 μm thickness at the bottom of the grooves close to the GaAs interface, and a highly disordered GaInP region (3 μm×2 μm) above the ordered region.[13] In Fig. 3, a sketch of the sample cross section is shown together with the ordered and disordered regions which have been identified by CL.[13]

NSOM PL spectra are obtained by locating the fiber probe on the cross sectional surface of sample AO-98. A pump beam (488 nm) is sent through the fiber providing approximately 1 nanowatt at the surface. The collected PL intensity is then measured as a function of wavelength with the spectrometer. The measured PL spectra are shown in Fig. 4. The peak wavelengths are 648 nm (for the disordered region) and 680 nm (for the ordered region) respectively. Both peaks are well separated. By fixing the monochromator bandwidth (10 nm) at 648 nm (to select PL from the disordered regions) or 680 nm (to select PL from ordered regions), a map of the disordered and ordered PL intensity is obtained. The results are shown in Fig. 5(a) and 5(b), respectively. The scan area is 9.5 μm by 12.0 μm and the full width at half maximum (FWHM) of the smallest feature in the NSOM images indicates a spatial resolution of 150 nm. The corresponding topographic images are also shown in Fig. 5(c) and 5(d), respectively. These topographic images show the same area except for a slight shift of 0.4 μm. The maximum height variation in the topographic image is less than 100 nm. The black regions in Fig. 5(a) and 5(b) indicate that no photoluminescence is detected from a large part of the GaInP epilayer or from the GaAs substrate. The small white region in Fig. 5(a) shows the photoluminescence signal from the disordered region, which corresponds to the highly disordered domain shown in Fig. 3. It is close to the sample surface (the growth surface). The size of the domain is approximately 2

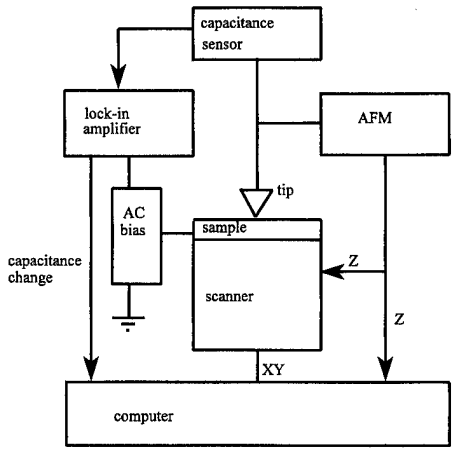


FIG. 2. Block diagram of the scanning capacitance microscope.

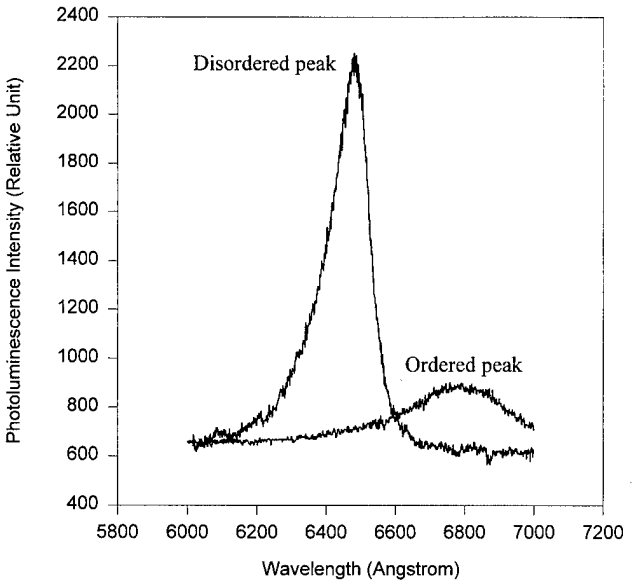


FIG. 4. Near-field photoluminescence spectra of ordered and disordered GaInP.

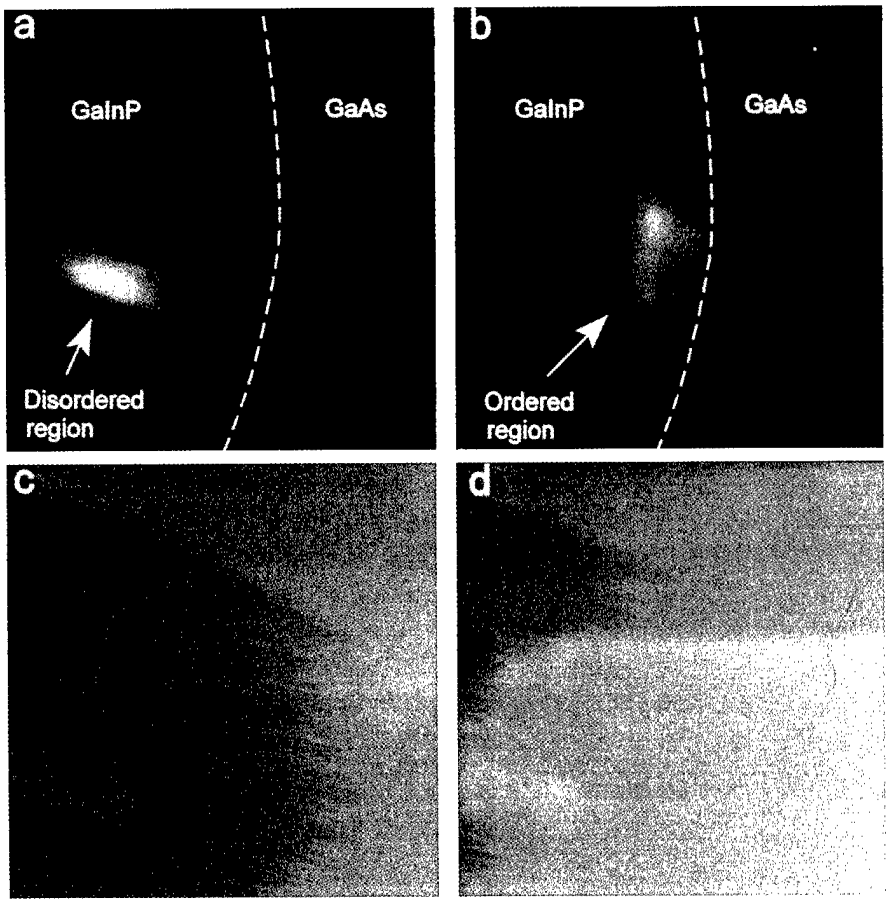


FIG. 5. (a) and (c) Near-field photoluminescence intensity map (λ=648 nm, 10 nm bandwidth) of cross sectioned GaInP/GaAs sample (AO-98) and simultaneously obtained topography (9.5 μm ×12.0 μm). (b) and (d) Near-field photoluminescence intensity map (λ=690 nm, 10 nm bandwidth) of AO-98 and simultaneously obtained topography (9.5 μm ×12.0 μm). All four images have been taken in the same sample area with the same tip.

μm wide and 3 μm long. The small white region in Fig. 5(b) represents the highly ordered domain, which is between the highly disordered domain and the GaAs substrate. The size of this domain is about 2 μm wide and a few micrometers long. The positions of these two regions are near the lateral center of the groove, which can be seen by comparison to the topographical images [Fig. 5(c) and 5(d)]. There also appears to be a halo of ordered material around the disordered region as seen in Fig. 5(b). Seven different places of sample (AO-98) have been measured by the NSOM, and it is found that the size and shape of these domains varies slightly from groove to groove. The NSOM results are in very good agreement with that of CL (Ref. 13) and KPFM (Ref. 5). A large part of the GaInP epilayer shown in Fig. 5(a) and 5(b) emits no photoluminescence. This region (outside the ordered and disordered domains identified in Fig. 3) is not well characterized by other methods and requires further study.

Figure 6 shows two NSOM photoluminescence line cuts, approximately perpendicular to the GaInP/GaAs interface, through the center of both disordered and ordered domains shown Fig. 5(a) and Fig. 5(b). The PL intensity in the ordered domain is smaller than that of the disordered domain, as can also be seen in Fig. 4. Our previous results, using the KPFM,[5] have shown that the surface of ordered GaInP is electrically pinned, indicating a large surface state density. This is consistent with the fact that the PL intensity is reduced in the highly ordered region, since it is expected that a large surface state density should decrease PL yield. Note that the PL intensity drops relatively sharply at the GaAs/GaInP interface in Fig. 6, because the carriers quickly diffuse into the GaAs substrate.

The SCM is also used to image the sample (AO-98) in cross section. A capacitance change image and a topographic

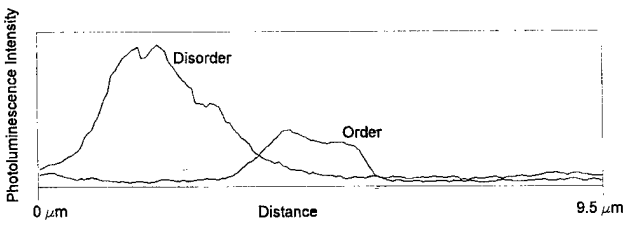


FIG. 6. Line cuts, approximately perpendicular to GaInP/GaAs interface, through the center of both ordered and disordered regions shown in Fig. 5(a) and 5(b).

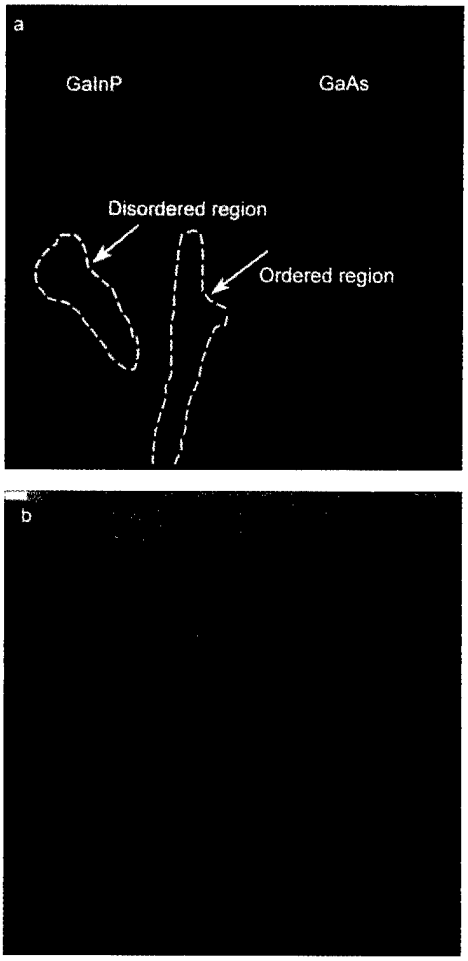


FIG. 7. (a) and (b) simultaneously obtained capacitance change image and topographic image (12 μm $\times$12 μm) of cross sectioned AO-98 sample. This image is not of the same groove as that of Fig. 5.

image obtained simultaneously by the SCM are shown in Fig. 7(a) and 7(b). The image size is 12 μm$\times$12 μm. The applied ac bias voltage between the tip and a sample is 2.5 Volts peak at 89.9 kHz. The capacitance change is measured by a lock-in amplifier at that frequency. In Fig. 7(a), the black color represents a negligibly small capacitance change signal. Because the GaAs surface is electrically pinned, there is a negligible amount of capacitance change with voltage. Just to the left of the GaAs substrate and at the bottom of the GaAs groove, there is an approximately 2 μm wide and several μm long region (black color) which corresponds to the highly ordered GaInP. The capacitance change signal of the ordered domain is the same as that of GaAs, indicating that the ordered domain is also electrically pinned. To the left of the highly ordered region, there is a region approximately a 2 μm wide and 3 μm long (gray color) which corresponds to the highly disordered region identified by NSOM. This region has the largest capacitance change. Compared with the PL images by NSOM, the size, shape, and position of these two domains are very similar. We believe that the SCM and NSOM data strongly confirm our previous observations that surface Fermi-level pinning is directly related to ordering in GaInP.

Since the atomic structure of the ordered and disordered GaInP is different in the bulk, it is not unreasonable to assume that ordering may also modify the electronic structure at the surface [in this case the cleaved (110) plane]. Our observation that the Fermi-level of ordered GaInP domains is pinned at the surface, while highly disordered GaInP domains are unpinned, supports the idea that ordering causes an increase in the density of surface states. We are not aware of the existence of any detailed model of the electronic structure of GaInP (ordered or disordered) at the surface.

The SCM image also shows several other clear variations in the GaInP layer which are not seen by NSOM and KPFM. We believe that the contrast seen by SCM may be directly related to the degree of order in the GaInP epilayer. Further study is required to substantiate this claim.

In conclusion, NSOM and SCM have been used to study GaInP in which ordering was controlled by substrate misorientation. The measurements have been made in cross section and at room temperature. The results clearly show that NSOM and SCM are capable of distinguishing between ordered and disordered GaInP. A previous result (KFPM) and these results (NSOM and SCM) all indicate that the Fermi-level in ordered GaInP is pinned at the surface in contrast to that of disordered GaInP.

Acknowledgment: The authors thank L. C. Su and M. C. DeLong for their general contribution to this work.

[1]G. B. Stringfellow and G. S. Chen, J. Vac. Sci. Technolo. B **9**, 2182 (1991).
[2]A. Gomyo and T. Suzuki, S. Iijima, Phys. Rev. Lett. **60**, 2645 (1988).
[3]D. J. Friedman, J. G. Zhu, A. E. Kibbler, J. M. Ohlson, and J. Moreland, Appl. Phys. Lett. **63**, 1774 (1993).
[4]D. J. Friedman, G. S. Horner, Sarah R. Kurtz, K. A. Bertness, and J. M. Olson, Appl. Phys. Lett. **65**, 878 (1994).
[5]Y. Leng, C. C. Williams, L. C. Su, and G. B. Stringfellow, Appl. Phys. Lett. **66**, 1264 (1995).
[6]L. C. Su, S. T. Pu, G. B. Stringfellow, J. Christen, H. Selber, and D. Bimberg, Appl. Phys. Lett. **62**, 3496 (1993).
[7]A. Gomyo, T. Suzuki, K. Kobayashi, S. Kawata, I. Hino, and T. Yuasa, Appl. Phys. Lett. **50**, 673 (1987).
[8]M. J. Gregor *et al.*, Appl. Phys. Lett. **67**, 3572 (1995).
[9]E. Betzig and J. K. Trautman, Science **257**, 189 (1992).
[10]Y. Huang and C. C. Williams, J. Vac. Sci. Technol. B **12**, 369 (1994).
[11]Y. Huang, C. C. Williams, and J. Silnkman, Appl. Phys. Lett. **66**, 344 (1995).
[12]J-K. Leong and C. C. Williams, Appl. Phys. Lett. **66**, 1432 (1995).
[13]A. Krost, N. Esser, H. Selber, J. Christen, W. Richter, D. Bimberg, L. C. Su, and G. B. Stringfellow, J. Vac. Sci. Technol. B **12**, 2558 (1994).

Papers from the First AVS Topical Conference on Magnetic Surfaces, Interfaces, and Nanostructures

16–20 October 1995
Minneapolis Convention Center
Minneapolis, Minnesota

Sponsored by
Office of Naval Research
American Vacuum Society

Co-Chairmen for the Conference:
S. D. Bader
B. T. Jonker
B. P. Tonner

Published for the American Vacuum Society by
the American Institute of Physics, New York, 1996

PREFACE

The First AVS Topical Conference on Magnetic Surfaces, Interfaces, and Nanostructures was held in conjunction with the 42nd AVS National Symposium in Minneapolis. It provided an excellent forum to address fundamental relationships between structure and magnetic properties. There were 15 invited and 40 contributed talks that spanned three and one half days of sessions. The highlights included presentations on (i) the fabrication of novel, low-dimensional magnetic heterostructures; (ii) state-of-the-art scanning probe, microscopic and holographic characterization of magnetic domains and electronic structure; and (iii) synchrotron radiation-based studies utilizing undulators and elliptically polarized light, including the first experiments to emerge from new beamlines at the Advanced Light Source. Electronic structure theory and computational results for systems with rough as well as ideal interfaces were compared to experiment. A variety of magnetic phase diagrams and transitions were discussed in terms of dynamic and equilibrium concepts. In addition, overviews were presented of the opportunities provided by new magnetoresistive materials ("giant" and "collosal"), and other diverse areas ranging from spin-polarized device structures to quantum well behavior in metals.

Attendance at the MCIN sessions was excellent, with standing room only for several of the invited talks. A straw vote at one of the sessions revealed that a sizable fraction of the audience (between 33% and 50% of the 116 session attendees) was attending the AVS meeting either for the first time or specifically for this topical conference, revealing the strong interest in this area and the need for a continuing presence in future AVS symposia to serve both current and new AVS members. The organizers (Berry Jonker—Naval Research Laboratory; Sam Bader—Argonne National Laboratory; and Brian Tonner—University of Wisconsin, Milwaukee) received many positive comments on the program itself and encouragement to include a focus on magnetism in future AVS programs.

As a result, this topical conference will be held again in 1996 at the 43rd AVS National Symposium in Philadelphia, with joint sessions planned with the Nanoscale Science and Technology, Thin Film, and Surface Science Divisions. The 1996 MSIN chairs will be Berry Jonker, Lennie Klebanoff (Lehigh University), Brian Tonner, and Dieter Weller (IBM Almaden). Finally, special thanks are due to the Office of Naval Research for the generous financial support provided for this topical conference in 1995 and 1996!

Berry Jonker
Co-Chair, Program Committee

X-ray magnetic microscopy and spectroscopy using a third generation synchrotron radiation source

A. T. Young,[a] H. A. Padmore, and N. V. Smith
Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, California 94720

(Received 2 October 1995; accepted 20 May 1996)

Applications of x-ray magnetic circular dichroism (XMCD) to the study of magnetic materials are described. XMCD spectra can be used to quantitatively determine magnetic properties on an element-specific basis. These spectra are also sensitive to the chemical state and environment of the element being probed. The dichroism effect can also be used to produce images of microscopic magnetic structures and domains. Third generation synchrotron light sources are well suited to these experiments. Current and planned facilities at the Advanced Light Source, the first of the new light sources in the U.S., are described, focusing on a new facility with specialized undulators which will directly produce high flux, high brightness beams of circularly polarized x-rays. With new beamlines which have been optimized for either spectroscopy or microscopy, this facility will provide the capability to provide detailed information about magnetic materials. © *1996 American Vacuum Society.*

I. INTRODUCTION

The application of circularly polarized x rays to studies in chemistry and physics has recently become a field of intense interest. Studies of systems as diverse as biological molecules[1] and monolayer/thin film structures[2] have been performed using circularly polarized x rays. In particular, studies of magnetic materials have been of special interest. Starting with the theoretical study by Erskine and Stern[3] and on to the first experimental studies by Schutz *et al.*[4] and Chen and co-workers,[5] x-ray magnetic circular dichroism (XMCD) has been used to study a variety of magnetic phenomena such as the measurement of magnetic moments and the determination of magnetic coupling. XMCD has also been used to produce images. Magnetic microscopy has been used to create element specific images of the magnetic domains and structures of interest to the semiconductor/magnetic recording industry.[6]

The Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory is ideally suited for both spectroscopy and microscopy using circularly polarized x rays. In addition to existing beamlines which have capabilities for circularly polarized x rays, the ALS is designing and building a complement of new insertion devices and beamlines to facilitate this research. Using *insertion devices* which directly produce circularly polarized light makes available high flux, high brightness beams of x rays at *application-specific* beamlines.

This article first reviews a few general properties of XMCD spectra. This review will be brief and will treat the physics only qualitatively. More detailed reviews of the theory of XMCD are available.[7,8] Section III then describes some of the applications of XMCD to magnetic materials research. Again, only a few examples of interest to magnetic measurements are presented. The present capabilities of the ALS with circularly polarized x rays are described. New facilities being built at the ALS are discussed; in particular, a new facility with circularly polarized undulators and optimized beamlines is described. Finally, the end stations will be described. For more detailed information about synchrotron radiation sources and optics, the reader is directed to several recent reviews.[9–11]

II. GENERAL PROPERTIES OF XMCD

XMCD uses the property that some materials, when magnetized, exhibit a different absorption coefficient for left circularly polarized x rays than for right circularly polarized x rays. The effect is dependent on the direction of both the magnetization vector and the polarization vector of the x rays, which for circular polarization is either parallel or antiparallel to the direction of photon propagation. Two spectra are possible: one where the magnetization and the photon polarization vectors are parallel, and the other with the magnetization and the polarization antiparallel. Switching between these two spectra can be done by changing either the magnetization direction or the photon polarization.

Excitation by x rays produces atoms in which one of the inner shell (core level) electrons are promoted to an empty valence shell energy level. X-ray absorption probes the population and state density of both the lower state and the upper state. For $3d$ transition metal elements, the upper levels are the $3d$ energy levels, which are the states responsible for magnetic properties. The strongest absorption to these levels are the $L_{2,3}$ lines (edges), which have the $2p$ levels as their lower state. The transition energies lie between 574 eV for vanadium to 952 eV for copper. Rare earth elements have $4f$ levels as the upper states. The strongest lines for these levels are the $M_{4,5}$ lines, which have the $3d$ levels as their lower states. These transition energies 880 to 1600 eV.

Figure 1 shows typical absorption spectra for the $L_{2,3}$ lines of nickel. In the upper panel, the solid and dashed lines are the spectra obtained with the magnetization vector and the photon helicity parallel and antiparallel. The lower panel

[a]Electronic mail: atyoung@lbl.gov

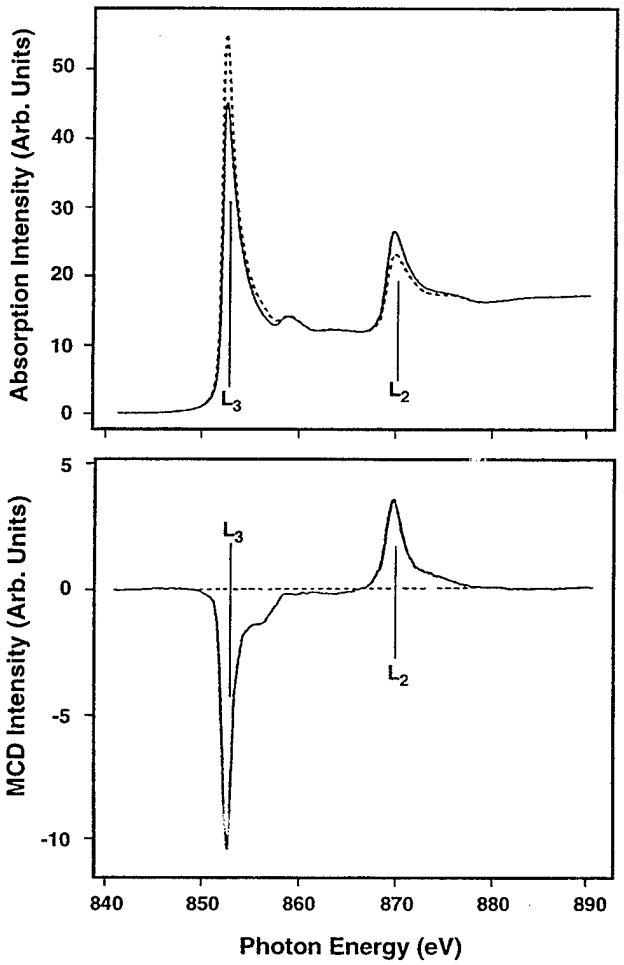


FIG. 1. X-ray absorption and XMCD spectra of nickel at the $L_{2,3}$ core level edges. Upper panel: absorption spectra with photon helicity parallel (solid) and antiparallel (dashed) Lower panel: XMCD spectrum.

shows the resulting XMCD spectra, i.e., the difference between the two absorption spectra. The intensity ratios of both the absorption spectra and the XMCD spectra are informative. Sum rules which relate the integral of the XMCD spectra to the orbital and spin magnetic moments have been developed.[12,13]

To use XMCD to produce images, the energy of the exciting x-ray beam is held fixed at the center of the XMCD peak. With a given photon helicity, the intensity of the response of the material will depend upon the magnetization of the material. If the response from the sample can be spatially resolved, a map of the magnetization of the material can be produced.

There are two primary schemes to image the surface. The first technique uses the x-ray beam to illuminate a small but finite area. An electron microscope is then used to create an image from the photoelectrons that have been liberated. This technique has the advantage of acquiring the entire image at once. Ultimate resolution is determined by the resolution of the electron microscope. The second technique focuses the x-ray beam to as small a focal spot as possible and recording the signal. Spatial resolution is set by the size of the focal spot. The spot is then scanned over the entire sample, either by translating the focusing element or by moving the sample. This technique has the advantage that a variety of detection techniques can be used, including transmission, fluorescence, and photoelectron detection.

III. APPLICATIONS OF XMCD

There have been a large number of experiments using XMCD. These can be divided into two major categories: spectroscopy and microscopy. A variety of systems has been studied with spectroscopy. Magnetic multilayer thin films have been a frequent topic of interest. Monolayer (ML) films of iron on single crystal copper have been studied,[14] with large dichroism effects detected. Nickel ultrathin films grown on Cu(100) have been studied,[15] and it was determined that the easy axis of magnetization varied with the thickness of the layer. It was determined that the axis changed from parallel to the surface for 5–9 ML to perpendicular for 10–75 ML. In another experiment, the exchange coupling between two layers has been studied as a function of the width of an intervening layer.[16] In this experiment, the coupling between a cobalt layer and a nickel-iron layer with an intervening wedge of ruthenium was determined using XMCD. It was found that the direction of magnetization oscillates with the thickness of the ruthenium.

Quantitative measurements of magnetic properties have also been made. For example, element specific magnetization curves have been obtained.[17] Iron/copper/cobalt trilayers were measured using XMCD. Figure 2 shows the hysteresis curves obtained with XMCD and a conventional method. As can be seen, the iron and cobalt hysteresis curves are very different, with different saturation and coercive fields. This is in contrast to measurements made on bulk alloys, in which the hysteresis curves for the two elements are identical. The sum of the two curves yields the curve obtained conventionally. This experiment also illustrates the fact that XMCD measurements are element specific. Another experiment illustrates that the magnetic moment can be decomposed into its orbital and spin contributions. Using the sum rules mentioned above, Chen and co-workers[18] determined that the orbital moment for a cobalt thin film is 0.14 μ_B, and the spin moment is 1.52 μ_B.

Much less work has been done on magnetic microscopy. Stohr *et al.* have imaged the magnetic bit structure of a magnetic recording disc.[6] Figure 3 shows a disc which has domains with alternating magnetization direction. The substrate is a cobalt/platinum/chromium alloy, and the XMCD spectrum was taken at the cobalt absorption edges. The light and dark areas correspond to regions where the magnetization direction is either parallel or antiparallel to the polarization direction. The narrowest domains are 0.5 μm wide. Another example of imaging is the work by Tonner *et al.* who measured iron/terbium/cobalt thin films.[19] In this experiment, the images were obtained at the iron and terbium edges.

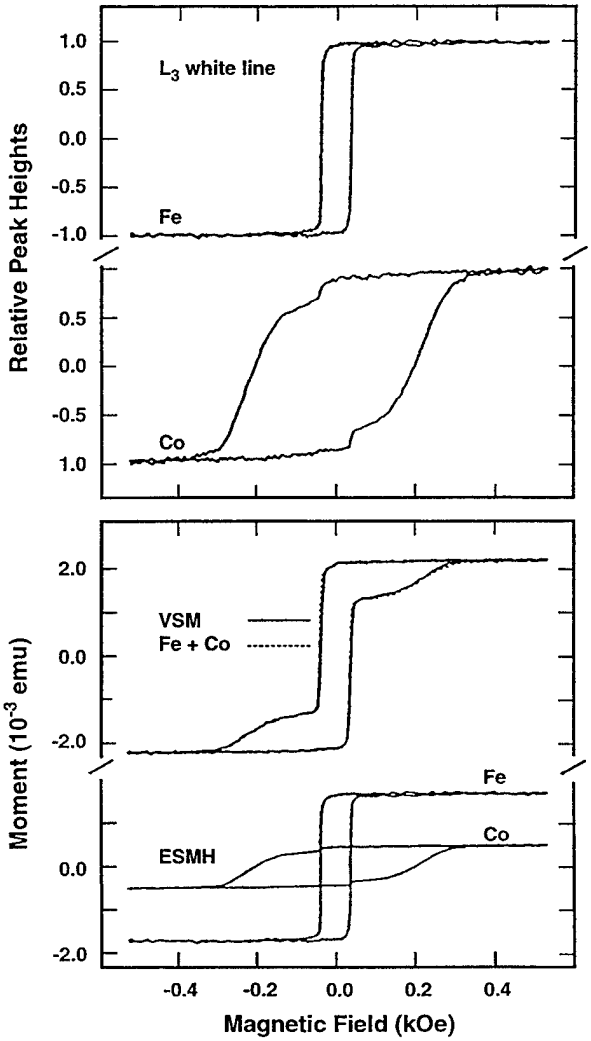


FIG. 2. Element specific magnetometry using XMCD. The top panel shows the Fe and Co L_3 line intensities as a function of applied magnetic field for a Fe/Cu/Co trilayer. The bottom panel shows the comparison between an overall hysteresis curve obtained using conventional methods (solid) and a linear combination fit of the two XMCD hysteresis curves (dotted).

IV. CIRCULAR POLARIZATION CAPABILITIES AT THE ALS

X-ray science has enjoyed a renaissance in recent years with the introduction of synchrotron light sources. These light sources use the fact that charged particles emit light when bent by a magnetic field. The spectrum of light that is emitted is a function of the mass and energy of the particle and the strength of the magnetic field. To produce high energy photons such as x rays, electrons accelerated to the GeV energy range are used in combination with magnetic fields of about 1–5 T. This produces light from the far infrared to x rays beyond 10 keV (a wavelength of 1.2 Å). In addition, by using these high energy electrons, which are relativistic at this energy, the emitted photons are directed in a very narrow, well directed beam, propagating in the same direction as the electron. Pictorially, this is analogous to the searchlight (the x-ray beam) sweeping around an arc.

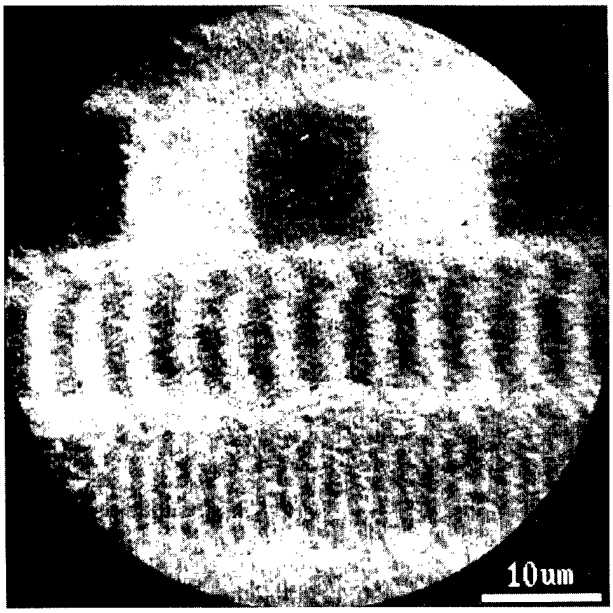


FIG. 3. XMCD images of the magnetic domains on a magnetic hard disc. The light and dark regions correspond to the domains having a magnetization direction either parallel or antiparallel to the x-ray polarization.

Third generation light sources, of which the ALS is the first in the world designed for soft x rays, have optimized the design of the storage ring to produce high quality photon beams. Third generation sources use an electron beam which has a small cross sectional area and small divergence, giving *low emittance*. This makes it possible to focus the emitted photon beam to small sizes. In addition, the design of the electron storage ring is in the shape of polygon with many sides or straight sections, rather than a circle or oval. In the case of the ALS, there are twelve straight sections. At the vertices of the polygon, bend magnets are placed which produce synchrotron radiation as described before. In the straight sections, special periodic arrays of magnets can be placed. These periodic arrays, or insertion devices, produce a much higher photon flux than a single bend magnet.

There are two types of insertion devices, wigglers, and undulators. Wigglers use magnets with a high magnetic field and produce a high flux photon beam with a broad energy spectrum. Undulators use periodic magnet arrays of lower field; however, because of constructive interference between the photons in the undulator, the beam is much narrower than the bend magnet beams, and is quasimonochromatic. Undulators thus produce, within a narrow bandwidth, more flux and have higher spectral brightness (the number of photons per cross sectional area, angular divergence, and photon energy spread per second) than bend magnets. The spectral brightness of typical ALS undulators is four to five orders of magnitude higher than that from bend magnets sources.

Synchrotron radiation is emitted with specific polarization properties. From bend magnets, the light emitted in the plane of the electron orbit is polarized horizontally. As one moves out of the orbit plane, though, the polarization of the beam becomes elliptical, becoming more and more circular the fur-

ther out-of-plane one goes. The flux of photons drops rapidly, however. The variation of flux and polarization with angle is a function of the photon energy. Using apertures, the degree of circular polarization can be selected.

Undulators and wigglers also produce horizontally polarized light. However, conventional wigglers and undulators do not produce circularly polarized radiation. To produce circularly polarized radiation, a variety of special insertion devices have been proposed, although not all designs allow for complete control of the energy or helicity of the radiation. The ALS is designing and building a set of these special undulators which allows for user-controlled energy and polarization.

A. Presently operating facilities

Beamline 9.3.2 is a bend magnet beamline which has been optimized for high resolution spectroscopy.[20] With an energy range from 30–1500 eV, it enables research with the most interesting $3d$ transition metals and $4f$ rare earths, including nickel, iron, cobalt, and neodymium. Equipped with a monochromator which has a resolving power greater than 7000, it is capable of determining shifts in the absorption spectra which occur due to differences in the chemical environment of the element probed.

To use the circularly polarized x rays, an aperture is used to select a portion of the out of plane radiation. At a modest reduction in photon flux, the beam can have a high degree of circular polarization. Typically, a degree of circular polarization of about 0.8 is used, corresponding to collecting radiation from about 0.2 to 0.5 mrad. Measurements of the polarization using a reflection polarimeter have verified the performance of this system.

Multiple end stations can be accommodated at this beamline. These include the advanced photoelectron spectrometer diffractometer (APES) which is capable of a variety of surface science experiments, such as high resolution photoelectron diffraction; the angle-resolved photoemission spectrometer (ARPES) with a rotatable electrostatic electron-energy analyzer; and the applied materials chamber (AMC). Using the AMC, XMCD measurements can rapidly be made on a variety of samples.

In addition to this facility, beamline 6.3.2 at the ALS also has a limited ability to produce circularly polarized x rays. It, too, is a bend magnet beamline and can collect a portion of the out-of-plane radiation.

B. Facilities under construction

1. Beamline 7.3.1 bend magnet endstation optimized for microscopy

Beamline 7.3.1 is also a bend magnet beamline. It is optimized for full field photoelectron microscopy applications.[21] Because the electron beam emittance from third generation synchrotrons is so low, with careful optical design, bend magnet radiation can be focused to produce very small illuminated areas, typically 50 μm diam. Microscopy experiments in general require a high photon flux. Because of this, the optical design of the beamline has been optimized to maximize the photon flux through the beamline to the endstation. The microscopy applications generally do not require high spectral resolution. This leads to a simplified monochromator design which improves the photon throughput. Spectral resolution of at least 1000 will be achieved throughout the energy range of the system. This range, 275–1500 eV, covers the carbon $1s$ absorption edge at the low end, to the M edges of the rare earths at the high end. The performance is optimized at approximately 800 eV, the energy of the L edges of iron, cobalt, and nickel.

Selection of the helicity of the circular polarization will be performed using a chopper/aperture to collect out of plane radiation. To improve the XMCD performance, the chopper can be used to switch the helicity at rates up to 0.5 Hz. This capability will allow for reductions in noise, as well as a modest capability to measure time dependent processes.

The endstation for beamline 7.3.1 will be a photoemission electron microscope (PEEM). Details of this type of microscope have been described elsewhere.[19] With this type of microscope, the photoelectrons ejected from the sample by the x rays are used by an electron microscope to image the surface. Sub-micron spatial resolution has been demonstrated with this instrument, and with planned improvements, resolution of <100 nm structures should be achievable. This experimental facility is currently under construction and will be ready for use by June of 1996.

2. Beamline 4.0 undulator facility for spectroscopy and microscopy

Beamline 4.0 is an undulator-based facility designed specifically for circular polarization applications. The insertion devices will be located in a single insertion device straight section, where two *undulator stations* will be placed end-to-end. Small bending magnets will produce a *chicane* in the straight section, directing the electron beam first through one station, then through the other, producing a 1.65 mrad angle between the two optical axes. Undulators can be positioned and operated at both stations simultaneously. In addition, at each station, a translation mechanism will allow either of two undulators to be placed in the beam. With this arrangement, a total of four different undulators will be available for use at the two beamlines, allowing for great flexibility in the energy range or helicity switching. The output from either station can be directed to either of the two branchlines, or the output from both stations can be directed to the same branchline.

To directly produce circularly polarized radiation, an undulator design similar to that of Sasaki[22] and Carr[23] has been adopted. The magnetic design of this insertion device has been modified to produce higher magnetic fields on the axis of the undulator, leading to an expanded photon energy range. The device is capable of producing polarized light of any ellipticity: horizontal, vertical, circular, or elliptical. Users will be able to select the mode used. In addition they will be able to change the helicity of the circularly or elliptically polarized light every few seconds.

TABLE I. Energy range of a 5 cm period elliptically polarizing undulator.

Mode	Peak Bx (T)	Peak By (T)	Min. energy (eV)	Max. energy (eV)
Elliptical	0.228	0.502	100	1500
Helical	0.424	0.424	140	680
Horizontal	0	0.793	85	1500
Vertical	0.502	0	180	1500

The figure of merit considered for the evaluation of performance is defined as

$$M = P_C^2 * F,$$

where P_C is the degree of circular polarization, and F is the flux. The merit function brightness is also considered (where brightness is substituted for flux in the above equation). Flux, brightness, and degree of circular polarization have been calculated using the formalizm described by Kim[24] for the planar and helical cases and generalized by Marks[25] for the elliptical case.

The first undulator to be installed in the ALS circular polarization facility will be a 5 cm period, 1.95 m long device with 37 full strength periods. The energy ranges of the various modes of polarization are summarized in Table I. These correspond to an electron energy of 1.9 GeV. Figure 4 shows the calculated performance for the merit function flux and brightness for this device for pure helical mode and for the first, third, and fifth harmonics when the undulator is operated in an elliptical mode. As can be seen, operation in pure helical mode gives slightly higher merit function performance, but with a very restricted energy range, as pure helical motion produces no higher harmonics. In elliptical mode, this undulator produces usable output from 100 to 1800 eV. This energy range covers the important core levels for magnetic materials—the $L_{2,3}$ edges of $3d$ transition metals and the $M_{4,5}$ edges of rare earths. Performance has been optimized for coverage of the iron, nickel, and cobalt absorption edges, near 800 eV. These calculations do not include

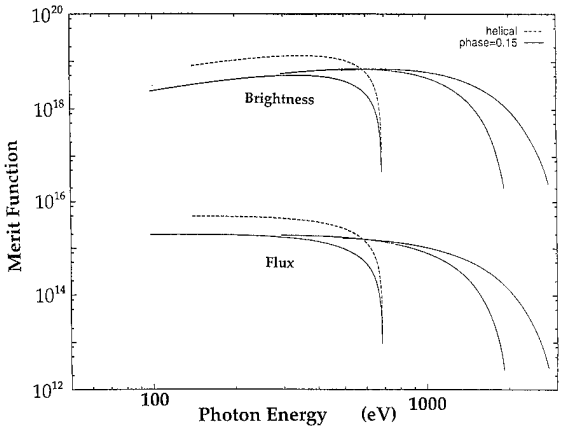


FIG. 4. Calculated performance curves of a 5.0 cm period elliptical polarized undulator. The flux (brightness) merit function is plotted, which is the flux (brightness) x degree-of-polarization (Ref. 2). Electron beam energy is 1.9 GeV.

electron beam energy spread, which will decrease the performance of the higher harmonics. For this reason, in determining the high energy performance, only the first and third harmonics have been considered. To reach even lower energies, a 7.5 cm period device has also been designed. With this device, energies as low as 20 eV will be available.

Emerging from the undulators, the two photon beams are separated by 1.65 mrad. After passing through the shield wall, the beams will enter a *mirror switchyard.* Using this set of mirrors, the output from either (or both) undulator station(s) can be directed to either of the branchlines. In the *standard* mode, each undulator beam would be directed to one of the beamlines by a translating mirror. To send both beams to the microscopy branchline, the mirror for the spectroscopy line would be retracted and the beam allowed to propagate to a second mirror which would then direct the beam down the microscopy line. A similar procedure is used to direct both beams to the spectroscopy line. Sending the output from both undulator stations down the same beamline will allow for experiments which require rapid changes in polarization. For these experiments, undulators in the two stations will be set to the same photon energy and equal but opposite helicity. Although both beams propagate down the same beamline, in some places they are physically separated. At one of these points, a mechanical chopper is placed which will alternately block one polarization or the other.

The conceptual design of the circular polarization beamlines is shown in Fig. 5. Two main beamlines will be built for this facility. The optical design of each branchline is tailored to suit one of the two major research areas, microscopy or high resolution spectroscopy. For spectroscopy studies, a wide energy range will be covered from 20 to 1800 eV, with a resolution of up to 10 000 at 100 eV. For the microscopy beamline, a higher-throughput, lower resolution monochromator has been designed. It is entrance-slitless, with an energy range of 100–1800 eV and a resolution of about 1500.

The microscopy beamline will have provisions for two end stations, each with its own monochromator exit aperture. One end station will be a PEEM, a similar but improved version of the instrument used on beamline 7.3.1. The current generation of PEEMs suffer from resolution degradation caused by chromatic aberration of the lens, caused by the spread in the kinetic energy of the photoelectrons. The new PEEM will focus all electrons to the image plane, independent of their kinetic energy. Ultimate resolution of an advanced PEEM should be better than 20 nm.

The other microscope end station will be equipped for scanning Fresnel zone plate microscopy. Zone plates are focusing elements which rely on diffraction rather than refraction, like lenses, or reflection, like mirrors. They consist of a series of concentric rings with a decreasing spacing as the radius increases. The radii of the rings is chosen so that the light passing through all the rings is diffracted to a single point. Zone plates are excellent optical elements for x-ray applications as, because of absorption, there are no materials which can be used as lenses. Mirrors for x-ray applications also suffer from absorption as well as the requirement that

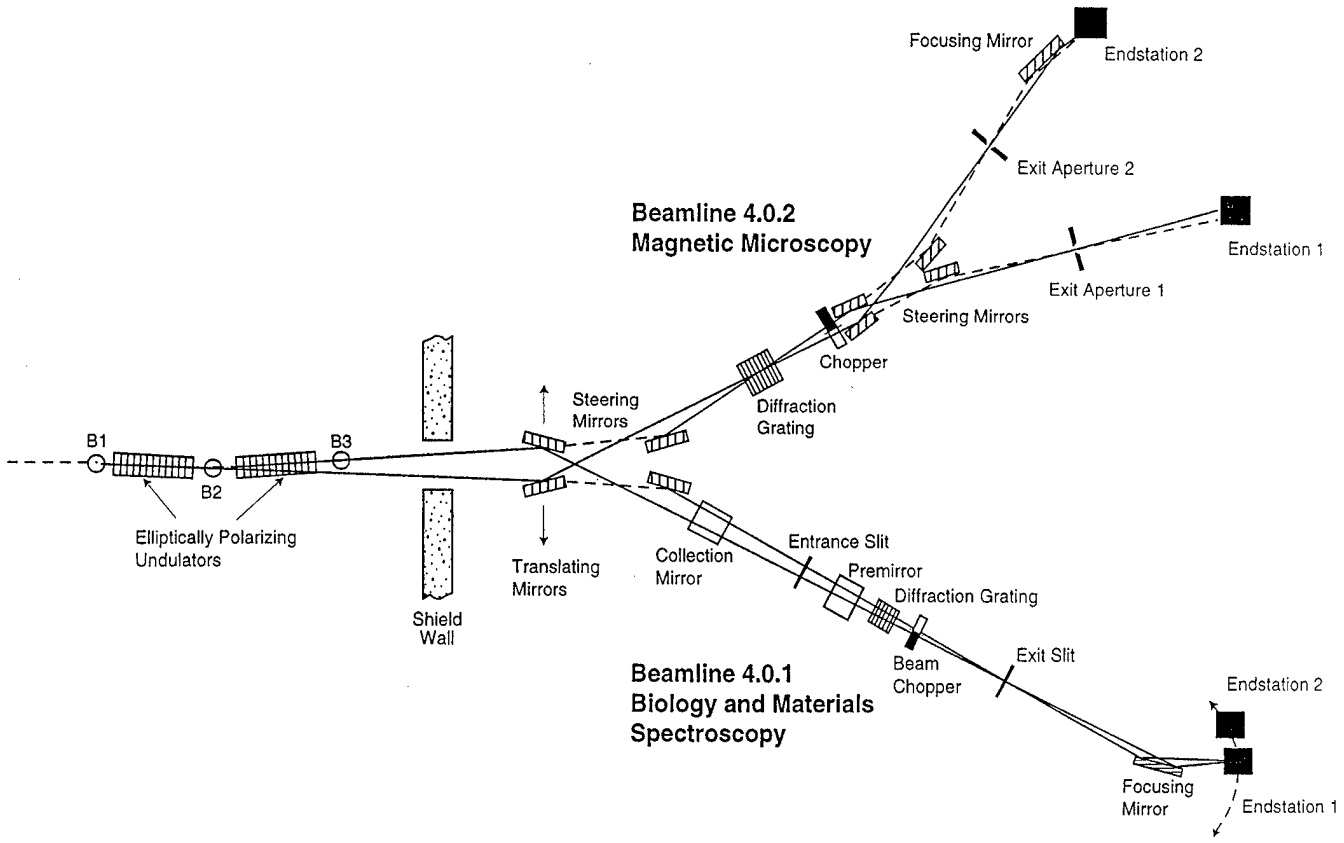


FIG. 5. Conceptual design of the ASL beamline 4.0, a new facility with elliptically polarizing undulators and beamlines optimized for spectroscopy and microscopy.

the mirrors need to be used at grazing incidence. A typical zone plate in use at the ALS has 100 zones (rings), a diameter of 100 μm and an outer zone width of 30 nm.

The resolution of a zone plate microscope is derived from the size of the focused spot. This, in turn, is approximately equal to the width of the outer zone. This is currently limited by microfabrication technology to about 30 nm.

Detection in a zone plate microscope can be via a variety of techniques, depending upon the sample. Using x-ray detectors, transmission through the sample can be measured, as can fluorescence excited by the synchrotron beam. Alternatively, photoelectrons can be collected. Images are collected by scanning the x-ray spot across the sample, either by translating the zone plate or by moving the sample.

The spectroscopy beamline will also have provisions for two end stations. For this beamline, a rotating experimental platform will allow two end stations to be attached and aligned to the beamline at a time. Many of the types of experiments to be run on this beamline are similar to the experiments performed on beamline 9.3.2, thus, many of the endstations will be similar. However, the much greater photon flux available from the undulators will enable experiments which are not currently possible. For example, experiments with samples which are very dilute in the species of interest or which have a finite lifetime will benefit from the increased photon flux. Typical end stations which will be used on this beamline will include UHV angle-resolved photoelectron spectroscopy chambers and general purpose chambers equipped for surface science and materials research. For XMCD measurements, chambers with magnets which can alternate their polarity will be used. Systems with permanent magnets or electromagnets can be used. In particular, superconducting magnet chambers for measurements at fields above 1 T will be available.

V. SUMMARY

XMCD has shown, in a short time, to be a valuable technique in characterizing magnetic materials. Third generation synchrotron light sources like the ALS can provide high flux, high brightness beams of x rays for this research. At the ALS, in addition to presently available bend magnet beamlines, two new facilities are being constructed. A new microscopy beamline is being built, and a facility with circularly polarizing undulators and beamlines optimized for high resolution spectroscopy and microscopy are being constructed. These new facilities will provide a new resource for probing magnetic materials.

ACKNOWLEDGMENT

This work is supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. DOE, under Contract No. DE-AC03-76SF00098.

[1]J. Christiansen, G. Peng, A. T. Young, L. B. LaCroix, E. I. Solomon, and S. P. Cramer, Inorg. Chim. Acta (submitted).
[2]J. G. Tobin and G. D. Waddill, J. Appl. Phys. **75**, 6369 (1994).
[3]J. L. Erskine and E. A. Stern, Phys. Rev. B **12**, 5016 (1975).
[4]G. Schutz, W. Wagner, W. Wilhelm, P. Kienle, Zeller, R. Grahm, and G. Materlik, Phys. Rev. Lett. **58**, 737 (1987).
[5]C. T. Chen, F. Sette, Y. Ma, and S. Modesti, Phys. Rev. B **42**, 7262 (1990).
[6]J. Stohr, Y. Wu, B. D. Hermsmeir, M. G. Samant, G. R. Harp, S. Koranda, D. Dunham, and B. P. Tonner, Science **259**, 658 (1993).
[7]D. B. McWhan and J. Synchrotron Radiat. 1 **83** (1994).
[8]J. Stohr and Y. Wu, in *New Directions in Research with Third Generation Soft X-ray Synchrotron Sources*, edited by A. S. Schlachter and F. J. Wuilleumier (1994), p. 221.
[9]Synchrotron Radiation Sources, edited by H. Winick (World Scientific, Singapore, 1994).
[10]*Handbook On Synchrotron Radiation*, edited by Ernst-Eckhard Koc (Elsevier, Co., New York, 1983–1991).
[11]*New Directions In Research With Third-Generation Soft X-Ray Synchrotron Radiation Sources*, edited by A. S. Schlachter and F. J. Wuilleumier (Kluwer Academic, Dordrecht, 1994).
[12]B. T. Thole, P. Carra, F. Sette, and G. van der Laan, Phys. Rev. Lett. **68**, 1243 (1992).
[13]P. Carra, B. T. Thole, M. Altarelli, and X. Wang, Phys. Rev. Lett. **70**, 694 (1993).
[14]J. G. Tobin and G. D. Waddill Phys. Rev. Lett. **68**, 3642 (1992).
[15]W. L. O'Brien and B. P. Tonner, Phys. Rev. B **49**, 15370 (1994).
[16]Y. Wu *et al.* Appl. Phys. Lett. **63**, 263 (1993).
[17]V. Chakarian, Y. U. Idzerda, G. Meigs, and E. W. Chaban, Appl. Phys. Lett. **66**, 3368 (1995).
[18]C. T. Chen, Y. U. Idzerda, H. J. Lin, and N. V. Smith, Phys. Rev. Lett. **75**, 152 (1995).
[19]B. P. Tonner and D. Dunham, Nucl. Instrum. Methods A **347**, 436 (1994).
[20]Z. Hussain, W. R. A. Huff, S. A. Kellar, E. J. Moler, P. A. Heimann, W. McKinney, C. Cummings, T. Lauritzen, J. P. McKean, F. J. Palomares, H. Wu, Y. Zheng, A. T. Young, H. A. Padmore, C. S. Fadely, and D. A. Shirley, Rev. Sci. Instrum. (in press).
[21]T. Warwick, H. A. Padmore, and T. Renner, LSBL-257, Lawrence Berkeley National Laboratory Internal Report.
[22]S. Sazaki, Nucl. Instrum. Methods A **347**, 83 (1994).
[23]R. Carr and S. Lydia, Proc. SPIE **2013**, 56 (1993).
[24]K.-J. Kim, Opt. Eng. **34**, 342 (1995).
[25]S. Marks, *Flux and Brightness Calculations for an Elliptically Polarizing Undulator*, LBNL ALS Note #LSBL-282, September 1995.

Imaging of ferromagnetic domains using photoelectrons: Photoelectron emission microscopy of neodymium-iron-boron ($Nd_2Fe_{14}B$)

M. Mundschau[a] and J. Romanowicz[b]
Center for Materials Science, Bowling Green State University, Bowling Green, Ohio 43403-0213

J. Y. Wang, D. L. Sun, and H. C. Chen
Institute of Crystal Materials, Shandong University, Jinan 250100, People's Republic of China

(Received 2 October 1995; accepted 22 April 1996)

Ferromagnetic domains of a single crystal of neodymium-iron-boron, $Nd_2Fe_{14}B$ (one of the strongest permanent magnetic materials known) are imaged by focusing a beam of photoelectrons with electrostatic optics in a photoelectron emission microscope. Photoelectrons emitted from the surface are deflected laterally into two opposite directions by stray magnetic fields that exist above the domains. The photoelectron beam is partially split into two. Magnetic contrast is produced by blocking part of the beam and imaging with an edge of the beam. The magnetic contrast mechanism appears to be similar to the type I magnetic contrast mechanism known from scanning electron microscopy, in which stray magnetic fields above the ferromagnetic domains deflect secondary electrons either towards or away from the electron detector. Upon heating the sample above the Curie temperature, the ferromagnetic domains gradually disappear, as expected for a second order phase transition. They reappear upon cooling. © *1996 American Vacuum Society.*

I. INTRODUCTION: PHOTOELECTRON EMISSION MICROSCOPY

Photoelectrons ejected from a solid sample into vacuum become free electrons. As free electrons, they are deflected by electric and magnetic fields, and a beam of photoelectrons can be focused by electrostatic and electromagnetic lenses to produce an electron optical image of a sample. Brüche[1] published photoelectron images in 1933—just one year after the first transmission electron microscopy images of Knoll and Ruska[2] and the first thermionic electron emission images of Brüche and Johannson.[3] Already in the 1930s, studies on the absorption of gases on Pt (Refs. 4 and 5) and Ni (Ref. 6) and of K and Ba on various metals, suggested that a single monolayer of adsorbate could radically alter the photoemissive properties of a surface, and that a single adsorbed monolayer could be imaged by photoelectron microscopy. However, because of poor vacuum, it is assumed that surfaces were quickly contaminated by residual gases.

A breakthrough occurred with the introduction of bakable, ultrahigh vacuum instruments.[7–10] These allowed controlled studies of monolayer films and the combination of photoelectron microscopy with other standard UHV techniques of surface science. If a substrate is chosen with a work function greater than the maximum photon energy of the light source, no photoelectrons are emitted from the substrate. If, onto such a substrate, a substance is deposited that lowers the work function below the maximum energy of the lamp, photoelectrons are emitted from areas covered with adsorbate. If the surface energy of the adsorbate is less than that of the substrate, the adsorbate can wet the surface and spread as a monolayer, and single monolayers can be imaged by photoelectron microscopy. For example, the epitaxial growth of a single monolayer of copper evaporated onto a (110) face of a single crystal of molybdenum was followed from the initial stages of nucleation at atomic steps to the completion of the first monolayer.[11–13] Three-dimensional epitaxial island growth has been seen as well.[14] An important application of photoelectron microscopy in surface science has been the observation of chemical reaction-diffusion fronts in monolayer films on single crystal surfaces during catalytic reactions.[15–17] Perhaps the most spectacular application has been the observation of chemical kinetic oscillations, chemical waves, and chemical chaos in monolayer films.[18–27] Numerous reviews on photoelectron emission microscopy have been published.[7,19,28–32]

In 1957 Spivak, Dombrovskaia, and Sedov[33] published photoelectron images of ferromagnetic domains of the (0001) basal plane of cobalt. The patterns of these domains are quite distinct and unmistakable and were directly compared to images produced by the magnetic colloid decoration technique of Francis Bitter.[34] Schur and Pfefferkorn[35] and Pfefferkorn *et al.*[36] used photoelectron microscopy to image ferromagnetic domains on prismatic planes of cobalt. They concluded that contrast was produced when part of the photoelectron beam was blocked by an edge, analogous to Töpler's schlieren technique, in which part of a beam of light is blocked by a knife edge to produce contrast. Magnetized regions of a magnetic recording disk have also been imaged using circularly polarized x rays from a synchrotron source.[37–39]

In this article we report the observation of ferromagnetic domains of a single crystal of neodymium-iron-boron ($Nd_2Fe_{14}B$) using photoelectron microscopy. Neodymium-iron-boron is one of the strongest permanent magnets.[40–45] Moreover, despite the use of the rare-earth element neodymium, it is cost effective and produces a higher magnetic field per unit cost compared to other materials. It has revo-

[a]Electronic mail: mmundsc@bgnet.bgsu.edu
[b]Electronic mail: jromano@bgnet.bgsu.edu

lutionized the permanent magnet industry.[46] Neodymium-iron-boron forms tetragonal crystals with 68 atoms in the unit cell.[44] It is a uniaxial magnetic material. The direction of easy magnetization is along the c axis.[45] The Curie temperature is reported as 585 K.[44] Magnets used in practice have small crystal grain sizes that are near the single domain size for this material.[40,42] It is the small single domain size crystallites that give the material much of its high coercivity.[40]

II. EXPERIMENT

A single crystal boule, 10 mm in diameter, was grown at the Institute of Crystal Materials at Shandong University, P.R. China, using the Czochralski method from the melt in a Hukin cold crucible.[47] X-ray diffraction studies[47] verified that the crystal had the well-known tetragonal structure of the magnetic $Nd_2Fe_{14}B$ phase.[44] A Laue pattern implied that much of the sample was single crystal,[47] although microscopy has revealed subgrain boundaries and inclusions. Energy dispersive x-ray fluorescence analysis performed in a scanning electron microscope (SEM) showed Si as a minor impurity.[47] The boule was cut and observations of ferromagnetic domains were made on {100} (a nonbasal) plane parallel to the c axis and perpendicular to an a axis of the tetragonal crystal. The accuracy of the orientation was assumed to be between 0.3 and 0.5° and the presence of some miscut was inferred from the fact that SEM observations of this crystal revealed type I magnetic contrast that would require slight misorientation and stray magnetic fields above domains on a nonbasal plane. To minimize oxidation, the cut crystal was shipped to the United States under a protective layer of grease. The sample was degreased in boiling toluene and in acetone. A final rinse in reagent grade absolute alcohol was made before the sample was placed in the vacuum chamber of the photoelectron microscope, which was baked overnight to a pressure of 2×10^{-9} Torr (3×10^{-7} Pa). No further attempts were made to clean the sample or to keep the surface free from residual gases. It was assumed that the surface contained carbon and other impurities.

Electrostatic electron optics, which were purchased from Staib GmbH, and assembled by Senftinger, were used in the photoelectron microscope. Use of electrostatic lenses, rather than electromagnetic lenses is a great advantage in the study of magnetic materials, because electrostatic lenses produce no magnetic field that might alter the magnetization of the sample. The Staib design was developed by Engel and Kordesch[10] at the Fritz Haber Institute in the groups of Bradshaw and Zeitler. It is the same lens design used by the groups of Ertl and Block to study reaction-diffusion phenomenon.[19–27] The system uses standard UHV components, and a Galileo-type channel plate. Detailed drawings of the electrostatic lenses are given in the work by von Oertzen.[48]

Because magnetic fields from $Nd_2Fe_{14}B$ magnetized to near saturation deflect the emitted electrons to too great an extent to form useful images, the sample was demagnetized by heating above the Curie temperature, 585 K. Radiation from a tungsten filament located below the sample was used as the heater. The sample was cooled in the earth's magnetic field to room temperature. As with all ferromagnetic materials, the ferromagnetic domains spontaneously reform upon cooling below the Curie temperature and partially cancel the external field. Ultraviolet light was used from a common Hg lamp (OSRAM HBO 100 W/2), which produces a maximum useful photon energy of ~5 eV (253 nm). It was partially focused with a parabolic mirror (in a PTI model A1010 lampholder) to a ~3 mm diam spot size. The light struck the surface 75° from normal. No attempt was made to polarize the light. Careful adjustment of the sample tilt relative to the electron optical axis was critical for the observation of the domains. To produce contrast it was necessary to tilt the sample slightly off axis. Deflection of the photoelectron beam with a hand-held magnet outside of the vacuum chamber showed that contrast did not occur if the center of the beam was used and that contrast was achieved only by using an outer edge of the beam. Observations by scanning electron microscopy showed that the features were not due to a topographical effect. Ferromagnetic domains were observed in the SEM by type I magnetic contrast and proved that the features seen by a photoelectron emission microscope (PEEM) were ferromagnetic domains. Furthermore, heating of the sample above the Curie temperature in the SEM led to the disappearance of the domains that reappeared upon cooling (unpublished). Images in the PEEM were photographed from the circular channel plate screen with a Canon Eos Elan 35 mm camera with exposure times of under 20 s using ASA 100 film.

III. RESULTS

Figure 1 shows photoelectron images of four regions of the sample. Although nominally a single crystal, the crystal also contained subgrain boundaries and inclusions. On regions of the crystal that were far from defects, alternating dark and bright parallel domains were seen as in Fig. 1(a). Most of the crystal contained these approximately parallel domains. Near subgrain boundaries and defects more complex domain structures were seen as in Figs. 1(b)–1(d). Images of domains were diffuse. As the sample was heated in the PEEM to above the Curie temperature the domains gradually disappeared. They gradually reappeared upon cooling. The heating and cooling cycle could be repeated indefinitely. By deflecting the photoelectron beam with a hand-held magnet, it was clear that no domains were seen if the center of the beam was used and that contrast was achieved only if one of the opposite edges of the beam was used. This is represented schematically in Fig. 2.

IV. DISCUSSION

Ferromagnetic domains of $Nd_2Fe_{14}B$ of various orientations have been imaged by other types of microscopy, especially magneto-optical Kerr microscopy.[49–56] Comparison of the images of Fig. 1 to the images taken by the Kerr microscopy technique, especially those of Luo and Zhang[54] show that the dark and bright stripes are quite similar to known ferromagnetic domain structures of $Nd_2Fe_{14}B$. They are also

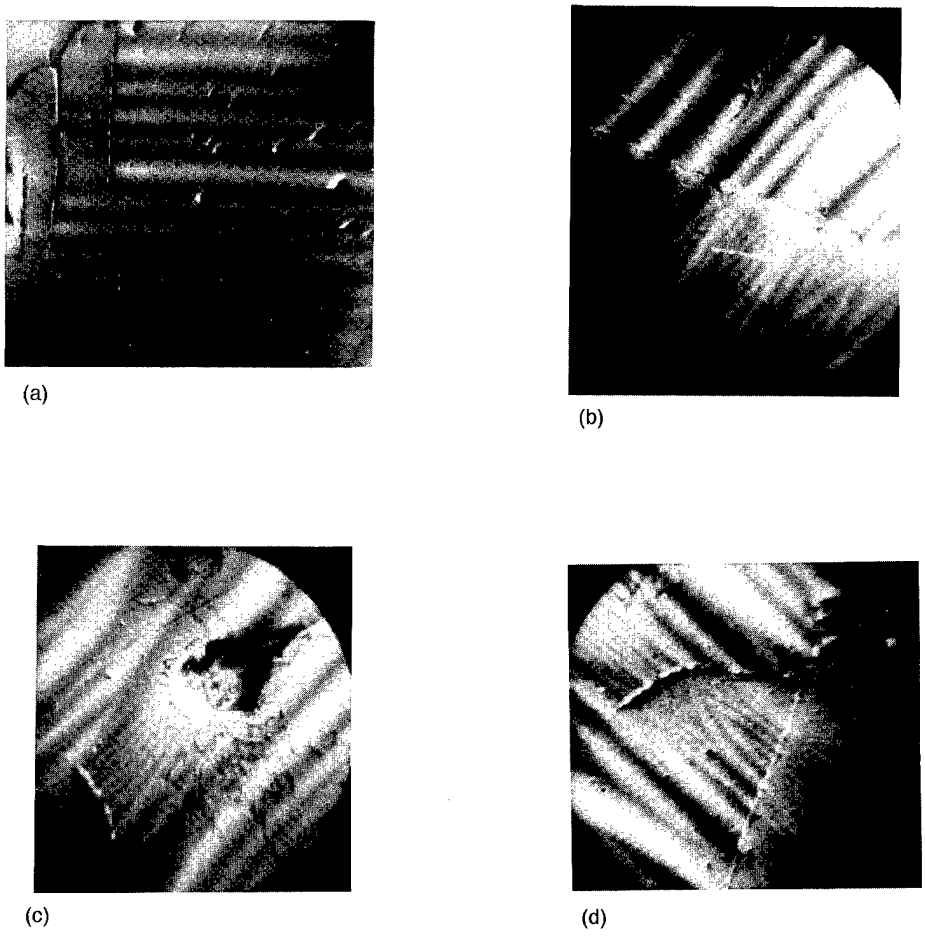


FIG. 1. Photoelectron images of ferromagnetic domains of $Nd_2Fe_{14}B$ (Hg lamp light source). Magnetic contrast is caused by lateral deflection of photoelectrons out of the aperture by stray magnetic fields that exist above the surface. (a) Parallel domain structure seen on much of the crystal (diameter 500 μm). (b)–(d) Complex domain structures typical near subgrain boundaries, defects, and inclusions [diameters (b) and (c) 135 μm; (d) 140 μm].

similar to ferromagnetic domain structures seen near grain boundaries on prismatic planes of cobalt[57] and other ferromagnetic materials.[58–60] The gradual disappearance of the domains (unpublished) as the sample is heated above the Curie temperature and their reappearance upon cooling is similar to that seen by Luo and Zhang[54] using Kerr microscopy and is as expected for a second order phase transition of ferromagnetic materials.

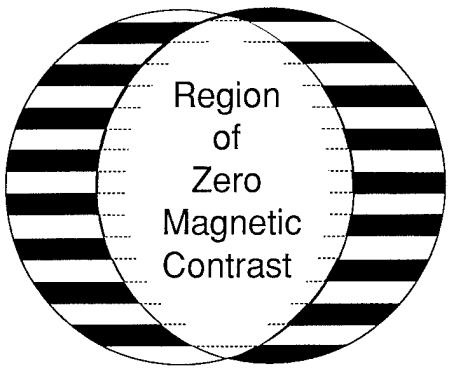


FIG. 2. Schematic representing the partial splitting of the emitted photoelectron beam. Maximum magnetic contrast is produced by using an edge of the beam. Little or no magnetic contrast is noted if the center of the beam is used to form the image.

Ferromagnetic domains have long been imaged by conventional scanning electron microscopy.[61–66] Two contrast mechanisms are recognized.[66] In type I contrast, secondary electrons are deflected by stray magnetic fields above the sample.[63] Stray magnetic fields can be present if the sample is miscut. This is shown very schematically in Figs. 3(a) and 3(b). In the type II mechanism,[66] backscattered electrons are deflected by the magnetic field within the specimen as shown schematically in Fig. 3(c). In this case no stray field is necessary. These mechanisms are well established and are discussed in textbooks on SEM.[66] We have examined the $Nd_2Fe_{14}B$ sample in the SEM and have found that the domains are imaged by the type I mechanism.[67] The direction of deflection of the secondary electrons toward the detector is as in Fig. 4. The images of domains are diffuse in both SEM and PEEM, which is consistent with deflection of both secondary electrons and photoelectrons by a stray magnetic field.[63] Finally, we have heated the sample in the SEM above the Curie temperature and observed the disappearance of the domains. This leaves little doubt that the structures shown in

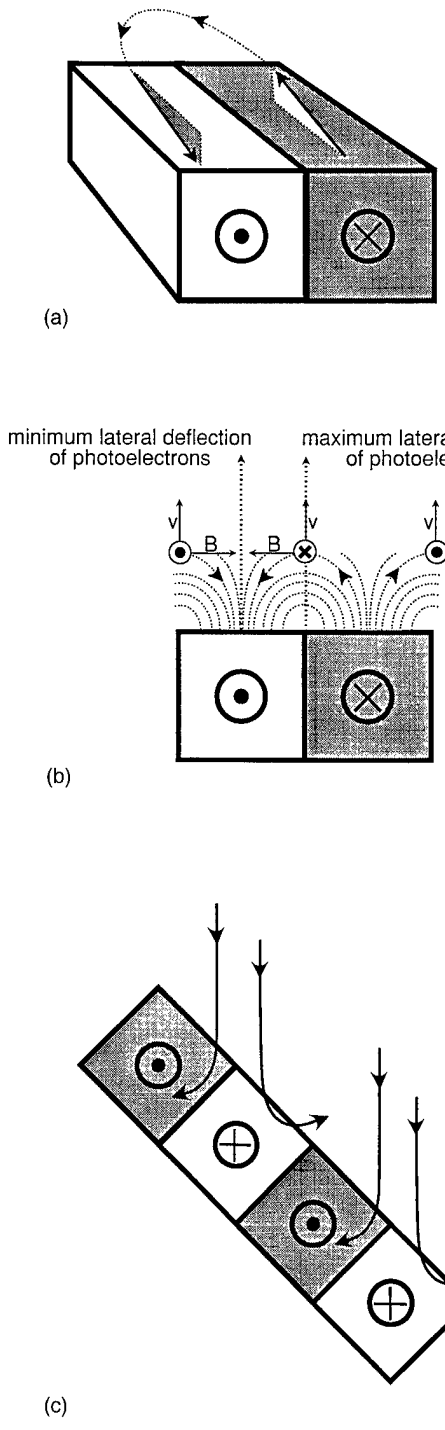


FIG. 3. Schematic representing type I and type II magnetic contrast known from scanning electron microscopy. The same principles can be applied to explain deflection of photoelectrons. (a) Schematic of assumed stray fields. If a crystal is not cut exactly on axis, the magnetization vector is not in the plane of the surface and stray fields will exist above the sample. The straight arrows point along the directions of easy magnetization, which is the c axis of this tetragonal crystal. (b) Projected side view of the assumed stray fields. Maximum lateral deflection of photoelectrons occurs above domain boundaries. Opposite directions of stray fields deflect photoelectrons in two opposite directions, causing partial splitting of the photoelectron beam. Minimum lateral deflection will occur for photoelectrons ejected from the center of domains. Variation of the lateral Lorentz force above the domains is responsible for the diffuse image of the domains in Fig. 1. (c) In type II contrast magnetic fields within the sample cause lateral deflection of backscattered electrons in the SEM. No stray field is necessary.

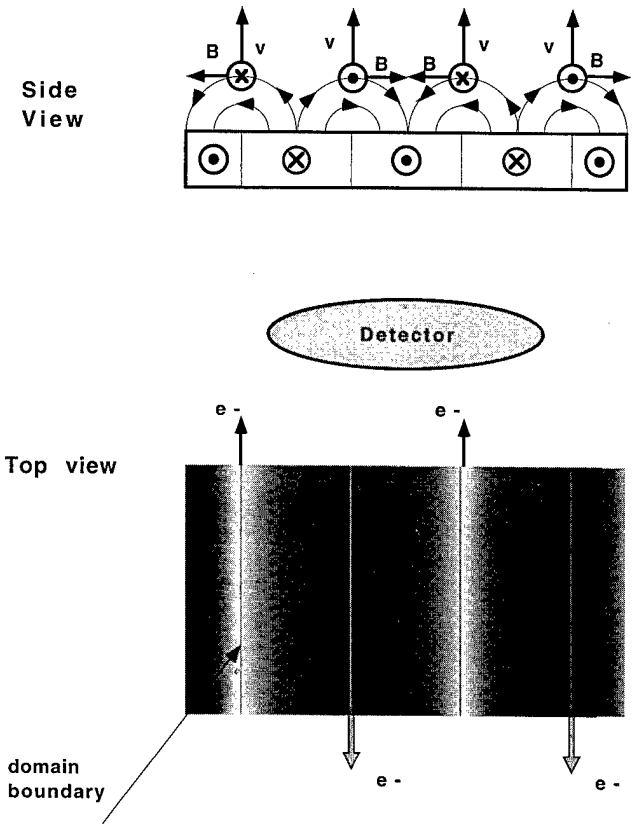


FIG. 4. Stray fields giving rise to diffuse domain structure and schematic of an image. In type I contrast, lateral deflection of emitted electrons is attributed to stray magnetic fields above the sample, which explains the diffuse images of the domains and explains the relative position of the secondary electron detector. This mechanism occurs for uniaxial magnetic materials which have stray fields above the sample that close the magnetic flux. Note that the position of the detector in the SEM requires a stray field to explain the direction of deflection.

Fig. 1 are ferromagnetic domains and that they can be imaged by photoelectron microscopy.

The magnetic contrast mechanism in photoelectron microscopy appears to be straight forward. Upon ejection from the sample, the photoelectrons become free electrons. The photoelectrons are deflected laterally by the Lorentz force as the electrons travel through the stray magnetic field above the sample. Opposite directions of the stray field above the opposite domain types deflects the photoelectrons in two opposite directions, partially splitting the photoelectron beam as shown in Fig. 2. Imaging with an edge of the beam reveals the domains. Because the photoelectrons must encounter the same magnetic fields as the secondary electrons, the deflection must be similar and the well-established mechanisms applied to the SEM (Ref. 66) can be applied to the photoelectron images. If a stray field configuration due to miscut of the crystal is assumed as shown in Fig. 3(a) (as adapted from computer simulations of Harasko, Pfützner, and Futschik[68]), maximum lateral deflection of the photoelectrons will occur above the domain boundaries. The domain boundaries must then be at positions of maximum and minimum photoelectron intensity in Fig. 1. Images of domains

will not be sharp but will be diffuse because the force of lateral deflection will vary with position as the photoelectrons travel through different regions of the stray field.

Because stray magnetic fields are an essential component in the function of magnetic recording media,[65] photoelectrons can also be deflected by the magnetized regions in magnetic video tapes, recording tapes, computer disks, etc. We have also imaged the magnetized regions in magnetic recording media by photoelectron microscopy. The electrostatic lens system appears well suited for such applications, since a strong magnetic field could erase or distort the magnetized regions of a recording medium.

ACKNOWLEDGMENTS

The authors thank the McMaster Foundation for major purchase of the equipment and for a McMaster Fellowship. The authors thank the Engineering Foundation for operating expenses. The authors also thank C. Heckman and D. Schwab for use of SEM facilities at Bowling Green State University.

[1]E. Brüche, Z. Phys. **86**, 448 (1933).
[2]M. Knoll and E. Ruska, Z. Phys. **78**, 318 (1932).
[3]E. Brüche and H. Johannson, Naturwissenschaften **20**, 353 (1932).
[4]J. Pohl, Phys. Z. **35**, 1003 (1934).
[5]H. Mahl and J. Pohl, Z. Tech. Phys. **16**, 219 (1935).
[6]H. Gross and G. Seitz, Z. Phys. **105**, 734 (1937).
[7]O. H. Griffith and G. F. Rempfer, in *Advances in Optical and Electron Microscopy*, edited by R. Barer and V. E. Cosslett (Academic, London, 1987) Vol. 10, pp. 269–337.
[8]H. Bethge and M. Klaua, Ultramicroscopy **11**, 207 (1983).
[9]W. Telieps and E. Bauer, Ultramicroscopy **17**, 57 (1985).
[10]W. Engel, M. E. Kordesch, H. H. Rotermund, S. Kubala, and A. von Oertzen, Ultramicroscopy **36**, 148 (1991).
[11]M. Mundschau, E. Bauer, and W. Święch, Surf. Sci. **203**, 412 (1988).
[12]M. Mundschau, E. Bauer, and W. Święch, Metall. Trans. A **22**, 1311 (1991).
[13]M. Mundschau, E. Bauer, and W. Święch, J. Appl. Phys. **65**, 581 (1989).
[14]M. Mundschau, E. Bauer, W. Telieps, and W. Święch, J. Appl. Phys. **65**, 4747 (1989).
[15]M. Mundschau, M. E. Kordesch, B. Rausenberger, W. Engel, A. M. Bradshaw, and E. Zeitler, Surf. Sci. **227**, 246 (1990).
[16]M. Mundschau and B. Rausenberger, Platinum Met. Rev. **35**, 188 (1991).
[17]B. Rausenberger, Ph.D. dissertation, *In-situ Investigations of Heterogeneous Catalytic Reactions on Single Crystal Surfaces with Emission and Reflection Microscopy of Low Energy Electrons* (Verlag Köster, Berlin, 1993) (in German).
[18]S. K. Scott, *Oscillations, Waves and Chaos in Chemical Kinetics* (Oxford University, New York, 1994).
[19]M. E. Kordesch, in *CRC Handbook of Surface Imaging and Visualization*, edited by A. T. Hubbard (Chemical Rubber, Boca Raton, FL, 1995), Chap. 56.
[20]H. H. Rotermund, W. Engel, S. Jakubith, A. von Oertzen, and G. Ertl, Ultramicroscopy **36**, 164 (1991).
[21]H. H. Rotermund, W. Engel, M. Kordesch, and G. Ertl, Nature **343**, 355 (1990).
[22]G. Ertl, Science **254**, 1750 (1991).
[23]S. Jakubith, H. H. Rotermund, W. Engel, A. von Oertzen, and G. Ertl, Phys. Rev. Lett. **65**, 3013 (1990).
[24]H. H. Rotermund, S. Jakubith, A. von Oertzen, and G. Ertl, Phys. Rev. Lett. **66**, 3083 (1991).
[25]M. Bär, S. Nettesheim, H. H. Rotermund, M. Eiswirth, and G. Ertl, Phys. Rev. Lett. **74**, 1246 (1995).
[26]M. D. Graham, I. G. Kevrekidis, K. Asakura, J. Lauterbach, K. Krischer, H. H. Rotermund, and G. Ertl, Science **264**, 80 (1994).
[27]M. Ehsasi, Appl. Surf. Sci. **76/77**, 89 (1994).
[28]L. Wegman, J. Microsc. **96**, 1 (1972).
[29]R. A. Schwarzer, Microsc. Acta **84**, 51 (1981); Beitr. Elektronenmikroskop. Direktabb. Oberfl. **12/2**, 3 (1979).
[30]E. Bauer and W. Telieps, in *Surface and Interface Characterization by Electron Optical Methods*, edited by A. Howie and U. Valdré (Plenum, New York, 1988), p. 185.
[31]M. Mundschau, Synchrotron Radiat. News **4**, 29 (1991).
[32]M. Mundschau, in *Handbook of Microscopy*, edited by S. Amelinckx, D. Van Dyck, J. F. Van Landuyt, and G. Van Tendeloo (VCH, Weinheim, 1996), Chap. 6.1.
[33]G. V. Spivak, T. N. Dombrovskaia, and N. N. Sedov, Sov. Phys. Dokl. **2**, 120 (1957).
[34]F. Bitter, Phys. Rev. **41**, 507 (1932).
[35]K. Schur and G. Pfefferkorn, Beitr. Elektronenmikroskop. Direktabb. Oberfl. **4/2**, 235 (1971).
[36]G. Pfefferkorn, L. Weber, K. Schur, and H. R. Oswald, *Proceedings of the Ninth Annual Scanning Electron Microscope Symposium* (SEM Inc., AMP O'Hare, Chicago, 1976), p. 129.
[37]J. Stöhr, Y. Wu, B. D. Hermsmeier, G. R. Harp, and B. P. Tonner, Science **259**, 658 (1993).
[38]B. P. Tonner, D. Dunham, J. Zhang, W. L. O'Brien, M. Samant, D. Weller, B. D. Hermsmeier, and J. Stöhr, Nucl. Instrum. Methods A **347**, 142 (1994).
[39]W. L. O'Brien and B. P. Tonner, J. Vac. Sci. Technol. A **13**, 1544 (1995).
[40]J. J. Croat, J. F. Herbst, R. W. Lee, and F. E. Pinkerton, Appl. Phys. Lett. **44**, 148 (1984).
[41]M. Sagawa, S. Fujimura, H. Yamamoto, and Y. Matsuura, J. Appl. Phys. **55**, 2083 (1984).
[42]J. J. Croat, J. F. Herbst, R. W. Lee, and F. E. Pinkerton, J. Appl. Phys. **55**, 2078 (1984).
[43]M. Sagawa, S. Fujimura, N. Togawa, H. Yamamoto, and Y. Matsuura, J. Appl. Phys. **55**, 2083 (1984).
[44]J. F. Herbst, J. J. Croat, and W. B. Yelon, J. Appl. Phys. **57**, 4086 (1985).
[45]J. F. Herbst, Rev. Mod. Phys. **63**, 819 (1991).
[46]T. Abraham, J. Met. **47**, 16 (1995).
[47]S. Daliang, J. Minhua, C. Huancha, and C. Fusheng, Kexue Tongbao **33**, 848 (1988).
[48]A. von Oertzen, Diploma thesis, Fritz Haber Institute, Berlin, 1990 (in German).
[49]J. D. Livingston, J. Appl. Phys. **57**, 4137 (1985).
[50]D. Li and K. J. Strnat, J. Appl. Phys. **57**, 4143 (1985).
[51]D. Plusa, J. J. Wyslocki, B. Wysłocki, and R. Pfranger, Appl. Phys. A **40**, 167 (1986).
[52]R. K. Mishra and R. W. Lee, Appl. Phys. Lett. **48**, 733 (1986).
[53]J. Pastushenkov, K. D. Durst, and H. Kronmüller, Phys. Status Solidi A **104**, 487 (1984).
[54]Y. Luo and N. Zhang, J. Appl. Phys. **61**, 3445 (1987).
[55]Y. Luo, Q. Ji, and N. Zhang, J. Appl. Phys. **64**, 5525 (1988).
[56]L. Folks, R. Street, and R. C. Woodward, Appl. Phys. Lett. **65**, 910 (1994).
[57]W. C. Elmore, Phys. Rev. **53**, 757 (1938).
[58]C. Kittel, Rev. Mod. Phys. **21**, 541 (1949), p. 548.
[59]B. W. Roberts and C. P. Bean, Phys. Rev. **96**, 1494 (1954).
[60]H. J. Williams, in *Modern Techniques in Physical Metallurgy*, edited by M. Cohen (American Society for Metals, Cleveland, 1953), p. 251.
[61]J. R. Banbury and W. C. Nixon, J. Sci. Instrum. **44**, 889 (1967).
[62]D. C. Joy and J. P. Jakubovics, Philos. Mag. **17**, 61 (1968).
[63]D. J. Fathers, J. P. Jakubovics, D. C. Joy, D. E. Newbury, and H. Yakowitz, Phys. Status Solidi A **20**, 535 (1973).
[64]D. J. Fathers, J. P. Jakubovics, and D. C. Joy, Philos. Mag. **27**, 765 (1973).
[65]J. B. Elsbrock and L. J. Balk, *Scanning Electron Microscopy 1984 I* (SEM Inc., AMF O'Hare, Chicago), p. 131.
[66]J. I. Goldstein, H. Yakowitz, D. E. Newbury, E. Lifshin, J. W. Colby, and J. R. Coleman, *Practical Scanning Electron Microscopy* (Plenum, New York, 1975), p. 180.
[67]M. Mundschau and J. Romanowicz, J. Appl. Phys. (submitted).
[68]G. Harasko, H. Pfützner, and K. Futschik, IEEE Trans. Magn. **31**, 938 (1995).

Effect of silicon processing on giant magnetoresistance

S. L. Burkett,[a] J. Yang, D. Pillai, and M. R. Parker[a]
Department of Electrical Engineering, Center for Materials for Information Technology, University of Alabama, Tuscaloosa, Alabama 35487-0286

(Received 2 October 1995; accepted 20 November 1995)

The primary objective of this study is to determine the effect of silicon wafer processing parameters on giant magnetoresistance (GMR). Thermally oxidized silicon wafers serve as substrates for alternating layers of NiFeCo and Cu (2.0–2.5 nm). SiO_2 films prepared in the laboratory are loaded into the sputtering chamber along with commercially oxidized wafers for comparison purposes. Four process parameters are systematically varied—oxidation method, postoxidation anneal (POA), oxide thickness, and oxidation temperature. Preliminary results indicate the two most influential processing parameters are the method of oxidation (dry oxygen or wet steam) and a POA. In contrast, GMR values do not seem to vary greatly for various oxidation or POA temperatures in the range 800–1000 °C. We present GMR results of NiFeCo/Cu multilayers on oxidized silicon substrates formed by various processing methods. GMR is observed for all films, with the magnitude of the effect increasing for substrates oxidized by "wet" methods and oxidized substrates that include a POA (1000 °C, nitrogen atmosphere, 20 min) previous to sputter deposition of the multilayers. Our data suggest that the GMR of the multilayer stack is sensitive to changes at the Si–SiO_2 interface. © *1996 American Vacuum Society.*

I. INTRODUCTION

A study of the recent literature clearly indicates greatly increased effort in the development of sensors employing ferromagnetic thin films with a magnetoresistance (MR) effect, for use as magnetic field sensors. "Giant" magnetoresistance (GMR) has been studied extensively since the first observations of large resistance changes induced by magnetic fields in Fe/Cr superlattices with antiferromagnetic coupling between adjacent Fe layers.[1,2] In Fe/Cr multilayers, it is generally accepted that spin-dependent scattering causes the GMR effect and is related to interface roughness.[1,3] Many other systems have been found to behave similarly to Fe/Cr, such as Fe/Cu, Fe/Mo, Co/Ru, Co/Cr, and Co/Cu. All of these systems exhibit oscillatory behavior between antiferromagnetic and ferromagnetic coupling.[4–8] GMR observed in sputtered Co/Cu multilayers suggests strong AFM coupling.[9,10]

The multilayer stacks that exhibit GMR are typically composed of alternate thin layers of magnetic material (e.g., Co, Ni, Fe, NiFe) and nonmagnetic, conducting materials (e.g., Cu, Ag, Au). One of the first research groups to observe GMR in NiFeCo/Cu multilayers observed room temperature MR values of 35% and 13% for 30 bilayer films with Cu thicknesses of approximately 0.8 and 2.2 nm, respectively.[11] The addition of Co to the NiFe alloy was expected to increase the MR magnitude over that of NiFe/Cu multilayers originally investigated by Parkin.[12] The room temperature MR values reported by Parkin for NiFe/Cu 14 bilayer films were 16% with a Cu thickness of 0.8 nm and under 2% for 2.0 nm Cu spacers, with basically no coupling observed at the latter thickness. Parkin's later work in NiFe/Cu multilayers with a thin Co layer at each interface showed a dramatic increase in MR to nearly 20% at room temperature with 2.0 nm Cu spacers.[13] In our laboratory, we have observed MR values of 10% in NiFeCo/Cu multilayers with a sensitivity of approximately 0.3% MR/Oe and near zero hysteretic profiles with a Cu spacer layer thickness of 2.3 nm.[14]

In our laboratory, we first observed a small increase in the MR ratio on a set of oxidized silicon substrates (5000 Å) prepared by "wet" oxidation methods; one sample included a postoxidation anneal. The multilayer materials consisted of four bilayers of sputter deposited Co/Cu (2.0 nm/2.3 nm). A commercially oxidized silicon wafer was included in the sputtering chamber for reference and resulted in a MR value of 6.1%. A MR value of 5.35% was obtained for the laboratory sample, which included a POA compared to a value of 4.5% for the sample with no extra annealing step. A POA step resulted in an increase in the MR ratio of approximately 20% as compared to the sample with no anneal.

This study was motivated by an interest in validating these experimental results and to further investigate the effect of substrate processing on GMR. NiFeCo/Cu were chosen as the materials for the multilayers. The ternary alloy NiFeCo exhibits low magnetostriction[15] and low magnetocrystalline anisotropy.[16] In candidate materials for magnetic field detection elements, both low magnetostriction and magnetocrystalline anisotropy are desired to reduce magnetic switching noise.

II. EXPERIMENTAL METHODS

Single crystal (111) oriented 0.005 Ω cm n-type silicon wafers are cleaned by a standard RCA cleaning procedure foregoing the oxidation process. Silicon substrates are thermally oxidized in a quartz tube using two oxidation methods. High purity oxygen provides the ambient for "dry" oxidation and steam produced from boiling deionized (DI) water

[a]Electronic mail: {sburkett, mparker}@coe.eng.ua.edu

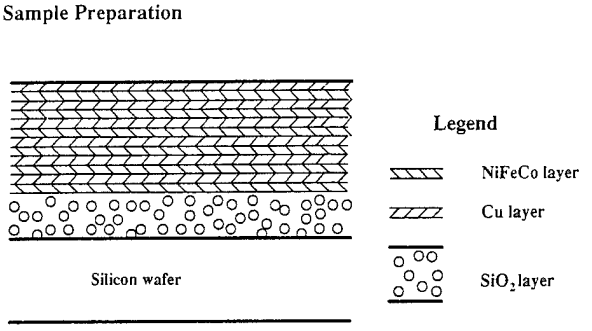


FIG. 1. Schematic illustrating sample preparation.

provides oxygen for "wet" oxidation. An important processing technique routinely used to improve the quality of the interface between silicon and thermally grown silicon dioxide is the postoxidation anneal (POA). The effect of a high temperature POA on the characteristics of the $Si–SiO_2$ interface has been widely studied.[17–19] The interest is primarily due to device scaling for integrated circuits utilizing metal–oxide–semiconductor (MOS) technology which involves reducing the thickness of gate oxides to approximately 100 Å. POAs have an important effect on device characteristics and have been found to improve transistor properties.[20] A high temperature POA can also increase Si–O bond strength at the interface due to densification, which leads to lower interface state density (D_{it}) generation.[21] A POA at various temperatures in a nitrogen ambient is performed to purposely modify the Si/SiO_2 interface properties. Unless otherwise noted, a POA is done at 1000 °C for 20 min in a nitrogen atmosphere.

NiFeCo multilayers are sputter deposited by dc magnetron sputtering on oxidized silicon substrates in a Vac-Tec Model 250 horizontal batch sputtering system using ultrahigh purity (99.999%) argon. The magnetic field from the permanent magnet results in higher sputtering rates for the same power compared to other dc sputtering systems due to higher current densities near the target surface.[22] This enhanced current density with magnetron sputtering is possible due to the higher concentration of electrons kept near the target surface, providing more opportunity for argon ionization. Additionally, a lower sputtering pressure is made possible by magnetron sputtering. The predeposition base pressure of the system is in the range of 2×10^{-7} Torr. Samples are deposited under an argon pressure of ~2.3 mTorr. The films are all deposited at ambient temperature with a permanent magnet behind each substrate supplying an aligning field of 80 Oe parallel to the substrate surface during deposition. A schematic illustrating sample preparation is shown in Fig. 1.

Commercially oxidized silicon wafers (SiO_2 thickness =10 000 Å) are loaded into the sputtering chamber along with wafers thermally oxidized in our laboratory. A $Ni_{66}Fe_{16}Co_{18}$ alloy target is used in the sputtering chamber. Electron microprobe results on granular $Ni_{66}Fe_{16}Co_{18}$–Ag sputtered films reveal a ferromagnetic portion composition of approximately $Ni_{76}Fe_{10}Co_{14}$.[14] The NiFeCo/Cu films consist of six sputtered bilayers. The Cu layers are 2.3–2.5 nm thick

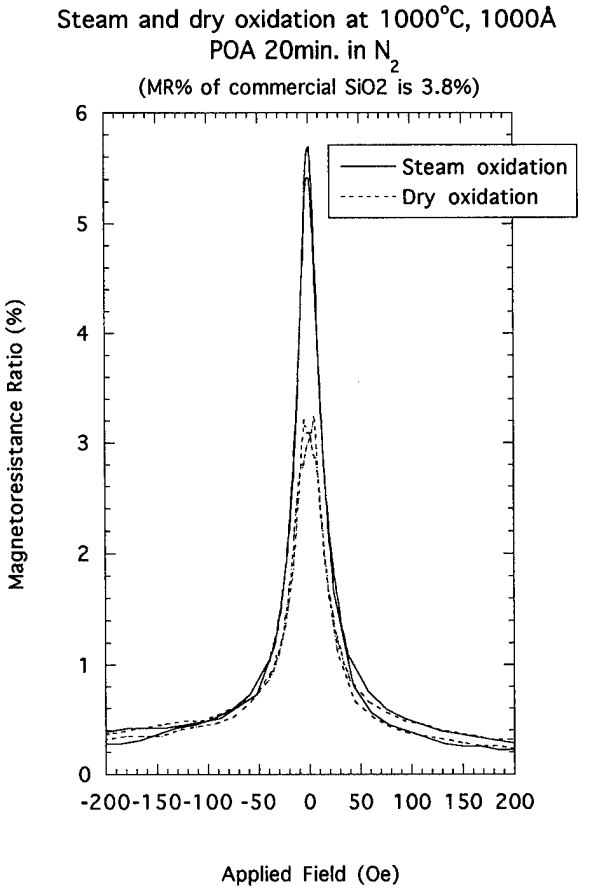


FIG. 2. Magnetoresistance ratio (%) vs applied field (Oe) for wet and dry oxidation methods.

and the alloy layers are approximately 2.0–2.5 nm thick. Sputtering rates are calculated from thickness measurements of reference films made with a Dektak IIa surface profilometer. Film thicknesses are controlled by controlling sputtering time. In this study the NiFeCo and the Cu spacer layer thickness is held constant as well as the number of bilayers for comparison purposes.

An important parameter for evaluating GMR materials is the normalized change in resistance, $\Delta R/R$, the MR ratio. MR measurements are made with a linear four point probe technique with both current and magnetic field in the film plane but perpendicular to each other. All measurements are made at room temperature on as-deposited samples unless otherwise noted. The MR percentage is calculated using the equation

$$\mathrm{MR}\% = [(\rho_{\max} - \rho_{\min})/\rho_{\min}] \times 100,$$

where $\rho_{\max}$ is the film resistivity at (or near) zero field, and $\rho_{\min}$ is the resistivity at the maximum applied field.

The variability observed in MR values for multilayers prepared at different times makes direct comparison of results difficult. This is remedied somewhat by comparing only MR magnitudes for substrates prepared in the same multilayer sputter deposition. Commercially oxidized wafers are also included in each sputter deposition as reference samples.

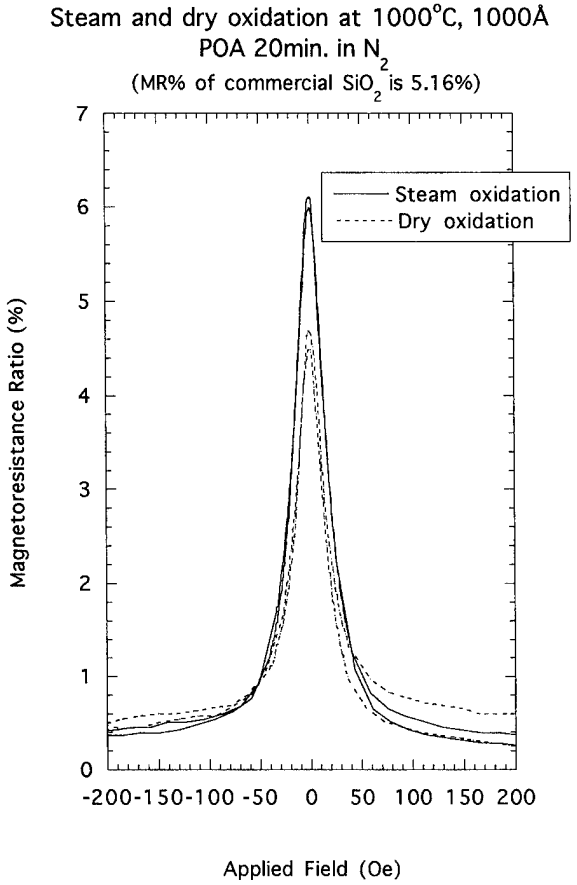


FIG. 3. Magnetoresistance ratio (%) vs applied field (Oe) for substrates prepared by wet and dry oxidation methods.

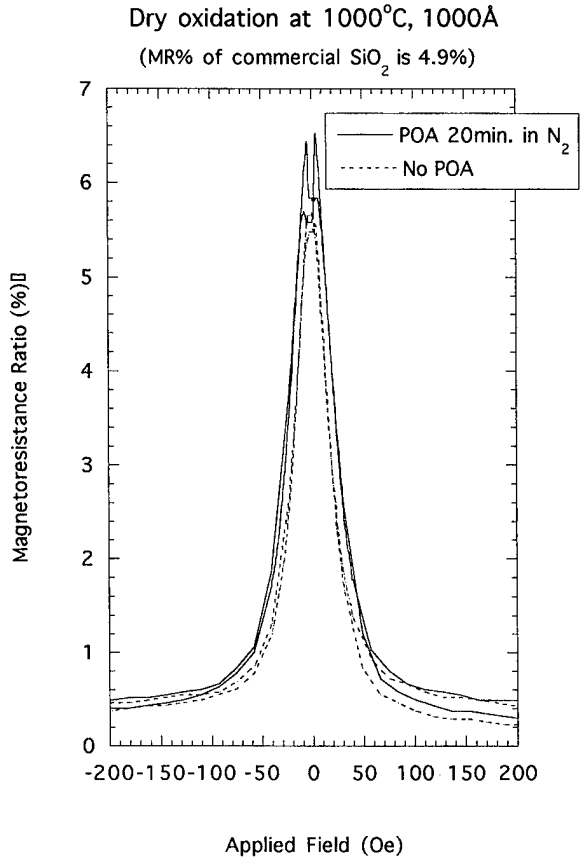


FIG. 4. Magnetoresistance ratio (%) vs applied field (Oe) for substrates prepared by dry oxidation methods and POA.

III. EXPERIMENTAL RESULTS

NiFeCo and Cu are the chosen multilayer materials and samples are prepared as described in Sec. II. Four process parameters are systematically varied. They are oxidation method, postoxidation annealing, oxide thickness, and oxidation and annealing temperature. The two most influential processing parameters are the method of oxidation and the performing of a POA.

A. Oxidation method

Multilayers sputtered on substrates prepared by wet methods consistently produce higher MR values. As stated before, due to the variability seen in samples prepared in different sputter depositions and the strong dependence of film thickness in the multilayer stack on MR values, experimental results are compared to results obtained from samples prepared in the same run. In one experiment, a SiO_2 film oxidized by dry oxygen (1000 °C, 1000 Å, POA), gives a MR value of 3.22% and one oxidized by wet methods gives a value of 5.7%. For reference, a commercially oxidized silicon wafer yielded an MR value of 3.8% for that particular sputter deposition. Typical MR profiles showing the dependence on oxidation method are shown in Figs. 2 and 3.

This effect is believed to be related to reduced interface roughness effects. Atomic force microscopy (AFM) studies are in progress, but preliminary analysis indicates very smooth surfaces and no distinct, distinguishable features between substrates prepared by wet or dry oxidation methods or substrates that include an additional annealing step. Independent AFM and ellipsometry studies reveal that interface roughness decreases with the extent of oxidation, for samples prepared with dry oxygen, and is relatively independent of

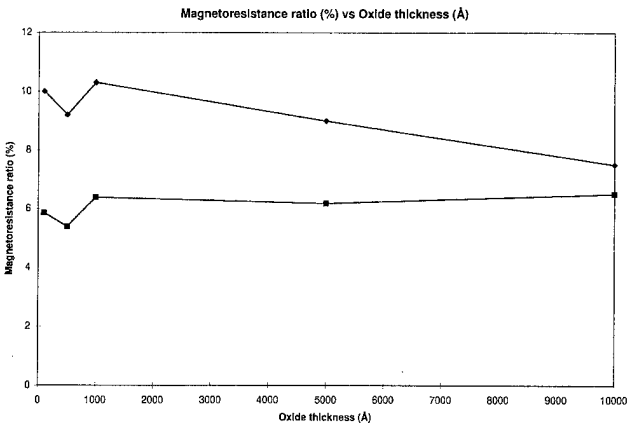


FIG. 5. Magnetoresistance ratio (%) vs silicon oxide thickness (Å) for substrates prepared by wet oxidation methods.

TABLE I. Summary of MR values for various silicon wafer processing conditions.

Silicon processing variation	Processing variation	MR value (%)
Oxidation method		
(1000 °C, 1000 Å), no POA	Dry	8.2
Commercial MR value 8.7%	Wet	10.0
(1000 °C, 1000 Å) POA	Dry	4.68 4.34
Commercial MR value 5.16%	Wet	6.1 6.14
(1000 °C, 1000 Å) POA	Dry	3.22
Commercial MR value 3.8%	Wet	5.7
POA		
Wet (1000 °C, 5000 Å)	POA	12.3
Commercial MR value 11.3%	No POA	11.6
Dry (1000 °C, 1000 Å)	POA	(a) 9.1 (b) 6.52
(a) Commercial MR value 8.7%	No POA	(a) 8.2 (b) 5.69
(b) Commercial MR value 4.9%		
Oxide thickness		
Wet (1000 °C), no POA		
(a) Commercial MR value 10.2%	100 Å	(a) 10.0 (b) 5.85
(b) Commercial MR value 5.16%	500 Å	(a) 9.2 (b) 5.39
	1 000 Å	(a) 10.3 (b) 6.38
	5 000 Å	(a) 9.0 (b) 6.18
	10 000 Å	(a) 7.5 (b) 6.5
Oxidation temperature		
Wet (1000 Å) no POA	800 °C	9.8
Commercial MR value 8.7%	900 °C	9.6
	1 000 °C	10.0
POA temperature		
Wet (1000 °C, 1000 Å)	600 °C	4.0
Commercial MR value 5.42%	800 °C	3.15
	1 000 °C	4.82
Dry (1000 °C, 1000 Å)	600 °C	5.0
Commercial MR 5.42%	800 °C	4.42
	1 000 °C	4.65

crystal orientation of the substrate.[23] In that work, attempts to measure changes in interface roughness were futile because the level of roughness was below the noise limits of the AFM. They examined purposely roughened silicon substrates to allow AFM studies to reveal the decrease in roughness with increasing oxidation.[23,24]

B. Postoxidation anneal

The presence of a POA also consistently increases MR values. POA is well known for reduction of the density of interface states at the $Si–SiO_2$ interface. For one set of samples, a MR value of 5.69% was obtained for a sample oxidized by dry oxygen; a similar sample undergoing a POA yielded a value of 6.52%. Similar results are obtained for substrates prepared by wet oxidation methods and annealed for 20 min in nitrogen at 1000 °C. For reference, a commercially oxidized wafer with no anneal gave a MR value of 4.9%. A typical MR profile for a sample oxidized by dry oxygen is shown in Fig. 4. The changes in MR ratios for the substrates that include postoxidation annealing is consistently observed.

C. Oxide thickness

Another interesting behavior is observed for substrates with various SiO_2 thicknesses prior to sputter deposition of the multilayers. Oxide thickness is varied from 100 to 10 000 Å in an attempt to understand the effect of an ultrathin oxide film. Thickness of the underlying oxide buffer will be very important in ultrahigh density magnetic recording applications where intershield distances are extremely small and insulating layers extremely thin. Samples are prepared by wet oxidation methods at 1000 °C and no POA is performed. Oxide thickness is varied by controlling oxidation time. Two sets of data for samples from two different multilayer sputter depositions reveal only slight changes in the magnitude of the MR ratio as a function of oxide thickness. The MR ratio dependence on oxide thickness is shown in Fig. 5. Comparison of the data shows that a distinct kink in the curve is observed at thickness in the range 100–1000 Å. This phenomenon is currently under detailed investigation but is quite reproducible as evidenced by Fig. 5. Both sets of substrates are prepared using the same methods but differ in their relative magnitude. The reason for the differences in MR values

between the two sets is merely due to the variations observed in separate depositions and arise from the variation in calibration of the sputtering target.

D. Oxidation and anneal temperature

MR values do not vary greatly for various oxidation temperatures in the range 800–1000 °C. For this particular experiment, samples are prepared by wet methods and thickness is held constant at 1000 Å. MR ratios of 9.8%, 9.6%, and 10% are obtained for oxidation temperatures of 800, 900, and 1000 °C, respectively, while 8.7% is obtained for the commercially oxidized wafer included in that deposition. Similar results are seen for samples with POA temperatures in the range 600–1000 °C. Table I summarizes the experimental results for each silicon wafer processing variation in this study.

IV. DISCUSSION

MR values are strongly dependent on the thickness of the multilayer films. The experimental results shown in this article come from measurements at room temperature on as-deposited multilayers. For one large sample set, MR values of approximately 1% were observed for a variety of samples. When the same samples were measured again after 1 month, a slight increase is observed for most samples. Some, however, had doubled in value, and a few isolated samples gave MR values of 4%–5%. This room temperature magnetic after effect has been seen before in our laboratory[25] and is consistent with gradual change from ferromagnetic to antiferromagnetic interlayer exchange coupling. After annealing the multilayers at 150 °C, a dramatic increase is observed and MR values range from approximately 3% to 6% overall with some samples as high as 7%. Further annealing at 250 °C brings about another slight increase. Because there is an optimum thickness for efficient exchange coupling between the multilayers, the annealing step allows one to obtain increased MR values for multilayers that are slightly thicker than optimal.[26]

V. SUMMARY

The sensitivity of GMR to processing conditions of a silicon dioxide substrate previous to sputter deposition of multilayer thin films is studied in the present work by monitoring comparative magnitudes of resulting magnetoresistance values. These values are compared to values obtained from multilayers sputter deposited on commercially oxidized silicon wafers. The magnitude of MR ratios increase for oxidized substrates that include a POA (1000 °C, nitrogen atmosphere, 20 min) and substrates oxidized by wet methods. MR values appear to be relatively independent of the thickness of the oxide layer, and even films as thin as 100 Å produce significant MR values. Our data suggests that GMR of the multilayer stack is sensitive to changes at the $Si–SiO_2$ interface. Future plans include a more detailed study of the variation in thickness of the underlying oxide film and investigation of multilayers deposited on alumina substrates. This work can be useful in finding optimum processing conditions and determining the role of the substrate on GMR.

ACKNOWLEDGMENTS

This research was supported in part by NSF Materials Research Science and Engineering Center (MRSEC) Cooperative Agreement No. DMR-9400399 and by a grant from ATP (Department of Commerce) administered by the National Storage Industrial Consortium.

[1]M. N. Baibich *et al.*, Phys. Rev. Lett. **61**, 2472 (1988).
[2]G. Binasch, P. Grünberg, F. Saurenback, and W. Zinn, Phys. Rev. B **39**, 4828 (1989).
[3]B. A. Gurney, P. Baumgart, D. R. Wilhoit, B. Dieny, and V. S. Speriosu, J. Appl. Phys. **70**, 5867 (1991).
[4]W. R. Bennett, W. Schwarzacher, and W. F. Egelhoff, Jr., Phys. Rev. Lett. **65**, 3169 (1990).
[5]M. E. Brubaker, J. E. Mattson, C. H. Sowers, and S. D. Bader, Appl. Phys. Lett. **58**, 2306 (1991).
[6]S. S. P. Parkin, N. More, and K. P. Roche, Phys. Rev. Lett. **64**, 2304 (1990).
[7]S. S. P. Parkin, R. Bhadra, and K. P. Roche, Phys. Rev. Lett. **66**, 2152 (1991).
[8]D. Pescia, D. Kerrkmann, F. Schumann, and W. Gudat, Z. Phys. B **78**, 475 (1990).
[9]A. Fert *et al.*, J. Magn. Magn. Mater. **104-107**, 1712 (1992).
[10]S. Honda, S. Ohmoto, R. Imada, and M. Nawate, J. Magn. Magn. Mater. **126**, 419 (1993).
[11]M. Jimbo, T. Kanda, S. Goto, S. Tsunashima, and S. Uchiyama, Jpn. J. Appl. Phys. **31**, L1348 (1992).
[12]S. S. P. Parkin, Appl. Phys. Lett. **60**, 512 (1992).
[13]S. S. P. Parkin, Appl. Phys. Lett. **61**, 1358 (1992).
[14]S. Hossain, D. Seale, G. Qiu, J. Jarratt, J. A. Barnard, H. Fujiwara, and M. R. Parker, J. Appl. Phys. **75**, 7067 (1994).
[15]C. H. Tolman, J. Appl. Phys. **38**, 3409 (1967).
[16]L. W. McKeehan, Phys. Rev. **51**, 136 (1937).
[17]B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, J. Electrochem. Soc. **114**, 266 (1967).
[18]F. Montillo and P. Balk, J. Electrochem. Soc. **118**, 1463 (1971).
[19]S. K. Lai, D. R. Young, J. A. Calise, and F. J. Feigl, J. Appl. Phys. **52**, 5691 (1981).
[20]P. J. Wright, A. Kermani, and K. C. Saraswat, IEEE Trans. Electron Devices **ED-37**, 1836 (1990).
[21]M. Shimaya and N. Shiono, Extended Abstracts of the 18th International Conference on Solid State Devices and Materials, Tokyo (1986), p. 471.
[22]R. K. Waits, *Thin Film Processes*, edited by J. L. Vossen and W. Kern (Academic, San Diego, 1978).
[23]Q. Liu, J. F. Wall, and E. A. Irene, J. Vac. Sci. Technol. A **12**, 2625 (1994).
[24]Q. Liu, L. Spanos, C. Zhao, and E. A. Irene, J. Vac. Sci. Technol. A **13**, 1977 (1995).
[25]S. A. Hossain (private communication).
[26]S. A. Hossain and M. R. Parker, in *Proceedings of the 2nd International Symposium Metallic Multilayers, Cambridge* (IOP, London, 1995), p. 50.

Electronic magnetic and structural coupling in colossal magnetoresistive (La,Ca)MnO_3

Warren E. Pickett and David J. Singh[a)]
Complex Systems Theory Branch, Naval Research Laboratory, Washington, DC 20375

(Received 2 October 1995; accepted 18 December 1995)

The local spin density approximation (LSDA) is used to investigate the interrelationships between electronic properties, magnetism, structure, and transport in $La_{1-x}Ca_xMnO_3$. The LSDA is shown to provide a realistic description of the endpoint compounds, including magnetic and structural symmetry. Mn–O hybridization is strong and spin dependent. Supercell calculations for $x=\frac{1}{4}$ and $x=\frac{1}{3}$ yield ferromagnetic configurations. Strong local La–Ca cation effects are found, which lead to localization of the low density of minority spin carriers in this regime. Implications of this effectively half-metallic electronic structure are discussed.

I. INTRODUCTION

The discovery of colossal magneto-resistance (CMR) in films of $La_{1-x}Ca_xMnO_3$ and related Mn containing perovskites[1–4] raises the question of what is special about these materials to permit extremely large resistivity reductions in laboratory fields. The endpoints of this alloy system are antiferromagnetic (AF) insulators, with different structural and magnetic symmetries. $CaMnO_3$ is cubic perovskite with a lattice parameter of 3.73 Å and G-type (rocksalt) AF ordering,[5] while $LaMnO_3$ occurs in a distorted perovskite structure ($GdFeO_3$ type, $Pnma$, No. 62), with A-type AF order.[5,6] In the concentration range, $0.2 < x < 0.5$, the alloy occurs in the cubic perovskite structure, and is metallic and ferromagnetic (F) at low temperatures. It is in this regime ($x \approx \frac{1}{3}$) that the CMR effect is observed.[7,8]

Goodenough[9] addressed these materials early on with a discussion of Mn d–O p hybridization leading to "semicovalent exchange." His theory led to connections between magnetic order, crystal structure, and electronic transport. The endpoints formally contain Mn^{3+} (d^4) and Mn^{4+} (d^3) ions. Based on the neutron data,[5] these are in high spin configurations, though possibly reduced from Hund's rule values due to hybridization. In the perovskite crystal field, the t_{2g} orbitals are lower in energy than the e_g, so that for $CaMnO_3$ the t_{2g} majority orbitals hold the three electrons, and the minority spin orbitals on the Mn are unoccupied. In $LaMnO_3$ there is one additional Mn d electron, leading to partial occupation of the two majority spin e_g states. This is presumed to result in a Jahn–Teller instability, accounting for the $Pnma$ structural distortion.

It should be noted that there are additional driving forces that may be operative. Rotational instabilities of oxygen octahedra due to ion size mismatch resulting in this symmetry are common in perovskites and the La and Mn ionic radii are consistent with this. However, such a mechanism does not generally lead to the substantial changes in transition metal–O distance occurring in $LaMnO_3$.

Zener[10] proposed a double exchange process to explain the occurrence of a metallic ferromagnet in the phase diagram. In his picture, the O ions are closed shell (O^{2-}). Nonetheless, hopping must occur through the O. This occurs through a correlated electron motion in which an electron jumps from O to a neighboring Mn while an electron hops from a different Mn neighbor onto the O. This process requires that both hopping electrons have the same spin. Since the Mn are in the high spin configuration, the implication is that for conduction, the active Mn ions must have the same spin. While this does nicely account for the relationships between transport and magnetic order in the (La,Ca)MnO_3 alloy, it leaves certain questions unanswered, particularly, the role of hybridization and how the very low resistivities that have been observed ($\sim 100\ \mu\Omega$ cm in films[11]) are consistent with a correlated hopping conduction process.

Here, local spin density approximation (LSDA) calculations are used to elucidate the extent and role of the hybridization and to establish the main features of the electronic structure in the CMR regime. Other electronic structure studies are those of Sarma *et al.*,[12] Satpathy *et al.*,[13] and Butler *et al.*[14] The LSDA description is found to account quite well for established experimental facts. Substantial, spin dependent hybridization is found along with effectively half-metallic character. This is related to the CMR phenomena.

II. METHOD

The calculations were performed with a generalized linearized augmented plane-wave method.[15,16] No shape approximations are made to the potential or the charge or spin densities in this method, and additionally a flexible basis is employed in all regions of space. This makes the method suitable for open structures with low site symmetries, as considered here. Well converged linear augmented plane wave (LAPW) basis sets of approximately 400 functions per formula unit were employed, with the addition of local orbitals to relax linearization errors and to treat the semicore states (Ca $3s$, $3p$; La $5s$, $5p$; O $2s$) on the same footing as the valence states. The La $4f$ states were also treated self-consistently with the valence states, and were found 2–3 eV above the Fermi energy (E_F). Sphere radii of 1.55 and 2.0 a.u. were used for O and the cations, respectively. All calculations were performed self-consistently. Scalar relativistic

a)Electronic mail: singh@dave.nrl.navy.mil

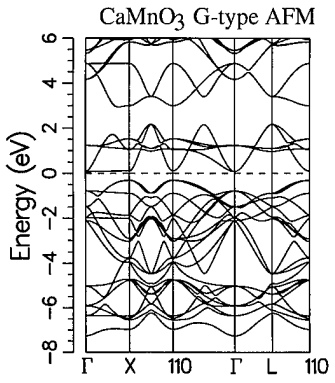


FIG. 1. Band structure of $CaMnO_3$ with G-type (rocksalt) AF ordering. The designation 110 indicates a second X point lying outside the first Brillouin zone.

effects were retained for the valence states, while core states were treated fully relativistically in an atomiclike fashion.

The Brillouin zone sampling was performed using the special **k** points method. A dense set of 42 points in the irreducible wedge was used for the supercells in the crucial $x=\frac{1}{3}$ CMR regime. Use of a smaller set of 20 points yielded the identical moments to within 1%. The calculations were performed for a volume/formula unit corresponding to that at the CMR composition, $x=\frac{1}{3}$. However, because of the near Hund's rule spin moments and the stiffness, which we find in fixed-spin-moment calculations, only weak and continuous variation with volume is anticipated.

III. $CaMnO_3$

As mentioned, $CaMnO_3$ is an AF insulator occurring with the cubic perovskite structure. Calculations were performed for this structure with various spin orderings: A-type AF, G-type AF (observed), and F. The energies of these configurations relative to a constrained nonspin-polarized calculation are -0.92 (A-type AF), -0.98 (G-type AF), and -0.86 eV (F). These large energy gains, which are relatively independent of the ordering, show the robustness of the magnetic solutions, leading to the expectation that the moments, although disordered and fluctuating, should persist relatively undisturbed above the ordering temperature T_N. The Mn d exchange splitting is $\Delta_{ex}\sim 3$ eV for each ordering.

The observed G-type AF magnetic ordering has the lowest energy (-0.98 eV), in agreement with the experimental symmetry. Furthermore, with this ordering an insulating band structure with a 0.42 eV gap is obtained (Fig. 1). This direct gap, which occurs at the X point, is between majority and minority spin Mn t_{2g} states and therefore is accessible only through spin-flip transitions. The measured optical gap should be larger.

The Mn t_{2g} manifold is essentially completely polarized, in accord with Hund's rule. However, the band structure and projections of the density of states (DOS) show substantial hybridization with O $2p$ states, primarily affecting the majority spin Mn states, which are below E_F and overlap O $2p$ bands. One measure of the hybridization is through the reduction from integral values of the moment. For the G-type ordering, the moment in the Mn sphere is 2.48 μ_B. Bearing in mind that 5% of the Mn d moment may lie outside the 2.0 a.u. sphere, this is in reasonable accord with the neutron value of 2.65 μ_B.[5]

The result for $CaMnO_3$ with an imposed F ordering is remarkable. The calculated magnetization is precisely 3 μ_B in spite of considerable hybridization. This occurs because the band structure is half-metallic ($\frac{1}{2}M$), with metallic majority spin bands crossing E_F but a gap about E_F in the minority spin bands. The lower lying majority Mn t_{2g} states are occupied and, as a result, overlap and strongly mix with O $2p$ states. The minority Mn bands are 3 eV higher, and lie above the O bands.

IV. $LaMnO_3$

$LaMnO_3$ occurs in a distorted Pnma crystal structure. However, before describing calculations for the actual structure, we discuss the cubic perovskite structure, which may be more directly related to the CMR alloys. For this structure, neither the F- nor the A-type (observed) or G-type magnetic orderings yielded an insulating band structure and the F ordering was lowest in energy. The bands for the F ordering are very similar to those of $\frac{1}{2}M$ $CaMnO_3$ but due to the increased band filling the minority spin gap occurs below E_F, and both spin channels are metallic. The F spin magnetization is 3.38 μ_B, which is a considerable drop from the Hund's rule value (4 μ_B). The moment in the Mn sphere is only 2.89 μ_B reflecting again the importance of hybridization.

The distortion to the Pnma structure involves a rotation of the oxygen octahedra and an additional strong Jahn–Teller-like distortion that changes the Mn–O bond lengths, which then differ by as much as 0.27 Å. The energy gain due to this distortion is calculated as 0.23 eV per formula unit. This energy difference is close to that obtained by Sarma *et al.*[12] Including the distortion the A-type AF ordering (observed) is lowered below the F state, and a 0.12 eV insulating gap develops in the band structure. The removal of minority spin carriers from E_F results in an increase of the moment to 3.31 μ_B within each Mn sphere. There is an additional polarization of 0.07 μ_B within those O spheres where it is allowed by symmetry. This is, however, a severe underestimate of the O contribution, since the $2p$ states in O ions are very extended and are not contained by a 1.55 a.u. sphere. In any case, the results show that the Pnma distortion is necessary to describe $LaMnO_3$. However, when it is included, the LSDA gives a quite reasonable description of this material.

V. CMR REGIME

As shown above, the LSDA describes both endpoints of the alloy system including the magnetic ordering, insulating character, and structural distortion. Accordingly, it is reasonable to proceed to the CMR alloys. In this regime, $(La,Ca)MnO_3$ has the cubic perovskite structure and ferromagnetic order. We studied this range using ordered supercells with $x=\frac{1}{3}$ and $x=\frac{1}{4}$.

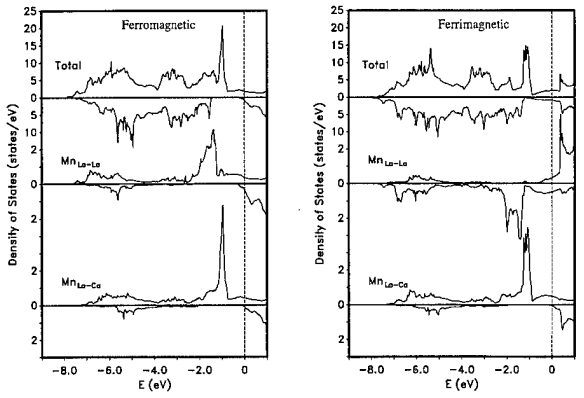


FIG. 2. DOS and projections onto the Mn spheres for the $x=\frac{1}{3}$ supercell with the actual F ordering (left) and with an imposed ferrimagnetic ordering (right). Majority (minority) spin is shown as positive (negative). The projections are on a per atom basis. E_F is denoted by the dashed vertical lines.

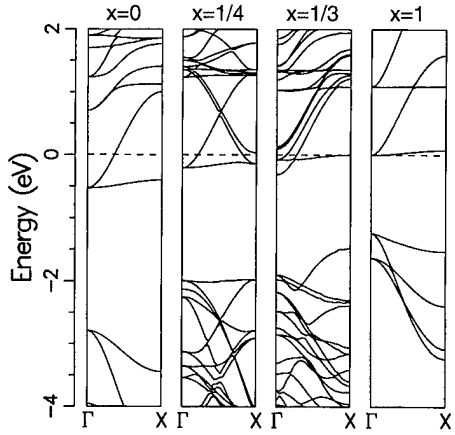


FIG. 3. Minority spin band structure along $\Gamma-X$ for F ordered $La_{1-x}Ca_xMnO_3$, $x=0$, $\frac{1}{4}$, $\frac{1}{3}$, and 1. Of the five bands near E_F at $x=\frac{1}{3}$, the lowest is purely Mn_{La-La} derived and the upper pair is derived from the two Mn_{La-Ca} sites; the middle pair involves all Mn ions. Note the decrease in minority gap with Ca content. This arises from raising of the O p bands as the neighboring cation charge decreases.

For an ordered $x=\frac{1}{3}$ supercell, $La_2CaMn_3O_9$ with periodically repeating [–La–Ca–La–] layers, F order is obtained with a moment of 3.51 μ_B per Mn. The artificial cation ordering leads to two distinct Mn sites, one with eight La neighbors (Mn_{La-La}) and the other with four La and four Ca neighbors (Mn_{La-Ca}). The moment on the Mn_{La-La} site is larger by 0.05 μ_B than that on the Mn_{La-Ca} site, as might be expected. This difference, however, is much less than would be expected in a purely ionic picture. The DOS and projections are shown in Fig. 2. Flipping the Mn_{La-La} moment, leads to a metastable "ferrimagnetic" configuration that is 0.07 eV higher in energy.

An ordered body-centered-cubic symmetry supercell at $x=\frac{1}{4}$, $La_3CaMn_4O_{12}$ (each Mn has six La and two Ca neighbors) with F ordering displays similar behavior, with a spin magnetization of 3.53 μ_B per Mn. As for the endpoint compounds the Mn d exchange splittings are $\Delta_{ex}\sim 3$ eV.

The DOS and projections show a certain constancy of behavior. The differentiation of the Mn ions is a local environment effect, that will have its counterpart in the disordered solid. The projected DOS (PDOS) shows hybridization with O $2p$ states in the majority bands at and below E_F. The PDOS are strongly spin differentiated in the range -2–0 eV (relative to E_F), and in this region the moment flip results simply in a flip in the PDOS of the affected atom, while leaving its shape and the PDOS of the neighboring Mn ions practically unchanged.

In the $x=\frac{1}{3}$ supercell, the small number of minority carriers are associated practically exclusively with the Mn_{La-La} site; this Mn is surrounded by trivalent La, and so it is not surprising that it is more attractive for electrons. The other Mn ions remain $\frac{1}{2}M$ with no minority spin carriers at E_F, similar to F ordered $CaMnO_3$. In the $x=\frac{1}{4}$ supercell, the minority spin band edge is again slightly below E_F, and is intermediate in position between the edges in the minority spin PDOS for Mn_{La-La} (8 La neighbors) and Mn_{La-Ca} (4 La neighbors). The dependence of the minority edge on local environment is illustrated in Fig. 3, where the minority spin band structure along $\Gamma-X$ is shown for the various calculations.

VI. DISCUSSION

CMR occurs in the cubic perovskite F ordered regime. The above calculations show that in this case the majority bands have strongly metallic, well hybridized Mn d–O p character. This is entirely consistent with the low observed resistivities as temperature is reduced, since ordinary metallic conduction is expected.

The minority spin band structures show strong variations in the position of the band edge and number of carriers with x. As mentioned this is a local environment effect. F ordered $CaMnO_3$ is a true $\frac{1}{2}M$ ferromagnet. The ordered $x=\frac{1}{4}$ and $x=\frac{1}{3}$ supercells have low, site dependent carrier densities. In reality, the A cation site is disordered, and this disorder will couple to the position of the minority edge through the local effects discussed above, and through local strain effects. The result will be localization of minority carriers, where E_F occurs below a mobility edge in this spin channel. Thus F ordered (La,Ca)MnO_3 is predicted to be effectively $\frac{1}{2}M$ in the CMR regime. For F alignments the majority spin bands are strongly covalent and metallic, while the minority bands are insulating. This implies the lack of a true Stoner continuum, although there is the possibility of new low energy excitations involving a hole in the majority spin bands and an electron in a localized minority spin state, or vice versa.

The local environment dependence of the electronic structure, discussed above, implies that the effective $\frac{1}{2}M$ behavior will persist near flipped Mn spins, such as would occur as temperature is raised, near abrupt domain boundaries, or in locally Ca-rich regions where AF order is favored. This provides a way to account for high resistivity in zero field near T_c. Due to the $\frac{1}{2}M$ conduction, hopping of carriers between F regions with opposite magnetization directions will be

strongly suppressed. This is similar to physics discussed for heterogeneous giant magnetoresistive (GMR) materials like Co–Cu,[17,18] but would be stronger because of the maximum $\frac{1}{2}M$ asymmetry. Above T_c there is no net moment. Whether the moments become disordered on an atomic scale or an a larger scale corresponding to several F coupled spins, hopping between the these units will be strongly suppressed when the relative orientation differs. We speculate that the conduction may become nonmetallic due to a combination of bandwidth narrowing (Mott localization) and disorder (Anderson localization). If the locally aligned regions become larger, as expected near T_c, it may be feasible to rotate them using applied fields, providing a mechanism for magnetoresistance.

VII. SUMMARY

Clearly, there is a considerable distance to go in understanding the CMR effect in manganites. Nonetheless, our calculations do provide some insight into the (La,Ca)MnO_3 system. The LSDA has been shown to provide a realistic description of the materials. Effectively $\frac{1}{2}M$ transport, with strongly hybridized conducting majority bands, and local environment effects are found for the CMR alloys. This provides a ready explanation for the low resistivity at low T, and suggests possible mechanisms for achieving high resistivity and field dependence near and above T_c.

ACKNOWLEDGMENTS

The authors are grateful for the helpful discussions with W. H. Butler, J. M. Buyers, J. Eckstein, N. C. Koon, D. A. Papaconstantopoulos, M. Rubenstein, G. M. Stocks, and X.-G. Zhang. Computations were performed at the Alaska Region Supercomputing Center and the DoD centers at CEWES and NAVO.

[1]R. von Helmolt, J. Wecker, B. Holzapfel, L. Schultz, and K. Samwer, Phys. Rev. Lett. **71**, 2331 (1993).
[2]M. McCormack, S. Jin, T. H. Tiefel, R. M. Fleming, J. M. Phillips, and R. Ramesh, Appl. Phys. Lett. **64**, 3045 (1994).
[3]S. Jin, T. H. Tiefel, M. McCormack, R. A. Fastnacht, R. Ramesh, and L. H. Chen, Science **264**, 413 (1994).
[4]R. Mahendiran, A. K. Raychaudhuri, A. Chainani, D. D. Sarma, and S. B. Roy, Appl. Phys. Lett. **66**, 233 (1995).
[5]E. O. Wollan and W. C. Koehler, Phys. Rev. **100**, 545 (1955).
[6]J. B. A. A. Elemans, B. van Laar, K. R. van der Veen, and B. O. Loopstra, J. Solid State Chem. **3**, 238 (1971).
[7]G. H. Jonker and J. H. van Santen, Physica **16**, 337 (1950).
[8]J. H. van Santen and G. H. Jonker, Physica **16**, 599 (1950).
[9]J. B. Goodenough, Phys. Rev. **100**, 564 (1955).
[10]C. Zener, Phys. Rev. **82**, 403 (1951).
[11]J. Eckstein (private communication).
[12]D. D. Sarma, N. Shanthi, S. R. Barman, N. Hamada, H. Sawada, and K. Terakura, Phys. Rev. Lett. **75**, 1126 (1995).
[13]S. Satpathy, Z. S. Popovic, and F. R. Vukajlovic, Phys. Rev. Lett. **76**, 960 (1996).
[14]W. H. Butler (unpublished).
[15]O. K. Andersen, Phys. Rev. B **12**, 3060 (1975); D. J. Singh, *Planewaves, Pseudopotentials, and the LAPW Method* (Kluwer, Boston, 1994), and references therein.
[16]D. Singh, Phys. Rev. B **43**, 6388 (1991).
[17]A. E. Berkowitz, J. R. Mitchell, M. J. Carey, A. P. Young, S. Zhang, F. E. Spada, F. T. Parker, A. Hutten, and G. Thomas, Phys. Rev. Lett. **68**, 3745 (1992).
[18]J. Q. Xiao, J. S. Jiang, and C. L. Chien, Phys. Rev. Lett. **68**, 3749 (1992).

Spin-resolved x-ray photoemission studies of ferromagnetic metals

L. E. Klebanoff[a)]
Department of Chemistry and Zettlemoyer Center for Surface Studies, Lehigh University, Bethlehem, Pennsylvania 18015

(Received 25 September 1995; accepted 29 January 1996)

Recent spin-resolved x-ray photoelectron spectroscopy (SRXPS) studies of ferromagnetic metals are reviewed. SRXPS studies of metallic Fe, Co, $Co^{66}Fe_4Ni_1B_{14}Si_{15}$, and Ni demonstrate that core-level photoemission, and the itinerant electron response to core-hole creation, are highly spin-dependent. The exchange splitting of the Fe $2p_{3/2}$ level is found to be 0.48 ± 0.05 eV. Lifetime broadening results for the Fe $2p_{3/2}$ $N\uparrow$ (majority spin) and $N\downarrow$ (minority spin) components indicate conservation of spin in core-hole filling processes involving the valence band. SRXPS study of the Fe $2p_{3/2}$ peak asymmetry α reveals a dependence of electron–hole excitation on the spin of the core hole. Spin analysis of the Fe $3s$ XPS line shape shows it to be a three-component spectrum, rather than the two-component line shape assumed previously. A photon energy dependence of one of the Fe $3s$ components explains disagreement among previous Fe $3s$ XPS results. Comparisons of SRXPS from Co metal and $Co_{66}Fe_4Ni_1B_{14}Si_{15}$ directly demonstrate the effect of a reduced atomic magnetic moment on the spin dependence of core-level XPS. The behavior of lifetime broadenings for the $N\uparrow$ and $N\downarrow$ Co $2p_{3/2}$ components show that the reduced Co magnetic moment found in the $Co_{66}Fe_4Ni_1B_{14}Si_{15}$ amorphous glass is due to the transfer of $\uparrow$-spin valence electron density to the $\downarrow$-spin valence band upon glass formation. SRXPS also allows investigation of spin-dependent core-hole screening processes and satellite production, as demonstrated in SRXPS studies of ferromagnetic Ni. Future directions of SRXPS are also explored. © *1996 American Vacuum Society.*

I. INTRODUCTION

Spin-resolved x-ray photoelectron spectroscopy (SRXPS) constitutes a local convergence of two independent fields of research: electron spin polarization studies and XPS investigations of core and valence electronic structure. Pierce and Celotta[1] have reviewed electron spin polarization research, emphasizing spin-polarized electron scattering from surfaces. By the early 1980s, ultraviolet photoelectron spectroscopy (UPS) had been performed with "spin resolution" via the coupling of electron spin detectors to electron energy analyzers.[2,3] Spin-resolved UPS studies probed the occupied spin-dependent valence electronic structure of ferromagnets, providing unprecedented information on the quantum mechanical origin of magnetism. However, the emphasis on spin-resolved UPS tended to concentrate photoemission instrumentation on synchrotron UV beam lines that did not extend to high enough energy to allow spin-resolved core-level photoemission studies.

When the construction of the SRXPS laboratory began at Lehigh University in 1988, the importance of conducting spin-resolved photoemission in the XPS regime was twofold. First, the resolution of electron spin in multiplet-split line shapes would provide fundamentally new information about the nature of the intra-atomic exchange interaction, the role of spin in XPS satellite production, and spin-dependent effects in core-hole screening. For many years, these fundamentally spin-dependent phenomena had been intensively investigated,[4] albeit without the benefit of spin analysis.

Second, SRXPS would allow element-specific magnetic measurement—a capability sorely lacking in almost all probes of magnetism. Magnetic films, alloys, and interfaces composed of two or more magnetic elements (heterogeneous magnetic systems) are not amenable to spin-resolved UPS measurements because of the overlap of valence band features from different elements. In contrast, core-level XPS, affords element specificity. A measurement of core-level spin polarization by SRXPS could, via the core-valence intra-atomic exchange splitting, provide a measurement of the local magnetic moment magnitude, while providing the magnetic moment direction directly from the observed spin polarization direction.

In order for SRXPS to provide even a semiquantitative probe of magnetism, the fundamental physical processes governing the photoelectron spin polarization at the instant of excitation needed to be understood. Additionally, the effects on electron spin of elastic and inelastic scattering during photoelectron transport out of the solid required study. This article reviews recent SRXPS results that focus on these foundational issues. Section II describes the SRXPS apparatus and the method of data acquisition and analysis. Section III presents and discusses recent SRXPS results from ferromagnetic Fe, Co, $Co_{66}Fe_4Ni_1B_{14}Si_{15}$, and Ni. The future of SRXPS will be examined in Sec. IV.

II. SRXPS APPARATUS

A diagram of the SRXPS apparatus is shown in Fig. 1. The sample is located in an UHV chamber constructed of mu metal. The x-ray source is an unmonochromatized single-

a)Electronic mail: lek@lehigh.edu

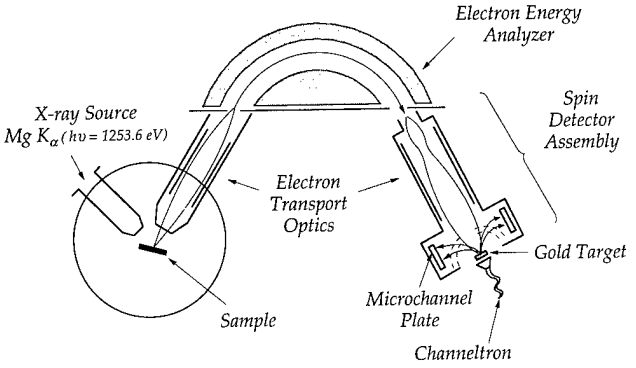


FIG. 1. Schematic diagram of the SRXPS apparatus.

anode Mg $K\alpha$ [$h\nu$= 1253.6 eV, 0.7 eV FWHM (full width at half-maximum) natural linewidth] x-ray tube designed to operate continuously at 800 W power. For SRXPS it is typically operated at 510 W. The overall instrumental energy resolution is 1.6 eV FWHM. The electron energy analyzer is a commercial V.G. ESCALAB MkII electron spectrometer (152 mm mean radius, 150° spherical sector), and delivers energy resolved photoelectrons to the electron spin detector.

The spin polarization P of electrons photoemitted from a ferromagnetic metal can be defined as[5]

$$P=\frac{N\uparrow-N\downarrow}{N\uparrow+N\downarrow}, \tag{1}$$

where $N\uparrow$ and $N\downarrow$ are the number of photoelectrons with magnetic moments oriented parallel and antiparallel, respectively, to the sample magnetization. Due to the negative gyromagnetic ratio of the electron, the electron magnetic moment lies antiparallel to the electron spin angular momentum. Although the polarization P defined in Eq. (1) is really a "magnetic moment polarization," the term "spin polarization" is conventionally used for the quantity P.[5] The polarization P refers to the projection of the "spin" along the sample magnetization direction.

Low-energy diffuse scattering from gold is the basis of the SRXPS instrumentation. The spin detector relies on the spin–orbit interaction ($V_{SO}\propto L\cdot S$) in electron scattering to produce a spin-dependent left–right asymmetry in the intensity scattered off a gold target.[6] The orbital angular momentum L associated with the scattering lies normal to the scattering plane at the gold target, and is the spin analysis axis of the detector. The sample is magnetized in a direction parallel to the spin analysis axis.

The photoelectron spin polarization P can be related to the observed scattering asymmetry A via the relation[6]

$$P=\frac{N_L-N_R}{S(N_L+N_R)}=\frac{A}{S}, \tag{2}$$

where N_L is the number of counts registered in the "left" microchannel plate detector, and N_R is the number of counts registered in the "right" microchannel plate detector (Fig. 1). The factor S is the effective Sherman function. The Sherman function characterizes the sensitivity of the detector to spin polarization and must be determined through calibration. The scattering energy for the diffuse scattering detector is 150 eV.[6] A complete description of the SRXPS spin detector construction and performance can be found in Ref. 7.

The SRXPS data are collected into four channels N_L^+, N_L^-, N_R^+, N_R^-. Here, N_L^+ represents the number of electrons diffusely scattered to the left (L) from the Au target in the spin detector when the sample magnetization is positive (+). N_R^- is the number of electrons scattered to the right (R) from the Au target when the sample magnetization has been reversed to the negative (−) direction. The electron beam polarization P can then be expressed as[5]

$$P=\frac{1}{S}\left(\frac{\sqrt{N_L^+N_R^-}-\sqrt{N_L^-N_R^+}}{\sqrt{N_L^+N_R^-}+\sqrt{N_L^-N_R^+}}\right). \tag{3}$$

SRXPS measurements using both (+) and (−) magnetizations removes from the polarization data apparatus asymmetry effects unrelated to the spin of the electron beam, and also averages over any magnetic dichroism effects. The polarization data can be separated[7] into individual $N\uparrow$ and $N\downarrow$ SRXPS spectra for the majority-spin (↑-spin) and minority-spin (↓-spin) photoelectrons via the following equations: $N\uparrow=2N_{tot}(1+P)$, $N\downarrow=2N_{tot}(1-P)$, where $N_{tot}=(N_L^++N_L^-+N_R^++N_R^-)/4$.

For SRXPS to reveal the role of spin in photoionization, ideally the spectral spin polarization should be unaffected by photoelectron transport[8] through the magnetic solid. Alternatively, a spin-dependent transport process might provide useful geometric structure information.[9] In SRXPS studies of Fe(011),[10] the spin-resolved Fe $2p_{3/2}$ line shape (kinetic energy=547 eV) showed no detectable variation on scanning the emission angle through forward scattering intensity maxima, corresponding to zeroth-order diffraction through nearest-neighbor atoms. In addition, for the Fe/W(011) and Fe/Ag(001) overlayer systems, the nominally unpolarized W $4f$ photoelectrons (kinetic energy=1222 eV) and Ag Auger electrons (kinetic energy=353 eV) acquired no detectable spin polarization as they traversed 30 Å of ferromagnetic Fe.[10] These unpublished results indicate that above 350 eV kinetic energy, ↑-spin and ↓-spin electrons elastically scatter with essentially the same cross section. Therefore, the spin polarizations observed in SRXPS spectra arise from the initial optical excitation, and not from electron scattering during photoelectron transport out of the solid.

III. RESULTS AND DISCUSSION

A. Fe metal

Ferromagnetic iron was the first subject of SRXPS study. Figure 2 shows SRXPS results[11] for the Fe $2p_{3/2}$ level. The lines through the data are simplex fits to each spin component using a single Doniach–Sunjic (DS) line shape[12] convoluted with a Gaussian of 1.6 eV FWHM (the instrumental energy resolution). The Fe $2p_{3/2}$ core-level binding energy, lifetime width, and solid state line shape asymmetry all show a strong spin dependence.

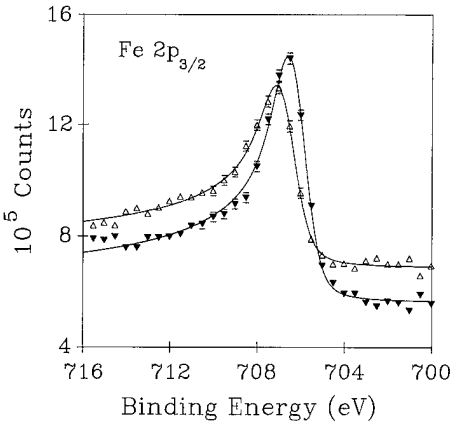


FIG. 2. Separate $N\uparrow$ and $N\downarrow$ SRXPS spectra for the Fe $2p_{3/2}$ majority-spin (△) and minority-spin (▼) photoelectrons, respectively. The lines through the raw data are the result of a simplex fit to each spin component using a single Doniach–Sunjic line shape convoluted with a Gaussian of 1.6 eV.

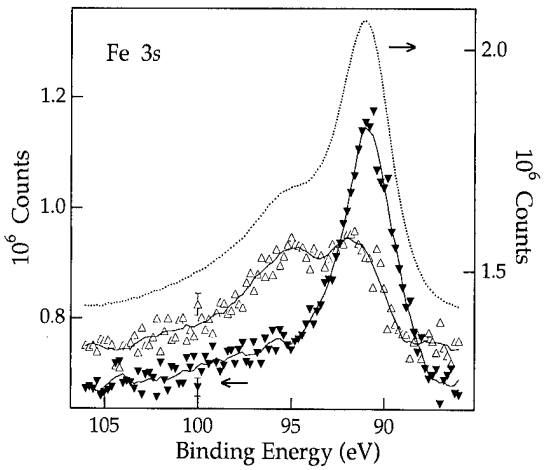


FIG. 3. Separate $N\uparrow$ (△) and $N\downarrow$ (▼) SRXPS spectra for the Fe $3s$ photoelectrons, respectively. Binding energy is plotted against photoelectron counts (left axis). The line through the raw $N\downarrow$ data is a second-order 11-point Savitzky–Golay (SG) smooth. The line through the raw $N\uparrow$ data is a second-order, 15-point SG smooth. The dotted line above the SRXPS data is the spin-summed spectrum ($N\uparrow+N\downarrow$). The spin-summed binding energy is plotted against intensity (right axis).

The exchange splitting between the $N\uparrow$ and $N\downarrow$ Fe $2p_{3/2}$ components is 0.48 ± 0.05 eV. Clearly, intra-atomic exchange is important even for deep core-level photoemission, despite the fact that the $2p-3d$ exchange integral is limited by the small (0.15 Å) radial extent of the $2p$ level. The $N\uparrow$ SRXPS core-level components generally have larger Lorentzian broadenings than the associated $N\downarrow$ components. The Fe $2p_{3/2}$ $N\uparrow$ and $N\downarrow$ lifetime widths, obtained from the DS line shapes in Fig. 2, are 0.56 ± 0.04 eV and 0.44 ± 0.04 eV, respectively. This spin dependence indicates that core-hole filling processes involving the valence band conserve spin, with the larger $N\uparrow$ Fe $2p_{3/2}$ width attributable to the preponderance of ↑-spin electrons in the Fe valence band. There have been no spin-resolved calculations of core-hole lifetimes. Such calculations would be an interesting area for future theoretical research.

The DS line shape singularity index α is consistently larger for the $N\uparrow$ SRXPS components than for the $N\downarrow$ components. For the Fe $2p_{3/2}$ level, the α values are 0.46 ± 0.02 for $N\uparrow$ and 0.42 ± 0.02 for $N\downarrow$. The $N\uparrow$ SXRPS peaks are therefore more asymmetric toward higher binding energy than the $N\downarrow$ peaks. This result is the first indication that the electron–hole ($e-h$) pair excitation spectrum accompanying core-hole creation is intrinsically spin-dependent and is stronger for the majority-spin core hole. A speculative mechanism has been offered for the spin dependence,[11] but the problem needs theoretical attention. Previous calculations[13–15] of the itinerant $e-h$ pair response to core-hole creation treat the core hole as spinless.

The spin dependence of the Fe $3s$ spectrum was first measured by Hillebrecht, Jungblut, and Kisker[16] using 250-eV synchrotron radiation. However, Qiu and co-workers[17] reported that the Fe $3s$ photoemission line shape observed at $h\nu=250$ eV differs significantly from that observed at high photon energy. These workers speculated that the use of a low-photon energy enhances the surface sensitivity of the measurement, with the varying Fe $3s$ line shapes reflecting differences between the Fe surface and bulk magnetic properties.

Figure 3 shows SRXPS data[18] for the Fe $3s$ level. Figure 4 directly compares high-photon energy SRXPS Fe $3s$ results with the 250-eV photon energy spin-resolved data of Hillebrecht, Jungblut, and Kisker.[16] For clarity, only the

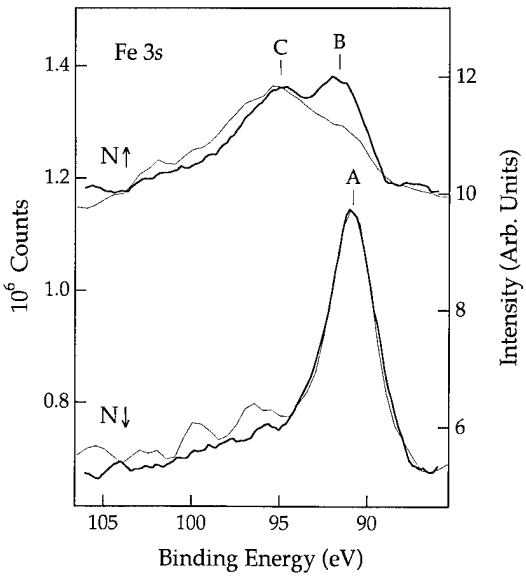


FIG. 4. Thick lines: SG smooth curves for the Fe $3s$ SRXPS data of Fig. 3, obtained at $h\nu=1253.6$ eV. The left axis is applicable to the SRXPS data. Thin lines: SG smooth curves for the Fe $3s$ 250-eV synchrotron data of Ref. 16. The right axis is applicable for the data of Ref. 16. The $N\uparrow$ curves have been rigidly shifted vertically (offset) to clarify the presentation. The right axis was scaled so that the two $N\downarrow$ peaks were of equal height. However, the relative intensities within a data set are accurate (i.e., $N\uparrow/N\downarrow$ intensity ratios for the SRXPS data are accurately represented, as are the $N\uparrow/N\downarrow$ intensities from Ref. 16).

smoothed $N\uparrow$ and $N\downarrow$ curves from Fig. 3 are shown in Fig. 4. It is clear from Fig. 4 that the overall Fe $3s$ line shape *does* vary with photon energy. As the photon energy increases, component B evolves from a weak shoulder at low-photon energy to a distinct and prominent $N\uparrow$ component in the Fe $3s$ SRXPS spectrum. The agreement between Fe $3s$ XPS spectra collected at normal emission and 70° off normal indicate that the differences observed in Fig. 4 are a true photon energy dependence, and cannot be attributed to a varying surface sensitivity.

The increase in the intensity of component B at larger photon energies explains puzzling line shape results from spin-integrated photoemission studies.[17,19] The SRXPS data demonstrate that the "satellite to main peak" intensity ratio, as gleaned from spin-integrated XPS data, is *not* a measure of the $N\uparrow/N\downarrow$ intensity ratio, since $\uparrow$-spin component B lies on the high binding energy side of $\downarrow$-spin component A. The increased prominence of component B at high photon energy causes the "satellite to main peak" intensity ratio to *decrease* with increasing photon energy as indicated in previous two-peak fits to the spin-integrated Fe $3s$ line shape.[17,19] It is probable that experimental variations in the Fe ΔE_{3s} values (from 4.2 to 4.9 eV) inferred[17,19,20] from two-peak fits to the Fe $3s$ spectrum are attributable in part to the photon energy dependence of component B. The correct value of ΔE_{3s} for metallic Fe is 4.5±0.1 eV.

Since component B lies only 0.83 eV above component A, we *do not* believe peak B represents a $3d$ "shakeup" transition in which the Fe $3d$ electron count is decreased. Rather we believe peak B corresponds, qualitatively, to the excited ^1G atomic term of the atomic $4s^23d^6$ valence configuration. If we couple $(3s^1, {}^2S)$ with $(4s^23d^6, {}^1G)$, we form a single ^2G final-state configuration. The theory of Dembczynksi[21] places the $(4s^23d^6, {}^1G)$ term 0.75 eV higher in energy than the $(4s^23d^6, {}^3G)$ configuration, close to the energy separation between the Fe $3s$ components A and B. One would expect[18] this ^2G final state to be accessible only by the photoemission of $\uparrow$-spin $3s$ electrons, consistent with the $\uparrow$-spin polarization of component B. Since peak B represents a higher-energy term within the $4s^23d^6$ valence configuration, its intensity may turn on more slowly with photon energy than components A and C which derive from multiplet splitting from the lowest-energy final-state term.

B. Co metal

Ferromagnetic cobalt metal has a magnetic moment of $1.72\mu_B$ per atom.[22] The amorphous ferromagnetic metallic glass $Co_{66}Fe_4Ni_1B_{14}Si_{15}$ belongs to a class of important technical ferromagnetic materials extensively used in transformer technology.[23] The Co atoms in the magnetic glass have a reduced magnetic moment of $1.0\mu_B$ due to chemical bonding with the metalloid (B,Si) components. Thus a comparison of Co core-level SRXPS spectra from Co metal and the Co-based glass can reveal the role of a varying valence spin polarization on an element's core-level SRXPS spectra.

Figures 5 and 6 display SRXPS spectra[24] for the Co $2p_{3/2}$ level for metallic Co and the Co glass, respectively. The solid line represent single DS fits to the $N\uparrow$ and $N\downarrow$ components over the limited binding energy range 781–774 eV. The $2p_{3/2}$ fit is limited in energy because the low intensity Mg $K\alpha_{3,4}$ x-ray satellites extending up to 10 eV from the Co $2p_{1/2}$ peak at ~793 eV interfere slightly with the spectral intensity from the $2p_{3/2}$ level above 782-eV binding energy. The single-component DS fits are quite satisfactory and for the metal (Fig. 5) yield the following spin-dependent values for the spectral binding energy (BE$\uparrow,\downarrow$), the singularity index ($\alpha\uparrow,\downarrow$), and the FWHM Lorentzian broadening ($\Gamma\uparrow,\downarrow$): $N\uparrow$ component: BE$\uparrow$=777.84±0.030 eV, $\alpha\uparrow$=0.39±0.020, and $\Gamma\uparrow$ =0.73±0.10 eV. For the $N\downarrow$ component we derive the spectral values: BE$\downarrow$=777.66±0.030 eV, $\alpha\downarrow$=0.36±0.02, and $\Gamma\downarrow$ =0.49±0.10 eV. The SRXPS data reveal a Co metal $2p_{3/2}$ spin-dependent (exchange) splitting of 0.18±0.05 eV. The $2p_{3/2}$ SRXPS spectra for the Co-based glass are shown in Fig. 6. The DS fit results can be summarized as follows: For $N\uparrow$: BE$\uparrow$=777.88±0.03 eV, $\alpha\uparrow$=0.35±0.02, and $\Gamma\uparrow$=0.65

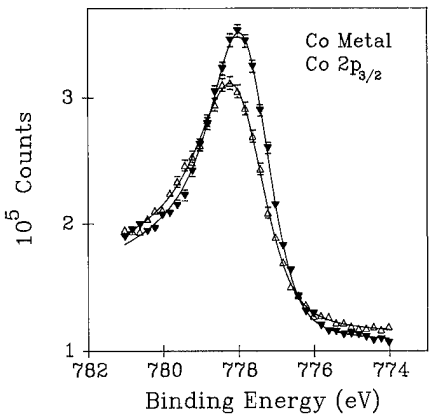


FIG. 5. Separate $N\uparrow$ (△) and $N\downarrow$ (▼) SRXPS spectra for the Co $2p_{3/2}$ level of Co metal. The lines through the data are the result of a simplex fit to each spin component using a single Doniach–Sunjic line shape convoluted with a Gaussian of 1.6 eV FWHM.

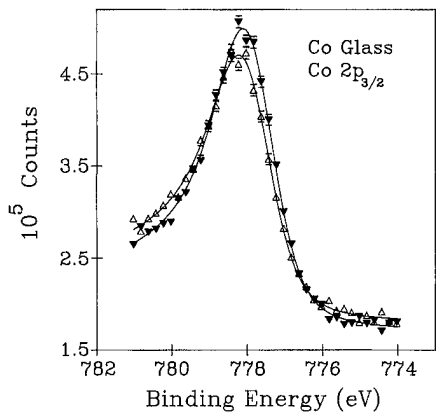


FIG. 6. Separate $N\uparrow$ (△) and $N\downarrow$ (▼) SRXPS spectra for the Co $2p_{3/2}$ level of $Co_{66}Fe_4Ni_1B_{14}Si_{15}$. The lines through the data are the result of a simplex fit to each spin component using a single Doniach–Sunjic line shape convoluted with a Gaussian of 1.6 eV FWHM.

± 0.1 eV. For $N\downarrow$: BE$\downarrow$=777.82$\pm$0.03 eV, $\alpha\downarrow$=0.31$\pm$0.02, and $\Gamma\downarrow$=0.64$\pm$0.1 eV. The reduced exchange splitting for the Co-glass $2p_{3/2}$ level, 0.060$\pm$0.05 eV compared with that of the metal, 0.18$\pm$0.05 eV, clearly demonstrates that as the valence electron spin polarization decreases, the Co $2p_{3/2}$ exchange splitting decreases.

As the Co $3d$ magnetic moment decreases in going from the metal to the glass, the spin dependence of the $2p_{3/2}$ core-hole lifetime Γ diminishes, with the spin-resolved values $\Gamma\uparrow$ and $\Gamma\downarrow$ converging toward the average of $\Gamma\uparrow$ and $\Gamma\downarrow$ for the metal. Since the absolute spin-resolved core filling rates are determined by the $\uparrow$-spin and $\downarrow$-spin populations in the $3d$ valence band, the convergence of $\Gamma\uparrow$ and $\Gamma\downarrow$ indicates that there exists (per Co atom) fewer $\uparrow$-spin $3d$ electrons and more $\downarrow$-spin $3d$ electrons in the glass than found in the metal. This is, of course, consistent with a lower magnetic moment for the Co atoms in the glass, but also suggests that the reduction of the magnetic moment is achieved by the transfer of $\uparrow$-spin $3d$ valence electron density into the $\downarrow$-spin $3d$ valence band, thereby producing fewer $\uparrow$-spin electrons and more $\downarrow$-spin electrons per Co atom. This is direct confirmation that Co $3d$ hybridization with the metalloid valence orbitals weakens the $3d$ electron–electron interaction in the glass,[25] thereby causing spin pairing in the valence band. The data are *inconsistent* with a reduced moment caused by the transfer of charge into the unoccupied minority-spin band from the metalloid components,[26] since this mechanism would not vary the $\uparrow$-spin valence electron population, and therefore would not affect the $\Gamma\uparrow$ value. In this way, SRXPS linewidths can provide valuable element-specific information on mechanisms by which valence electronic structure is modified in alloys.

C. Ni metal

SRXPS is a powerful probe of the itinerant-electron response of the $N-1$ electron system to core-hole creation. This is clearly shown in recent SRXPS studies of ferromagnetic Ni.[27] The core and valence XPS spectra of metallic Ni display a prominent "main" component that is accompanied by satellites located at higher binding energies.[27] The existence of satellites suggests a variable final-state $3d$ electron count, influenced in part by core-hole screening. Theoretically,[28] the strong electron–electron interaction in Ni leads to an increased probability for atomiclike valence-level excitations. Hufner and Wertheim[29] reported the absence of exchange splitting in the main component of Ni $3s$ XPS, leading these workers[29] to attribute the main $3s$ core-level XPS component of Ni to an atomiclike final-state electron configuration of $c3d^{10}$ (c=core hole). Metallic Ni has a ground-state valence electronic configuration of $\sim(4sp)^1 3d^9$.[30,31] However, a $c3d^{10}$ final state could be theoretically generated from the $\sim(4sp)^1 3d^9$ initial state upon the creation of a core hole. This could occur by extra-atomic "d-wave" core-hole screening in which a previously unoccupied $3d$ level were pulled below the Fermi level E_f by the core-hole potential and occupied.[32,33] The main components of core-level XPS spectra of Ni metal have been extensively assigned[28] to the $c3d^{10}$ final-state configuration.

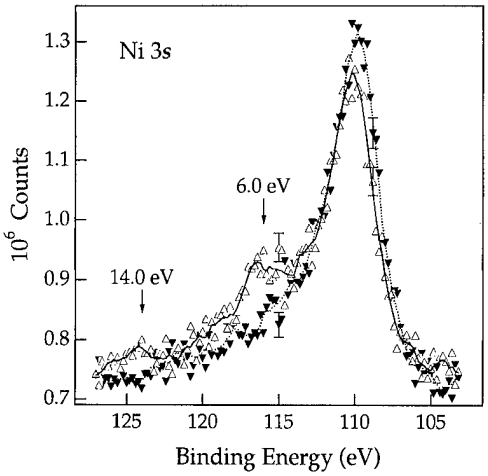


FIG. 7. Separate $N\uparrow$ (△) and $N\downarrow$ (▼) SRXPS spectra for the Ni $3s$ photoelectrons, respectively. The lines through the data are the result of an 11-point second-order SG smooth through the raw $N\uparrow$ and $N\downarrow$ data points. The arrows mark energy positions 6.0 and 14.0 eV from the main $3s$ $N\uparrow$ component. Sample error bars are shown for both the $N\uparrow$ and $N\downarrow$ spectra.

More recently, many-body $3d$ configuration interaction (CI) calculations[34–38] employing the Anderson impurity model have described the Ni ground state as a mixture of $3d^{10}$, $3d^9$, and $3d^8$ valence configurations with relative weights of ~32%, ~53%, and ~15%, respectively. The relative weights can vary when a core-level electron is photoemitted. The final-state configuration associated with the "main" core-level component has not been calculated, but a simple estimate suggests that ~10% $c3d^9$ and ~90% $c3d^{10}$ could be typical for this model.[39]

SRXPS data[27,40] for the Ni $3s$ level are shown in Fig. 7. The main Ni $3s$ $\uparrow$-spin component ($N\uparrow$) is displaced 0.38 $\pm$0.06 eV to higher binding energy and has a smaller intensity than the $\downarrow$-spin $3s$ component ($N\downarrow$). This $3s$ data demonstrate that the final-state Ni valence configuration upon core-hole creation is *not* purely $3d^{10}$, in disagreement with previous assignments,[28,29] since the *nonmagnetic* $3s^1 3d^{10}$ configuration cannot produce the Ni $3s$ main-component exchange splitting evident in Fig. 7.

Although the Ni SRXPS data demonstrate conclusively that the main-line component is not *purely* $c3d^{10}$, the main Ni core-level XPS components could have a mixed $c3d^9$, $c3d^{10}$ character. An empirical argument can be constructed to qualitatively evaluate this possibility. As discussed previously,[27] there is strong evidence that the $3d$ electron count (and therefore the spin polarization) in metallic Fe and Co *does not* change upon the creation of a core hole. Figure 8 plots ΔE_{3s} versus the initial-state spin polarization (in electrons) for Cu, Ni, Co, and Fe metals. The solid line is a linear least squares fit to the Cu, Co, and Fe data, excluding the Ni $3s$ data. Figure 8 can be used *qualitatively* to derive by interpolation the expected metallic Ni $3s$ exchange splittings were the Ni $3d$ spin polarization to remain the same as in the $\sim 3d^9$ initial state (0.56 electrons). For the Ni $3s$ level,

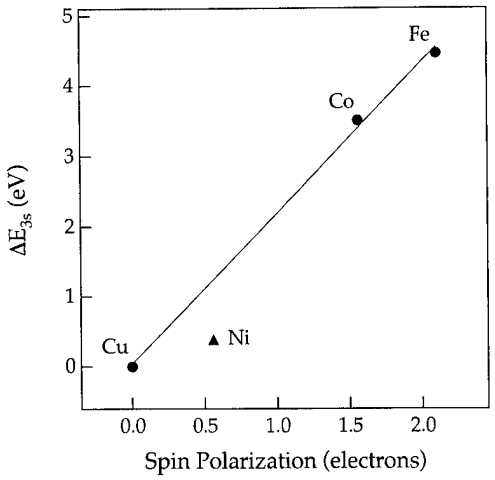


FIG. 8. A plot of the $3s$ multiplet splitting ΔE_{3s} for Cu, Ni, Co, and Fe metals vs the initial-state spin polarization in electrons. The line is the result of a linear least-squares fit to the Cu, Co, and Fe data, excluding the Ni data. The Fe and Co metal ΔE_{3s} values are taken, respectively, from Ref. 16 and from D. G. Van Campen *et al.*, Phys. Rev. B **49**, 2040 (1994).

this interpolated splitting is 1.18 eV. The measured Ni $3s$ exchange splitting (0.38 eV) is ~30% of the expected value. If this decrease in core-valence exchange is *completely* attributed to a reduced valence spin polarization in the photoemission final state, then a final-state $3d$ spin polarization of 0.17 electrons is inferred, corresponding to a 30% $c3d^9$ and 70% $c3d^{10}$ final-state mixture. This estimate is probably an upper limit to the nonmagnetic $c3d^{10}$ contribution, since the Ni $3p$ peak has a surprisingly large spin-dependent splitting and the magnitudes of the SRXPS spin polarizations for Ni are [when compared to Co (Ref. 24) and Fe (Ref. 11)] more consistent with a Ni final-state spin characteristic of the $(4sp)^1 3d^9$ initial state. Therefore, we believe that the true final-state Ni $3d$ valence composition following core electron ionization is one of reduced (but not zero) magnetization, and is most likely ~50% $c3d^9$ and ~50% $3d^{10}$. This final-state composition estimate is consistent with an itinerant-electron screening model proposed in Ref. 40.

The spin polarization of the Ni main components may, alternatively, have an explanation in localized models employing Anderson impurity theory.[34–38] In these models, a main-component spin dependence could arise from the admixture of $3d^9$ and $3d^{10}$ "atomiclike" valence configurations in the final state associated with the Ni main component. The composition of this admixture in the final state for the "main" Ni core-level component has not been calculated, but a simple estimate suggests that it is primarily (~90%) $3d^{10}$ in character, with a "small" (~10%) $3d^9$ admixture.[39] The itinerant-electron screening model[40] suggests that the $c3d^9$ configuration survives at the ~50% level. Since the two models differ quantitatively, but not qualitatively, in their predictions for the main components, the main-component SRXPS data *alone* do not indicate a decisive preference for either model. The main-component and satellite SRXPS data are, as a whole, more consistent with the itinerant-electron screening model.[40]

The "6-eV satellite" and a secondary satellite located ~14 eV from the main $3s$ peak (referred to hereafter as the "14-eV satellite"), both display a very large (nearly 100%) majority-spin polarization. The large ↑-spin polarizations for the Ni $3s$ 6- and 14-eV satellites, combined with the prominence of these satellites in Ni and the relative weakness of satellites in Cu (a known $3d^{10}$ system), indicate that a $3d$ shakeup *from* a $3d^{10}$ configuration *cannot* generate strong and discrete satellite structure. Thus, we believe that the $c3d^9$ final-state component in the main line, which results from extra-atomic screening via the $3d$–$4sp$ states above E_f (Fig. 4 of Ref. 40), is the starting point for the 6- and 14-eV satellites in Ni.

A central issue that distinguishes the core-level satellite problem from the valence satellite problem is the extent to which the $3d$ valence electron participating in the "shakeup" is coupled to the remaining and highly localized core-level electrons after core ionization. The SRXPS data suggest that upon core-level ionization, the excitation of a Ni valence $3d$ electron (initially in a ~$3d^9$ configuration) to an unoccupied state above E_f *does* reduce the $3d$ valence charge within the radial extent of the remaining core electrons. From the "point of view" of the remaining electrons in the ionized core level, the $3d$ shakeup reduces the "local" $3d$ valence charge from ~$3d^9$ to ~$3d^8$. This mechanism differs conceptually from valence $3d$ ionization (which literally produces a $3d^8$ final state) for which the $3d$ shakeup is thought *not* to change the total $3d$ count in the final state,[28] and the 6-eV valence satellite is interpreted with the $3d^8$ final-state valence configuration.

In the core-level scenario above, the 6-eV satellite appearing in Ni $3s$ SRXPS can be understood (using atomic terminology only qualitatively appropriate for the metallic problem) as derived from the dominant ^1G term[41–43] of the Ni $3d^8$ final-state valence configuration. For $3s$ photoemission, the 6-eV satellite would then be associated with the ^2G final-state term resulting from the coupling of $3s^1(^2S)$ with $3d^8(^1G)$. In cases where LS coupling is valid (i.e., spin–orbit interactions are zero or negligible), the spin component along the magnetization direction S_z is a conserved quantity. Conservation of this spin component[41,44] suggests that

$$\underset{(3d^9)}{M_s^{\text{in}}} = \underset{(3d^8)}{M_s^{\text{fin}}} + \underset{(3d^1)}{M_s^{\text{su}}} + \underset{(3s^1)}{M_s^{\text{pe}}}. \tag{4}$$

Here, M_s^{in} is the initial-state spin component along the magnetization direction (i.e., the eigenvalue for S_z in the initial state), which in the $3d^9$ atomic case would be $+\frac{1}{2}$ in units of $\hbar$. M_s^{fin} is the spin component of the final ionic state ($3d^8$) excluding the shakeup electron; M_s^{su} is the spin component of the $3d$ electron ($3d^1$) involved in the shakeup (su) and M_s^{pe} is the spin component of the detected $3s$ photoelectron (pe). If we assume that the $3d$ electron participating in the shakeup is a ↓-spin electron ($M_s^{\text{su}} = -\frac{1}{2}$), then the ^2G term (with $M_s^{\text{fin}} = +\frac{1}{2}$) can only be produced by ↑-spin $3s$ photoelectrons ($M_s^{\text{pe}} = +\frac{1}{2}$), i.e., $+\frac{1}{2} = +\frac{1}{2} - \frac{1}{2} + \frac{1}{2}$ from Eq. (4).

These considerations explain the nearly 100% ↑-spin polarization observed for the Ni $3s$ 6-eV satellite in Fig. 7.

We believe the secondary satellite at 14 eV is qualitatively derived from the ^1S term of the $3d^8$ atomic multiplet.[42] The Ni $3s$ 14-eV satellite would then be associated with the ^2S final-state term arising from the coupling of $3s^1(^2S)$ with $3d^8(^1S)$. Given the assumptions above, the ^2S peak (with $M_s^{\text{fin}} = +\frac{1}{2}$) is also allowed only by the photoemission of ↑-spin ($M_s^{\text{pe}} = +\frac{1}{2}$) Ni $3s$ electrons, as suggested by Eq. (4) and observed in Fig. 7. The ^2S assignment for the 14-eV satellite explains its large ↑-spin polarization and the higher binding energy for the ^2S state relative to the ^2G term. However, its displacement of ~8 eV beyond the 6-eV satellite is rather large. Calculations for atomic Ni by Combet-Farnoux and Ben Amar[45] predict a $^1S-^1G$ ($3d^8$) term separation of 4.7 eV. Theoretical results by Liebsch[42] for metallic Ni place the ^1S peak only ~2 eV beyond the 6-eV satellite. The gas phase measurements of Schmidt *et al.*[46] suggest that the ^1S peak lies 2.3 eV to higher binding energy than the ^1G state in atomic Ni. Thus the large $^2S-^2G$ separation is not adequately understood.

Calculations[34] using Anderson impurity theory have considered the Ni $3s$ 6-eV satellite to consist of ^3D and ^1D terms of the $3s^13d^9$ final state configuration. The essentially 100% ↑-spin polarization observed for the Ni $3s$ 6-eV satellite is inconsistent with a $3s^13d^9$ final-state configuration assignment for the satellite, since ^1D and ^3D terms would lead to both ↑-spin and ↓-spin components in the Ni $3s$ 6-eV satellite, which are not observed.

IV. FUTURE OF SRXPS

SRXPS has already provided valuable information on the spin-dependent nature of core-level and valence photoemission, as well as on the magnetism of bulk ferromagnets. To fully exploit the information content of SRXPS, comparison must be made with advanced theory of the SRXPS process. The paper presented by Dr. Jose Menchero at this conference will report remarkable calculations of Ni SRXPS that agree almost quantitatively with the Ni SRXPS experimental results. Similar theoretical effort extended to other ferromagnetic metals and to magnetic interfaces are needed and will greatly increase the utility of SRXPS measurements. Calculations of core-level exchange splittings, lifetimes, and asymmetries will be particularly helpful.

To date, these high photon energy SRXPS measurements have been conducted with a high power Mg $K\alpha$ x-ray anode. This apparatus remains the only SRXPS facility that allows spin analysis of deep core-level electrons while maintaining a reasonably high photoelectron kinetic energy and a decent energy resolution (1.6 eV FWHM). SRXPS studies will greatly benefit from the third-generation synchrotron sources being commissioned that will allow high flux and high-energy resolution SRXPS in the photon energy range 1200–1800 eV.

SRXPS will be very profitably extended to investigations of heterogeneous magnetic interfaces involving one monolayer of a magnetic metal deposited on a magnetic substrate. The element specificity of SRXPS will allow the monolayer properties to be probed selectively, and the magnitude and direction of the overlayer magnetic moment to be quantitatively determined. Such SRXPS studies, when combined with determinations of interfacial geometric structure, will greatly clarify questions of magnetic moment enhancement and creation for monolayer magnetic systems.

The high kinetic energies of SRXPS spectra afford two important advantages: (1) the spin polarization is unaffected by photoelectron transport through the solid, and (2) the inelastic mean free path[47] of the photoelectrons is 10–20 Å. These attributes allow SRXPS to probe magnetic layers "buried" beneath up to 30–40 Å of another material. The effective depth sensitivity of SRXPS can be varied by changing the angle of electron detection with respect to the surface plane, providing SRXPS with a magnetic depth profiling capability. These attributes should allow SRXPS to probe internal interfaces in multilayer magnetic films of technological importance.

ACKNOWLEDGMENTS

With great pleasure the author thanks his former graduate students Douglas G. Van Campen and Roland J. Pouliot, as well as his former post-doc Alex See, for their hard work, perseverance, and fellowship in the establishment of the SRXPS laboratory and conducting the SRXPS measurements reviewed here. It is also with gratitude that the author thanks the funding agencies who have supported this work, including the National Science Foundation under Grant No. CHE-9117138, The Petroleum Research Fund, and The Camille and Henry Dreyfus Foundation.

[1]D. T. Pierce and R. J. Celotta, Adv. Electron. Electron. Phys. **56**, 219 (1981).
[2]E. Kisker, K. Schroder, M. Campagna, and W. Gudat, Phys. Rev. Lett. **52**, 2285 (1984).
[3]W. Schmitt, H. Hopster, and G. Guntherodt, Phys. Rev. B **31**, 40 351 (1985).
[4]For example, D. G. Van Campen, M. L. Knieriem, and L. E. Klebanoff, Phys. Rev. B **43**, 11 668 (1991).
[5]J. Kessler, in *Polarized Electrons* (Springer, Berlin, 1985), pp. 16, 197, and 234.
[6]M. R. Scheinfein, D. T. Pierce, J. Unguris, J. J. McClelland, R. J. Celotta, and M. H. Kelly, Rev. Sci. Instrum. **60**, 1 (1989).
[7]L. E. Klebanoff, D. G. Van Campen, and R. J. Pouliot, Rev. Sci. Instrum. **64**, 2863 (1993).
[8]J. A. D. Matthew, Phys. Rev. B **25**, 3326 (1982).
[9]E. M. E. Timmermans, G. T. Trammell, and J. P. Hannon, Phys. Rev. Lett. **72**, 832 (1994).
[10]A. K. See and L. E. Klebanoff (unpublished).
[11]D. G. Van Campen, R. J. Pouliot, and L. E. Klebanoff, Phys. Rev. B **48**, 17533 (1993).
[12]S. Doniach and M. Sunjic, J. Phys. C: Solid State Phys. **3**, 285 (1970).
[13]P. Nozieres and C. T. DeDominicis, Phys. Rev. **178**, 1084 (1969).
[14]G. D. Mahan, Phys. Rev. B **11**, 4814 (1975).
[15]R. P. Gupta and S. K. Sen, Phys. Rev. B **10**, 71 (1974).
[16]F. U. Hillebrecht, R. Jungblut, and E. Kisker, Phys. Rev. Lett. **65**, 2450 (1990).
[17]S. L. Qiu, R. G. Jordan, A. M. Begley, X. Wang, Y. Liu, and M. W. Ruckman, Phys. Rev. B **46**, 13004 (1992).
[18]A. K. See and L. E. Klebanoff, Phys. Rev. B **51**, 7901 (1995).
[19]J. F. Van Acker, Z. M. Stadnik, J. C. Fuggle, H. J. W. M. Hoekstra, K. H. J. Bushow, and G. Stroink, Phys. Rev. B **37**, 6827 (1989).

[20]D. J. Joyner, O. Johnson, and D. M. Hercules, J. Phys. F: Metal Phys. **10**, 169 (1980).
[21]J. Dembczynski, Physica C **100**, 105 (1980).
[22]C. Kittel, in *Solid State Physics* (Wiley, New York, 1976), p. 465.
[23]D. Raskin and C. H. Smith, in *Amorphous Metallic Alloys*, edited by F. E. Luborsky (Buttersworth, London, 1980), p. 381.
[24]L. E. Klebanoff, D. G. Van Campen, and R. J. Pouliot, Phys. Rev. B **49**, 2047 (1994).
[25]For a discussion, see R. C. O'Handley, in *Amorphous Metallic Alloys*, edited by F. E. Luborsky (Buttersworth, London, 1983), p. 381.
[26]K. Yamauchi and T. Mizoguchi, J. Phys. Soc. Jpn. **39**, 541 (1975).
[27]A. K. See and L. E. Klebanoff, Phys. Rev. B **51**, 11002 (1995).
[28]For an excellent review, see L. C. Davis, J. Appl. Phys. **59**, R25 (1986).
[29]S. Hufner and G. K. Wertheim, Phys. Lett. A **51**, 301 (1975).
[30]Results from unpublished electronic structure calculations of R. H. Victora (private communication).
[31]T. H. Upton and W. A. Goddard III, Phys. Rev. Lett. **42**, 472 (1979).
[32]L. Ley, S. P. Kowalczyk, F. R. McFeely, R. A. Pollak, and D. A. Shirley, Phys. Rev. B **8**, 2392 (1973).
[33]A. Kotani and Y. Toyozawa, J. Phys. Soc. Jpn. **37**, 912 (1974).
[34]G. van der Laan, M. Surman, M. A. Hoyland, C. F. J. Flipse, B. T. Thole, Y. Seino, H. Ogasawara, and A. Kotani, Phys. Rev. B **46**, 9336 (1992).
[35]B. T. Thole and G. van der Laan, Phys. Rev. Lett. **67**, 3306 (1991).
[36]G. van der Laan, M. A. Hoyland, M. Surman, C. F. J. Flipse, and B. T. Thole, Phys. Rev. Lett. **69**, 3827 (1992).
[37]G. van der Laan and B. T. Thole, J. Phys. Condensed Matter **4**, 4181 (1992).
[38]A. Tanaka, T. Jo, and G. A. Sawatsky, J. Phys. Soc. Jpn. **61**, 2636 (1992).
[39]J. W. Allen (private communication).
[40]A. K. See and L. E. Klebanoff, Phys. Rev. Lett. **74**, 1454 (1995).
[41]L. A. Feldkamp and L. C. Davis, Phys. Rev. Lett. **43**, 151 (1979).
[42]A. Liebsch, Phys. Rev. Lett. **43**, 1431 (1979).
[43]E. Antonides, E. C. Janse, and G. A. Sawatsky, Phys. Rev. B **15**, 1669 (1977).
[44]S. F. Alvarado and P. S. Bagus, Phys. Lett. A **67**, 397 (1978).
[45]F. Combet-Farnoux and M. Ben Amar, Phys. Rev. A **21**, 1975 (1980).
[46]E. Schmidt, H. Schroder, B. Sonntag, H. Voss, and H. E. Wetzel, J. Phys. B: At. Mol. Phys. **16**, 2961 (1983).
[47]S. Tanuma, C. J. Powell, and D. R. Penn, Surf. Interface Anal. **17**, 911 (1991).

Magnetic circular dichroism in reflection electron energy loss spectroscopy?

G. R. Harp[a)]
Department of Physics and Astronomy, Ohio University, Athens, Ohio 45701

R. F. C. Farrow[b)] and R. F. Marks[c)]
IBM Research Division, Almaden Research Center, San Jose, California 95120-6099

(Received 26 September 1995; accepted 19 February 1996)

We evaluate the possibility of using dichroic electron energy loss spectroscopy (DEELS) as an alternative to x-ray magnetic circular dichroism (XMCD). It is well known that electron energy loss spectroscopy and x-ray absorption spectroscopy provide similar information. A simple semiclassical model suggests that reflection DEELS might have a magnetic sensitivity similar to that of XMCD. This sensitivity will be reduced, however, by multiple scattering of the probe electron before and after the energy loss event. Thus, it is difficult to predict the magnitude of the DEELS effect. Experiments were performed at the L edges of polycrystalline Fe, Co, and Ni thin-film samples prepared *in situ* with a uniaxial magnetic bias. Even in these most favorable cases, the DEELS effect is limited to less than one-tenth of related effects in XMCD. © *1996 American Vacuum Society.*

Electron energy loss spectroscopy (EELS) is in many ways highly analogous to x-ray absorption spectroscopy (XAS), as is well known.[1,2] For example, near-edge reflection EELS spectra have been shown to provide the same chemical information (projected densities of states) that can be obtained from x-ray absorption near-edge spectroscopy.[3] Similarly, the well-known structural analysis technique of extended x-ray absorption fine structure spectroscopy has its EELS analog: extended energy loss fine structure.[4] In this article, we explore yet another reflection EELS variant. This is analogous to x-ray magnetic circular dichroism (XMCD), which we call dichroic electron energy loss spectroscopy (DEELS).

To understand how EELS techniques are similar to XAS techniques, consider that XAS measures the probability that a photon will be absorbed by a given material as a function of the photon energy. In comparison, EELS measures the probability that a relatively high-energy electron will lose a quantum of energy when it interacts with a material, as a function of the loss energy. Put another way, photons in XAS induce transitions of core level electrons to energies above the Fermi level. In EELS, high-energy electrons induce similar transitions, and in so doing, lose energy. But recall that electron–electron interactions are mediated by virtual photons in the Coulomb interaction. Thus EELS may be considered a photoabsorption process, where the photon is *created* very close to the absorbing atom. This is at the heart of the XAS/EELS analogy.

There are two main characteristics which distinguish EELS techniques from XAS techniques: the excitation source and the selection rules for energy absorption. In XAS, the excitation source must be an energy-tunable photon source: usually a synchrotron radiation source. By comparison, EELS requires only a simple laboratory electron gun source, which is relatively inexpensive and widely available. The other distinguishing characteristics, the selection rules, determine which transitions are observed. In XAS, dipole selection rules apply to a very good approximation, and current techniques permit complete control of the photon polarization. In EELS, however, the selection rules are not dipole limited, and the selection rules and virtual photon polarization depend on the electron scattering geometry.[3] This latter aspect can make the interpretation of EELS more complicated than XAS. Nevertheless, by manipulation of the EELS scattering geometry, it is sometimes possible to simulate dipole selection rules to a very good approximation,[5] and, as we explain below, also to control the polarization of the virtual photons.

In recent years, a powerful new magnetic and element-specific XAS technique known as XMCD has been developed[6] and has been highly successful in the study of magnetic materials. Here, the absorption coefficient of circularly polarized x rays shows large changes depending on the relative orientations of photon helicity and the local magnetic moment of a particular atomic species in a material. XMCD thus provides information about those local magnetic moments. In the present article, we propose that similar effects should be observable in EELS: i.e., that for certain EELS scattering geometries, there is a net circular polarization of the virtual photons that are exchanged between electrons. Thus the EELS scattering cross section should depend on the relative orientation of local atomic moments relative to this circular polarization.

A semiclassical model elucidates this process. Consider a classical point-charge electron with energy E_0 incident on an Fe atom with impact parameter b. The electron interacts with the screened Coulomb potential of the atom nucleus and undergoes scattering, as shown in Fig. 1. We are interested only in electrons which undergo 90° scattering. We wish to esti-

[a)]Electronic mail: harp@helios.phy.ohiou.edu
[b)]Electronic mail: farrow@almaden.ibm.com
[c)]Electronic mail: marks@almaden.ibm.com

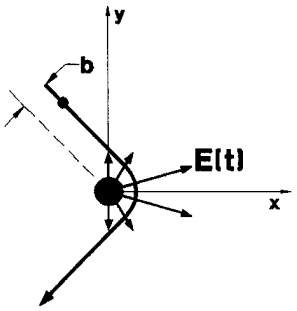


FIG. 1. Scattering geometry for the semiclassical model. An electron is incident upon an atom at an angle of 135° with respect to the x axis and scatters into an angle of 225°. The magnitude and direction of the electric field at the atom vary in both magnitude and direction as a function of time.

mate the virtual photon amplitude caused by this incident electron near the atom core.

At this point, it is already possible to see how this scattering geometry gives rise to circular polarization. At the time when the electron makes its closest approach to the atom ($t\equiv 0$), the electron is moving perpendicular to its radius from the atom center, along a path well approximated by a circular arc. At this moment, the electric field at the atom position is constant in magnitude but is rotating with some angular velocity, ω_c. This is very similar to the situation when an Fe atom is immersed in a circularly polarized photon field with energy $\hbar\omega_c$.

It was shown by Slater[7] that the Fe $2p$ electrons move in a potential which is reasonably approximated as Z'/r, where $Z'=Z-3.15=22.85$ since for the $2p$ electrons the nuclear potential, Z/r, is somewhat screened by the $1s$, $2s$, and other $2p$ electrons. Because we are interested in transitions of $2p$ electrons, we assume in our model that the incident electron moves in the same potential. In any case, the exact form of the scattering potential does not qualitatively change the arguments presented here.

For a particle moving in a $1/r$ potential, the scattering path of the particle can be integrated exactly,[8] provided it scatters elastically. To make the problem more tractable, therefore, we assume that the scattering path of an electron with energy E_0 and which loses energy ΔE can be approximated by the elastic scattering path of an electron with energy E^*, where $E^*=(E_0+E_0-\Delta E)/2$. This approximation is most valid when $\Delta E \ll E_0$.

Now, the core electron experiences a time-dependent electric field $\mathbf{E}(t)$ due to the incident electron with

$$\mathbf{E}(t)=\frac{e}{R(t)}\,\hat{\mathbf{R}}(t), \tag{1}$$

where $\mathbf{R}=[\mathbf{r}_{\rm mc}(t)-\mathbf{r}_{\rm core}]$, and $\mathbf{r}_{\rm inc}(t)$ and $\mathbf{r}_{\rm core}$ are the position vectors of the incident and core electrons, respectively. The core electron is assumed to be at rest at the atom center.

To obtain the "virtual photon amplitude," $\mathscr{A}(\hbar\omega)$, we Fourier transform $\mathbf{E}(t)$:

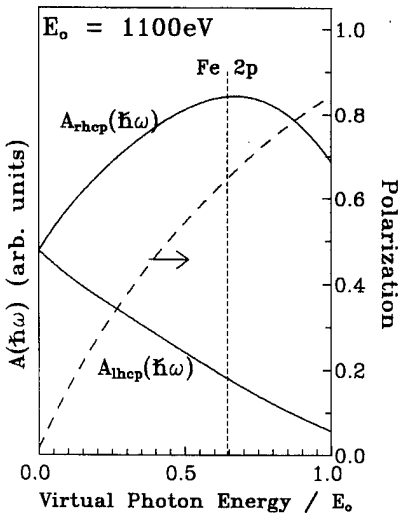


FIG. 2. Plot of the "virtual photon amplitudes" of right- and left-handed circular polarization (solid lines), and the net circular polarization (dashed line) of the electric field at the atom. For an incident electron kinetic energy of 1100 eV, the position of the Fe $2p$ core energy loss feature is marked.

$$\mathscr{A}_x(\hbar\omega)=\int_{-\infty}^{\infty} dt\,\cos(\omega t)\mathbf{E}(t),$$
$$\mathscr{A}_y(\hbar\omega)=\int_{-\infty}^{\infty} dt\,\sin(\omega t)\mathbf{E}(t). \tag{2}$$

From these x and y components, which have a well-defined phase relationship, we construct the right- and left-handed circularly polarized virtual-photon amplitudes.

The results of this simple calculation are shown in Fig. 2 for $E_0=1100$ eV. The abcissa is chosen to be in units of E_0 since we find that the curves are almost independent of E_0 (for simpler geometries this is rigorously true). It is seen that low-energy virtual photons (i.e., $\Delta E \ll E_0$) have almost equal right- and left-handed amplitude (linear polarization). This comes from the fact that in the low-wavelength limit, $\mathscr{A}_x\sim$constant, while $\mathscr{A}_y\sim 0$. The circular polarization, $P=(\mathscr{A}_r-\mathscr{A}_l)/(\mathscr{A}_r+\mathscr{A}_l)$, increases monotonically with increasing virtual photon energy.

For a 710 eV ($0.65\,E_0$) energy loss, sufficient to excite an Fe $2p$ core electron, the effective photon polarization is >60%. Although for such large energy losses our approximation $\Delta E \ll E_0$ breaks down, we believe that Fig. 2 at least qualitatively describes the polarization behavior for such scattering processes. It is known that with 100% circularly polarized light, the XMCD signal at the $2p$ core energy of $3d$ transition metals can vary by as much as 100%.[9,10] Based on this simple semiclassical model, one might hope to see magnetization-dependent changes of an EELS cross section as large as 60%. Such changes should be easily measurable using standard techniques.

One obvious flaw with this model is that it neglects elastic scattering from all the other atoms which surround the atom of interest. In a reflection EELS experiment at these kinetic energies, it is well known[11] that the incident electron undergoes multiple elastic scattering before and after the inelastic EELS scattering event. What this means is that the incident

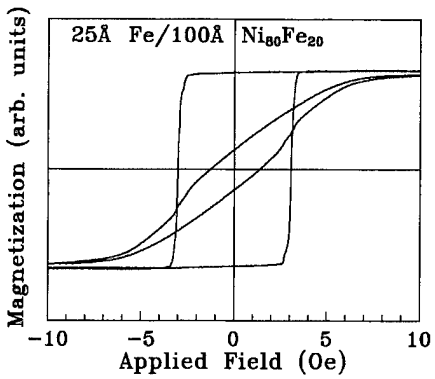


FIG. 3. Magnetization loops of a Si/ZrO_2 1000 Å/$Ni_{80}Fe_{20}$ 100 Å/Fe 25 Å/Au 5 Å sample taken along the easy and hard axes of its uniaxial bias.

and outgoing directions of the incident electron are no longer well defined by the position of the electron gun and detector. Clearly, multiple scattering will reduce the magnitude of the dichroic effects, "blurring" them out as a function of scattering angle. Such blurring may not be complete, however, and a residual dichroic effect probably remains. *A priori*, it is difficult to predict the magnitude of such blurring effects. While it is possible, in principle, to make such predictions using state of the art multiple-scattering calculations, it is in fact easier to simply attempt the experiment and measure the effect.

To test these ideas, experiments were performed on thin films of Fe, Co, and Ni. All the experiments were performed at the IBM Almaden Research Center in San Jose, CA. In order to see dichroic effects, it is necessary to ensure that the films are uniformly magnetized. This was accomplished by depositing them on magnetically biased $Ni_{80}Fe_{20}$ films, 100 Å thick, deposited on polycrystalline 1000 Å ZrO_2 films on Si wafers. The $Ni_{80}Fe_{20}$ films were biased by growing them in a magnetic field using a small permanent magnet mounted directly behind the sample. Such polycrystalline films have a uniaxial magnetic anisotropy field of about 10 Oe, which ensures single-domain magnetization. On top of these $Ni_{80}Fe_{20}$ films, elemental films of ≈25 Å Fe, Co, and Ni were deposited. Preliminary experiments showed that over the course of the EELS experiments, Ni films became contaminated with C and O due to electron-beam decomposition of residual chamber gases. For this reason, subsequent depositions of Fe and Co films were capped with ≈5 Å Au and Ag films, respectively, before transfer to the analysis chamber. Then the backing magnet was removed, and the samples were transferred, *in vacuo*, into an EELS analysis chamber with a base pressure of 2×10^{-10} Torr.

Figure 3 shows typical magnetization loops from a sample with structure Si/ZrO_2 1000 Å/$Ni_{80}Fe_{20}$ 100 Å/Fe 25 Å/Au 5 Å, after removal from the ultrahigh vacuum system. These loops were taken along the easy and hard directions of the film's uniaxial bias. The easy-axis loop shows almost 100% remanence, with a coercive field of ≈3 Oe. This coercivity is larger than any expected stray magnetic fields in our apparatus. Thus, the samples were effectively 100% magnetized during the DEELS measurements.

We present data from two experiments involving Fe and Co films, taken with incident electron energies of ≈1110 and ≈1210 eV, respectively. These incident kinetic energies were chosen to (a) maximize the right handed polarization amplitude (following Fig. 2) and (b) eliminate any interference from nearby Auger features. The EELS intensity was measured using a Vacuum Generator's hemispherical electron energy analyzer with an angular acceptance of a 5° half-angle cone, 1 eV energy resolution, and using standard lock-in techniques. Measurements were made in both first and second derivative modes, and we present second derivative data since in this case peak positions are more easily compared with XAS spectra. Comparing derivative spectra should preserve the magnitude of observable dichroism effects.

To see dichroism, it is necessary to vary the relative directions of circular polarization and magnetization. This was accomplished through a rotation of the sample before the electron energy analyzer. Electrons are incident on the sample at about 20° with respect to the surface plane and with an in-plane projected direction perpendicular to the sample magnetization **M**. Electrons are measured at normal emission. This configuration puts the scattering plane perpendicular to **M**, and thus the net circular polarization (helicity) is parallel to **M**. After a spectrum is taken at this position (0°), the sample is rotated 180° about the surface normal (under computer control), and a second spectrum is taken with the helicity antiparallel to **M**. These two orientations should show the greatest changes due to magnetic dichroism. In addition, spectra were taken in geometries with **M** parallel to the scattering plane (90° and 270°). The latter two spectra should show no circular dichroism effects.

Figure 4 shows the results of our experiments. We plot only the spectra measured at 0° and 180°. It is seen that no significant differences are observed between the spectra for either the Fe or the Co film. Furthermore, these spectra are not significantly different from the spectra obtained along either the 90° or 270° directions. In addition to these measurements, we have obtained similar results for incident electron kinetic energies up to 2.5 keV on various Fe, Co, and Ni films, and also at the $3p$ absorption edge of Ni. The results from the Ni films were similar, but of somewhat lower quality due to the sample contamination mentioned above. Based on these data, we estimate that while magnetic dichroism effects probably do exist in the EELS of these materials, it must amount to less than 5% of the total absorption coefficient.[12]

This is the main result of the present article. A semiclassical calculation suggests the possibility that EELS can provide a convenient chemically sensitive probe of local atomic magnetization. However, the absolute magnitude of these effects is difficult to predict. We have shown empirically that these effects are small (<5%) even for the most advantageous cases (i.e., Fe, Co). We believe it is likely that some dichroic effects *do* exist for these materials, and more sensitive measurements of EELS would likely prove their exist-

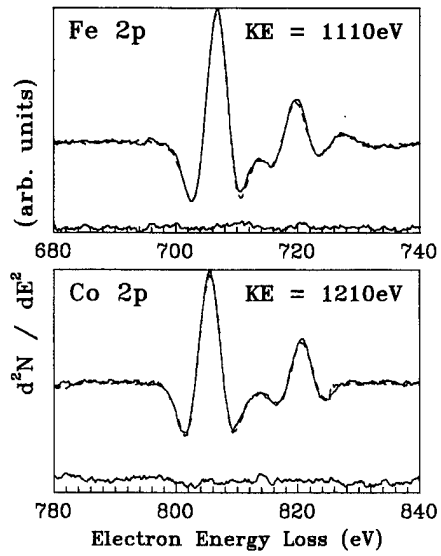


FIG. 4. Electron energy loss spectra from Fe and Co films taken with the magnetization parallel (solid) or antiparallel (dashed) to the net helicity of the electric field for this scattering geometry. Difference spectra are plotted below, in each case. It is seen that any dichroism effects are smaller than 5% of the total EELS cross section for these films.

ence. However, because these effects are small, and because of the complications caused by electron multiple scattering, we feel it is unlikely that the DEELS technique can effectively compete with XMCD as a local probe of atomic moments.

Acknowledgments: The authors gratefully acknowledge A. Cebollada for technical assistance during the initial phases of this work. One of the authors (G. R. H.) wishes to thank the IBM Almaden Research Center for its hospitality and support during his visit for the completion of this work.

[1]M. Inokuti, Rev. Mod. Phys. **43**, 297 (1971).
[2]W. L. Schiach, Phys. Rev. B **29**, 6513 (1984).
[3]A. P. Hitchcock and T. Tyliszczak, Surf. Rev. Lett. (in press).
[4]M. De Crescenzi, Surf. Sci. **162**, 838 (1985).
[5]D. K. Saldin and J. M. Yao, Phys. Rev. B **41**, 52 (1990); D. K. Saldin and Y. Ueda, *ibid.* **46**, 5100 (1992); Y. Ueda and D. K. Saldin, *ibid.* **46**, 13 697 (1992).
[6]G. Schütz, W. Wagner, W. Wilhelm, P. Keinle, R. Zeller, R. Frahm, and G. Materlik, Phys. Rev. Lett. **58**, 737 (1987).
[7]J. C. Slater, Phys. Rev. **36**, 57 (1930).
[8]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1980), p. 96.
[9]C. T. Chen, F. Cette, Y. Ma, and S. Modesti, Phys. Rev. B **42**, 7262 (1990).
[10]G. van der Laan and B. T. Thole, Phys. Rev. B **43**, 13 401 (1991).
[11]D. K. Saldin, Phys. Rev. Lett. **60**, 1197 (1988).
[12]As mentioned above the incident electron kinetic energies were chosen to balance the competing requirements of high circular polarization (maximized at lower energies) and high overall cross section (maximized at higher energies). Unfortunately, our chosen energy range happens to be in a regime where the dipole approximation for EELS selection rules is not particularly good (see Ref. 5). We must therefore assume that monopole and quadrupole excitations could contribute to our spectra. Excitations to $4p$-derived states are probably the leading nondipole contribution, and such transitions would likely dilute observed dichroic effects since these states are nearly all unoccupied. However, just above the Fermi level, the unoccupied states of these metals are dominated by s and d states, which are just those permitted by dipole selection rules. Hence, the inclusion of nondipole excitations in our analysis would not strongly affect our conclusions.

Probing surface and thin film magnetic structure with circularly polarized synchrotron radiation

G. D. Waddill[a)]
University of Missouri-Rolla, Physics Department, Rolla, Missouri 65401

J. G. Tobin
Lawrence Livermore National Laboratory, Livermore, California 94550

X. Guo and S. Y. Tong
University of Wisconsin-Milwaukee, Laboratory for Surface Studies and Department of Physics, Milwaukee, Wisconsin 53201

(Received 2 October 1995; accepted 13 May 1996)

A new group of techniques predicated on the availability of a tunable source of circularly polarized soft x-rays show tremendous potential for advancing the understanding of magnetic nanostructures. These techniques involve use of circularly polarized radiation as probes in x-ray absorption and x-ray photoemission and photoelectron diffraction studies, and provide unique information about the element specific magnetic moments and magnetic structure of thin films, interfaces, and surfaces. Examples of each technique will be discussed with emphasis given to the recently demonstrated first observation of spin-dependent photoelectron diffraction using circularly polarized x-rays. This technique promises the direct, element specific determination of local magnetic structure on an atomic scale, and will be of tremendous utility in the effort to establish structure-property relationships in magnetic nanostructures. Circularly polarized x-rays are used to produce spin-polarized photoelectrons from the Fe $2p$ spin-orbit doublet. Intensity asymmetries of up to $\sim 3\%$ are observed for the $2p_{3/2}$ sublevel. These asymmetries depend on the relative orientation of the x-ray polarization vector and the sample magnetization as well as upon the photoelectron kinetic energy and emission direction. Multiple scattering calculations are found to reproduce both the observed energy and angular asymmetry variations. Strategies to increase the measured effect are also discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

The requirements of the magnetic storage industry coupled with the development of modern thin film growth techniques have led to the production of new artificially structured nanoscale magnetic materials with unique magnetic properties.[1] The preparation and ultimately the control of the properties of these systems depends critically on the development of a better understanding of the magnetism of surfaces, interfaces, and thin films. A critical issue in this development is the subtle but critical relationship between structure and magnetic properties. In order to better understand these issues, a number of new sophisticated sample characterization techniques have been developed.[1–3] One such set of techniques is based upon the recent availability of circularly polarized x-rays over a wide energy range at synchrotron radiation sources. These techniques fall generally under the category of magnetic x-ray circular dichroism (MXCD) techniques which include the application of circularly polarized x-rays in absorption[2] and photoemission studies.[3]

In x-ray absorption studies MXCD provides element specific information about the magnetic moments (both spin and orbital) of magnetic materials.[2,4] Also, pioneering spin-polarized x-ray absorption fine structure studies using MXCD techniques have demonstrated a sensitivity to magnetic structure in bulk magnetic systems.[5] Core level photoemission, utilizing circularly polarized x-rays, probes the spin-orbit and exchange interactions and their effect on the core level spectral features.[3] This technique is complimentary to spin-polarized photoemission studies that measure the spin of the emitted electrons.[6] All these experimental techniques draw heavily from calculational support for proper interpretation.

This article details the application of MXCD techniques to the field of photoelectron diffraction. Specifically, the circularly-polarized x-rays are used to create an "internal" source of spin polarized photoelectrons. By comparing the angular and energy dependent intensity variations of oppositely spin-polarized photoelectrons it is possible to isolate the *magnetic* contribution to the scattering cross section for the photoelectrons. This technique can therefore be viewed as spin-polarized photoelectron diffraction (SPPED), and, in the same way that conventional photoelectron diffraction has been demonstrated to be sensitive to local atomic structure,[7] is sensitive to local *magnetic* structure. Furthermore, SPPED is surface sensitive and element specific, and as such will ultimately prove to be a valuable tool in the study of nanoscale magnetic systems.

The remainder of this article will detail the general development of MXCD techniques particularly where they impact the development of MXCD-SPPED. The issues critical to the

[a)]Electronic mail: Waddill@physics.umr.edu

MXCD-SPPED experiment will be discussed extensively, and first results from fcc Fe films grown on Cu(001) will be presented. The interpretation of these results hinges on comparison to spin-polarized, multiple-scattering calculations. The calculational details along with comparison to experimental results are presented. Sensitivity to magnetic structure is demonstrated by varying structural parameters in the model calculations. Finally, the results of calculations for the complete, or double, polarization experiment in which circularly polarized x-rays are used to produce an internal source of spin-polarized photoelectrons *and* the spin of the emitted photoelectron is measured are presented, and the implications for future SPPED studies are discussed.

II. MAGNETIC X-RAY CIRCULAR DICHROISM

All MXCD techniques are predicated on the availability of a tunable source of circularly or elliptically polarized x-rays. The most common usage of circularly polarized x-rays is in x-ray absorption spectroscopy where the photon energy is scanned through an absorption threshold exciting a core electron into an unoccupied valence state.[2] For transition metals, magnetic properties are determined primarily by the d-electrons, so most MXCD absorption studies of transition metal magnetism have concentrated on the L_2 and L_3 absorption edges that involve $2p$ to d transitions, and yield information about the empty d-states in the material. In MXCD absorption, the polarization vector, or helicity, of the incident photons (along the photon propagation direction for left circularly polarized photons and opposite the propagation direction for right circularly polarized photons) is aligned or anti-aligned with the sample magnetization. In this arrangement, it is a simple matter to show that the important dipole selection rules are $\Delta m_j = +1$ for right circularly polarized x-rays and $\Delta m_j = -1$ for left circularly polarized x-rays. If the d-band is not exchange split and an equal number of majority and minority spin vacancies exist, then excitation with circularly polarized x-rays produces 62.5% (37.5%) minority spin electrons from the $2p_{3/2}$ (L_3 edge) and 25% (75%) minority spin electrons from the $2p_{1/2}$ (L_2 edge) for excitation by right (left) circularly polarized radiation.[2] In ferromagnetic materials, the exchange interaction leads to an imbalance in the occupation of majority and minority spin states and will therefore favor transitions involving one spin state. Thus the absorption intensities at the L_3 and L_2 edges will vary for different relative orientations of photon helicity and magnetization, and these variations contain information about the spin and orbital components of the magnetic moment of the sample.[2,4]

In MXCD applied to core level photoemission, one again relies on the effects of proper orientation of the photon helicity and magnetization. Here, however, instead of exciting the core electrons to empty levels just above the Fermi energy, the photon energy is selected so that the excited photoelectron escapes the solid with kinetic energy typically 100 eV or greater. Thus, there is no spin imbalance in the available final states, and the large intensity variations observed in MXCD absorption are absent. Instead one observes (small) shifts in the kinetic energy of the emitted electrons that depends upon the relative orientation of the photon helicity and magnetization.[3] The origin of these shifts again rests upon the production of oppositely spin-polarized photoelectrons from the $2p_{3/2}$ and $2p_{1/2}$ levels of a ferromagnetic material by circularly polarized radiation. For example, for antiparallel orientation of helicity and magnetization, the $2p_{3/2}$ photoemission peak is derived from primarily majority spin electrons and that in the $2p_{1/2}$ peak from primarily minority spin electrons. The opposite spin polarization in each peak is achieved for a parallel orientation of helicity and magnetization. The effect of the magnetic exchange interaction is then to shift the $2p_{3/2}$ photoemission peak to lower kinetic energy and the $2p_{1/2}$ peak to higher kinetic energy than is observed for the parallel orientation. These shifts have been observed in both bulk and monolayer Fe systems.[3] The initial investigations of MXCD core level photoemission concentrated on angle integrated effects, but subsequent studies, both experimental and theoretical, demonstrated that the spectral features were also dependent upon photoelectron emission angle.[3] These angular variations can be used to further elucidate details of the spin-orbit and magnetic exchange interactions, however, the primary interest of the studies described herein is the spin dependence of the photoelectron scattering processes in magnetic materials. Ultimately it is this information that provides the sensitivity to the local magnetic structure.

III. SPIN-POLARIZED PHOTOELECTRON DIFFRACTION

The first studies of SPPED involved the use of s core level multiplet splitting in transition metal antiferromagnetic salts as a source of internally spin-polarized photoelectrons.[8] For example, the relative intensities of the oppositely spin-polarized Mn $3s$ doublet states were observed to exhibit temperature and emission direction dependent fluctuations which were used to examine short range antiferromagnetic ordering at the surface of $KMnF_3$.[8] The extension of this technique to metal surfaces, interfaces, and thin films has proved a difficult task that has suffered from small $3s$ cross section, and the absence of sufficiently resolved multiplet states in many materials. Thus, the present approach to SPPED was developed as an alternative. In this approach, excitation of a ferromagnetic transition metal $2p$ core level with circularly polarized radiation is used to produce a source of internally spin-polarized photoelectrons. The mechanism is the same as that outlined in the previous section and relies on the proper relative orientation of photon polarization and sample magnetization. Recall that in an atomic picture with equal number of minority and majority spin state vacancies the $2p_{3/2}$ and $2p_{1/2}$ peaks consist of 62.5% (37.5%) and 25% (75%) minority spin electrons for excitation using right (left) circular polarized x-rays. In this approach, the photoelectron spin polarization can be reversed by reversing either the sample magnetization direction or the helicity of the exciting radiation.

Following photoexcitation, the spin-polarized electrons scatter from near neighbor atoms. This scattering contains spin-dependent terms that vary according to the orientation of the photoelectron spin with respect to the net spin of the scattering atom. The resulting emission pattern of photoelectrons will therefore reveal details of the local *magnetic* as well as geometric structure due to the presence of exchange terms in the interaction Hamiltonian. The spin-dependent part of the scattering process can then be isolated by reversing the relative orientation of the photon helicity and sample magnetization. This is accomplished by either reversing the photon helicity or the sample magnetization. Comparison of the results of both methods provides a crucial internal consistency check of the results since the spin-dependent diffraction processes depend *only* on the relative orientation of the photon polarization and sample magnetization, and not on the *absolute* orientation of either.

IV. EXPERIMENT

The experiments were performed at the Stanford Synchrotron Radiation Laboratory (SSRL) using a spherical grating monochromator capable of delivering ~90% circularly polarized photons over the energy range from approximately 100 eV to 1000 eV. Four monolayer (ML) Fe films grown on a Cu(001) substrate held at 150 K resulted in a relatively poorly ordered metastable fcc Fe overlayer as evidenced by a diffuse $p(1\times1)$ low-energy electron diffraction (LEED) pattern consistent with previous studies of this system.[9] These films exhibit a magnetic easy axis along the sample normal,[9] and were magnetized *in situ* by means of an electromagnetic coil capable of producing fields up to 3 kOe. All measurements were made in remanance. The Fe $2p$ spectra were collected with an angle-resolving hemispherical analyzer with angular acceptance of $\pm 3°$. The perpendicular magnetic axis dictated that the photoelectron polarization is maximized for normal photon incidence. The analyzer position was then adjusted to the desired electron emission angle. Emission along two high symmetry directions was investigated. In the first, the scattering was defined by the sample normal and the (110) azimuth with emission along the [111] direction, and in the second the scattering plane was defined by the sample normal and the (100) azimuth with emission in the [110] direction. Both geometries are in mirror planes where, as will be discussed shortly, *nonmagnetic* photon polarization effects vanish. Also, in both cases the scattering plane coincided with the electron orbital plane in the storage ring (horizontal). Photon energies between 800 and 860 eV were employed giving rise to Fe $2p_{3/2}$ kinetic energies of about 90–150 eV. Angular alignment was determined using combined LEED and laser reflection measurements.

Typical Fe $2p$ spectra for emission along the [110] (lower panel) and [111] (upper panel) directions are shown in Fig. 1. These spectra were collected using a photon energy of 835 eV giving a kinetic energy of ~125 eV at the Fe $2p_{3/2}$ peak. The spectral differences discussed below were found to depend *only* on the relative orientation of photon polarization and sample magnetization, and not on the absolute orientation of either. In both sets of spectra, the apparent spin-orbit splitting is reduced on going from a parallel to an antiparallel orientation of photon and minority electron spin. As discussed previously, these energy shifts are due to the magnetic exchange interaction which, for the anti-parallel orientation, shifts the predominantly majority spin photoelectrons in the $2p_{3/2}$ peak to higher binding energy and the predominantly minority spin electrons in the $2p_{1/2}$ peak to lower binding energy. Also apparent for emission in the [111] direction is an *intensity* variation in the $2p_{3/2}$ peak. These intensity variations are found to be angle and energy dependent, and, as we will demonstrate, are consistent with calculations of spin-dependent diffraction.

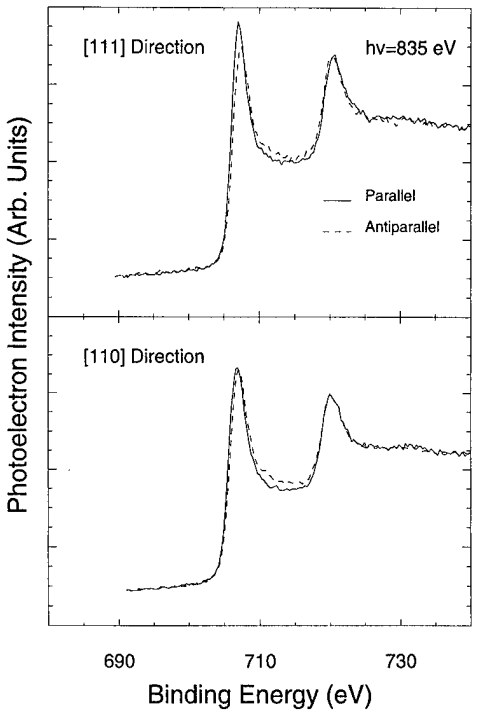


FIG. 1. Fe $2p$ spectra for 4 ML Fe/Cu(001) along the [110] direction (lower panel) and the [111] direction (upper panel) for parallel (solid curves) and antiparallel (dashed curves) photon and minority electron-spin orientations. The energy shifts are due to the magnetic exchange interaction on the oppositely spin-polarized electrons in each $2p$ sublevel. Note the intensity variation apparent in the $2p_{3/2}$ peak for the [111] direction and its absence in the [110] direction.

In order to isolate spin-dependent diffraction effects in our data, we have used the following scheme to eliminate exchange interaction energy shifts from the data and to concentrate exclusively on the intensity variations in the $2p_{3/2}$ peak for parallel and antiparallel photon and minority electron spin orientations. The energy shifts are removed by artificially separating the Fe $2p$ spectra for the antiparallel orientation in the region between the $2p_{3/2}$ and $2p_{1/2}$ peaks. Each peak is then displaced in energy (the $2p_{3/2}$ to lower binding energy and the $2p_{1/2}$ to higher binding energy) until their centroids align with the like $2p$ sublevel centroid in the spectrum obtained for parallel orientation of photon and minority electron spin. This is shown schematically in Figs. 2(a) and 2(b) for the data from the upper panel of Fig. 1. Here the two

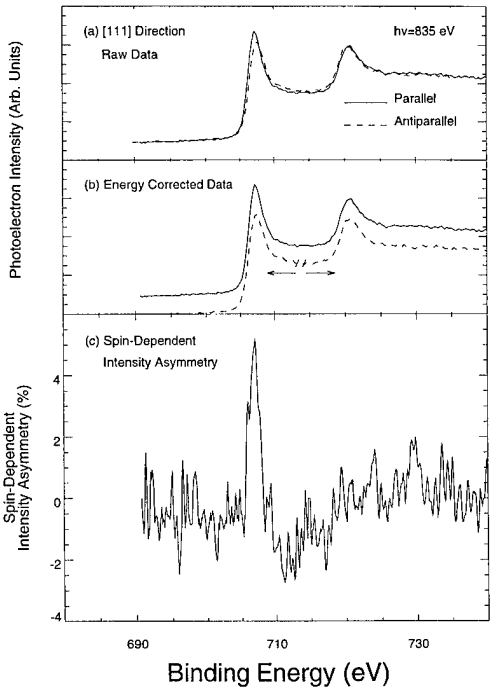


FIG. 2. A demonstration of the removal of exchange shifts in the data to isolate diffraction derived intensity variations. The data in the [111] direction from Fig. 1(b) are reproduced in (a) and the energy correction to the spectrum for the antiparallel orientation is demonstrated in (b). The intensity asymmetry for this spectral pair is shown in (c). The asymmetry at the $2p_{3/2}$ maximum is $\sim 5\%$.

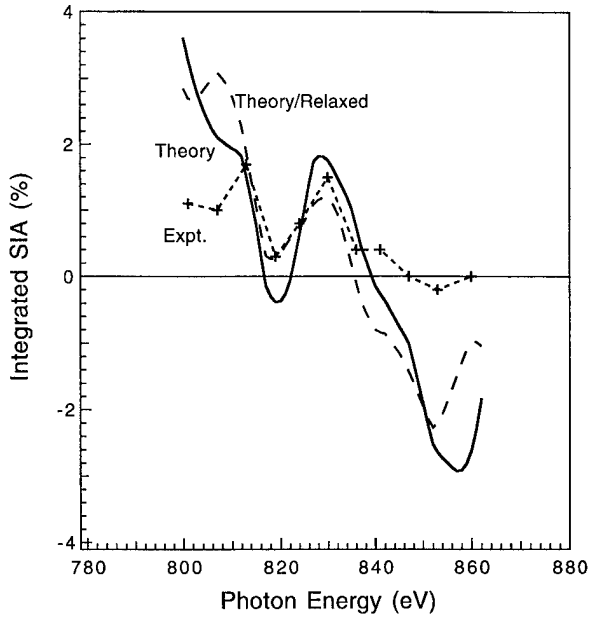


FIG. 3. Calculated (solid curve: $d_{12} = d_{23} = 1.8$ Å, $r_P = 0.19$; dashed curve: $d_{12} = 1.9$ Å, $d_{23} = 1.7$ Å, $r_P = 0.17$) and measured asymmetries along the [111] direction as a function of photon energy. The oscillatory behavior in the curves is due to spin-dependent photoelectron diffraction.

spectra are offset for clearer presentation. The spectra [in Fig.1 and in Fig. 2(a)] were also normalized to the low- and high-energy backgrounds in order to extract the proper $2p$ photoemission peak intensity differences. Following peak alignment, a difference spectrum is formed by subtracting the two spectra and normalizing to the sum of the two on a channel-by-channel basis [see Fig. 2(c)]. Because the spin polarization in each peak is reversed upon going from a parallel to antiparallel photon and minority electron spin orientation, this difference spectrum isolates the spin-dependent scattering intensities or diffraction effects. We are effectively isolating the following term,

$$\frac{I^{\uparrow}_{3/2} - I^{\downarrow}_{3/2}}{I^{\uparrow}_{3/2} + I^{\downarrow}_{3/2}}.$$

We note that this data extraction procedure is equivalent to that obtained by independently integrating each $2p$ sublevel (after background removal) for both spin orientations and forming the normalized difference of these values.

The spin-dependent intensity asymmetry (SIA) is calculated for each of the two emission directions investigated for all photon energies employed. This is shown for the [111] direction and $h\nu = 835$ eV in Fig. 2(c) where the observed asymmetry varies from -2 to $+5\%$ over the entire $2p$ manifold. We note that the normalization scheme described above essentially equates the intensity in the Fe $2p_{1/2}$ peak for both spin orientations and places all spin-dependent asymmetry effects within the $2p_{3/2}$ peak. The effective equalization of the intensities of the $2p_{1/2}$ features (which should also exhibit spin-dependent scattering effects) is due to both the higher secondary electron background underlying this peak together with the Coster–Kronig broadening of the peak. This combination of effects makes extraction of the spin-dependent $2p_{1/2}$ photoelectron scattering prohibitively difficult.

Figure 3 shows the results of integrating the SIA curves over a 10 eV energy window centered on the $2p_{3/2}$ peak for 11 different photon energies and emission in the [111] direction. The rationale for this integration is for proper comparison with calculations of spin-dependent scattering intensities for the entire $2p_{3/2}$ and $2p_{1/2}$ manifolds that will be discussed in the following section. Here the experimental SIA (broken curve connecting data points, +) varies from about -0.5% to 2% and exhibits oscillations as a function of photon energy. These oscillations are analogous to those observed in traditional (spin-integrated) photoelectron diffraction, and represent the first observation of spin-dependent photoelectron diffraction using circularly polarized photons to produce an internal source of spin-polarized photoelectrons. Due to the rather weak nature of these oscillations (down roughly one order of magnitude from spin-integrated photoelectron diffraction oscillations[7]) proper comparison to spin-dependent multiple scattering calculations are critical in assessing the experimental data. This is the subject of the following section.

V. THEORY

The calculation of spin-dependent multiple scattering photoemission intensities combines conventional photoemission and spin-polarized LEED methods. Calculation of the

excitation matrix element, $\langle \Phi_E | H | \Phi_E \rangle$, uses the nonrelativistic approximation with the dipole approximation used for the interaction Hamiltonian. In this article, the Dirac matrix, α, is replaced by the Pauli matrices, σ. The selection rules for excitation by circularly polarized light restrict transitions to $\Delta m_j = +1$ for right circular polarization and $\Delta m_j = -1$ for left circular polarization. Transitions from the spin-orbit split Fe $2p$ core sublevels are governed by these selection rules.

Following excitation, the internally spin-polarized photoelectrons are multiply scattered inside the crystal in a manner similar to that experienced by spin-polarized LEED electrons. The single site scattering matrix is calculated using the Dirac equation with spin-polarized potentials generated by a self-consistent linear augmented plane wave band calculation. This scattering matrix is converted to the (*lms*) representation and used to construct layer diffraction matrices. The calculation then proceeds as in conventional photoemission calculations except that the dimension of the layer diffraction matrices is doubled to include spin-dependent scattering and spin flip effects. Note that both spin-orbit coupling and exchange effects are accounted for since off-diagonal matrix elements are nonvanishing. In the calculation, the inner potential is set to 10 eV and inelastic scattering is simulated by an imaginary potential of 4.5 eV. Due to the low kinetic energies utilized in the experiment only terms up to $l=4$ are used in most calculations, and convergence has been checked using terms up to $l=6$ with no significant differences found.

The calculated spin-dependent intensities are for an entire $2p_{3/2}$ or $2p_{1/2}$ manifold, and comparison to the experimentally extracted SIA requires the formation of a calculated SIA defined as follows:

$$\mathrm{SIA}(h\nu) = \frac{[I^+_{3/2}(h\nu)/I^+_{1/2}(h\nu) - I^-_{3/2}(h\nu)/I^-_{1/2}(h\nu)]}{[I^+_{3/2}(h\nu)/I^+_{1/2}(h\nu) + I^-_{3/2}(h\nu)/I^-_{1/2}(h\nu)]}, \quad (1)$$

where, for example, $I^+_{3/2}$ is the integrated intensity in the $2p_{3/2}$ manifold for a parallel (+) orientation of photon and minority electron spin. In Fig. 3, we show the calculated asymmetry (solid curve) along the [111] direction for a model with fcc Fe spacing ($d_{12}=d_{23}=1.8$ Å) as a function of incident photon energy compared to the integrated SIA extracted from the experimental data. The agreement is qualitative with oscillations as well as the change in sign ca. 840 eV well reproduced. We stress that it is the position of the intensity oscillations rather than the magnitude that is of most fundamental importance in photoelectron diffraction. Also, we demonstrate in Fig. 3 the sensitivity of the data to variations in structural parameters. Varying the values of d_{12} and d_{23} yields a best fit to the data for $d_{12}=1.9$ Å and $d_{23}=1.7$ Å (dashed curve). This fit has a Pendry r factor of 0.17 which is 9% better than the fit for the unrelaxed overlayer (solid curve). Variations of greater than ± 0.2 Å destroys the agreement between calculation and experiment. These results agree with our previous determinations of the structure of the Fe overlayer using conventional photoemission.[10]

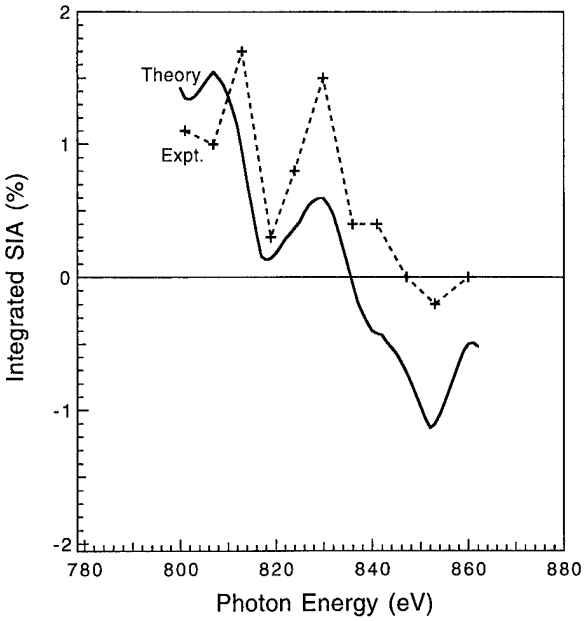


FIG. 4. The effect of inclusion of disorder on the calculation of the asymmetry along the [111] direction for the relaxed (dashed curve in Fig. 3) Fe layer. The calculated values are scaled by a factor of $(1-x)$ where x is the percentage disorder present in the Fe layer. The experimental results are shown with crosses connected by a broken line.

Further improvement of the agreement between calculation and experiment can be gained by introducing the effects of partial disorder within the overlayer. Figure 4 shows a comparison of calculations of the relaxed Fe overlayer with 50% disorder compared to the experimental data. As can be seen, the primary effect of the inclusion of disorder on the calculated integrated SIA is a reduction of the asymmetry. This brings the calculation into better agreement with the experiment. The precise origin of the disorder is not investigated here since its introduction at this level essentially represents a rescaling of the calculated SIA, however previous investigations of the Fe/Cu(001) system suggest substantial structural disorder together with Fe/Cu intermixing at room temperature and below.[9] The good agreement between our calculations and experimental data coupled with the consistency of those results with a number of previous studies of the Fe/Cu(001) system using a variety of techniques confirm that the intensity variations observed in the experiment are indeed due to spin-dependent photoelectron diffraction.

VI. DISCUSSION

At this point it is necessary to address a number of issues that influence the details of proper experimental configurations as well as affect the proper interpretation of the experimental results. These discussions are intended to outline the known limitations of this technique as well as to speculate about possible extensions and future investigations.

First, dichroism has been observed in *nonmagnetic* systems with reversal of incident photon polarization.[11] The calculated $2p_{3/2}$ intensity for excitation from fcc Fe assuming no magnetization by left and right circularly polarized x-rays

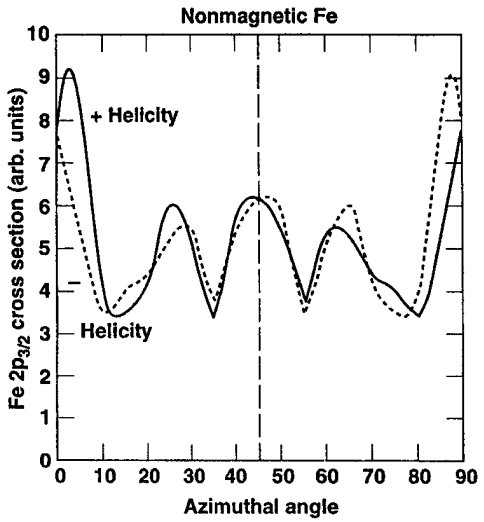


FIG. 5. Circular dichroism in non-magnetic fcc Fe. The calculated dichroism away from mirror planes arises from the effects of experimentally induced chirality by orientation of the crystallographic axes, photoelectron emission direction, and photon helicity.

is shown in Fig. 5. Away from high symmetry mirror planes one sees the presence of large photon polarization dependence in the excitation cross section that is obviously not of magnetic origin. In studies of magnetic properties using circular dichroism it is necessary to avoid this complication in data interpretation by performing measurements *only* in mirror planes where the nonmagnetic dichroism vanishes. The origin of the nonmagnetic dichroism lies in the imposition of an experimentally induced chirality in which the photoelectron wavevector, the sample normal, and the photon helicity define a coordinate system with a definite handedness. When all three vectors are coplanar *and* lie in a mirror symmetry plane, this induced chirality vanishes as can be seen, for example, for $\theta=0°$, 45°, and 90° [(110) and (001) planes] in Fig. 5. It is important to note that the experimentally induced dichroism depends upon the absolute orientation of photon helicity while the magnetic dichroism depends *only* on the *relative* alignment of helicity and magnetization. MCXD diffraction asymmetries will necessarily be equivalent for all possible parallel versus antiparallel configurations.

Next, we discuss the variation of photoelectron spin-polarization with emission angle. The simple atomic model discussed previously predicts a spin-polarization of $\pm 25\%$ and $\mp 50\%$ from the $2p_{3/2}$ and $2p_{1/2}$ sublevels respectively for excitation to empty d-levels. In angle-resolving experiments, the d-character of the photoelectron final state and therefore the photoelectron polarization will depend upon the photoelectron emission angle and kinetic energy. Thus, a proper interpretation of the experimental data must account for this variable spin-polarization. In Fig. 6 we show the results of multiple scattering calculations of the Fe $2p_{3/2}$ spin-polarization versus polar emission angle in the [110] plane at a kinetic energy of 125 eV. Here the spin-up and spin-down intensities for parallel (solid curve) and antiparallel (dashed curve) orientation of helicity and sample magnetization are shown. The top panel is the so-called direct term which is essentially atomic emission incorporating the appropriate crystalline symmetry as well as proper surface boundary conditions.[12] In the bottom panel, the effects of multiple scattering are included. One sees that there is a variable photoelectron spin-polarization even in the absence of diffraction effects (upper panel), but that this spin-polarization is strongly modulated by diffraction effects (bottom panel). It is this modulation that gives spin-dependent photoelectron diffraction its sensitivity to local magnetic structure. It is, however critical to properly account for these angular photoelectron spin polarization variations when interpreting the MXCD-SPPED data.

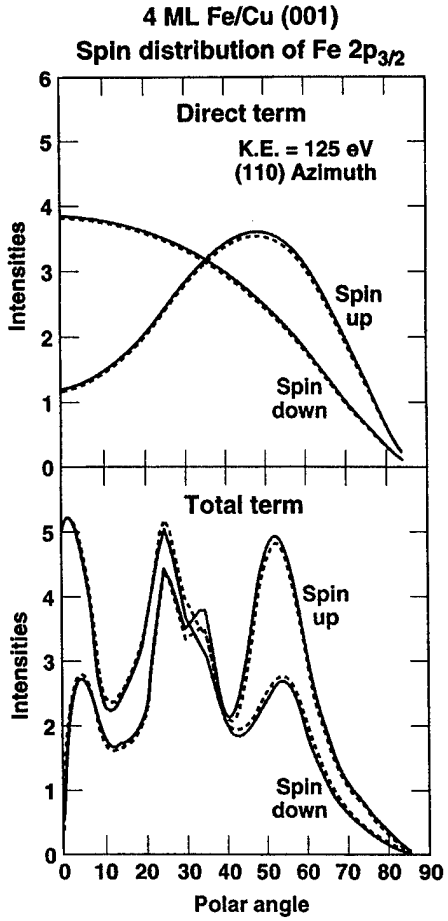


FIG. 6. Photoemission cross section calculations for the Fe $2p_{3/2}$ level showing the spin polarization of the $2p_{3/2}$ peak for excitation with right circularly polarized light as a function of polar angle in the [110] plane for the direct (upper panel) and total (lower panel) terms. All calculations are for 4 ML Fe/Cu(001) assuming pseudomorphic growth, and for a photoelectron kinetic energy of 125 eV. Note the additional diffraction related structure in the total term. The solid and dashed curves are for parallel and antiparallel orientation of helicity and sample magnetization, respectively.

Finally, we turn to a discussion of possible future extensions for this technique. Obviously the magnitude of the spin dependent scattering asymmetries observed herein are sufficient to limit future studies using this technique. In order to generate a more widely applicable technique some mechanism for increasing the spin dependent asymmetry in the

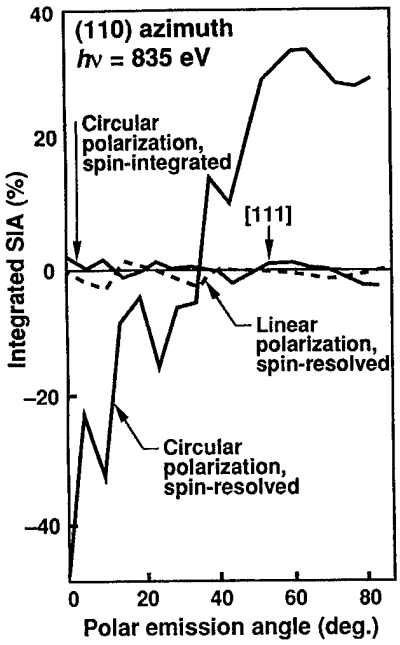


FIG. 7. Integrated SIA using Eq. (1) for scattering along the (110) azimuth (solid curve) as a function of polar emission angle. For the dashed curve the results for linear polarized incident photons and external photoelectron spin detection are shown. Note the similarity in the size of the spin-dependent oscillations to those for excitation using circularly polarized light and no external photoelectron spin detection (solid curve). Finally, the spin polarization of the Fe $2p_{3/2}$ photoemission peak for excitation by right circularly polarized light is shown. This arrangement requires the double or complete polarization experiment that uses excitation by circularly polarized photons together with external photoelectron spin detection.

diffraction data must be arrived at. In Fig. 7 we show the results of calculations that suggest that such improvements may well be within reach in the very near future. This figure shows calculations of the spin dependent asymmetry as defined in Eq. (1) (solid line) as a function of polar emission angle along the (110) azimuth with $h\nu=835$ eV for fcc Fe. This result is obtained by varying the incident photon helicity and detecting *all* photoelectrons emitted in a given direction. Also shown are the results of calculations using linearly-polarized photon excitation at the same energy, but now detecting the spin-polarization of the photoelectrons emitted in a given direction (dashed line). Here too, one sees polarization oscillations that reflect a sensitivity to the local magnetic structure, however these oscillations are again only about ± 1–2%. This is further complicated by the relative inefficiency of photoelectron spin-detection which is probably at least two orders of magnitude less efficient than spin-integrated photoelectron detection. In comparison, the emitted photoelectron intensity for excitation by circularly polarized radiation produced at a bending magnet monochromator is only reduced from that for linear polarization by about a factor 5. However, if one combines *both* excitation with circularly polarized photons and external photoelectron spin detection in a so-called complete or double polarization experiment the spin dependent oscillations increase to approximately 8–10%. These oscillations are on top of a smoothly increasing background which arises from angle-dependent spin polarization in the direct term (see Fig. 6). Here we plot the following:

$$\frac{I^R_{3/2}(\uparrow)-I^R_{3/2}(\downarrow)}{I^R_{3/2}(\uparrow)+I^R_{3/2}(\downarrow)}.$$

In order to perform the double polarization experiment, the reduced efficiency of both the production of circular versus linear polarized x-rays and the requirement of external photoelectron spin detection must be overcome. This capability, using an insertion device for the production of circularly polarized x-rays, has been demonstrated recently by Starke *et al.*[13] for tungsten $4f$ photoemission. The magnitude of the calculated spin-dependent diffraction oscillations in the double polarization experiment of Fig. 7 are very similar to those observed in spin-independent photoelectron diffraction,[7] and should prove tractable in extracting valuable information about local magnetic structure from any number of surface and thin film magnetic systems.

VII. CONCLUSIONS

We have presented the results of the first measurement of spin-dependent photoelectron diffraction using circularly polarized x-rays in investigations of fcc Fe films grown on Cu(001). The resulting spin-polarized Fe $2p$ photoemission peaks exhibit angle- and energy-dependent intensity variations due to spin-dependent diffraction which provides sensitivity to the details of local magnetic structure. The importance of relativistic, multiple scattering calculations in the proper interpretation of the experimental data is demonstrated. The results demonstrate sensitivity to magnetic structure details at the sub-Ångstrom level. Factors which currently limit the experimental technique as well as dictate details of experimental configurations were discussed. Finally, the results of calculations for double polarization experiments which combine excitation by circularly polarized x-rays with external photoelectron spin detection suggest that this will become a powerful technique for the study of surface, interface, and thin film magnetism in the near future.

ACKNOWLEDGMENTS

Work was partially performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-ENG-48. SSRL is supported by the Chemical Sciences Division of DOE/BES. The work at the University of Wisconsin-Milwaukee was supported by DOE DE-FG02-84ER45076.

[1]L. M. Falicov *et al.*, J. Mater. Res. **5**, 1299 (1990); S. D. Bader, Proc. IEEE **78**, 909 (1990).

[2]See, for example, C. T. Chen, N. V. Smith, and F. Sette, Phys. Rev. B **43**, 6785 (1991); Y. Wu, J. Stöhr, B. D. Hermsmeier, M. G. Samant, and D. Weller, Phys. Rev. Lett. **69**, 2307 (1992); J. G. Tobin, G. D. Waddill, and D. P. Pappas, Phys. Rev. Lett. **68**, 3642 (1992).

[3]See, for example, L. Baumgarten, C. M. Schneider, H. Petersen, F. Schäfers, and J. Kirschner, Phys. Rev. Lett. **65**, 492 (1990); H. Ebert, L. Baumgarten, C. M. Schneider and J. Kirschner, Phys. Rev. B **44**, 4406 (1991); D. Venus, L. Baumgarten, C. M. Schneider, C. Boeglin, and J. Kirschner, J. Phys. C **5**, 1239 (1993); G. D. Waddill, J. G. Tobin, and D.

P. Pappas, Phys. Rev. B **46**, 552 (1992); E. Tamura, G. D. Waddill, J. G. Tobin, and P. A. Sterne, Phys. Rev. Lett. **73**, 1533 (1994).
[4]B. T. Thole, P. Carra, F. Sette, and G. Van der Laan, Phys. Rev. Lett. **68**, 1943 (1992); P. Carra, B. T. Thole, M. Altarelli, and X. Wang, Phys. Rev. Lett. **70**, 694 (1993); R. Wu, D. Wang, and A. J. Freeman, Phys. Rev. Lett. **71**, 3581 (1993).
[5]G. Schütz, Physica Scripta T **29**, 172 (1990); G. Schütz, R. Frahm, P. Mautner, R. Wienke, W. Wagner, W. Wilhelm, and P. Kienle, Phys. Rev. Lett. **62**, 2620 (1989).
[6]See, for example, P. D. Johnson, A. Clarke, N. B. Brookes, S. L. Hulbert, B. Sinkovic, and N. V. Smith, Phys. Rev. Lett. **61**, 2257 (1988).
[7]See, for example, S. A. Chambers, Surf. Sci. Rep. **16**, 261 (1992); C. S. Fadley, Surf. Sci. Rep. **19**, 231 (1993).
[8]B. Sinkovic, B. Hermsmeier, and C. S. Fadley, Phys. Rev. Lett. **55**, 1227 (1985); B. Sinkovic and C. S. Fadley, Phys. Rev. B **31**, 4665 (1985); and M. T. Johnson, H. I. Starnberg, and H. P. Hughes, J. Phys. C **20**, 4385 (1987).
[9]See, for example, D. D. Chambliss, R. J. Wilson, and S. Chiang, J. Vac. Sci. Technol. A **10**, 1993 (1992); P. Xhonneux and E. Courtens, Phys. Rev. B **46**, 556 (1992); M. T. Kief and W. F. Egelhoff, Jr., Phys. Rev. B **47**, 10 785 (1993); and P. Bayer, S. Müller, P. Schmailzl, and K. Heinz, Phys. Rev. B **48**, 17611 (1993).
[10]G. D. Waddill, J. G. Tobin, X. Guo, and S. Y. Tong, Phys. Rev. B **50**, 6774 (1994); and J. Appl. Phys. **76**, 6465 (1994); and references therein.
[11]C. Westphal, J. Bansmann, M. Getzlaff, and G. Schönhense, Phys. Rev. Lett. **63**, 151 (1989); G. Schönhense, Phys. Scripta T **31**, 255 (1989); C. Westphal, A. P. Kaduwela, C. S. Fadley, and M.A. Van Hove, Phys. Rev. B **50**, 6209 (1994).
[12]S. Y. Tong and H. C. Poon, Phys. Rev. B **37**, 2884 (1988).
[13]K. Starke, U. Liu, P. D. Johnson, V. Chakarian, E. E. Chaban, G. Meils, and C. T. Chen, BAPS **40**, 334 (1995).

Theoretical investigation of structural instabilities of Fe layers on face-centered-cubic Cu

Lars Nordström[a)]
Institute for Computational Sciences and Informatics, George Mason University, Fairfax, Virginia 22030-4444 and Complex Systems Theory Branch, Naval Research Laboratory, Washington, DC 20375-5345

David J. Singh
Complex Systems Theory Branch, Naval Research Laboratory, Washington, DC 20375-5345

(Received 2 October 1995; accepted 11 December 1995)

Calculations of structural stabilities for thin iron layers as grown on Cu(001) are reported. It is found that a free-standing fcc(001) monolayer is unstable with respect to 3×1 and 4×1 distortions. The former instability is directly related to the Pitsch transformation from fcc(001) to bcc(110), and calculations along the path find bcc(110) as the stable structure. Inclusion of the Cu substrate in the calculation tends to stabilize fcc(001). The calculated instabilities are of similar form to the superstructures observed for Fe grown on Cu(001). The relation to the experimental situation is discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

The fact that iron was observed to grow in a fcc structure on Cu(001)[1] has attracted a large number of experimental studies on this system. It is now well established that Fe grown on the fcc(001) Cu surface actually has a very rich phase diagram.[2–5] As a function of overlayer thickness and growth conditions, it has been conveniently divided into three regions (I–III) according to the overlayer structure.[2] All of these phases are characterized by superstructures of the basic square structure of a pure (001) layer: Phase I shows 4×1 or 5×1, phase II 2×1, and phase III 3×1-like superstructure patterns in low-energy electron diffraction analyses. Changes in magnetic ordering,[2] e.g., ferro- (F) or antiferromagnetic (AF) ordering, and spin reorientations[4,6,7] (i.e., changes of easy axis) are closely connected to the structural changes. This phase diagram indicates that Fe is never stable as a pure fcc or fct (strained 1×1 fcc) overlayer on Cu(001). This article reports on a theoretical study of this structural instability for some simple systems by means of density-functional based first-principles calculations.

The experimental situation seems to be quite consistent. The magnetic orderings of phases I and III are ferromagnetic. It has been observed that phase III, which occurs at thicker Fe films, is actually bcc-like.[3,8,9] The 3×1 pattern arises from the fact that the phase transformation with growing layer thickness, from the fcc-like structures of phase I and II to phase III, is martensitic of the type described first by Pitsch.[10] In this transformation, an fcc(001) layer transforms into a bcc(110) one, with the latter as an approximate 3×1 superstructure of the former. The structure of the very thin iron layers of phase I, less than 5 monolayers (ML), has very recently been determined.[5] It is found that the 5×1 phase is characterized by a distortion of the fcc(001) layers due to sinusoidal shifts and bucklings of [110] atomic rows. This is also accompanied by an increase of the layer spacing. At intermediate thicknesses, phase II has been observed, but the range for which it exists is sensitive to the growth temperature. In fact, when the Fe layer is grown at 100 K, there seems to be a direct transformation from phase I to III without any intermediate phase.[5] When phase II has been observed, it is found to have an AF ordering between all layers except between the two top layers, which are F aligned. The 2×1 pattern reflects a reconstruction of the surface layer.[2,11]

There have been some theoretical calculations for Fe/Cu(001). The first study was for 1 and 2 ML of Fe on Cu(001) where large magnetic moments were found.[12] These magnetic moments are much larger than what is calculated for fcc Fe with the same lattice constant and even larger than those of bcc Fe. Recently, calculations have been reported for thicker Fe layers, 4–11 ML, which are in good accord with what is experimentally observed for phase II.[13] There has also been a theoretical study of different magnetic orderings within fct Fe for different c/a, incorporating both bcc and fcc as two special cases.[14] Although it was suggested to be of direct relevance for Fe/Cu(001), this Bain path of the fcc–bcc transformation is not what is experimentally observed. None of this work allowed for any reconstructions of the in-plane structure. The present work elucidates the instability against such distortions and the effects on the electronic and magnetic structure.

II. COMPUTATIONAL METHOD

The calculations were performed with the linearized augmented plane wave (LAPW) method[15] within the local-spin density approximation (LSDA).[16] This method makes no approximation to the shape of the densities or the potential. The surface was simulated by a supercell technique, where the gap between the slabs corresponds to three layer distances. Brillouin zone integrations were performed using 54 special **k** points in the irreducible wedge for the unreconstructed ML. For the larger structures, this number was reduced in a way which keeps the **k**-point density in reciprocal space constant.

a)Electronic mail: nordstro@dave.nrl.navy.mil

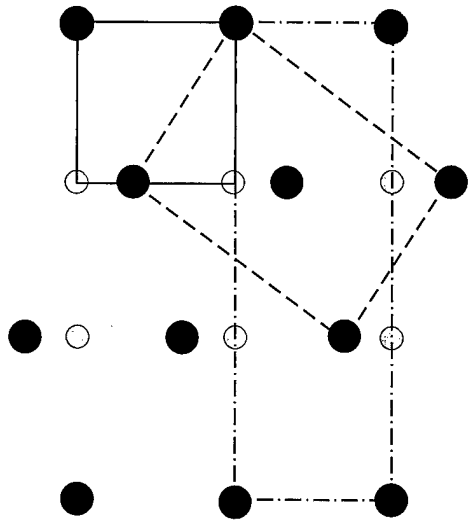


FIG. 1. The fcc(001) (small gray circles) and the corresponding reconstructed 3×1 bcc(110)-like (large black circles) structures are shown. The unit cells of the unreconstructed square lattice (full line), the fully reconstructed bcc(110)-like (dashed line), and the distorted lattice used in the calculations (dash–dotted line) are indicated.

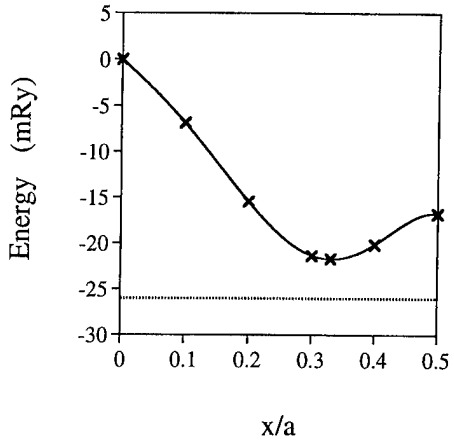


FIG. 2. The total energy per atom is plotted against the atomic displacement x for a 3×1 reconstruction. The energy is put to zero for the unreconstructed layer. Also shown is the total energy for a true bcc(110) layer (dashed line) with the same area per atom as the fcc(001) layer.

III. RESULTS

A. Structure

From the description of the different phases in Sec. I, it is clear that they are mainly formed from reconstructions within the fcc(001) layers. For instance, in Fig. 1 a 3×1 superstructure is displayed. This corresponds closely to the Pitsch-transformed bcc(110) layer, although it is not yet rectangular. Instead of having a square angle and a ratio between the sides of $\sqrt{2}$, the reconstructed unit cell indicated in Fig. 1 has the angle and ratio given in Table I. This reconstruction can be formed by a continuous distortion from the square lattice by sinusoidal shifts of atoms in a [01] row, which corresponds to the [110] row of the fcc lattice. The bcc(110)-like structure shown in the figure corresponds to a shifting of atoms a third of the square lattice constant a long [10]. It is possible in a similar way to construct bcc(110)-like structures corresponding to 4×1 and 5×1 superstructures. These are of a similar closeness to a rectangular form as the 3×1 structure, as can be seen in Table I. However, they are not continuous in the same way, i.e., they are not martensitic, so the corresponding bcc(110)-like structure is not directly related to any sinusoidal shifts of 4×1 or 5×1 character along the [110] rows.

TABLE I. Geometrical data for the bcc(110)-like distortions of the fcc(100) layer compared to those of truly rectangular bcc(110) layer.

Structure	Angle	Ratio of the sides
3×1	86.8°	1.387
4×1	88.4°	1.442
5×1	89.0°	1.473
bcc(110)	90.0°	1.414

A three-dimensional phase transformation of Pitsch nature also involves, in addition to the sinusoidal shifting and the subsequent skewing to get a rectangle within each layer, a sliding of every second layer $a/2$ along [10]. The interlayer distance is, however, kept constant if there is no accompanying volume change.

B. Free-standing monolayer

To start, we studied the simplest possible system which still has some relevance to the experimental observations: a free-standing monolayer (FSM) of Fe with a lattice matching Cu(001). This is a reasonable model system since the hybridization effects due to the Cu substrate are known to be quite small.[12,13] The calculated magnetic moments are, however, slightly larger than for an overlayer, 3.1 vs 2.8 μ_B/atom.

The results of total energy calculations for different atomic shifts corresponding to the 3×1 reconstruction of Fig. 1 are shown in Fig. 2. The square FSM is highly unstable with respect to a Pitsch transformation. The energy minimum is given by a distortion $x=a/3$, which corresponds exactly the bcc(110)-like structure shown in Fig. 1. Also given in Fig. 2 is the extra energy gained in completing the Pitsch path to form a true bcc(110) layer with the same unit area. The energy of a 4×1 bcc(110)-like structure is found to be 4.2 mRy higher than the 3×1 version and 17.1 mRy lower than the undistorted fcc(001) layer.

In order to illustrate the interplay of magnetism with this instability, the second derivative of the total energy with respect to the amplitude of the distortion is calculated when the magnetic moment is constrained[17] to different values. This is done both for a small 3×1 sinusoidal distortion as well as for a 4×1 one, and the results are tabulated in Table II. It is found that the FSM is unstable for a wide range of magnetic moments for both types of distortions. For smaller magnetic moments the inequivalent atomic moments of the larger unit cells start to arrange in a ferrimagnetic or even an AF fashion in order to fulfill the constraint on the total magnetic moment

TABLE II. The second derivative of the total energy per Fe atom with respect to the amplitude of two different sinusoidal distortions is given in units of mRy/(a.u.)^2. A positive value indicates an instability. The values are for a free-standing ML with different magnetic moments m, for an overlayer, and for bulk fcc Fe. The magnetic moments are given in μ_B per atom.

Structure	m=3.1	m=2.7	m=2.4	m=2.1	Overlayer	Bulk
3×1	5.18	4.83	3.99		−12.78	−5.30
4×1	3.90	4.03	3.04	0.90		

while retaining large local moments, instead of decreasing in a smooth fashion. This leads to an enhanced instability, so these cases have been excluded from Table II. The 3×1 instability is larger over the full range but is decreasing with decreasing magnetic moment, while the 4×1 instability has a maximum at a lower moment. These two sinusoidal distortions are identical to transverse phonons with different wave vectors. The instabilities might hence be viewed as due to softening of these phonons.

The nature of the instability is not clear, other than that it relates to the stability of bulk bcc Fe over fcc. It is known that such instabilities can be driven by Fermi surface properties. In order to see whether this is the case here, we show the Fermi surface of the FSM. It is formed by two bands, one of each spin. For the majority spin for which the 3D states are filled, the Fermi surface is close to circular, i.e., free-electron-like. The Fermi surface of the minority spin is shown in Fig. 3. This shows some rather squarelike features, which would indicate good nesting. In fact, the nesting vectors are actually quite close to 1/3 of a reciprocal lattice vector, in accordance with the 3×1 instability calculated. Although this is suggestive, more careful analysis is needed before it can be concluded that this is the driving force responsible for the instability. The shape of this minority spin Fermi surface inevitably has to change with the magnetic moment leading to a less squarelike structure or other nesting vectors. But as we have seen, the instability is rather insensitive to the exact value of the magnetic moments, which indicates that this nesting only can be a part of the full story.

A major simplification made when ignoring the substrate is the neglect of repulsive forces when the displaced Fe atoms approach substrate atoms. The FSM might, however, still be viewed as the limiting case of very large adsorbate–substrate distance.

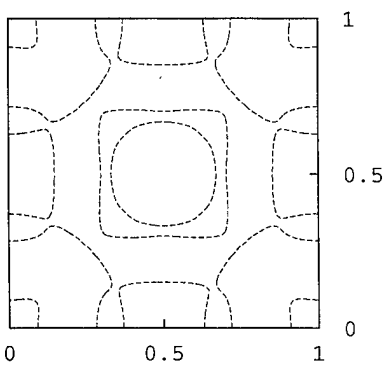


FIG. 3. The majority spin Fermi surface of a free-standing ML with its self-consistent magnetic moment.

C. Overlayer and bulk

When going to more realistic systems, reconstructions inevitably involve more structural parameters. We present here some results when the distortions are only allowed in a subspace of this full parameter space. First, to study the effect of the Cu substrate, calculations for the same 3×1 distortions as for the FSM have been done for a slab of Fe/Cu_3/Fe. The layer distances between the Cu layers are taken to be given by the Cu lattice constant, while the Fe–Cu distance is relaxed. The calculated Fe height, 1.74 Å, and magnetic moments, 2.82 μ_B, are in good agreement with earlier calculations,[12] 1.75 Å and 2.85 μ_B. As can be seen in Table II, we find the square overlayer to be clearly stable. So the Cu substrate is able to stabilize the Fe overlayer. This is probably due to a combination of the fact that the hybridization stiffens the Fe layer and that the substrate atoms repel displaced Fe atoms. However, there are two things to notice: (1) the LSDA-theoretical Fe–Cu distance is smaller than that experimentally observed which means that the effect of the hybridization is overestimated in the calculation, and (2) any buckling out of the plane of the distorted atoms, which would be a way to avoid the underlying Cu atoms, has not been allowed for. Any attempt to correct for these two facts would work in favor of an instability, as observed.

It was suggested in Ref. 5 that the superstructure of phase I implies that ferromagnetic bulk fcc Fe is intrinsically unstable. To test this possibility we have performed calculations for ferromagnetic bulk fcc Fe at the Cu lattice constant and find that it is stable with respect to a 3×1 distortion. It should be noted that the magnetic moment is comparatively small in this case, 1.1 μ_B, which means that, in contrast to the cases above, the majority d band is far from filled and hence contributes significantly to the bonding. If the full Pitsch path were followed it would of course have been found that the end-point bcc Fe is more stable, but it should be remembered that the present calculation only involves a small in-plane sinusoidal distortion. The sliding of the (001) planes was ignored.

IV. DISCUSSION AND CONCLUSIONS

These calculations have been performed within LSDA. This suggests some caution since it is well known that this approximation incorrectly predicts bulk fcc Fe to be more

stable than bcc. However, if gradient corrected approximations work in this case in the same way as for bulk Fe,[18] i.e., favor a bcc structure, the tendency toward an instability would increase. A large problem for LSDA in the bulk case is that fcc is nonmagnetic while bcc is ferromagnetic, so a comparison between these two structures involves a large difference in spin-polarization energies. The same problem does not exist in this case where the two types of layer structures have similar spin-polarization energies.

It has been shown that calculations for a FSM give a strong Pitsch instability of the fcc(001) lattice toward a 3×1 reconstruction. This reconstruction is related by a shear to a bcc(110) layer which is slightly more stable. The fcc(001) FSM is also shown to be unstable toward sinusoidal 4×1 distortions. These two types of instabilities are similar to the experimentally observed superstructures of phases III and I, respectively: In phase III the full Pitsch transformation is observed, while in phase I only sinusoidal 4×1 or 5×1 distortions are observed. The 3×1 distortion is close in nature to the 4×1 or (5×1) distortion for small amplitudes, but it differs in that it continuously gives rise to a Pitsch reconstruction for large amplitudes.

Calculations for more realistic overlayers find that the Cu substrate stabilizes the 1 ML (001) Fe overlayer. This is due to the small Fe–Cu distance used in these calculations. Increasing this will lead to instabilities since the FSM, which corresponds to very large distances, is unstable. The question is for which interlayer distance these instabilities occur. Calculations are in progress for the Fe/Cu(001) system to investigate this. Also, the effect of more than one Fe layer should be studied in order to get closer to the experimental situation.

ACKNOWLEDGMENTS

Work at the Naval Research Laboratory is supported by the Office of Naval Research. The computations were performed at the CEWES and NAVO computer centers of the Department of Defense.

[1]W. A. Jesser and J. W. Matthews, Philos. Mag. **15**, 1097 (1967).
[2]J. Thomassen, F. May, B. Feldman, M. Wuttig, and H. Ibach, Phys. Rev. Lett. **69**, 3831 (1992).
[3]J. Giergel *et al.*, Surf. Sci. **310**, 1 (1994).
[4]D. Li, M. Freitag, J. Pearson, Z. Q. Qui, and S. D. Bader, Phys. Rev. Lett. **72**, 3112 (1994).
[5]S. Muller, P. Bayer, C. Reischl, K. Heinz, B. Feldman, H. Zillgen, and M. Wuttig, Phys. Rev. Lett. **74**, 765 (1995).
[6]C. Liu, E. R. Moog, and S. D. Bader, Phys. Rev. Lett. **60**, 2422 (1988).
[7]D. P. Pappas, K.-P. Kämper, and H. Hopster, Phys. Rev. Lett. **64**, 3179 (1990).
[8]M. Wuttig *et al.*, Surf. Sci. **291**, 14 (1993).
[9]D. D. Chambliss *et al.*, in *Multilayers and Surfaces; Interfaces and Characterization*, MRS Symposia Proceedings No. 313 (Materials Research Society, Pittsburgh, 1993), p. 713.
[10]W. Pitsch, Philos. Mag. **5**, 477 (1959).
[11]M. Wuttig and J. Thomassen, Surf. Sci. **282**, 237 (1993).
[12]C. L. Fu and A. J. Freeman, Phys. Rev. B **35**, 925 (1987).
[13]T. Kraft, P. M. Marcus, and M. Scheffler, Phys. Rev. B **49**, 11511 (1994).
[14]S. S. Peng and H. J. F. Jansen, J. Appl. Phys. **69**, 6132 (1991).
[15]O. K. Andersen, Phys. Rev. B **12**, 3060 (1975); see also D. J. Singh, *Planewaves, Pseudopotentials and the LAPW Method* (Kluwer, Boston, 1994).
[16]U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).
[17]K. Schwarz and P. Mohn, J. Phys. F **14**, L129, (1984); A. R. Williams *et al.*, Bull. Am. Phys. Soc. **29**, 278 (1984).
[18]D. J. Singh, W. E. Pickett, and H. Krakauer, Phys. Rev. B **43**, 11628 (1991).

Investigation of face-centered-cubic Fe thin films using wedged samples

R. K. Kawakami,[a] Ernesto J. Escorcia-Aparicio, and Z. Q. Qiu[b]
Department of Physics, University of California at Berkeley, Berkeley, California 94720 and
Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

(Received 2 October 1995; accepted 22 December 1995)

The structural and magnetic properties of metastable face-centered-cubic (fcc) Fe thin films on fcc Co(100) substrate were studied using wedged samples. fcc Co(100) was chosen for the substrate because it is structurally very similar to Cu(100) but is ferromagnetic at room temperature. Reflection high energy electron diffraction and low energy electron diffraction characterizations confirm that epitaxially grown (MBE) Fe on Co(100) is structurally very similar to Fe on Cu(100): face-centered-tetragonal (fct) for $d_{Fe}<6$ ML, fcc for 6 ML$<d_{Fe}<11$ ML, bcc for $d_{Fe}>11$ ML. *In situ* surface magneto-optic Kerr effect measurements show that at room temperature the fct and bcc regions are ferromagnetic, while the fcc region is nonferromagnetic with some magnetic live layers. All magnetizations are in-plane. Oxygen absorption experiments suggest that these live layers are at the Fe/Co interface. Low temperature growth Fe/Co(100) shows a Kerr signal that increases linearly with d_{Fe} and suggests that the magnetic moments for fcc Fe and bcc Fe are the same. To further study the magnetic properties of the nonferromagnetic "fcc" phase, we used metastable fcc Fe as a spacer layer between two Co layers. The Co/fcc Fe/Co on Cu(100) sandwiches exhibit ferromagnetic coupling, strong antiferromagnetic coupling (AFC) and weak AFC. An oscillation in the strong AFC was found by artificially lengthening the thickness range of the nonferromagnetic fcc phase. © *1996 American Vacuum Society.*

I. INTRODUCTION

Investigation of magnetic ultrathin films has developed rapidly in the last decade after the discovery of antiferromagnetic coupling (AFC)[1] and giant magnetoresistance[2] in Fe/Cr multilayers. One important experimental method is to perform a thickness dependent study. Related topics using this method include the three- to two-dimensional phase transition, spin-reorientation transition, and the oscillations in magnetic coupling, etc. Because of the short-range character of the magnetic exchange interaction, precise control of the film thickness is needed in order to explore the novel magnetic behavior that varies on the atomic scale. Prior to wedged samples, thickness-dependent studies were performed on uniform thickness samples. A large number of samples with different thicknesses are usually required for one project. There are a number of difficulties using the uniform thickness samples. First, such investigations are time consuming. Typically, a separate sample must be made for each film thickness. This is particularly significant for molecular beam epitaxy (MBE) systems, where the sample production is slow (~1–2 samples per day per substrate). Second, variability in the growth conditions from sample to sample introduces uncertainty. If two samples of differing thicknesses exhibit differing behavior, it might be due to the thickness. However, it might alternatively be due to slightly different growth conditions (e.g., chamber pressure, growth rate, temperature, substrate cleaning procedure). This makes it more difficult to definitively separate the thickness dependence from other effects. Third, when given two samples of nearly equal thickness, one cannot be sure which sample is actually thicker because of uncertainties in the growth rates.

To overcome the difficulties of the uniform thickness samples, wedged samples were developed and applied to the study of magnetic thin films in 1991.[3] In conjunction with local magnetization measurement techniques, wedged samples allow for the efficient exploration of thickness dependent phenomena on a single sample. The wedge technique basically removes all the difficulties of the uniform thickness samples. First, many thicknesses are present on a single wedge so that the time consumed in film growth is greatly reduced. Second, since all the thicknesses are on one sample, they are all subject to the same growth conditions. If the substrate has been tested for uniformity, any differences in behavior can be attributed to thickness. Third, even though there is uncertainty in the growth rate, the overall thickness always increases with position as one moves from the thin side to the thick side of the wedge. Hence, given two measurements, one always knows which corresponds to the thicker film.

Figure 1 illustrates how to make a wedged layer. In an MBE system, a wedge is made by putting a "knife edge" mask in the molecular beam and translating the substrate behind it with uniform velocity V. If the evaporation rate is r, the produced wedge will have a slope of r/V. With a typical evaporation rate of ~1 Å/min and a substrate translation speed of ~1 mm/min, the resulting slope is $\sim 10^{-7}$, or approximately a million atoms per monolayer step. Hence, the film can be considered locally flat. The thickness resolution using wedged samples depends on the beam size of the magnetic probe. For the surface magneto-optic Kerr effect (SMOKE) measurement, the beam size is ~0.1 mm so that the thickness resolution with the above slope can reach ~0.1

[a]Electronic mail: kawakami@physics.berkeley.edu
[b]Electronic mail: qiu@physics.berkeley.edu

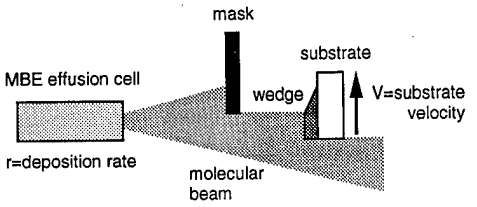


FIG. 1. For MBE systems, wedge samples are made by placing a "knife edge" mask in the molecular beam and translating the substrate behind it with a constant velocity V. The wedge slope is r/V, where r is the deposition rate.

Å. This is good enough to explore any fine structures on the atomic scale. A 1-cm-long wedge is then equivalent to ~100 uniform thickness samples.

Wedged samples were first applied to study the short-period oscillations of magnetic coupling in magnetic multilayers,[4] and then were extended to study the spin-reorientation transition[5] and many other topics. In this article, we report our results on the investigation of the metastable fcc Fe films by using wedged samples. fcc Fe exhibits rich structural and magnetic phases that sensitively depend on the film thickness. Thus, wedged samples will be perfect for this system to explore the fine features associated with each phase.

II. MOTIVATION

The unique property of transition metal magnets is that there is a high density of states of $3d$ magnetic electrons at the Fermi level. This property makes it possible to manipulate the magnetic phases of transition metals by changing the crystal structure. Fe is the best choice for this kind of study. First, Fe is the most common and the most intensively studied magnetic element. Second, Fe is between the antiferromagnetic Mn and the ferromagnetic Co in the periodic table. Thus, it has the intrinsic instability to fall into different magnetic phases. It is well known that bulk Fe has bcc phase at room temperature and fcc phase (γ-Fe) at high temperatures (>1180 K). The room temperature lattice constant (~3.59 Å) of fcc Fe (extrapolated from high temperature) is very close to Cu (3.61 Å). This lattice matching led to the idea of stabilizing fcc Fe at low temperatures, either as small γ-Fe precipitates in a Cu matrix[6] or as an epitaxial ultrathin film on Cu(100). While the antiferromagnetic γ-Fe precipitates show a low magnetic moment ($\sim 0.7\mu_B$) with a Néel temperature T_N=67 K, the magnetic phase of the epitaxial fcc Fe film, however, is very complicated and seems to depend strongly on the substrate temperature. For example, while the antiferromagnetic[7] and nonmagnetic[8] phases were reported for fcc-Fe/Cu(100), a ferromagnetic phase of fcc-Fe/Cu(100) was also identified.[9] Particularly striking is the recent report of the $(n\times 1)$ superstructure on the surface of Fe/Cu(100)[10] and a ferromagnetic live surface layer for ~5–11 ML Fe/Cu(100)[11] grown at room temperature. The sublayer below the "live layer" was also identified to be antiferromagnetic with T_N~200 K.[12] This rich variety of properties indicates that fcc Fe on Cu(100) may just be at a magnetic instability point. Band structure calculations showed[13] that fcc Fe has more magnetic structure than bcc Fe. For r_{ws} (Wigner–Seitz radius) in the range of 2.3–2.8 a.u., the total energy for bcc Fe has only one minima, corresponding to a ferromagnetic state; the total energy for fcc Fe has two minima, corresponding to a low-spin (LS) state and a high-spin state (HS). The instability point between the LS and HS states occurs at $r_{ws}\approx 2.66$ a.u. Noting that the Wigner–Seitz radius is 2.652 a.u. for fcc Fe and 2.667 a.u. for Cu, and that the substrate temperature determines the interlayer mixture, which in turn affects the overlayer lattice relaxation, it is therefore very interesting to ask if the complicated magnetic phases for fcc Fe/Cu(100) is driven by the LS–HS instability. The experimental difficulty is that the effect of the interfacial electronic hybridization mixes with the effect of the interfacial lattice relaxation. The fcc Fe overlayer shows a very complicated lattice structure,[14] making it extremely difficult to analyze the different effects.

In light of these results, there are two natural directions of investigation. The first possibility is to slightly alter the structure of the substrate while holding its magnetic properties fixed. This direction has already yielded attention.[15] The second possibility is to alter the magnetic properties of the substrate while holding the structural properties fixed. This is the subject of our present study. In an attempt to alter the magnetic properties of fcc Fe without further changing the structural properties observed in Fe/Cu(100), Fe was epitaxially grown on an fcc Co(100) substrate. The fcc Co(100) substrate was made by depositing several monolayers of Co on top of Cu(100). Since fcc Co has a similar lattice constant to Cu, they are structurally very similar. However, fcc Co is ferromagnetic at room temperature.

To further study the nature of fcc Fe, we used it as a spacer layer between two Co layers. This was made possible because of the existence of a nonferromagnetic phase of fcc Fe (described below). Since the discovery of the AFC of two ferromagnetic layers across a nonferromagnetic spacer layer, there has been extensive work on the coupling across many types of spacer layers. Antiferromagnetic materials (Cr,Mn),[16] semiconductors (Si),[17] and many nonmagnetic transition metals[18] have been used as spacer layers to provide AFC and coupling strength oscillations. The coupling of the Co/Fe/Co system could be compared to these systems to help determine the magnetic structure of the interlayer.

III. EXPERIMENT

All samples were epitaxially grown in an ultra-high vacuum (UHV) chamber with a base pressure of $\sim 2\times 10^{-10}$ Torr. The UHV system is equipped with four effusion cells for MBE, quartz thickness monitor, reflection high energy electron diffraction (RHEED), low energy electron diffraction (LEED), Auger electron spectroscopy (AES), Ar^+ ion sputtering, and an *in situ* SMOKE setup. A sample manipulator provides a temperature range of 120–900 K and permits the use of all the above techniques at different stages of the

chamber. With this system, we are able to perform the growth, characterization, and the magnetic measurement inside a single chamber.

A single crystal Cu(100) disk of ~1 cm diameter and ~2 mm thickness serves as the starting substrate. The substrate was mechanically polished down to 0.25 μm diamond paste and was finished with vibration polishing using 0.05 μm Al_2O_3 powder. The substrate was then cleaned in UHV with cycles of Ar^+ ions sputtering at 2–5 keV and annealing at 920 K for several hours. AES was used to determine the presence of impurities to insure the cleanliness of the surface. For the growth of Fe on a dirty substrate, we observed as many as 40 RHEED oscillations accompanied by a 2×2 reconstruction in the LEED pattern, consistent with the report of Li *et al.*[19] A uniform thickness Co layer was grown first on Cu(100) to serve as the substrate. Then, Fe wedges were grown on Co using the method shown in Fig. 1. A typical evaporation rate for Co and Fe is ~0.5–1.0 Å/min. A typical slope of the Fe wedge is ~2–10 Å/mm.

Magnetic properties of the films were measured by SMOKE with an intensity stabilized He–Ne laser. The SMOKE setup consists of two pairs of electromagnets that can apply a magnetic field up to 1.9 kOe, either perpendicular to the film plane to generate the polar signal, or in the film plane to generate the longitudinal signal. The incident angle of the laser is fixed at ~45° to allow for both measurements. The incident beam is usually *p* polarized and is focused by an optical lens onto the sample surface. After reflection, the beam intensity is measured by a photodiode behind a linear polarizer, which is set at ~1° from extinction, to generate the magnetic hysteresis loop. There is also a quarter-waveplate in the reflection path to cancel the birefringence of the UHV windows. The sample is transferred up and down along the wedge for the thickness dependent study.

IV. RESULTS

A. Fe/Co(100)

First we needed to determine how many layers of Co should be deposited onto Cu(100) to allow it to be considered a "Co(100) substrate." We did this by growing identical Fe wedges on different thickness of Co layers (2–20 ML). The longitudinal Kerr signals of two of these samples (6 and 10 ML Co) are shown in Fig. 2. The two curves have a similar shape, but are offset from each other due to the Kerr signal from the uniform Co layers. If the Co Kerr signal is subtracted from both, the curves coincide precisely. Therefore, 6 ML of Co is enough for all the features of the Fe/Co/Cu(100) system to be present. The choice of 6 ML as the Co thickness for most of our experiments was purely arbitrary. We will discuss the magnetic properties in detail later.

Our next concern was to test the validity of our conjecture that Fe/Co(100) should have similar structural properties as Fe/Cu(100). The structural properties of the room temperature growth Fe/Co(100) system were investigated by RHEED and LEED.

RHEED intensity was monitored during the growth of uniform thickness Fe/Co(100) and Fe/Cu(100) films. Figure 3 shows that the oscillations for both systems are similar. Three regions can be identified in both samples: a region of erratic oscillations (region I), followed by a region with fairly periodic and constant oscillations (region II), which gives way to a third region where the oscillations disappear and the RHEED intensity is greatly decreased (region III). The RHEED diffraction patterns for regions I and II are consistent with an fcc-like crystal structure, while the pattern for region III is consistent with bcc islandlike growth. Studies performed on the Fe/Cu(100) system[10,12,19,20] have yielded similar RHEED oscillation patterns. Other studies performed on the Fe/Cu system have shown that region I for that system has a pseudo-fct crystal structure, while regions II and III are fcc and bcc, respectively.[10,11,14,21]

A study of the surface structure in regions I and II was performed using LEED. A sample in which layers of uniform Fe were grown on top of Co(100) was used, taking a LEED

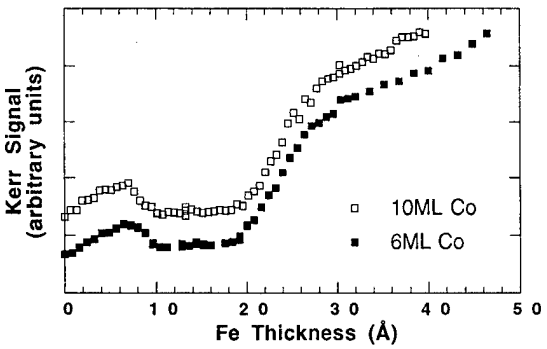


FIG. 2. Longitudinal Kerr signals of Fe wedges grown on different thicknesses of fcc Co(100) at room temperature. A study of the effect of Co thickness on the Fe/Co(100) system was performed by growing identical Fe wedges on different thicknesses of fcc Co(100). Two typical curves for 6 and 10 ML of Co are shown. Only in-plane magnetization was found in these samples throughout the whole thickness range.

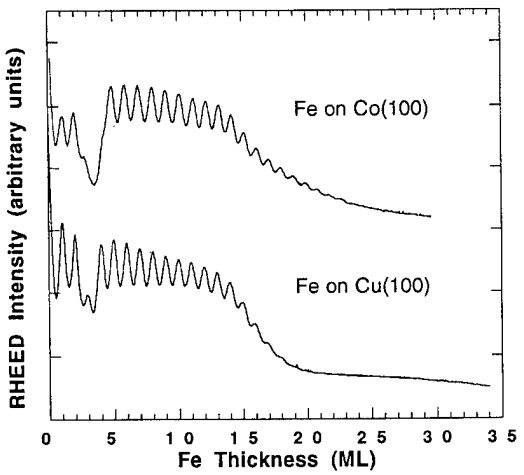


FIG. 3. Typical RHEED intensity variations during Fe deposition on fcc Co(100) and Cu(100). Samples were grown at room temperature with a growth pressure $<6\times10^{-9}$ Torr. Different oscillatory behaviors identify three distinct regions. RHEED and LEED patterns show no difference between the Fe/Co(100) and Fe/Cu(100) systems.

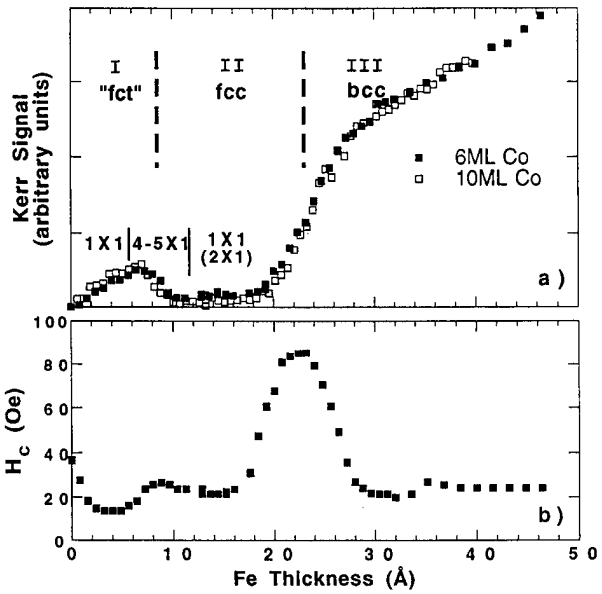


FIG. 4. (a) The same Kerr signals as in Fig. 2 with the Co backgrounds subtracted. The curves are identical for both Co thicknesses. Measurements were taken at room temperature. Note that the magneto-optic response in regions I and III falls roughly on the same line. (b) Cusps in the coercivity of the SMOKE loops were observed to correspond to the transitions between regions.

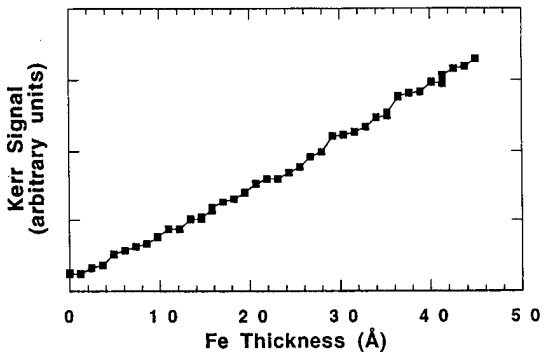


FIG. 5. Low temperature growth Fe/Co(100). Room temperature SMOKE measurement shows longitudinal Kerr signal increasing linearly with Fe thickness. The nonferromagnetic phase seen in the room temperature growth sample does not exist. The linear dependence across the fcc-to-bcc transition suggests that the magnetic moment for fcc Fe and bcc Fe are the same.

picture after each successive growth. LEED patterns for 0, 3.6 (2 ML), 6.3 (3.5 ML), and 10.8 Å (6 ML) Fe on Co(100) were carefully studied. 0 and 3.6 Å Fe (region I) showed a sharp 1×1 reconstruction. The LEED pattern for 6.3 Å Fe (region I/II transition), showed a 4×1 or 5×1 reconstruction. Finally, the pattern for 10.8 Å Fe (region II) showed a 1×1 or a weak 2×1 reconstruction. Wedged samples of Fe on Co(100) were also studied with LEED, yielding similar results. Figure 4 includes a summary of the results for the LEED study. Experiments performed on the Fe/Cu(100) system yield similar results: 1×1 reconstruction for region I, 4–5×1 reconstruction for the region I/II transition, and 2×1 reconstruction for region II.

The similarity between the Fe/Cu and Fe/Co systems in the RHEED and LEED studies is consistent with the assertion that Fe/Co(100) is structurally similar to Fe/Cu(100). Hence, we believe that regions I, II, and III of Fe/Co(100) have fct, fcc, and bcc crystal structures, respectively, as in the Fe/Cu(100) system. This also suggests that the magnetic properties of the substrate do not affect the growth of the fcc Fe film.

Magnetic measurements were performed using the *in situ* SMOKE technique (Fig. 4). Each of the three structural regions was observed to have a distinct magnetic behavior. In contrast to the room temperature growth Fe/Cu(100) system, where the magnetization is out-of-plane in the first two regions with an ordering temperature below room temperature,[10,12,19,20] the room temperature growth Fe/Co(100) system exhibits in-plane magnetization throughout the thickness range with an ordering temperature well above room temperature.

Region I, which goes from 0 to 10 Å (6 ML) of Fe, is ferromagnetic, with the Kerr signal increasing linearly with the Fe thickness (d_{Fe}). The Kerr signal drops to a small and constant level in region II, which goes from 10–20 Å (6–11 ML) of Fe. In region III, the Kerr signal increases and again acquires a ferromagnetic character, again increasing linearly with d_{Fe}. Since all three regions exhibit in-plane magnetization, we are able to compare directly the magnitudes of the Kerr signals in these three regions. It is interesting to note that the magneto-optic responses for regions I and III fall roughly on the same line, suggesting that the Fe film retains the magnitude of the magnetic moment as its structure changes from fcc to bcc.

Since both Co and Cu produce the same crystal structure for the Fe overlayer, we attribute the different magnetic behavior of the Fe/Co(100) and Fe/Cu(100) systems to the magnetic nature of the Co substrate. We believe that it is the direct polarization of the Fe by the Co moment at the interface that causes the in-plane magnetization and the enhanced ordering temperature in Fe/Co system as compared to the Fe/Cu system. To support our idea about direct coupling between Fe and Co, we investigated the shape of the magnetic hysteresis loop of the Fe/Co film. If the Fe overlayer were not coupled to the Co underlayer, a hysteresis loop for a Fe/Co film would have a step at the fields corresponding to the coercivities of Fe and Co. What we observe, however, is a single square loop with the coercivity shifted from that of Co to a new value as d_{Fe} increases. The coercivity of the Fe/Co system versus d_{Fe} is shown in Fig. 4. Similar to the Fe/Cu system, as reported by Li *et al.*,[12] we observed cusps in the coercivity at the transitions between the three structural regions. Recalling that the area enclosed by the hysteresis loop is the work done by the external magnetic field, the observed cusps in the coercivity are a signature of the energy dissipation associated with the phase transition.

We also studied low temperature growth of Fe/Co(100). After growing the Fe film at 120 K and warming up the sample to room temperature, a SMOKE measurement was performed (Fig. 5). Again, only the longitudinal signal is present. Different from the room temperature growth film,

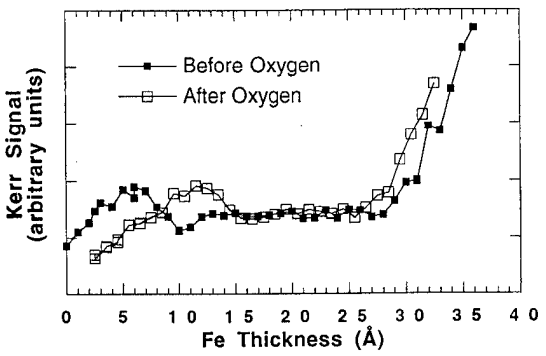


FIG. 6. Oxygen absorption experiment performed to ascertain the origin of the magnetic live layers in region II. SMOKE scans before and after oxygen absorption are shown. The shift in region I is evidence that the oxygen reacted with the surface. The magnitude of the Kerr signal in region II remains unchanged after absorption, implying that the magnetic live layers in this region are not at the surface.

the low temperature growth film shows a linear increase of the Kerr signal through the entire range of the Fe film thickness, indicating that only the ferromagnetic phase is present. This result is similar to that of the low temperature growth Fe/Cu(100) system,[12,14] further confirming our speculation that the structural properties of the fcc Fe overlayer are the same for both the Fe/Co(100) and the Fe/Cu(100) systems. In addition, the same linear dependence of the SMOKE signal as the Fe film crosses the fcc-to-bcc boundary suggests the same magnetic moment of fcc Fe as compared with bcc Fe.

B. Oxygen absorption on Fe/Co(100)

For the room temperature growth Fe/Co(100), region II seems to be the most interesting of the three regions found in this system. The fact that the Kerr signal is constant throughout the whole region seems to imply that there are only a few layers that are magnetically live. A similar live layer was also found in the Fe/Cu system below room temperature and seems to come from the surface layer of the Fe film. Since the live layer in our Fe/Co system was observed at room temperature and there exists a direct magnetic coupling between Fe and Co, we speculate that the live layer in the Fe/Co system originates from the interface of Fe and Co instead of from the surface layer of the Fe. A simple experiment was performed to determine whether or not the magnetic live layer is at the surface of the film. An Fe wedge grown on 6 ML Co was prepared and SMOKE measurements were performed just as described previously. A small amount of oxygen was leaked into the chamber so that it would react with the surface of our sample (a few minutes at 10^{-7} Torr). In principle, the oxygen would react with a few surface layers of our sample and if the magnetic live layers of the fcc region were located at the surface, we would see a change in the Kerr signal. A complete SMOKE scan across the wedge was performed after the oxygen absorption. The results of the SMOKE scans both before and after oxygen absorption are shown in Fig. 6. First, we notice that there is a reduction of magnetization in region I, indicating that the absorbed oxygen indeed kills the surface magnetism. Second, the transition thickness from region I to region II increases. This is not surprising, since both regions I and II are metastable phases and the transition between them is very sensitive to the surface structure, so that the oxygen atoms at the surface should effectively change the transition thickness. The fact that the transition to region III does not shift too much can be attributed to the fact that surface changes are less likely to cause a shift from the more stable bcc phase. The most salient result is that the Kerr signal in region II remains the same, supporting our speculation that the magnetic live layers for this region are from the Fe/Co interface rather than from the Fe surface.

C. Coupling in Co/fcc Fe/Co on Cu(100)

Fe normally cannot be used as the spacer layer to study interlayer coupling. Because Fe is a ferromagnetic material, the interlayer coupling will always be ferromagnetic and the coupling strength will not depend on the spacer thickness.

On the other hand, we have seen that fcc Fe possesses a thickness range where it is not ferromagnetic (region II). Thus, it is feasible to use it as a spacer layer. Because fcc Fe can be grown on fcc Co, the appropriate system to study is Co/Fe/Co on Cu(100). Using the extensive work on AFC as a reference, this system may provide clues as to the magnetic nature of the fcc Fe.

"Symmetric" Co/Fe/Co sandwiches on Cu(100) with Co thicknesses of 13, 15, 20, and 40 ML were fabricated at room temperature. The Co layers were of uniform thickness and the Fe layer was a wedge.

Figure 7 shows representative longitudinal SMOKE hysteresis loops observed in a Co(15 ML)/Fe/Co(15 ML) sample at room temperature. Figure 7(a) shows ferromagnetic coupling of the Co layers for d_{Fe} corresponding to region I of Fig. 4. Figure 7(b) shows strong AFC in the region I/II transition. The AFC is strong enough that we cannot saturate the magnetization with our 1.9 kOe magnetic field. Figure 7(c) shows weak AFC for d_{Fe} in region II. It was found that the splitting of the loops changed very little as d_{Fe} varied from ~10 to ~15 Å. Figure 7(d) shows a "stretched out" AFC loop, for d_{Fe} in the thicker part of region II. This loop will be discussed later.

Loops 7(a) and 7(b) were easily reproducible, but loop 7(c) was not easily reproducible. Typically the loop was not well split. We hypothesized that structural instabilities of the fcc Fe might allow restructuring of the Fe as the Co overlayer is deposited. We based this on several observations.

First, we believed that a structural instability would not arise from the Co underlayer because the structural and magnetic properties of Fig. 4 were easily reproducible. In fact, for each of our sandwiches, we performed quick SMOKE measurements before putting on the Co overlayer. We found that for all samples, the qualitative features of Fig. 4 persisted—a linear increase in region I, and a drop to a roughly constant signal in region II. In addition, the LEED patterns of the Fe/Co/Cu(100) systems were sharp. During

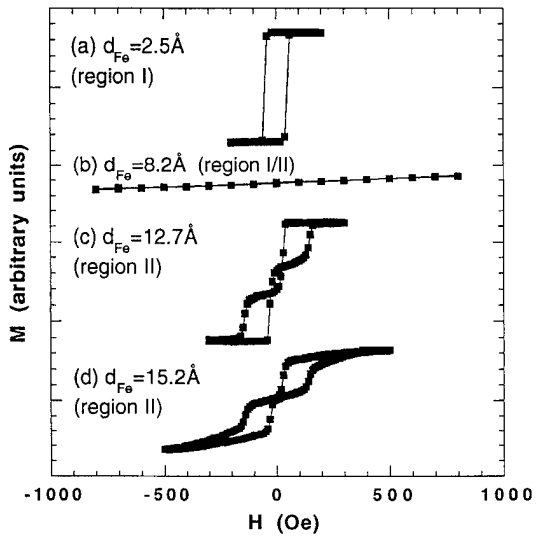


FIG. 7. Longitudinal SMOKE hysteresis loops for Co(15 ML)/Fe/Co(15 ML). Many Fe thicknesses (d_{Fe}) were studied by using a wedged Fe layer. Four representative loops are shown: (a) d_{Fe}=2.5 Å. Ferromagnetic coupling (region I). (b) d_{Fe}=8.2 Å. Strong AFC (region I/II transition). We were not able to saturate the magnetization with our electromagnets (H<1.9 kOe) (Note: H greater than 600 Oe caused our sample manipulator to move. We subtracted out the background signal due to this motion.) (c) d_{Fe}=12.7 Å. Weak AFC (region II). (d) d_{Fe}=15.2 Å. The weak AFC is beginning to get "stretched out," signifying either the beginning of the fcc-bcc transition or a strong AFC oscillation. Loops (a) and (b) were easily reproducible, while (c) and (d) were not.

the growth of the Co overlayer, however, we were never able to obtain RHEED oscillations. This contrasts with the Co underlayer growth, where we were able to get ~30 oscillations with very little amplitude drop. Also, after growth, the LEED patterns became diffuse.

To further support our idea, we investigated the properties of Co on fcc Fe without the Co underlayer. This allowed us to remove any effect that the Co underlayer may have had, thus isolating the effect of the Co overlayer. We first made an Fe wedge on Cu(100) and then grew a uniform Co layer on top of it. Figure 8 shows the result of a longitudinal SMOKE

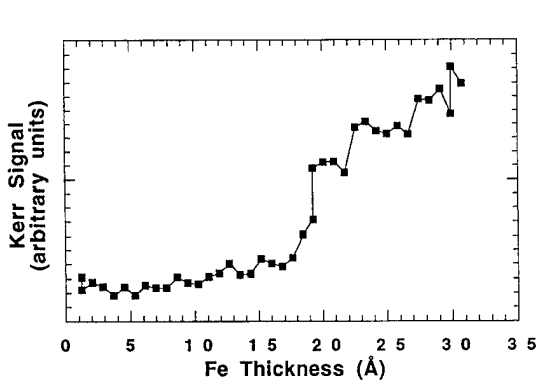


FIG. 8. Longitudinal Kerr signal vs Fe thickness for Co(15 ML)/Fe/Cu(100). Comparing with Fe/Co (Fig. 4) shows qualitative differences. For Fe/Co, the Kerr signal increases linearly with the Fe thickness in region I. For corresponding Fe thicknesses in Co(15 ML)/Fe, the Kerr signal is approximately constant.

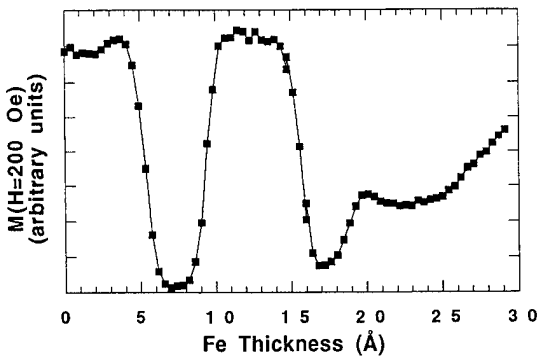


FIG. 9. Longitudinal SMOKE measurement of $M(H$=200 Oe) vs Fe thickness for the Co(15 ML)/Fe/Co(15 ML) sample. 200 Oe is sufficient to saturate the magnetization for the weak AFC and the ferromagnetic coupling across fcc Fe. However, it does not saturate the magnetization for the strong AFC or in the fcc–bcc transition region. Hence, dips in $M(H$=200 Oe) signify the onset of the strong AFC or the fcc–bcc transition. The first dip at ~6 Å is strong AFC. The second dip at ~16 Å is due to either strong AFC or the fcc–bcc transition.

measurement for 15 ML Co. The middle region is flat, like region II of Fe/Co/Cu(100) (Fig. 4). The thicker region has increasing Kerr signal, as in Fig. 4. However, the thinner region does *not* have a linearly increasing signal like region I of Fig. 4. This result is consistent with the idea that the Co overlayer somehow changes the structural properties of the underlying fcc Fe. However, more studies are needed to determine the effect of overlayers on fcc Fe.

D. Oscillations in coupling

Returning to the Co(15 ML)/Fe/Co(15 ML) sample, Fig. 9 shows the magnetization (Kerr signal) at H=200 Oe vs Fe thickness, taken from SMOKE hysteresis loops. This provides an efficient way of representing our data: 200 Oe is sufficient to saturate the magnetization for the weak AFC and for the ferromagnetic coupling across fcc Fe, but it does not saturate the magnetization for the fcc–bcc transition region (due to the cusp in the coercivity) or for the strong AFC. Hence, dips in $M(H$=200 Oe) signify either strong AFC or the fcc–bcc transition. The first dip at d_{Fe}~6 Å is strong AFC. The second dip is ambiguous because of the onset of the fcc–bcc transition.

Thus, loop 7(d), which lies at the beginning of the second dip, poses an interesting question. Does the "stretching out" of the loop signify the beginning of the fcc–bcc transition, or is it a superposition of a weak AFC loop and a strong AFC loop? If it is the latter, perhaps there will be an oscillation back to strong AFC. Unfortunately, we do not find out because the fcc–bcc-transition begins and the structure of the Fe changes drastically.

To examine this issue, we made a compromise. We know from Li *et al.*[19] and our own experience that if the Cu(100) substrate is "dirty" (i.e., not sufficiently cleaned by sputtering and annealing), the fcc phase of Fe can persist to a higher thickness. (For detailed information, please see Refs. 19 and 22.) Of course, the fcc Fe on a "dirty" substrate will be different from the fcc Fe on a clean substrate, but we assume

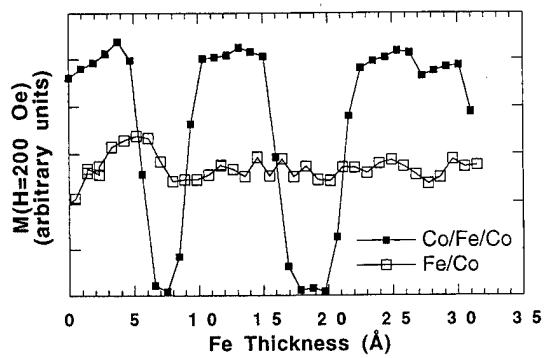


FIG. 10. Longitudinal SMOKE measurements of $M(H=200$ Oe) vs Fe thickness for Fe/Co(6 ML) and Co(6 ML)/Fe/Co(6 ML) on a "dirty" Cu(100) substrate. The "dirty" substrate allows the fcc structure to persist at higher Fe thicknesses, allowing us to search for an oscillation back to strong AFC as suggested in Fig. 9. The Fe/Co measurement verifies that region II is longer than usual; it normally ends at $d_{Fe}\sim 20$ Å. The Co/Fe/Co measurement shows a second dip at $d_{Fe}\sim 19$ Å, signifying an oscillation back to strong AFC.

the difference is negligible with regard to the *existence* of a second strong AFC peak. We do not know the amount of impurity on the "dirty" substrate since the impurity level is too low to be detected by our Auger electron spectroscopy. The signature of a "dirty" film in practice is that region II in Fig. 4 gets bigger and that the surface reconstruction is 2×2 instead of 2×1. Nevertheless, we studied a "dirty" Co(6 ML)/Fe/Co(6 ML) sandwich. Figure 10 shows the Kerr signal before and after the Co overlayer was grown. The Fe/Co signal shows that the fcc region is indeed longer. To study the oscillations of the magnetic coupling in the Co/Fe/Co sandwich, we measured the magnetization (Kerr signal) at a field of 200 Oe, as before. Two dips were observed, each signifying the strong AFC. We conclude that we have found an oscillation back to the strong AFC with a periodicity is ~12 Å. It is interesting to note that the first AFC occurs at the transition point between regions I and II. We do not know if this is a coincidence or if this is somewhat associated with the structural transition (fct–fcc).

V. SUMMARY

We have successfully grown and characterized fcc Fe on fcc Co(100). Structurally, Fe/Co(100) is very similar to Fe/Cu(100). The magnetic phases (e.g., ferromagnetic, paramagnetic, etc.) are similar in both systems, but Fe/Co(100) differs from Fe/Cu(100) because it has in-plane magnetization for all film thicknesses. The nonferromagnetic phase has a magnetic live layer. Oxygen absorption experiments show that the live layer is not at the surface, so it is most likely at the Fe/Co interface. Low temperature growth Fe/Co(100) exhibits a longitudinal Kerr signal that grows linearly with d_{Fe}, even across the fcc-to-bcc boundary. The same linear dependence suggests that the magnetic moments of fcc Fe and bcc Fe are the same. Studies of the coupling in Co/fcc Fe/Co/Cu(100) sandwiches yield ferromagnetic coupling, strong AFC, and weak AFC. We find that Co/Fe/Cu(100) is different from Fe/Co/Cu(100), so the Co layers do not act symmetrically; the difference is possibly caused by a structural instability. By using a "dirty" substrate to artificially lengthen region II, we find an oscillation in the strong AFC.

ACKNOWLEDGMENT

This work was supported in part by DOE under Contract DE-AC03-76SF00098.

[1]P. Grünberg, R. Schreiber, Y. Pang, M. B. Brodsky, and H. Sowers, Phys. Rev. Lett. **57**, 2442 (1986).
[2]M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Greuzet, A. Friedrich, and J. Chazelas, Phys. Rev. Lett. **61**, 2472 (1988).
[3]S. Demokritov, J. A. Wolf, and P. Grünberg, Europhys. Lett. **15**, 881 (1991).
[4]J. Unguris, R. J. Celotta, and D. T. Pierce, Phys. Rev. Lett. **67**, 140 (1991); S. T. Purcell, W. Folkerts, M. T. Johnson, N. W. E. McGee, K. Jager, J. ann de Stegge, W. B. Zeper, W. Hoving, and P. Grünberg, *ibid.* **67**, 903 (1991); Z. Q. Qiu, J. Pearson, A. Berger, and S. D. Bader, *ibid.* **68**, 1398 (1992).
[5]R. Allenspach and A. Bischof, Phys. Rev. Lett. **69**, 3385 (1992); Z. Q. Qiu, J. Pearson, and S. D. Bader, *ibid.* **70**, 1006 (1993).
[6]J. B. Newkirk, Trans. Am. Inst. Min. Metall. Pet. Eng. **209**, 1214 (1957).
[7]W. A. A. Macedo and W. Keune, Phys. Rev. Lett. **61**, 475 (1988).
[8]P. A. Montano *et al.*, Phys. Rev. Lett. **59**, 1041 (1987).
[9]C. Liu, E. R. Moog, and S. D. Bader, Phys. Rev. Lett. **60**, 2422 (1988).
[10]J. Thomassen, F. May, B. Feldmann, M. Wuttig, and H. Ibach, Phys. Rev. Lett. **69**, 3831 (1992).
[11]P. Xhonneux and E. Courtens, Phys. Rev. B **46**, 556 (1992).
[12]D. Li, M. Freitag, J. Pearson, Z. Q. Qiu, and S. D. Bader, Phys. Rev. Lett. **72**, 3112 (1994).
[13]V. L. Moruzzi, P. M. Marcus, K. Schwarz, and P. Mohn, Phys. Rev. B **34**, 1784 (1986).
[14]S. Müller, P. Bayer, C. Reischl, K. Heinz, B. Feldmann, H. Zilgen, and M. Wuttig, Phys. Rev. Lett. **74**, 765 (1995).
[15]F. Baudelet, M.-T. Lin, W. Kuch, K. Meinel, B. Choi, C. M. Schneider, and J. Kirschner, Phys. Rev. B **51**, 12 563 (1995).
[16]J. Unguris, R. J. Celotta, and D. T. Pierce, Phys. Rev. Lett. **67**, 140 (1991); S. T. Purcell, W. Folkerts, M. T. Johnson, N. W. E. McGee, K. Jager, J. ann de Stegge, W. B. Zeper, W. Hoving, and P. Grünberg, *ibid.* **67**, 903 (1991); S. T. Purcell, M. T. Johnson, N. W. E. McGee, R. Coehoorn, and W. Hoving, Phys. Rev. B **45**, 13 064 (1992).
[17]S. Toscano, B. Briner, H. Hopster, and M. Landolt, J. Magn. Magn. Mater. **114**, L6 (1992); E. E. Fullerton, J. E. Mattson, S. R. Lee, C. H. Sowers, Y. Y. Huang, G. Felcher, S. D. Bader, and F. T. Parker, *ibid.* **117**, L301 (1992).
[18]S. S. P. Parkin, Phys. Rev. Lett. **67**, 3598 (1991).
[19]Dongqi Li, M. Freitag, J. Pearson, Z. Q. Qiu, and S. D. Bader, J. Appl. Phys. **76**, 6425 (1994).
[20]M. Wuttig and J. Thomassen, Surf. Sci. **282**, 237 (1993); K. R. Heim, S. D. Healy, Z. J. Yang, J. S. Drucker, G. G. Hembree, and M. R. Scheinfein, J. Appl. Phys. **74**, 7422 (1993).
[21]S. Müller, P. Bayer, A. Kinne, P. Schmailzl, and K. Heinz, Surf. Science **322**, 21 (1995); H. Magnan, D. Chandesris, B. Villette, O. Heckman, and J. Lecante, Phys. Rev. Lett. **67**, 859 (1991).
[22]A. Kirilyuk, J. Giergiel, J. Shen, and J. Kirschner, Phys. Rev. B **52**, 11 672 (1995).

Magnetic x-ray dichroism in the spectroscopy of ultrathin magnetic alloy films

J. G. Tobin[a] and K. W. Goodman
Lawrence Livermore National Laboratory, Livermore, California 94550

G. J. Mankey[b] and R. F. Willis
Physics Department, Pennsylvania State University, University Park, Pennsylvania 16802

J. D. Denlinger, E. Rotenberg, and A. Warwick
Advanced Light Source, Lawrence Berkeley Laboratory, Berkeley, California 94720

(Received 2 October 1996; accepted 6 May 1996)

The magnetic structure of nanoscale alloy films has been probed using the magnetic x-ray linear dichroism in photoelectron spectroscopy and magnetic x-ray circular dichroism in x-ray absorption spectroscopy. FeNi and CoFe epitaxial films were grown on Cu(001), *in situ* and using MBE techniques. Because soft x-rays were used to probe the $2p$ and $3p$ core levels, both elemental selectivity and magnetic sensitivity were achieved simultaneously. Correlation of these magnetic techniques with compositional and structural information will be done. Ultimately, from studies such as this a complete determination of the structure-property relationships will be achievable.
© *1996 American Vacuum Society.*

I. INTRODUCTION

A full elucidation of the underlying principles driving magnetic properties in complex systems will require the applications of probes that couple elemental specificity and magnetic sensitivity. Two important classes of magnetic devices, spin valves,[1] and GMR materials[2] are typically composed of several nanoscale layers or aggregations[3] composed of different elements or alloys. The ultrathin nature of these films or clusters further complicates the picture by introducing the impact of interfacial effects, including pseudomorphic strain and spin-specific scattering at the interfaces.[4] One avenue to address such issues is to build epitaxial ultrathin alloy films and probe them directly using techniques that are both elementally selective and magnetically sensitive. Here we report the beginnings of such a study, using CoFe and FeNi films grown with molecular beam epitaxy (MBE) techniques and investigated with magnetic x-ray linear dichroism[5–7] (MXLD) in the core-level photoelectron spectroscopy of the $3p$ states and magnetic x-ray circular dichroism in the x-ray absorption[8] of the $2p$ states (Fig. 1).

II. EXPERIMENT

The x-ray absorption experiments (Fig. 2) were performed at the Stanford Synchrotron Radiation Laboratory (SSRL) using Beamline 8-2, part of the University of California/National Laboratories facilities.[9] Beamline 8-2, is based on a spherical grating monochromator with only three optical elements reflecting the x rays, including the gratings. Extraction of circularly polarized x-rays was achieved by either raising or lowering the first mirror or insertion of vertical slit for blocking part of the beam.[8,10] As illustrated in Fig. 1, in x-ray absorption an electron moves from a core level into the unoccupied density of states immediately above the Fermi energy. The combination of circularly polarized excitation (x rays), strong dipole selection rules, and spin-polarized valence bands (i.e., magnetized atom) gives rise to strong dichroic effects in these spectra, as we will see below. In this experiment, total yield detection was utilized. Here, the cascade of electrons, including Auger decay electrons, but more predominantly low-energy secondary electrons, causes a net loss of electrons by the sample. Hence, we can measure the absorption of ultrathin films simply by monitoring the replenishing current through a picoammeter (A). Because we are using ultrathin films of only multiple monolayer thickness, x-ray attenuation, and electron escape depth effects will be minimized if not totally eliminated. From these intensity variations, and within a localized picture, it is possible to extract, with varying degrees of precision and accuracy, the spin and orbital components of the elemental magnetic moments.[8] For more detail concerning MXCD absorption, please see Ref. 8 and references therein. Sample cleanliness and composition were monitored with photoelectron spectroscopy using photons of about 500 eV for excitation.

The MXLD photoemission measurements (Fig. 3) were made at the Spectromicroscopy Facility (Beamline 7) of the Advanced Light Source at Lawrence Berkeley Laboratory.[11] Extraordinarily bright, linearly polarized x-rays were generated by the U 5.0 undulator and wavelength selection was achieved using the spherical grating monochromator, with a resolving power of over 8000. The photoelectrons were detected using the angle-resolving, multichannel, 5.4 in. radius, Perkin–Elmer hemispherical deflector system. Sample alignment (including pseudomorphic growth), cleanliness, and composition were measured using the hemispherical deflector and a separate Mg $K\alpha$ (photon energy=1254 eV) source, thus freeing up the beamline for other uses during our periods of sample preparation. The actual MXLD measurements

[a]Electronic mail: Tobin1@LLNL.GOV
[b]Present address: Louisiana State University, Physics Department, Baton Rouge, LA 70803.

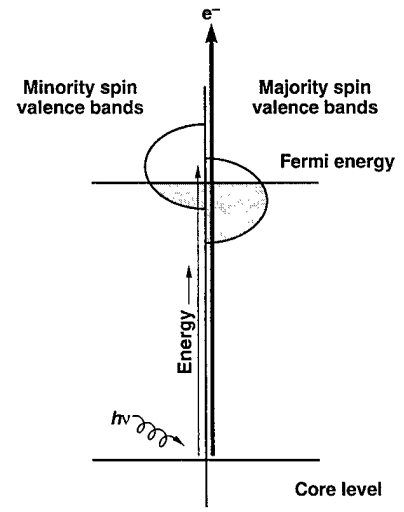


FIG. 1. Shown here is the excitation of a core-level electron in a magnetically polarized atom, i.e., the valence bands are spin polarized. For illustrative purposes only, a circularly polarized photon is shown, with a minority electron moving into the unoccupied valence (conduction) band and a majority electron being ejected. In x-ray absorption (light vertical arrow) near-edge measurements, the electron moves into the strongly polarized valence states (e.g., the $3d$ states), giving rise to strong dichroic effects at the $l+\frac{1}{2}$ and $l-\frac{1}{2}$ edges. In photoelectron spectroscopy (heavy vertical arrow with e^-), the electron moves into states well above the strongly polarized conduction bands and is collected in an energy analyzer.

used the highly linearly polarized synchrotron radiation and were performed with a total instrumental energy resolution of less than 100 meV and angular resolution of 2° or less. The angle of incidence of the x rays was 30° relative to the surface plane. The electrons were collected along the surface normal; i.e., "normal emission." Typically, the magnetic alloy was magnetized in the plane of the surface but perpendicularly to the plane containing the emission direction (surface normal) and the Poynting vector and electric polarization vector of the x rays. Thus, the "transverse-chiral" condition necessary for MXLD was achieved: reversing the magnetic field causes two mirror-image configurations that are equivalent but totally nonsuperimpossible. By comparing spectra from these configurations, it is possible to directly probe the magnetic perturbations of the elementally specific electronic structure of the systems.

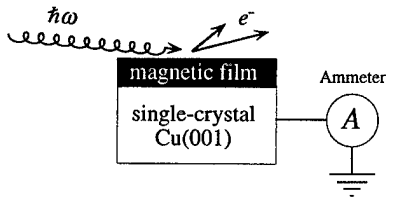


FIG. 2. Shown here is a schematic representation of the MXCD absorption experiment. A total yield method has been used. The picoammeter (A) monitors the grounding current, giving a measure of the absorption in an ultrathin film. It is necessary to align and antialign the magnetization and helicity or circular polarization, hence grazing incidence (15°) is used with in-plane magnetized samples (see Ref. 8).

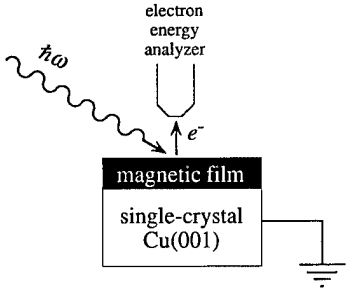


FIG. 3. The MXLD experimental configuration is schematically shown here. The grazing angle of incidence of the linearly polarized x rays was 30°, relative to the surface plane. Electrons ejected perpendicularly to the surface were collected by the angle resolving detector (i.e., normal emission). The plane containing the ejected electron and the Poynting vector of the x ray also contained the electric (linear polarization) vector of the x ray as well. The magnetization vector was perpendicular to this experimental plane, but in the plane of the surface. To achieve MXLD, comparisons were made between spectra in which only the direction of the magnetization was reversed; all else was kept the same.

III. OVERLAYER GROWTH: COMPOSITION, THICKNESS, AND ORDERING

The alloy systems under consideration were CoFe and FeNi. Ultrathin films of each were grown on Cu(001) at room temperature, using well-developed MBE techniques.[12] All samples were about 2–5 monolayers thick. Studies of the bulk electronic structure have been performed,[13–16] but the properties of these pseudomorphic overlayers are further complicated by epitaxial strain and other nanoscale effects. An important issue is growth control, including thickness, composition, and ordering. At the ALS,[11] we can monitor thickness and composition by performing x-ray photoelectron spectroscopy (XPS) and element specific ordering with x-ray photoelectron diffraction (XPD), using the Mg $K\alpha$ x-ray source. Examples of XPS spectra are shown in Fig. 4. The XPD analyses are illustrated in Figs. 5 and 6.

Figure 4 illustrates the type of quantitative information that can be gleaned from XPS. Here, three major sets of peak structure are immediately apparent: The Cu $2p$ ($B^F=933$, 952 eV), the Co $2p$ ($B^F=778$, 793 eV), and the Fe $3p$ ($B^F=707$, 720 eV) (B^F is the binding energy with respect to the Fermi level.) The relative intensities of the Fe and Co lines give a zeroth order estimate of composition, assuming chemical disorder within the alloy overlayer. The relative intensity of the Cu $2p$ lines versus the sum of the Co and Fe lines gives a measure of overlayer thickness, if there is no substrate intermixing or other such complications. A quantitative estimate of thickness requires assumptions concerning x-ray penetration depths and electron escape depths. Quantitative estimates in general necessitate corrections for photon energy dependent cross sections and angles of emission. After properly including such effects, it is possible to arrive at composition and thickness estimates such as those shown in Fig. 4.

A key question in the growth of epitaxial films is that of the degree and nature of ordering, including strain. A means to probe ordering in ultrathin films in an elementally specific

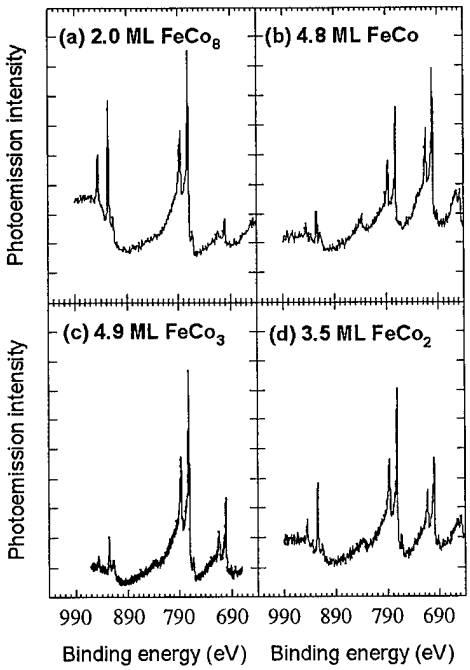


FIG. 4. XPS analysis of various $Fe(1-x)$ $Co(x)$ alloy overlayers on Cu(001). Spectral structure includes the Fe $2p$ ($B^F=707$, 720 eV), Co $2p$ ($B^F=778$, 793 eV) and Cu $2p$ ($B^F=932$, 952 eV). The sample in (c) may be slightly Fe rich at the interface; this effect was not included in the concentration estimate.

fashion is the application of forward focusing in x-ray photoelectron diffraction (XPD).[17–19] Here, buried atoms emit core-level photoelectrons when excited with x rays and intensity maxima correspond to positions of nearby atoms

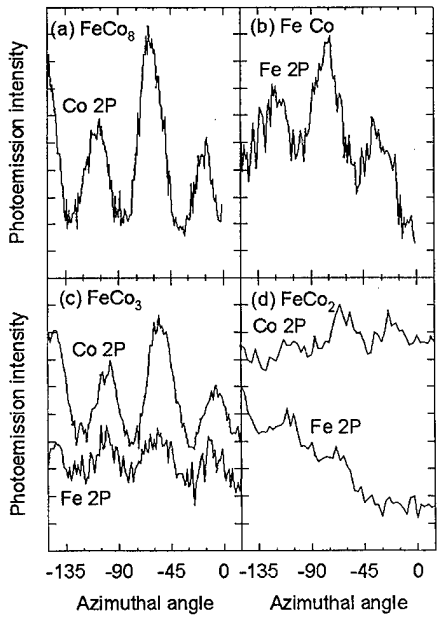


FIG. 5. XPD azimuthal scans are shown here, for a series of FeCo alloy overlayers, with the energy set to the core levels of either Co or Fe. The polar angle is 45°: the diffraction peaks demonstrate the degree to which the overlayer atoms are occupying ordered lattice sites.

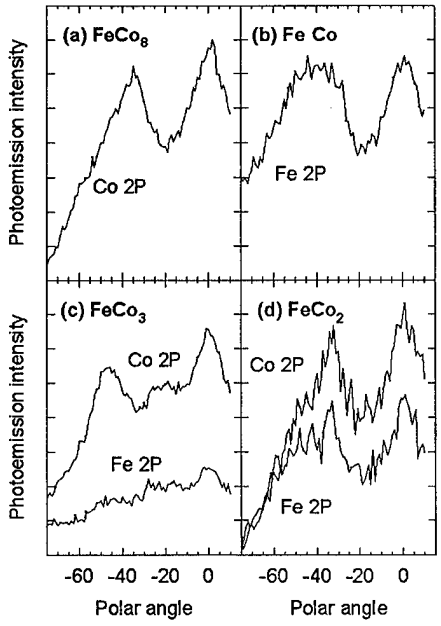


FIG. 6. XPD polar scans are plotted in this figure. The azimuth is set on a peak and the polar angle is scanned. For a, b, and d, the plane of rotation is a {110} plane. For c, it is a {100} plane.

above the emitters. This is because these nearby atoms can act as lenses to focus electrons of fairly high kinetic energies (hundreds of eV or more). Thus, one can immediately identify bond axes or nearest neighbor directions. Of course, complications exist: long chains of atoms can cause defocusing, multiple scattering events can arise, and bulk signals can overwhelm those of a near surface atom of the same elemental nature as the bulk. However, many of these complications are minimized in ultrathin overlayers. In fact, the elemental specificity of the source core level electrons allows for a separation of overlayer (here, Co and Fe) forward focusing from the bulk (Cu) effects and a separation of overlayer components from each other (e.g., Co vs Fe). Thus on can directly probe *overlayer* ordering on an element-by-element basis, independent of substrate ordering effects.

In FeCo alloy films, it is expected that the growth will be FCC-like with the Cu(001) serving as a template. Two tests of this ordering are shown: in Fig. 5 are azimuthal scans and in Fig. 6 are polar scans. FCC growth will give rise to strong forward focusing peaks at a 45° polar angle relative to the normal (0°): this is along the ⟨101⟩ and ⟨011⟩ directions (i.e., the face diagonal). Because of the width of the forward focusing peaks, at 45° polar angle one may also observe part of the ⟨112⟩ and ⟨111⟩ families of peaks, centered at a polar angle of 35° and 55° from the normal (0°), respectively. Thus, an azimuthal scan (polar angle of 45°) will give rise to a series of peaks separated by 45° azimuthally, but possibly with an intensity period of 90° azimuthally. That is what is observed for $FeCo_8$, FeCo, and $FeCo_3$ in Fig. 5, indicating a high degree of ordering. The $FeCo_2$ seems to be more disordered, with only weaker maxima. While a high degree of ordering is indicated, it still needs to be proven that it is

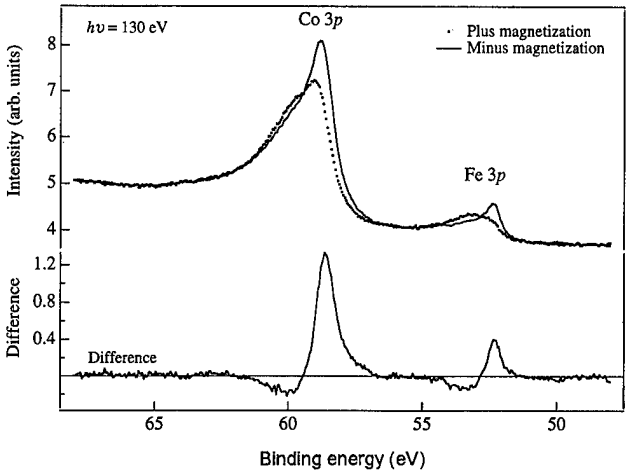


FIG. 7. Photoelectron spectra of 2 ML (Fe[1]Co[9])Cu(001), taken using a photon energy of 130 eV. Also shown is the difference for these two spectra, a measure of the MXLD effect.

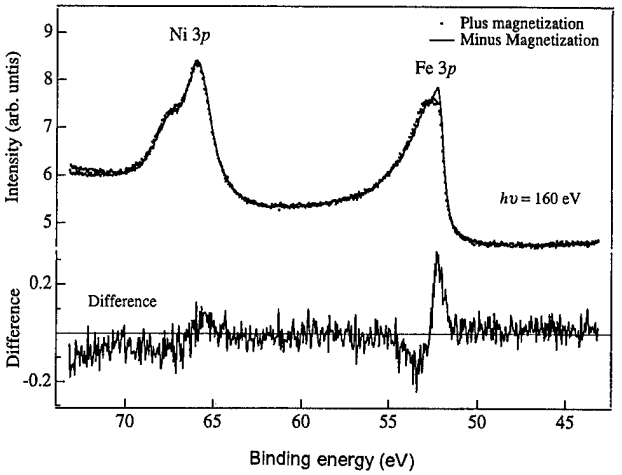


FIG. 8. Photoelectron spectra of 2 ML (Fe[1]Ni[4])Cu(001), taken using a photon energy of 160 eV. Again, a difference spectrum is also shown, as a measure of the MXLD effect.

FCC-like. Unfortunately, BCC growth could give rise to a similar azimuthal pattern, at least qualitatively. A further test of FCC growth is to perform polar scans in the {100} and {110} planes. In the {100} plane, the maximum associated with the ⟨101⟩ should be at 45°: An example of this result is shown in Fig. 6(c). In the {110} plane, the ⟨112⟩ maximum should be at 35° and the ⟨111⟩ maximum at 55°. Figures 6(a), 6(b), and 6(d) show examples of peaks at or near 35°, with possible shoulders at 55°. Normal emission peaks (0°) are also expected and evident in all of the subsections of Fig. 6. In the case of $FeCo_3$ [Fig. 6(c)], the order appears to be stronger around the Co emitters than the Fe emitters, consistent with Fig. 5(c). A complete analysis would, of course, include a full computational simulation and explanation of all features and effects. Nevertheless, the presence of strong maxima indicates a high degree of FCC ordering. It should be noted that while strong structural ordering is indicated here and the source of the photoelectrons can be specified elementally, the scattering events do not have such elemental specificity. Moreover, we assume that while these films are structurally ordered they are chemically disordered, a natural by-product of the simultaneous co-evaporation of Fe and Co onto Cu(001).

IV. MAGNETIC X-RAY DICHROISM

Now, let us turn our attention to the MXLD results. Examples of our spectroscopic MXLD photoemission results are shown in Figs. 7 and 8. In this energy regime, the cross sections for the Fe, Co, and Ni 3P are strongly photon-energy dependent: hence, the peaks do not scale with concentration without further correction. Along with the photo emission spectra are difference curves. From these we can calculate an asymmetry, using the definition $\mathrm{ASYM}=(I^+-I^-)/(I^++I^-)$. Setting each prepeak region equal to zero, asymmetry values of approximately 15% for Co and 40% for Fe in FeCo and 5% for Fe and 2% for Ni in NiFe are obtained.

Presently, we are pursuing studies of these systems where we vary the thickness and composition of the magnetic overlayers and probe the $3p$ levels of each element. In essence, we are attempting to use the MXLD measurements as element-specific, surface magnetometers. This is not quite as straightforward as it might seem: it is clear that simple perturbative (Fig. 9) models often fail.[20,21] For example, if it is possible to work in the regime where $\Delta E_{\mathrm{s.o.}}$, the spin orbit splitting, is large versus ΔE_{EX}, the exchange splitting, then ΔE_{EX} can be treated as perturbation upon $\Delta E_{\mathrm{s.o.}}$. Under these conditions, it seems possible that the variations in asymmetry will correlate with the elemental magnetic moment. A fair candidate for this condition is Ni, where the stronger spin-orbit splitting is evidenced by the peak and shoulder appearance in both spectra in Fig. 8. Fe and Co are probably more difficult cases, where the exchange and spin orbit splittings need to be dealt with on an equal footing, in a non-perturbative approach.[20] However, it may be possible to follow variations around a central value in a perturbative

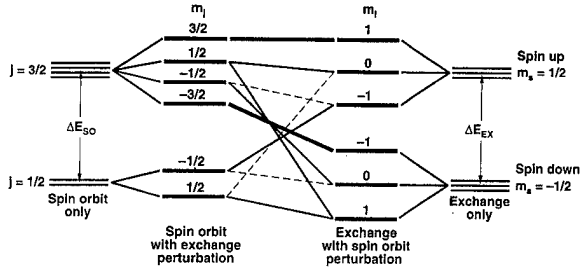


FIG. 9. The effect of spin orbit and exchange splitting upon a p core level is illustrated here. Two perturbative approaches are shown: large spin orbit and small exchange splitting (left) and small spin orbit and large exchange splitting (right). The Ni $3p$ would fall to the left while the Fe $3p$ is near the center, where spin orbit and exchange effects are nearly equal and a non-perturbative approach is required. State mixing is indicated by dashed and solid lines.

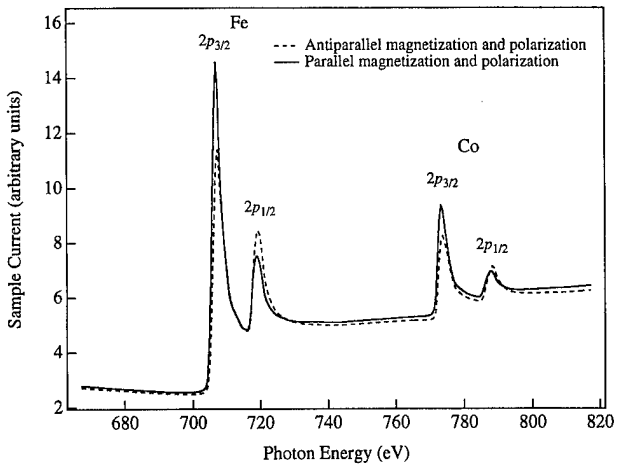


FIG. 10. An example of MXCD absorption is shown here. The 5 ML $(Fe_3Co_1)Cu(001)$ sample was magnetized in-plane. The degree of circular polarization was approximately 70%. This lower-than-usual value was chosen for ease of operation and enhanced counting rates. Corrections for non-ideal polarization (<100%) can be made following Ref. 8. Regardless, strong dichroisms are easily observable here.

fashion.[7] Moreover, we can correlate the MXLD photoemission by comparing that technique with the results from MXCD absorption. An example of our MXCD absorption results is shown in Fig. 10. Here, a strong dichroism is observed at both the Fe $2p$ and Co $2p$ peaks. Within certain limitations,[8] it is possible to extract from such spectra both the spin and orbital components of the projection of the elementally specific magnetic moments along the axis of quantization. (Again, it should be noted that the ultrathin nature of our samples minimizes the impact of electron escape or x-ray attenuation effects.) Although our initial results are only preliminary and further study is warranted, they are promising in terms of understanding key effects in magnetic nanoscale films, including such possibilities as strain-modified invar quenching.

V. SUMMARY

We are using the elementally specific and magnetically sensitive techniques of MXLD photoelectron spectroscopy and MXCD absorption to probe the structure-property relationships of nanoscale magnetic alloys. These ultrathin films are the building blocks of many new magnetic devices, including spin-valve and GMR materials. The variation of classical magnetic effects, such as invar quenching, due to pseudomorphic strain and other nanoscale perturbations, is being investigated.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-ENG-48. The Spectromicroscopy Facility and the Advanced Light Source were constructed with support from the U.S. Department of Energy. The ALS and SSRL are operated under support from the U.S. Department of Energy.

[1]B. A. Gurney, V. S. Speriosu, J. P. Nozieres, H. F. Lefakis, D. R. Wilhoit, and D. U. Need, Phys. Rev. Lett. **71**, 4023 (1993); B. Dieny, V. S. Speriosu, S. Metin, S. S. P. Parkin, B. A. Gurney, P. Baumgart, and D. R. Wilhoit, J. Appl. Phys. **69**, 4774 (1991); B. Dieny, V. S. Speriosu, S. S. P. Parkin, B. A. Gurney, D. R. Wilhoit, and D. Mauri, Phys. Rev. B **43**, 1297 (1991).
[2]S. S. P. Parkin, Phys. Rev. Lett. **71**, 1641 (1993); A. C. Ehrlich, Phys. Rev. Lett. **71**, 2300 (1993); V. Grolier *et al.*, Phys. Rev. Lett. **71**, 3023 (1993).
[3]R. Marks *et al.*, Mater. Res. Soc. Symp. Proc. **313**, 411 (1993).
[4]N. M. Rensing and B. M. Clemens, Mater. Res. Soc. Symp. Proc. **313**, 197 (1993); J. M. George *et al.*, Mater. Res. Soc. Symp. Proc. **313**, 737 (1993).
[5]C. H. Roth, F. U. Hillebrecht, H. B. Rose, and E. Kisker, Phys. Rev. Lett. **70**, 3479 (1993); Solid State Commun. **86**, 647 (1993).
[6]W. Kuch, M. T. Lin, W. Steinhogl, C. M. Schneider, D. Venus, and J. Kirschner, Phys. Rev. B **51**, 609 (1995).
[7]G. Rossi, F. Sirotti, N. A. Cherepkov, F. Cambet-Farnoux, and G. Panaccione, Solid State Commun. **90**, 557 (1994); F. Sirotti and G. Rossi, Phys. Rev. B **49**, 15682 (1994).
[8]J. G. Tobin, G. D. Waddill, A. F. Jankowski, P. A. Sterne, and D. P. Pappas, Phys. Rev. B **52**, 6530 (1995).
[9]K. G. Tirsell and J. Karpenko, Nucl. Intrum. Methods A **291**, 511 (1990).
[10]L. J. Terminello, G. D. Waddill, and J. G. Tobin, Nucl. Instrum. Methods A **317**, 271 (1992).
[11]J. D. Denlinger, E. Rotenberg, T. Warwick, G. Visser, J. Nordgren, J. H. Guo, P. Skytt, S. D. Kevan, K. S. McCutcheon, D. Shuh, J. Bucher, N. Edelstein, J. G. Tobin, and B. P. Tonner, Rev. Sci. Instrum. **66**, 1342 (1995).
[12]G. J. Mankey, M. T. Kief, and R. F. Willis, J. Vac. Sci. Technol. A **9**, 1595 (1991); F. Huang, M. T. Kief, G. J. Mankey, and R. F. Willis, Phys. Rev. B **49**, 3962 (1994).
[13]D. D. Johnson, F. J. Pinski, and J. B. Staunton, J. Appl. Phys. **61**, 3715 (1987).
[14]I. A. Abrikosov, O. Eriksson, P. Soderlind, H. L. Skiver, and B. Johansson, Phys. Rev. B **51**, 1058 (1995).
[15]P. Soderlind, O. Eriksson, B. Johansson, R. C. Albers, and A. M. Boring, Phys. Rev. B **45**, 12911 (1992).
[16]C. Kittel, *Introduction to Solid State Physics*, 6th ed. (Wiley, New York, 1953), p. 437.
[17]W. F. Egelhoff, Jr., Phys. Rev. B **30**, 1052 (1984).
[18]S. A. Chambers, T. J. Wagener, and J. H. Weaver, Phys. Rev. B **36**, 8992 (1987).
[19]H. C. Poon and S. Y. Tong, Phys. Rev. B **30**, 6211 (1984).
[20]E. Tamura, G. D. Waddill, J. G. Tobin, and P. A. Sterne, Phys. Rev. Lett. **63**, 3642 (1992).
[21]J. G. Tobin, G. D. Waddill, E. Tamura, P. A. Sterne, P. J. Bedrossian, X. Guo, and S. Y. Tong, Surf. Rev. Lett. (in press).

Wavelength dependence of the magneto-optical properties of the interfaces of a Au sandwiched (001) Fe film

Wim Geerts[a] and T. Katayama
Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba, Japan

Y. Suzuki
Joint Research Center for Atom Technology National Institute for Advanced Interdisciplinary Research, Higashi, Tsukuba, Japan

J. Childress
NHMFL/Department of Materials Science, University of Florida, Gainesville, Florida 32611

(Received 2 October 1995; accepted 20 November 1995)

The polar ellipticity spectra of the interfaces of epitaxial Fe/Au(001) wedges have been measured. The interface spectra resemble neither the reflection Faraday spectrum of Au nor the thin-film spectra of thin Au sandwiched Fe. The existence of low and high energy peaks suggest that magneto-optical transitions occur from and to the interface states. © *1996 American Vacuum Society.*

I. INTRODUCTION

The study of the magneto-optical (MO) properties of ultrathin films and multilayers is important both from an application as well as a theoretical point of view.[1,2] The dielectric properties of ultrathin films are different from that of bulk material. Strain caused by the difference in lattice parameters between thin film and substrate influences the band structure and can change the MO transitions. Calculation on Ni showed that the position of the lower sp band relative to the Fermi level depends on the lattice parameter.[3] Changing the effective lattice parameter by alloying with a paramagnetic element causes a shift of the corresponding MO peaks.[3,4] The difference in crystal potential of bulk and thin film causes the electrons to reflect at the interfaces,[5] because of the finite dimensions of the thin film perpendicular to the substrate, some electrons will be confined in quantum well states (QWSs).[6–8] Such discontinuous electron dispersion causes new peaks in the MO spectra[9] and the Kerr rotation of ferromagnetic[10] and nonferromagnetic thin films[11,12] oscillate as a function of thickness. Extended overviews on this subject can be found in Ref. 13. In addition to strain and finite size effects, it is also expected that the interfaces themselves will contribute to modifying the MO properties. The redistribution of the free carriers and/or the occurrence of interface electronic states[6,14] might change the effective dielectric properties of the entire thin-film system.

Measurements on polycrystalline Co films sandwiched between Pd and Pt have provided evidence of the existence of such an interface contribution.[15] A non-vanishing Kerr rotation and Kerr ellipticity exists when extrapolating the results to zero film thickness.[15] Furthermore the existence of an interface MO effect is evident from the MO spectra of multilayers and ultrathin film.[16–18] Xu found that the Kerr spectra of Fe/Ag multilayers[19,16] could only be understood by assuming a band splitting in the paramagnetic Ag.[16] Visnovsky *et al.* investigated the MO properties of Au-sandwiched ultrathin Co films. Although most of their results could be explained by the classical electromagnetic wave theory (Maxwell equations), evidence for the existence of interface MO contributions exists.[18] Their interface spectra resembled that of the reflection Faraday spectra of pure Au.

In this article, we report on a systematic investigation of the MO interface effects of Au-sandwiched (001) Fe films. To our knowledge, this is the first time the MO interface effect is studied on epitaxial thin films.[2]

II. EXPERIMENTAL PROCEDURE

Thin Fe wedges were grown by molecular beam epitaxy (MBE). Cleaning of the substrates and deposition conditions are extensively described in.[9,2,20] The total preparation procedure was optimized so that flat films with sharp interfaces were obtained. Because of the perfect matching of the lattice parameters of Au and Fe, the small difference in surface free energy of Au and Fe, and the possible surfactant-effect of the Au atoms, high-quality films could be grown. Two wedges were prepared with the following structure (from surface to substrate). Sample No. 1: Au(20 Å)/Fe(0–20 Å)/Au(2000 Å)/Ag(2000 Å)/polished MgO[(001) orientation]; Sample No. 2: Au(20 Å)/Fe(0–10 Å)/Au(20 Å)/Ag(3000 Å)/polished MgO [(001) orientation]. Figure 1 shows their structure. In both samples, the Fe layer is sandwiched by Au. By changing the thick buffer layer from Au (sample No. 1) to Ag (sample No. 2), we can check optical effects on the spectra[21] (the average penetration depth of the light is about 200 Å).

MO experiments were performed using a system based on the piezobirefringent modulator technique (Jasco-2500 system).[20] The polar Kerr ellipticity was determined along the wedge (at 80 different positions) for different photon energies (1.6,1.8,..,6 eV). The applied field exceeded 1.8 T to ensure saturation of the magnetization for all points along the

[a]Current address: NHMFL/Department of Materials Science, University of Florida, Gainesville, Florida; Electronic mail: wgeer@silica.msl.ufl.edu

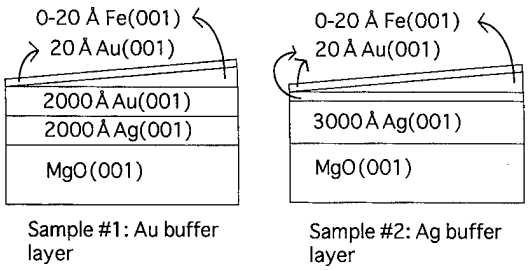


FIG. 1. (a) Wedge structure of sample No. 1 and (b) sample No. 2.

wedge. All measurement conditions were similar to those described in.[10,20]

III. MEASUREMENT RESULTS

Figure 2(a) shows the measured Kerr ellipticity (η_k) as a function of the film thickness (D) for two different photon energies. Three different contributions can be observed

(1) A damped oscillating term originating from the electron confinement in the thin-film structure.[10,22] In both samples, multiple oscillating terms occurred.[23] For most of the photon energies, a short period oscillation (3–5 Å) and a long period oscillation (7–10 Å) could be observed.

(2) An interface term: extrapolation of the curves to zero film thickness results in nonzero Kerr ellipticities[2] (see Fig. 2). It is believed that this MO contribution originates from transitions from or to states which are located near the interfaces of the thin-film structure. Three types of states can be considered.

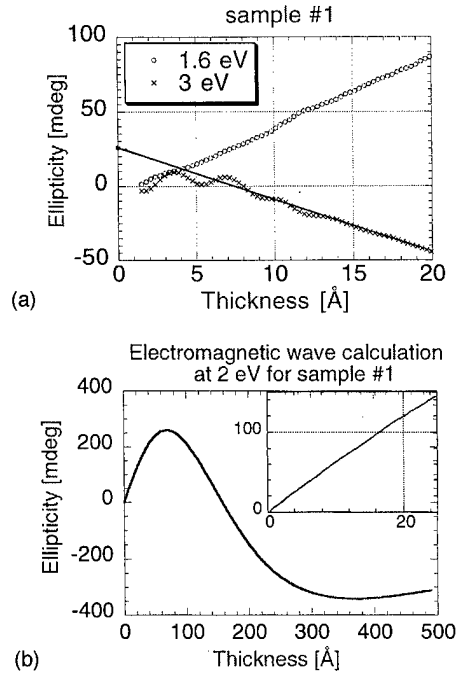


FIG. 2. (a) Measured Kerr ellipticity as a function of the film thickness for several photon energies (sample No. 1); (b) Calculated ellipticity as a function of the film thickness for sample No. 1 (bulk term only).

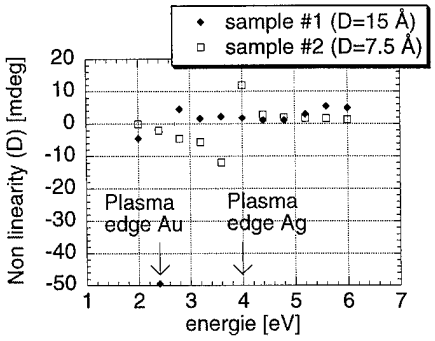


FIG. 3. Calculated photon energy dependence of the non linear bulk term for sample No. 1 (at 15 Å) and sample No. 2 (at 7.5 Å).

(a) The electronic wave function is limited to the interface; i.e., the envelope function of this kind of state has an exponential decay in both the Fe and the Au layers. The eigen energy will be in the band gap of both metals.

(b) The electronic wave function describes an extended state in the Fe layer and penetrates into the Au layer; only in the Au layer, the envelope function will have an exponential decay. The eigen energy of such a state will be in an Fe band but not in a Au band.

(c) The electronic wave function describes an extended state in Au and an exponential decaying one in Fe.

It is expected that the interface states of type (a) and (b) show exchange splitting. As the spin orbit coupling in Au is very large,[24] large interface MO effects are expected in the Au/Fe system.

(3) A *bulk* term originating from the bulk dielectric properties of the thin-film structure. Using optical[25] and magneto-optical[26] experimental data from the literature, we can calculate the expected thickness dependence of this bulk term. A calculation scheme similar to that given in Ref. 21 was used. Some results for sample No. 1 are given in Fig. 2(b). For the relevant thickness range (0–20 Å), the ellipticity can be considered to be linear proportional on the film thickness, which is in agreement with[9,27] [see inset in Fig. 2(b)]. It appears that this linear behavior occurs at all photon energies except for a small interval around the plasma edge of the Au buffer layer (2.3–2.7 eV). For these wavelengths, the adsorption in the Fe layer causes large second-order terms even below a film thickness of 20 Å.[28]

Calculations on sample No. 2 revealed essentially the same results. Except for the energy interval 3.8–4.2 eV (the plasma edge of the Ag buffer layer), the second-order term was negligible when compared to the linear term. Figure 3 shows the nonlinear bulk term (equal to the 2nd-order term at a specified thickness) of sample No. 1 and sample No. 2 at 15 and 7.5 Å, respectively.

The interface term was isolated from the measurement data by a linear regression algorithm. For sample No. 1 (sample No. 2) the data from film thicknesses between 3 and 20 Å (4 and 10 Å) were used. For lower thicknesses, the amplitudes of the finite size oscillations are very large and an accurate determination of the interface term becomes diffi-

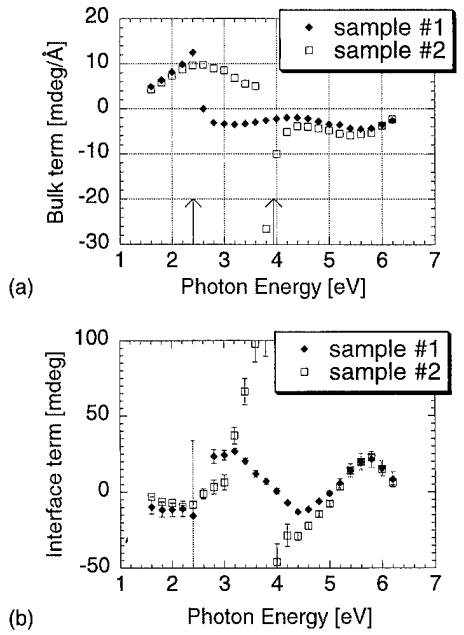


FIG. 4. Measured bulk (a) and interface term (b) as a function of the photon energy for sample No. 1 and sample No. 2.

cult. The results of this analysis are given in Fig. 4. Figure 4(a) gives the spectrum of the calculated bulk term. The shape of both spectra is in agreement with the results of the calculations.[2] The structures around 2.5 and 4 eV originate from a sudden change of the optical properties of the corresponding buffer layers.[21] A first-order linear approximation of the relation between dielectric properties and observable complex Kerr rotation is given by

$$\Phi_k = \phi_k + i\eta_k = 4\pi i D/\lambda \epsilon_{xy}/(1-\epsilon_s) \quad (1)$$

in which $\epsilon_s = \epsilon_{sr} + i\epsilon_{si}$ is the dielectric constant of the buffer layer, $\epsilon_{xy} = \epsilon_{xyr} + i\epsilon_{xyi}$ the nondiagonal component of the dielectric tensor of the thin Fe layer, λ the wavelength of the light, D the film thickness, Φ_k the complex Kerr rotation, ϕ_k the Kerr rotation, and η_k the Kerr ellipticity.

At the plasma edge, the complex dielectric constant (ϵ_s) changes from mainly real to mainly imaginary; we can conclude from Eq. (1) that below the plasma edge, the ellipticity spectra mainly represents the ϵ_{xyr} spectra while above the plasma edge, ϵ_{xyi} will have the largest contribution to the ellipticity.

In Fig. 4(b) the calculated interface MO contribution is presented. The spectra is completely different from the bulk term [Fig. 4(a)]. The sign change at 2.5 eV in the spectrum of sample No. 1 can be attributed to the aforementioned plasma edge effect of the buffer layer. As the spectrum of sample No. 2 changes sign for the same photon energy it can be also concluded that the non-diagonal component of the dielectric tensor (ϵ_{xy}) is changing drastically around 2.5 eV. The same can be said about the structure at 4 eV in the spectrum of sample No. 2.

The nonlinearities of the bulk term shown in Fig. 3 could give rise to a virtual interface term. The magnitude of this virtual interface term is one half the value shown in Fig. 3, and is thus negligible except for values around the plasma edges of the buffer layers. This nonlinear term is also shown in the error bars of Fig. 4(b). Thus, it can be concluded that the interface ellipticity spectra of Fe/Au (001) display peaks around 3 and 5.7 eV.

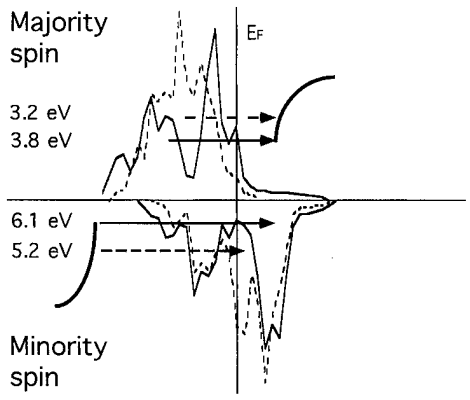


FIG. 5. Band structure model of the Fe-interface atoms in the Au sandwiched Fe(001) structure.

IV. DISCUSSION

The large difference in work function between Au and Fe (5.46 and 4.67 eV)[29] suggests that charge transfer will occur from the Fe to the Au layer. Because of the large density of states around the Fermi level of both materials, the charge redistribution will be limited to the interface, i.e., the dipole layer will be thinner than a monolayer (ML). In this case, it is not clear whether or not it is possible to approximate the band structure of the total system by bending of the bulk bands of the materials involved, and exact band calculations need to be performed. Wang *et al.*[30] and Ohnishi *et al.*[31] performed an *ab initio* self-consistent spin-polarized energy band study of a 7 ML Fe thin film. They found narrowing of the *d* bands near the surface which resulted in a decrease of the density of states of the majority spin electrons. Because of the existence of surface states around the Fermi level in the minority spin *d* band, the minority spin density was largely enhanced near the surface. A similar behavior was found by Li *et al.*[32] for a monolayer of Fe on Au (001): a narrow majority spin band completely under the Fermi level and a large minority spin density at the Fermi level.

Furthermore, Li *et al.* found that the Au interface atoms have a net magnetic moment of 0.03 μ_B (Bohr magneton). A closer look at Fig. 2 of their publication[32] learns that the spin polarization of the Au interface atoms is clearly inhomogeneous. Surface states occur along all symmetry directions. Similar results are expected for a real Fe/Au (001) interface. Figure 5 gives a schematic diagram of the electronic structure of the Fe-interface atoms. The bulk density of states for the *d* band is given for both spin-up and spin-down electrons. The dashed curves give the modifications for the interface area (we neglect the effect of band narrowing). The bold solid curves at the bottom and the top of the figure represent

the dispersion of the *sp* band. The solid arrows show the main optical transitions of bulk Fe.[33–36] The MO effect arises from the combination of spin–orbit interaction and exchange splitting, which results in different absorption coefficient for right-circularly polarized and left-circularly polarized light.[37] Bulk and interface values of both spin–orbit interaction and exchange splitting will be different. The results found in Ref. 3, however, justify that for an impression of the global shape of the spectra, it is sufficient to consider only the change in exchange splitting near the interface. So a global idea of the shape of the MO interface spectrum analog to Ref. 38 can be obtained from Fig. 5. The band structure suggests that, near the surface, transitions will occur to and from the surface states. Transitions from the surface states in the majority spin band to the *sp* band will cause a peak near 3.2 eV. Transitions from the deep *sp* band to the surface states in the minority spin band will cause a structure near 5.2 eV in the MO interface spectrum. This picture is in basic agreement with the experimental results of Fig. 4(b) which show a low and high energy interface peak.

By using the data of Ref. 32 and Ref. 39, a picture similar to Fig. 5 can be constructed for the Au-interface atoms. The calculation of Li *et al.*[32] shows a polarization of the density of states 2 eV below the Fermi level. Because of the large number of possible optical transitions in Au[39] however, an argumentation scheme similar to that used for the Fe-interface (mentioned previously), would lead to the prediction of a lot of peaks in the interface spectrum.

V. CONCLUSION

MO measurements on a Au sandwiched Fe(001)-wedge show the existence of a low and high energy peak in the interface ellipticity spectra. Although the changed electronic structure of the Fe interface can explain the existence of these peaks, the large spin orbit interaction[24] and the exchange splitting of the Au,[32] suggest a contribution from the Au-interface atoms. Experiments on thin Fe films sandwiched between other noble metals (Ag, Cu, etc.) could separate both contributions. Furthermore, theoretical calculations on the electronic and magneto-optical properties of the Fe/Au interface could reveal the exact origin of the presented measurement results.

ACKNOWLEDGMENTS

One of the authors would like to thank the Japan International Science and Technology Exchange Center and the National Institute for Advanced Interdisciplinary Research in Tsukuba.

[1]Y. Suzuki and T. Katayama, J. Magn. Soc. Jpn. **17**, 29 (1993).
[2]W. Geerts, T. Katayama, and Y. Suzuki, J. Magn. Soc. Jpn. **19**, 309 (1995).
[3]P. M. Oppeneer, J. Sticht, T. Maurer, and J. Kubler, Z. Phys. B-Condensed Matter **88**, 309 (1992).
[4]K. H. J. Buschow, P. G. van Engen, and R. Jongebreur, J. Magn. Magn. Mater. **38**, 1 (1983).
[5]For example, M. D. Stiles, Phys. Rev. B **48**, 7238 (1993); P. Bruno, Europhys. Lett. **23**, 615 (1993).
[6]N. B. Brookes, Y. Chang, and P. D. Johnson, Phys. Rev. Lett. **67**, 254 (1991).
[7]F. J. Himpsel, Phys. Rev. B **44**, 5966 (1991); J. E. Ortega, F. J. Himpsel, G. J. Mankey, and R. F. Willis, *ibid.* **47**, 1540 (1993).
[8]K. Koike, T. Furukawa, G. P. Cameron, and Y. Murayama, Jpn. J. Appl. Phys. **33**, L769 (1994).
[9]Y. Suzuki, T. Katayama, K. Tanaka, K. Ando, and S. Yoshida, Phys. Rev. B **50**, 12 581 (1994).
[10]Wim Geerts, Y. Suzuki, T. Katayama, K. Tanaka, K. Ando, and S. Yoshida, Phys. Rev. B **50**, 12581 (1994).
[11]W. R. Bennet, W. Schwarzacher, and W. F. Egelhoff, Jr., Phys. Rev. Lett. **65**, 3169 (1990).
[12]T. Katayama, Y. Suzuki, M. Hayashi, and W. Geerts, J. Appl. Phys. **75**, 6360 (1994).
[13]Y. Suzuki, T. Katayama, W. Geerts, P. Bruno, and H. Sawada, Materials Research Society Spring Meeting '95 San Francisco, CA (unpublished), p. 6.7.1.
[14]A. M. Turner, Yu Jeng Chang, and J. L. Erskine, Phys. Rev. Lett. **5**, 348 (1982).
[15]S. T. Purcell, M. T. Johnson, N. W. E. McGee, J. J. de Vries, W. B. Zeeper, and W. Hooving, J. Appl. Phys. **73**, 1360 (1993); N. W. E. McGee, M. T. Johnson, J. J. de Vries, and J. aan de Stegge, *ibid.* **73**, 3418 (1993).
[16]Y. B. Xu, Q. J. Jin, Y. Zhai, Y. Z. Miao, M. Lu, and H. R. Zhai, J. Magn. Magn. Mater. **126**, 541 (1993).
[17]A. Yamaguchi, S. Iwata, S. Tunashima, and S. Uchiyama, J. Magn. Magn. Mater. **121**, 542 (1993).
[18]S. Visnovsky, M. Nyvlt, V. Prosser, J. Ferre, G. Penissard, D. Renard, and G. Sczigel, J. Magn. Magn. Mater. **128**, 179 (1993).
[19]T. Katayama, H. Awano, and Y. Nishihara, Proceedings of ISPMM'87, Sendai, 1987 (unpublished), p. 283.
[20]Wim Geerts, T. Katayama, and Y. Suzuki, Phys. Rev. B (submitted).
[21] T. Katayama, Y. Suzuki, H. Awano, Y. Nishinara, and N. Koshizuka, Phys. Rev. Lett. **60**, 1426 (1988).
[22]Y. Suzuki and P. Bruno, J. Magn. Magn. Matter. **121**, 248 (1993).
[23] Y. Suzuki, P. Bruno, W. Geerts, and T. Katayama, Phys. Rev. B (submitted).
[24]N. E. Christensen and B. O. Seraphin, Phys. Rev. B **4**, 3321 (1971).
[25]*Physics Data: Optical Properties of Metals*, edited by J. H. Weaver, C. Krafka, D. Lynch, and E. Koch (Karlsruhe, Germany, 1981), Vols. 1 and 2.
[26] N. Okuzawa, T. Katayama, D. Fujitani, W. Geerts, and Y. Suzuki (unpublished).
[27]E. R. Moog, C. Lium S. D. Bader, and J. Zak, Phys. Rev. B **39**, 6949 (1989).
[28]R. Megy, A. Bounouh, Y. Suzuki, P. Beauvillain, P. Bruno, C. Chappert, B. Lecuyer, and P. Veillet, Phys. Rev. B **51**, 5586 (1995).
[29]*Handbook of Chemistry and Physics,* 66th ed. (Chemical Rubber, Boca Raton, FL, 1985).
[30]C. S. Wang and A. J. Freeman, Phys. Rev. B **24**, 4364 (1981).
[31]S. Ohnishi, A. J. Freeman, and W. Meinert, Phys. Rev. B **28**, 6741 (1983).
[32]Chun Li, A. J. Freeman, and C. L. Fu, J. Magn. Magn. Mater. **75**, 201 (1988).
[33]Yu A. Uspenskii and S. V. Khalilov, Sov. Phys. JETP **68**, 588 (1989).
[34] D. G. Laurent, J. Callaway, and C. S. Wang, Phys. Rev. B **20**, 1134 (1979).
[35]M. Alouani, J. M. Koch, and M. A. Kahn, J. Phys. F: Met. Phys. **16**, 473, (1986).
[36]P. M. Oppeneer, T. Maurer, J. Sticht, and J. Kubler, Phys. Rev. B **45**, 10924 (1992).
[37]H. S. Bennet and Edward A. Stern, Phys. Rev. **137**, A448 (1965).
[38] J. L. Erskine and E. A. Stern, Phys. Rev. B **8**, 1239 (1973).
[39]R. Lasser and N. V. Smith, and R. L. Benbow, Phys. Rev. B **24**, 1895 (1981).

Correlation between dynamic magnetic hysteresis loops and nanoscale roughness of ultrathin Co films

Q. Jiang, H.-N. Yang, and G.-C. Wang
Department of Physics, Applied Physics and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180-3590

(Received 13 October 1995; accepted 22 April 1996)

Ultrathin Co films were epitaxially grown on a Cu(001) surface with different initial roughness created by Ar ion sputtering. The roughness of the Cu substrate and the Co film were characterized by high resolution low energy electron diffraction. The measured angular profiles were compared with a diffraction theory for rough surfaces and the roughness parameters were extracted quantitatively. Magnetic hysteresis loops of these characterized films were measured by surface magneto-optic Kerr effect. The hysteresis loop shape and loop area can be related to the nanoscale roughness in the Co films. For the roughest film with interface width $\approx 1.2t$, where t is the single atomic step height, the magnetization is reduced several fold compared with that of smooth films with interface width $\approx 0.5t$. Also, the coercivity in the roughest film is the highest and there exists a wide range of nucleation centers and coercive fields for magnetic domain reversals. These are related to the high step density in the rough substrate as the pinning centers. The hysteresis loop changes its shape and area under a sinusoidal external magnetic field as functions of frequency and field amplitude for all films. For the smooth films in the low frequency and low field regimes the loop area shows a $\frac{2}{3}$ power law scaling behavior. The $\frac{2}{3}$ value of scaling exponents are consistent with the prediction of a dynamic mean field theory with a double-well energy barrier. For a film with the same interface width but different step density and lateral correlation length the scaling exponents deviate from $\frac{2}{3}$ value drastically. © *1996 American Vacuum Society.*

I. INTRODUCTION

The advancement of magnetic characterization in magnetospectroscopy, magnetic imaging (magnetic force microscopy, scanning electron microscopy with polarization analysis (SEMPA), spin polarized scanning tunneling microscopy, spin polarized low energy electron microscopy, x-ray dichroism photoemission microscopy), electron diffraction techniques (spin polarized electron diffraction, spin polarized photoelectron diffraction), material growth and characterization, and large scale computations has provided new opportunities for fundamental research[1] in surface, ultrathin, and thin film magnetism as well as applications to magnetic devices.[2] Issues such as proximity effects at the interface between a magnetic film and a nonmagnetic substrate, the exchange coupling across interfaces, coupling through nonmagnetic layers in trilayers, multilayers, and superlattices, and magnetic domain structures have been actively studied.[3,4] Applications of the giant magnetoresistance (GMR) effect in devices such as hard disk drive read heads are being researched and developed.[5] Activities in magnetoelectronics have also been initiated.[6] The understanding of the magnetic properties of nanoscale structures has grown exponentially in the past decade.

In this article, we present the results of our study of hysteresis loops of heteroepitaxial ultrathin Co films grown on a Cu(001) surface using surface magneto-optic Kerr effect (SMOKE). The SMOKE technique was pioneered by Moog and Bader in a study of Fe on Au(001) surface[7] and has become one of the most simple but powerful techniques used in ultrathin magnetic film characterization. The hysteresis loop provides information on magnetization, coercivity, and anisotropy. Although the ultrathin film is one of the simplest systems being studied, many variables such as substrate roughness, film roughness, film thickness, film temperature, applied field direction, and the amplitude and frequency of periodic applied magnetic field can be changed. These variations allow the study of equilibrium magnetic phase transitions, in-plane or out-of-plane anisotropy, effects of dimensionality on magnetic properties, effects of roughness on magnetic properties, dynamic magnetic phase transitions, etc.

In a heteroepitaxial ultrathin film, the deposited film material may have symmetry and lattice constant different from those of the substrate. This mismatch in lattice constants may induce strain at the interface and roughness at the surface of films. Often the surface free energy of the film is different from that of the substrate. This may cause compositional heterogeneities or intermixing at the interface. It is not unusual to form a metastable epitaxial film or an alloyed film. These possible microstructures may be examined using a structural technique such as high resolution low energy electron diffraction (HRLEED) developed by Henzler's group.[8] This technique is sensitive to atomic arrangements in the surface and ultrathin film regime and is also quantitative in the regime from sub nm to hundreds of nm.

One expects that the deviation of structural properties in the epitaxial film from those of the bulk material might change the magnetic moments, remanent magnetization, magnetic domain structure, anisotropy, coercivity, etc. Indeed, many prior results of structures and magnetic domain structures exist.[9,10] In this article we will present the magnetic properties of Co/Cu(001) obtained from the hysteresis

loops measured by SMOKE under a sinusoidal external magnetic field with various frequencies and field amplitudes. We found that the area and shape of a hysteresis loop depends on how the film is prepared. Different preparations of initial Cu substrates for the growth of films result in different nanoscale roughness in the films. Three kinds of Co films were grown: on a clean smooth Cu substrate, on a lightly sputtered Cu substrate, and on a heavily sputtered Cu surface. The goal of this study is to find out the effect of initial substrate imperfection, which causes interface imperfection and film roughness, on magnetic properties. The quantitative characterization of structural imperfection is obtained from the angular profile measurements of diffraction beams using HR-LEED. The correlation between hysteresis loop shape and nanoscale roughness will be presented.

In addition, the area of a hysteresis loop is measured as a function of either the amplitude or the frequency of an applied sinusoidal magnetic field for all films. The frequency- and amplitude-dependent loops obtained from the Co film on the lightly sputtered surface are qualitatively the same as that of the smooth film grown on smooth surface but differ quantitatively in resonance frequency, coercive field, and scaling exponents. For the smooth film, there exists a dynamic scaling behavior of power law form with well-defined exponents in the low field and low frequency regime.[11] A dynamic magnetic phase transition is observed at high frequencies. The values of the scaling exponents indicate that the dynamic behavior of Co belongs to the Ising universality class.

II. EXPERIMENT

All experiments were performed in an ultrahigh vacuum chamber (UHV). The Cu(001) surface was sputtered by Ar^+ ($\sim$2.0 μA/cm^2 or 1.0×10^{-4} Torr at 500 eV) at room temperature and then annealed at 600 °C for 30 min. The HR-LEED pattern showed a 1×1 pattern after this cycling process. The cleanliness was checked by Auger electron spectroscopy. This is the ordered smooth surface. For the lightly sputtered surface, Ar^+ sputtering was applied at $\sim$0.2 μA/cm^2 or 6.0×10^{-6} Torr partial pressure at 500 eV for 10 min and no annealing. For the heavily sputtered surface (HSS), the sputtering was held for 30 min at Ar^+ pressure of 1.6×10^{-4} Torr or $\sim$2.2 μA/cm^2 or at 500 eV. The surface perfection was then characterized by the analysis of measured angular profiles of the (00) and (01) diffraction beams.

The evaporation of Co is from an electron beam bombarded Co foil. The rate can be varied from 0.01 to 1.0 ML/min by adjusting the temperature of the foil and the electron bombardment voltage. The deposition rates for all films grown on ordered smooth surface, lightly sputtered surface, and heavily sputtered surface were fixed at 0.08 ML/min to minimize possible changes due to different rates. All Co films grown on a Cu(001) surface held at room temperature were epitaxial fcc films as evidenced from the HRLEED diffraction patterns. The pressure during evaporation was $<5.0\times10^{-10}$ Torr. The angular profiles of integer order diffraction beams, (00) and (01), as a function of energy were measured after each monolayer deposition at room temperature. The roughness parameters such as average terrace width, interface width, lateral correlation length, and roughness exponent can be extracted from the analysis of the energy dependent diffraction profiles.

After a certain thickness of Co film was grown at room temperature, magnetic hysteresis loops were measured at room temperature as functions of applied magnetic field amplitude and frequency using a UHV SMOKE setup.[12] The maximum magnetic field strength of the electromagnet can reach 1000 Oe. All the hysteresis loops presented have been saturated before the maximum field amplitude is reached. The magnetic field can be applied to either the [100] parallel or the [001] perpendicular direction of the Cu(001) surface by rotating the magnet inside the UHV. Note that the easy axis direction in bulk Co is along the [011] direction.[9] When measuring hysteresis loops from the Co film, the eddy current effect from the Cu substrate needs to be considered. The major eddy current effect comes from the magnetic core. All the calculations of loop areas have been corrected[11] by the frequency-dependent field strength calibration factor $g(\Omega)$, where $\Omega=2\pi f$ and f is the frequency of the applied sinusoidal magnetic field. Our UHV electromagnet provides a maximum magnetic field strength of 1000 Oe at a very low frequency. All the loops measured and presented here have been saturated.

III. RESULTS AND DISCUSSION

A. Growth of Co films and characterization of film roughness

1. Roughness parameters

To characterize a rough surface, one can employ several parameters: average terrace width η, interface width w, lateral correlation length ξ, local slope ρ, and roughness exponent α. The average terrace width is the average distance between two adjacent steps. The interface width w is defined as the square root of $\langle[h(\mathbf{r})-\langle h\rangle]^2\rangle$, where $h(\mathbf{r})$ is the surface height at position $\mathbf{r}$. $\langle h\rangle$ is the average height. w is the "amplitude" of surface fluctuations (rms roughness) and thus describes the property of the interface at a large distance. The ξ describes the "wavelength" of surface fluctuations. Two surfaces having the same interface width w and the same lateral correlation length ξ may possess very different local roughness structures, which are described by the local slope ρ and roughness exponent α, where $\rho\sim w^{1/\alpha}/\xi$. Details of how to extract these parameters from diffraction profiles can be found elsewhere.[13] We will briefly summarize them in the next paragraph.

2. Diffraction characteristics

All diffraction profiles contain a diffraction structure factor, which is related to the morphology of the growth front. In most cases all the important parameters related to the growth front, w, ξ, ρ, and α can be determined from the diffraction profile. In general, the specularly reflected inten-

TABLE I. Roughness parameters for three Co films with different initial roughness in Cu(001).

	Initial clean Cu terrace width η (Å)	Co film terrace width η (Å)	Interface width w	Lateral correlation length ξ (Å)
OS	300±10	···	···	408±10
Co/OS (3.0 ML)	···	31.4±2	0.54±0.02t	285±7
LSS	44±2	···	0.50±0.02t	218±5
Co/Cu (3.0 ML)	···	25±2	0.50±0.02t	190±5
HSS	10±1	···	1.81±0.05t	96±4
Co/HSS (2.5 ML)	···	12.5±1	1.18±0.05t	27±2

sity can be written as $I(k)\sim\exp[-(k_\perp w)^2]\,\delta(k_\parallel)+I_{\text{diff}}(k_\parallel,k_\perp)$, where $k=k_{\text{out}}-k_{\text{in}}$ is the momentum transfer with k_{out} and k_{in} being the outgoing and incoming wave vectors, and $k_\parallel$ and $k_\perp$ are the momentum transfers parallel and perpendicular to the surface, respectively. The central sharp intensity is the δ component that depends on the interface width and is used for the determination of w. The lateral correlation length ξ is inversely proportional to the full width at half maximum (FWHM) of the diffuse part of the intensity, $I_{\text{diff}}(k_\parallel,k_\perp)$, when $k_\perp$ is small (or near the in-phase condition). For large $k_\perp$ conditions (or near the out-of-phase conditions), the FWHM of the diffuse part of the intensity (diffuse line shape) is proportional to $\rho k_\perp^{1/\alpha}$. One therefore is able to determine both α and ρ by plotting the FWHM as a function of $k_\perp$. The FWHM at the out-of-phase condition is inversely proportional to the average terrace width η and the local slope or step density ρ is equal to $1/\eta$.

3. Angular profiles from ultrathin Co films grown on Cu surfaces with three different initial roughnesses in Cu and decomposition of profiles

Co films were grown on a Cu(001) substrate with three different initial roughnesses: an ordered smooth surface (OS) with a large average terrace width ≈300 Å, a lightly sputtered surface (LSS) with an average terrace width ≈44 Å, and a heavily sputtered surface (HSS) with an average terrace width ≈10 Å. These numbers are listed in Table I. The *in situ* growth of Co films was monitored by measuring the angular profiles of the (00) and (01) diffraction beams at the near out-of-phase condition of $E=150$ and 116 eV, respectively. Figures 1(a), 1(b), and 1(c) show the measured angular intensity profiles of the (00) beam at both in-phase (108 eV) and out-of-phase conditions (150 eV) for about 3-ML-thick Co films on three Cu surfaces with different roughnesses.

For the Co/OS shown in Fig. 1(a), the in-phase profile is basically the instrument function. The near out-of-phase profile is typical of a two level system, i.e., the profile has a central spike and a broad profile. The profile has been decomposed into a narrow instrument function (1D Lorentzian) and a broad profile (2D Lorentzian) plus a constant background shown as solid curves in Fig. 1(a). Note the peak intensity of the in-phase profile is a factor of 6 higher than the out-of-phase peak intensity.

For the Co/LSS, Fig. 1(b) shows the profiles near in-phase and out-of-phase conditions. The best fit indicates that the in-phase profile contains not only the instrument function but also an extra profile. This is an indication of reduced lateral correlation length ξ after sputtering compared with that of the Co/OS. The decomposed profiles near the out-of-phase profile consist of a broader profile with higher intensity compared with that of Fig. 1(a) and a very small component of the central narrow profile. This implies a multilevel film rather than a two-level film. The peak intensity ratio between the in-phase and out-of-phase data is about a factor of 5. The overall peak intensity of the Co film on a lightly sputtered surface is reduced by about a factor of 2 compared with that of the Co film on the ordered smooth surface.

For the Co/HSS, Fig. 1(c) shows two profiles. The in-phase profile broadens considerably and the peak intensity is reduced by a factor of 6 compared with that of Fig. 1(b). The central δ peak has almost disappeared. This implies that the surface after heavy sputtering has a very small lateral correlation length. The FWHM of the broad profile at the out-of-phase condition is further increased compared to the lightly sputtered surface. Again, the central narrow profile is almost gone, an indication of multiple levels in the Co film. The peak intensities for both in-phase and out-of-phase have the same order of magnitude.

4. Roughness parameters obtained from the analysis of diffraction profiles

Figures 2(a) and 2(b) show the peak intensity and the FWHM of broad diffuse profiles of the (10) beam versus Co coverage in a few monolayer (ML) regime for the three different initial substrates. The coverage is determined from the periodic oscillation of the peak intensity and FWHM versus deposition time for Co/OS. The FWHM oscillates periodically for the Co/OS but it only has one very obvious oscillation for both Co/LSS and Co/HSS. The absolute values of the FWHM around 3 ML are ≈0.20, ≈0.25, and ≈0.50 Å^{-1} for Co/OS, Co/LSS, and Co/HSS, respectively. If the FWHM is inversely proportional to the average terrace width in the Co film at the out-of-phase condition, then at about 3 ML, the average terrace widths listed in Table I for the Co films prepared on OS, LSS, and HSS are ≈31.4, ≈25, and ≈12.5 Å, respectively. This is due to the loss of long range order as more sputtering is applied to the Cu substrate. The (10) beam peak intensity versus coverage for both Co/LSS and Co/HSS has one obvious oscillation at the beginning, but the intensity oscillation is not obvious thereafter, in contrast to many nice periodic oscillations usually observed in the case of the layer-by-layer growth. The loss of periodic peak intensity oscillation and the broadening of diffuse profile are consistent with the buildup of a second layer[14] before the completion of the first layer, which results in multilevel films.

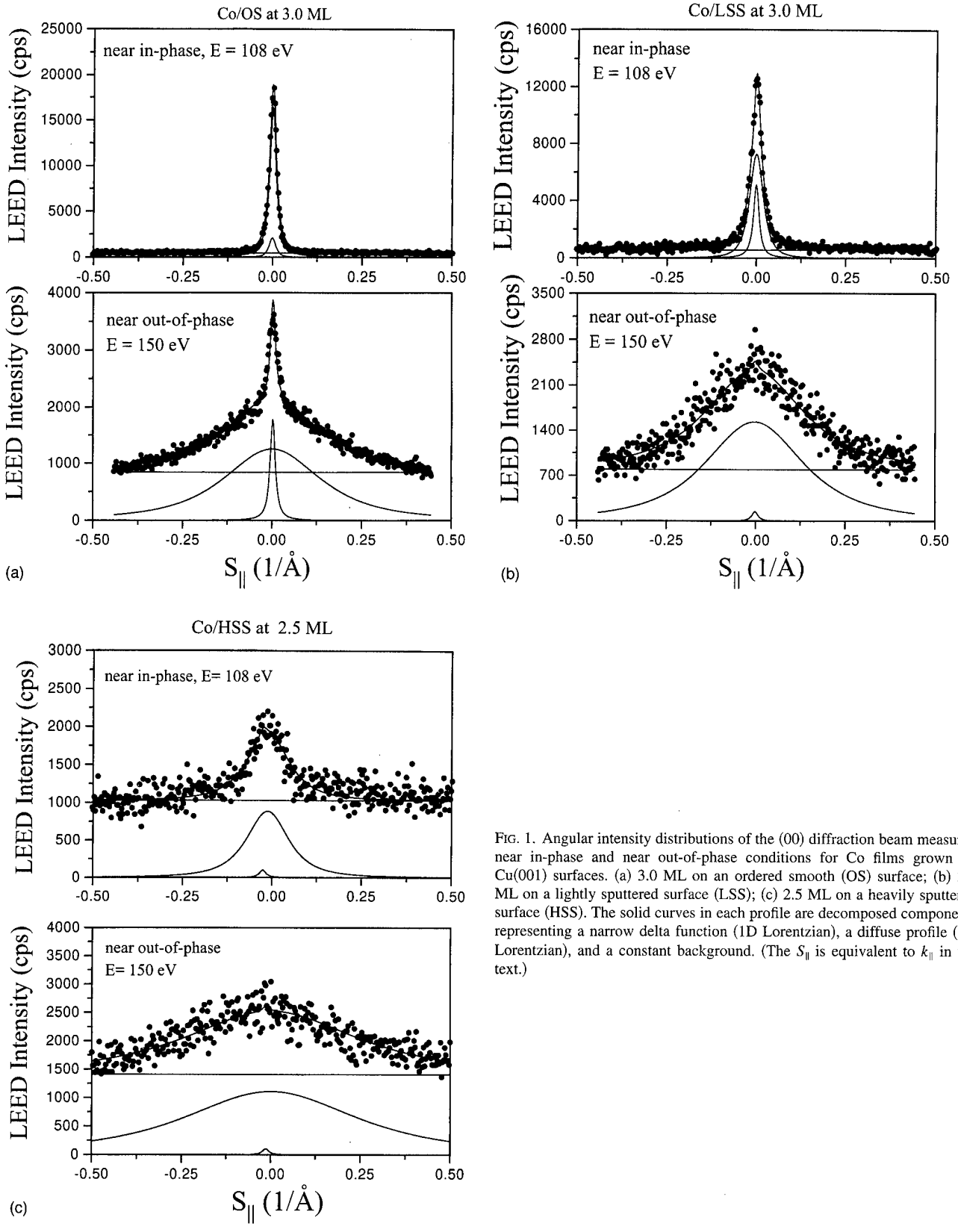


FIG. 1. Angular intensity distributions of the (00) diffraction beam measured near in-phase and near out-of-phase conditions for Co films grown on Cu(001) surfaces. (a) 3.0 ML on an ordered smooth (OS) surface; (b) 3.0 ML on a lightly sputtered surface (LSS); (c) 2.5 ML on a heavily sputtered surface (HSS). The solid curves in each profile are decomposed components representing a narrow delta function (1D Lorentzian), a diffuse profile (2D Lorentzian), and a constant background. (The $S_{\|}$ is equivalent to $k_{\|}$ in the text.)

Assumptions made in the derivation of diffraction profiles from a real space rough surface are that the roughness is random and the surface height distribution is a Gaussian. For the ultrathin Co films grown on an OS, the roughness is not random and there are multilevels. For the Co/LSS, the films are rougher but barely satisfy the assumptions. For the film grown on the HSS, the assumptions may be valid. If we assume that the assumptions made about the roughness are

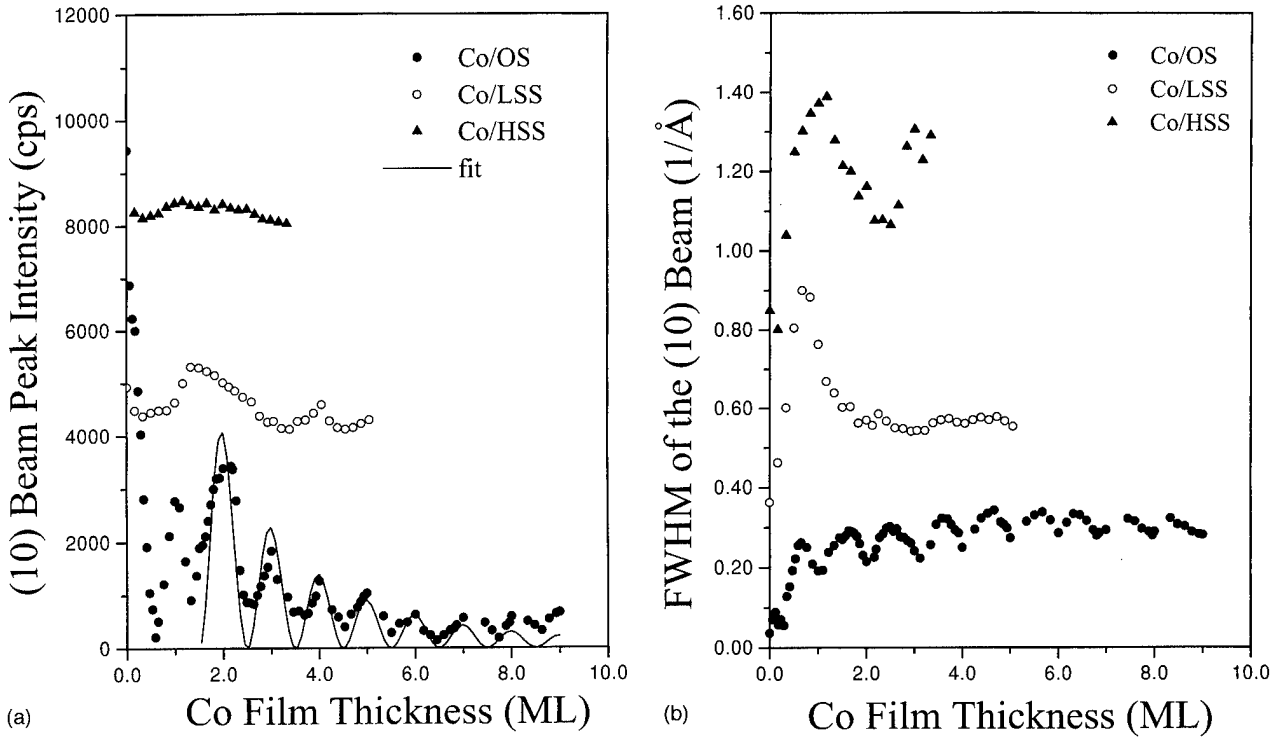


FIG. 2. (a) The peak intensity and (b) the FWHM of the (10) diffraction beam as a function of Co film thickness measured near the out-of-phase condition for Co films grown on ordered smooth, lightly sputtered, and heavily sputtered Cu(001) surfaces. The data for Co/LSS and Co/HSS in (a) are offset from zero by 4000 and 8000 cps, respectively. The data for Co/LSS and Co/HSS in (b) are offset from zero by 0.3 and 0.6 $Å^{-1}$, respectively. The solid curve through the peak intensity data for Co/OS is a fit using the formula $\exp\{-(k_\perp w)^2[1+\cos(2\pi d)]\}$.

valid on the HSS and approximately valid on the LSS, then the interface width is $1.81t$ after heavy sputtering and is $1.18t$ after about 2.5 ML Co film was deposited on the HSS. The t is a single atomic step height of 1.76 Å. This $1.18t$ value is in contrast to the value around $0.5t$ obtained in LSS. The roughness exponents are around 0.6 for Co/HSS and Co/LSS.

If the peak intensity oscillation is layer-by-layer-like, one can use a recently developed analytical formula,[15] $I_{peak}\sim\exp\{-(\pi w)^2[1+\cos(2\pi d)]\}$ where d is the thickness in units of monolayer, to extract the w from the decay of peak oscillations. Figure 2(a) shows the peak intensity of the (10) beam from Co/OS film and the solid curve is the fit. The value of interface width obtained from the fit changes from about $0.2t$ to $0.7t$ in about 10 oscillations. The w obtained from the Co/OS at 3 ML is $\approx 0.54t$ and is listed in Table I.

The surface height distribution can be obtained either from a real space imaging height profile scan or from a G factor analysis of diffraction profiles.[6] At the present time, we have not obtained the real space images. The G factor analysis was done on Co/Cu (001) films, which shows obvious peak intensity oscillations.[17] The vertical distribution or the probability P_i of finding a surface atom in level i decreases as the level number increases.

The chemical inhomogeneity at the interface can be obtained from the analysis of angular profiles if the atomic scattering factors of the two elements differ in strength by at least a factor of 2. Examples have been demonstrated for Co silicide, Ni silicide, Si/SiO_2,[18] and Fe/Au(001).[19] For the Co/Cu, the atomic scattering factors of Co ($Z=27$) and Cu ($Z=29$) are not much different, so an analysis of inhomogeneity is not attempted.

B. Hysteresis loops and analysis

An ultrathin ferromagnetic film exhibits a change in magnetization under an applied magnetic field. A closed magnetization curve under one cycle of applied magnetic field is called the hysteresis loop. It takes work to magnetize a ferromagnetic film. The energy supplied by the work is partially stored as the potential energy and partly lost as heat in the film. Because at the end of a cycle of the loop the potential energy returns to its original value, the work must be consumed as heat in the film. This is called hysteresis loss. The area surrounded by the hysteresis loop is $A=\oint M dH$. The remanent magnetization M_r is the magnetization at $H=0$. The squareness is the ratio of M_r to M_s. The value of M_r and therefore the squareness depends on the saturation magnetization M_s and also on how well the preferred axes of the domains are aligned with the direction in which the magnetization is being measured. For a symmetrical loop about the origin ($H=0$), the coercive force is defined as the value of H at which I_k is zero. If the field strength is high enough to magnetize the sample magnetization to saturation, then the coercive force becomes the coercivity H_c. Soft magnetic materials such as soft iron and Permalloy usually have low H_c, small hysteresis loss, and high permeability. This is dif-

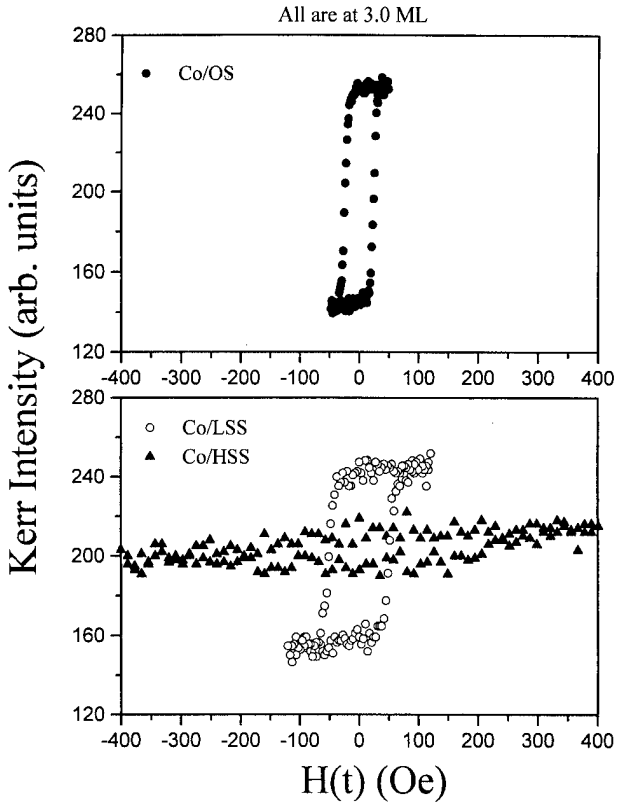


FIG. 3. Hysteresis loops for 3.0 ML-thick Co films grown on OS, LSS, and HSS Cu(001) surfaces at room temperature. The amplifier gains for the Kerr intensity were the same in three measurements.

ferent from hard magnetic materials such as permanent magnets, which have opposite magnetic properties to those just mentioned.

1. Hysteresis loop shapes from 3-ML-thick Co films grown on Cu substrates with three different initial surface roughnesses

The in-plane hysteresis loops measured around 3 ML Co films grown on three kinds of substrates are shown in Fig. 3. All loops including the Co/HSS have been saturated. The loop areas and shapes are very different for Co/OS, Co/LSS, and Co/HSS. The squareness (M_r/M_s) for Co/LSS and for Co/HSS are degraded from the Co/OS ($\approx$0.99). The rounded loop shape was also observed on a clean single crystal Fe(001) surface 50 Å deep after heavy sputtering.[20]

a. Kerr intensity or magnetization. When the loops were measured, the gain in the photodetector and all experimental conditions were identical except for the preparation of films. Because the remanent magnetization is proportional to the Kerr intensity,[7] for films with equivalent thickness one can compare the relative change in Kerr intensity as the relative change in magnetization as shown in Fig. 3. The value of M_r in arbitrary units are 57, 42, and 8.5 for Co/OS, Co/LSS, and Co/HSS. If we express the reduction of magnetization in terms of a ratio, one gets 57/42=1.36 and 57/8.5=6.7 for Co/LSS and Co/HSS. From the average terrace width listed in Table I, one can estimate the step density that is the highest in the Co/HSS. Higher step density means there are more edge Co atoms associated with the steps. It is believed that an edge atom has less coordinate numbers and the magnetic moment may change from its ideal value. If we take the ratio of step density as an indication of increased number of edge atoms, one can estimate the ratio of increase in edge atoms. Using the average terrace width in Table I, we take the ratio of terrace width between Co/OS and Co/LSS and between Co/OS and Co/HSS and then square the ratios by assuming isotropic terrace width to get the factors of increasing step density. That is, $(31.4/25)^2=1.59$ for Co/OS vs Co/LSS and $(31.4/12.5)^2=6.3$ for Co/OS vs Co/HSS. These ratios are consistent with the reduction of magnetization from Co/OS to Co/LSS and to Co/HSS. Experimentally, there is no Kerr intensity for smooth Co films thinner than 1.3±0.3 ML, i.e., the thickness of a magnetic dead layer. For a rough Co film, the thickness is not uniform and there exist many edge and step atoms. In some areas of the film, the thickness may be less than 1.3±0.3 ML, a reduction of magnetization (due to the dead layer) from a rough film is likely to occur. This result may indicate that the edge atoms dominate the reduction of magnetization.[21] The reduction in magnetization is also observed in other epitaxial systems such as ultrathin Ni films on a rough Cu(001) surface.[22]

b. Coercivity. The values of H_c are about 25, 45, and 160 Oe for 3-ML-thick Co/OS, Co/LSS, and Co/HSS, respectively. See Fig. 3. The increase of H_c for rougher films grown on LSS and HSS may be due to the increase number of pinning centers such as steps which pin the magnetic domain walls. A SEMPA study of ultrathin Co film ranging from 3 to 19.5 ML grown on a smooth Cu(001) surface has found multidomain structure with irregular shapes and the domain boundaries are bent or extremely rough.[9] We believe there exist much more multidomains in Co/LSS and Co/HSS compared with the multidomains in Co/OS. It has been estimated that for a smooth Co film, the configurational entropy at 300 K exceeds the wall energy when the thickness is less than 2 ML.[9,23] This results in rough domain walls as the state of lower free energy. The domain walls in the rough film could be even more irregular because sputtering induced roughness. A SMOKE study of Co on a terraced Cu(1 1 13) finds a step induced anisotropy favoring magnetization parallel to the steps.[21] Because of the sputtering, there is a high density of steps created in our Cu surface with random direction and height level, the magnetization may also oriented along random step direction in different height levels. To reverse the multidomains with irregular domain walls and anisotropic alignment of magnetization, a larger field is needed. If we take the ratio of H_c values from Co/LSS over Co/OS and from Co/HSS over Co/OS, we find 45/25=1.8 and 160/25=6.4, respectively. The ratios are consistent with the ratios found in the change of step density from OS to HSS (1.6) and HSS (6.3) mentioned in the last section.

Other heteroepitaxial film study also indicates film imperfections can act as hard pinning centers for domains and walls. In that case the domain structure is directly deter-

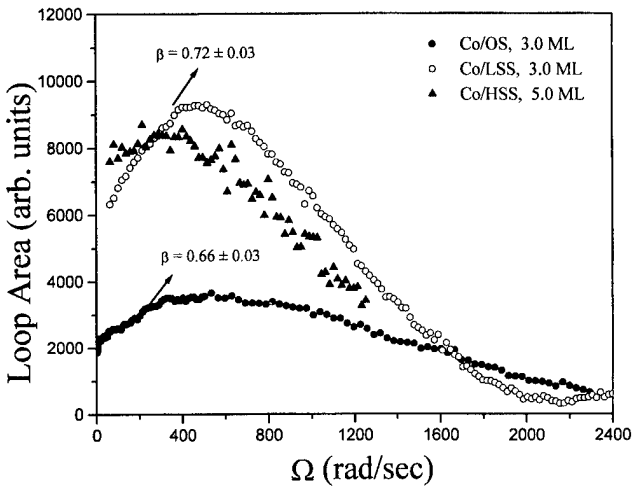


FIG. 4. Hysteresis loop areas as a function of applied magnetic field frequency for 3.0-, 3.0-, and 5.0-ML-thick Co films grown on OS, LSS, and HSS Cu(001) surfaces at room temperature, respectively. The β values obtained from the low frequency scaling regime are 0.66 ± 0.03 and 0.72 ± 0.03 for Co/OS and Co/LSS, respectively. No β value is obtained for Co/HSS due to noisy data from a rough film. Because the value of applied magnetic field H_0 depends on the magnitude of the current and is frequency dependence, we set the field strengths H_0 near the dc field to be 82, 124, and 310 Oe for Co/OS, Co/LSS, and Co/HSS, respectively. The effective field strength at a given frequency is $H_0 g(\Omega)$, where $g(\Omega)$ is the frequency-dependent field strength calibration factor. The effective field strengths at the resonance frequency for Co/OS, Co/LSS, and Co/HSS are 49, 74, and 246 Oe, respectively.

mined by the film morphology.[10e] Similar results in the substrate roughness dependent coercivity were observed in other clean ferromagnetic surfaces,[24] epitaxial systems,[22,25] and in multilayer thin films of Co/Pt on glass, Si, Si nitride, etc.[26]

2. Frequency Ω-dependent hysteresis loop area

All the hysteresis loop data presented so far were taken at low frequencies and low fields. In the following we will present the loop area over a wider range of frequency.

a. Scaling law in low frequency for smooth Co/OS films. The hysteresis loop area A as a function of frequency Ω for 3.0-, 3.0-, and 5.0-ML-thick Co films grown on three substrates OS, LSS, and HSS, respectively, are shown in Fig. 4. The reason to plot data for 5.0-ML-thick Co film on HSS is that the Kerr intensity is too low at 3.0 ML to do frequency-dependent measurements over a wide range of Ω. The area increases at low frequencies and reaches a maximum at a resonance frequency Ω_r. As frequency increases beyond Ω_r, the area decreases. This behavior can be understood using a mean field dynamics[27] with a double well barrier as follows: when the $\Omega\to\infty$ but $H_0<H_t$, where H_t, is the threshold field, the system's magnetization is unable to response to the fast alternating driving field and the magnetization stays at one state, i.e., the average magnetization $\langle M(t)\rangle=M_r$, or $-M_r$. When the $\Omega\to0$, the magnetization can be trapped in a metastable state where $\langle M(t)\rangle\neq0$, while $H_0<H_t$ if the fluctuations are suppressed. The system can return to its equilibrium state $\langle M(t)\rangle=0$ when the field is larger than the threshold field ($H_0>H_t$). This results in a finite loop area. When Ω increases, the magnetization cannot respond to the rapidly varying field and thus it results in a delay. Such a delay increases the loop area A and a scaling relation is found $A\sim\Omega^{\beta}$ with $\beta=\frac{2}{3}$ for $H_0>H_t$ and $\Omega<\Omega_r$. Typical loop shapes at different frequencies have been shown previously. At low frequencies and low fields, the area of hysteresis loops for Co/OS show scaling behavior.[12] The details have been published and will not be repeated here.

But is there scaling behavior for Co films grown on LSS or on HSS? We can apply a similar fit procedure as we did for Co on OS. We found the value of β increases from 0.66 to 0.72 for Co/LSS. For Co/HSS, the loop area data is noisy and a fit may not be meaningful. Physically the domains' reversal barriers are not so uniform as that of Co/OS, and one cannot use a single double well potential as the barrier. Therefore, the scaling may not exist and even if it exists one needs a more complicated potential barrier to describe the rough films.

b. Roughness-dependent resonance frequency Ω_r. It is shown in Fig. 4 that the Ω_r is the highest for Co film grown on OS. The Ω_r decreases as the film becomes rougher. The resonance frequency Ω_r can be identified with the inverse of the characteristic time for the particle to slide from an unstable maximum to one of the minima of the effective free energy in a double well potential barrier. The resonance frequency Ω_r as a function of film thickness also shows a decrease of value due to the buildup of roughness as the film becomes thicker and will be published elsewhere.[28] A recent simulation of submonolayer magnetic islands indicates that the magnetic moments of edge atoms that have less coordination number can affect the response of the entire magnetic system, especially when the surface to volume ratio is large.[29] The prediction is that for a fixed submonolayer coverage and fixed total island number density, the Ω_r increases monotonically as the fractal character of each island increases. Our results from ultrathin film shown in Fig. 4 is not consistent with this prediction. But the scaled area versus frequency plots of Luse and Zangwill[29] are consistent with the Co/OS and Co/LSS data shown in Fig. 4. The drastic change in the hysteresis loop shape from OS and LSS to HSS data suggests that a new physical mechanism is needed in the HSS case. For the Co/OS film, we observed that the resonance frequency Ω_r depends on the applied field strength and the result will be published elsewhere.[28]

3. Field strength H_0-dependent hysteresis loop area

The hysteresis loop area as a function of applied field strength H_0 at a fixed frequency (4 Hz) for 3.0-, 3.0-, and 5.0-ML-thick Co films grown on OS, LSS, and HSS are shown in Fig. 5. The field at which the loop area increases gradually is extended over about 150 Oe for Co/HSS compared with 30 Oe of Co/OS and Co/LSS. The threshold field H_t where a more rapid increase of area occurs is rounded between 200 to 400 Oe for Co/HSS. For Co/LSS and Co/OS, the H_t values are narrowed down to 45 and 25 Oe, respec-

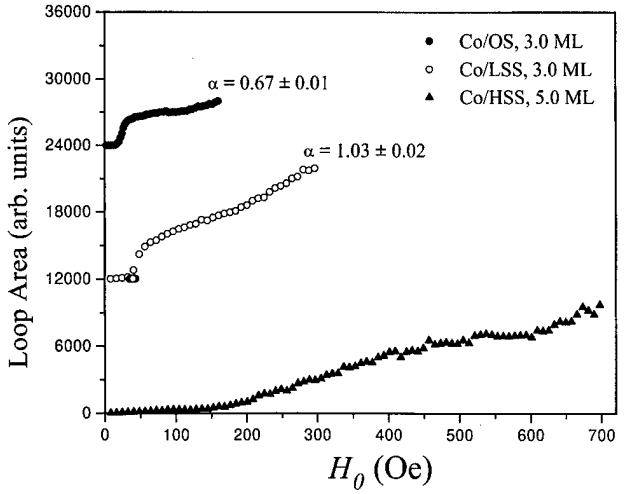


FIG. 5. Hysteresis loop areas as a function of applied magnetic field strength for 3.0-, 3.0-, and 5.0-ML-thick Co films grown on OS, LSS, and HSS Cu(001) surfaces at room temperature, respectively. The α values obtained from the low field scaling regime are 0.67 ± 0.01 and 1.03 ± 0.02 for Co/OS and Co/LSS, respectively. No α value is obtained for Co/HSS due to the uncertainty in determining the scaling regime. The data for Co/LSS and Co/OS are offset by 12 000 and 24 000 arbitrary units from zero.

tively. Again, this is consistent with the fact that there exists a wide range of nucleation fields for domains' reversals in rough films.

a. Scaling law in low field strength for smooth Co/OS films. For $H_0>H_t$ and $\Omega<\Omega_r$, it has been observed that the loop area A exhibits a dynamic scaling law in the low field strength regime, i.e., $A\sim H_0^{\alpha}$ with $\alpha=0.67\pm0.01\approx\frac{2}{3}$ for a smooth Co film on Cu(001) surface[11] and Fe/Au(001).[30] The details can be found elsewhere. That the α value for H_0 is the same as the β value for Ω can be understood from the result of mean field dynamics, where the hysteresis loop area is proportional to the product of Ω and H_0 with the same dimension. From the experimental point of view, increasing H_0 at a fixed Ω means increasing sweep rate dH_0/dt. The increase of dH_0/dt can be achieved by increasing Ω at a fixed H_0.

Judging from the data in Fig. 5, the value of α for Co/LSS and Co/HSS definitely deviates from the $\frac{2}{3}$. If we applied a fit to Co/LSS as we did for Co/Os, we obtain $\alpha=1.03\pm0.02$. This change is expected because the interface roughness changes the dynamic magnetic response of the whole film.[29]

b. Field strength H_0-dependent Kerr I_k. From the dynamic mean field model,[27] it is derived that $A\sim4M_rH_c$ or $\sim4I_kH_c$. We obtained the loop area by integrating $\oint I_k dH$ using the digitized data of I_k vs $H(t)$. In Figs. 6(a) and 6(b) we plot the I_k vs H_0 and H_c vs H_0, respectively. This allows us to understand the field dependent loop area in more detail.

Figure 6(a) shows I_k vs H_0 for Co/OS and Co/LSS. Below 40 Oe, the Kerr I_k increases rapidly in Co/OS films, indicating that stronger fields align the magnetic moments in better order. For $H_0>40$ Oe, the I_k is pretty constant up to 160 Oe (slope $dI_k/dH_0=0.0016$) for Co/OS and implies that a uniform alignment of magnetic moments in the domains has been reached. However, for Co/LSS, it takes a larger field of

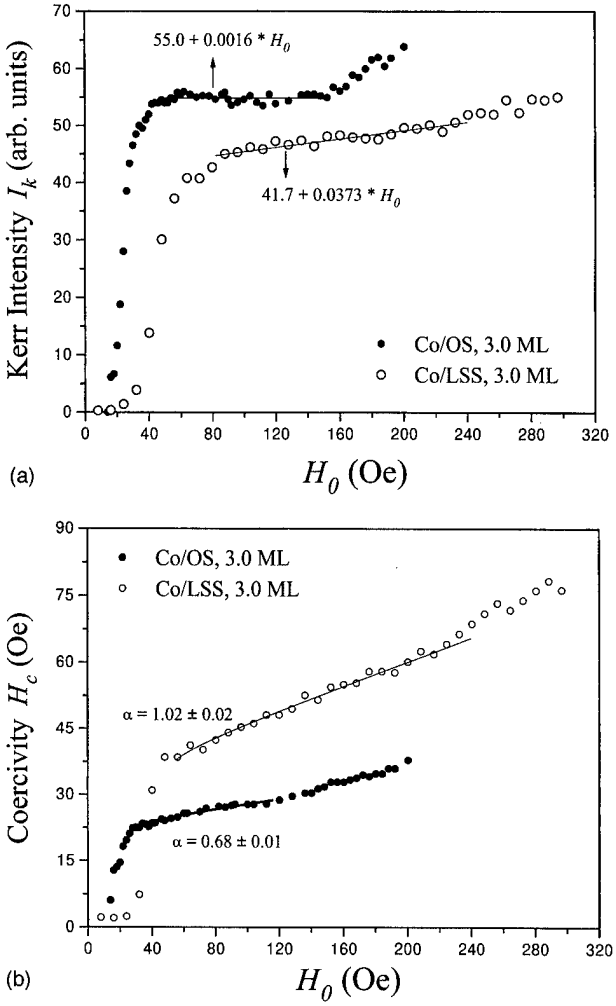


FIG. 6. (a) Kerr intensity I_k as a function of applied magnetic field strength H_0 for 3.0-ML-thick Co films grown on OS and LSS Cu(001) surfaces at room temperature. In the low field regime, $H_0<160$ Oe, the Kerr intensity is nearly a constant for Co/OS, whereas the Kerr intensity increases for Co/LSS. (b) Coercivity H_c as a function of applied magnetic field strength H_0 for 3.0-ML-thick Co films grown on OS and LSS Cu(001) surfaces at room temperature. In the low field regime, the values of the scaling exponent α obtained from fits are 0.68 and 1.02 for Co/OS and Co/LSS, respectively. These values are consistent with the fit from loop area vs field strength.

about 80 Oe to align majority of the magnetic moments. In order to align even more magnetic moments in-plane, an increasing field strength beyond 80 Oe is needed as seen from the slight positive slope ($dI_k/dH_0=0.0373$) up to 240 Oe in the I_k vs H_0.

The step densities ($1/\eta$) listed in Table I for 3.0-ML-thick Co/OS and Co/LSS are 1/31.4 Å ($=0.03$ Å^{-1}) and 1/25 Å ($=0.04$ Å^{-1}), respectively. Also, the lateral correlation lengths are 285 and 190 Å, respectively. These numbers indicate that the atoms in the Co/LSS film are less correlated compared with those of the Co/OS films, although the interface width is about the same ($\approx0.5t$) in both films. This is seen in Fig. 6(a), where a stronger field is needed to align the majority magnetic moments in the same in-plane direction and in Fig. 3 where the lower Kerr intensity is observed.

For the Co/OS film, we notice that the Kerr intensity in-

creases with a larger slope beyond the scaling regime (40 Oe$<H_0<$160 Oe). This may be related to the change of domain switching mechanism at higher fields and we do not have an explanation.

c. Field strength H_0-dependent coercivity H_c. Figure 6(b) shows the coercivity H_c versus field strength H_0 for Co/OS and Co/LSS. We have learned from Sec. II B 3 a that for Co/OS, $A \sim H_0^{\alpha}$ with $\alpha=0.67\pm0.01\pm0.01\approx\frac{2}{3}$. Because $A \sim 4M_r$, $H_c \sim 4I_k H_c$, and the I_k is nearly constant in between 40 Oe and 120 Oe, we expect the H_c has a scaling behavior as $H_c \sim H_0^{\alpha}$ Co/OS. By fitting the Co/OS data, we obtained $\alpha=0.68$ [see Fig. 6(b)]. This tells us that the area scaling in low field regime at a fixed low frequency is actually the coercivity scaling with H_0.

For Co/LSS, we know from Sec. III B 3 b and Fig. 6(a) that I_k vs H_0 is not a constant in the low field regime. In fact, the increase is positive. Therefore, a fit of H_c vs H_0 is not meaningful. If we do apply a similar scaling fit as we did for Co/OS, as shown in Fig. 6(b), the α value obtained is 1.02.

IV. CONCLUSION

SMOKE and HRLEED have been combined to study the hysteresis loops and nanoscale roughness at the interface of Co and Cu(001) and in the epitaxial Co films. Three different degrees of roughness in a Cu(001) substrate were created by Ar^+ sputtering. The initial Cu(001) substrate roughness determines the hysteresis loop shape and area of the Co films. A several fold increase of step density and an interface width of less than 2 ML in a rough Co film grown on a rough Cu(001) substrate results in several times higher coercivity with a wide range of coercive fields, and severalfold reduced magnetization as compared with that of a smooth Co film grown on a smooth Cu(001) substrate. These changes are related to the pinning of Co magnetic domains by sputtering induced steps. For a smooth 3.0-ML-thick Co film, the dynamic hysteresis loop area as functions of frequency and field amplitude of a sinusoidal magnetic field show a power law scaling behavior at low frequencies and low amplitudes. The value of scaling exponents ($\approx\frac{2}{3}$) is consistent with the dynamic mean field Ising model assuming a double well potential barrier. For rougher Co films with the same interface width but a few percent increase in step density and 30% decrease in lateral correlation length as compared with a smooth film, the scaling exponents deviates from the $\frac{2}{3}$ drastically. These results demonstrate that the stringent structural perfection in ultrathin films down to the nanoscale may be required when desired magnetic properties of ultrathin films are to be used in magnetic devices. It also cautions us when experimental data are to be compared with microscopic model predictions in order to better understand the fundamental magnetic properties of ultrathin magnetic films.

ACKNOWLEDGMENTS

This work is supported by ONR. The authors deeply appreciate K. Liang for the loan of the Cu(001) crystal. We also acknowledge the use of the HRLEED system from Leybold Vacuum Products.

[1]See many papers in these proceedings: J. Appl. Phys. **76**, No. 10 (1994).
[2]See, for example, IEEE Trans. Magn. **30**, No. 6 (1994).
[3]*Ultrathin Magnetic Structures I and II*, edited by B. Heinrich and J. A. C. Bland (Springer, Berlin, 1994).
[4]L. M. Falicov, D. T. Pierce, S. D. Bader, R. Gronsky, K. B. Hathaway, H. B. Hopster, D. N. Lambeth, S. S. P. Parkin, I. K. Schuller, and R. H. Victora, J. Mater. Res. **5**, 1299 (1990); L. M. Falicov, Thin Solid Films **216**, 169 (1992).
[5]See, for example, K. Derbyshire and E. Korczynski, Solid State Technology, Sept., Vol. 38, 57 (1995).
[6]Physics Today, April (1995).
[7]E. R. Moog and S. D. Bader, Superlatt. Microstruct. **1**, 543 (1985); S. D. Bader and C. Liu, J. Vac. Sci. Technol. A **9**, 1924 (1991); S. D. Bader, D. Li, and Z. Q. Qiu, J. Appl. Phys. **76**, 6419 (1994); S. D. Bader and J. L. Erskine, in Ref. 3, Vol. 2, p. 297.
[8]U. Scheithauer, G. Meyer, and M. Henzler, Surf. Sci. **178**, 441 (1986); M. Henzler, *ibid.* **298**, 369 (1993).
[9]H. P. Oepen, M. Benning, H. Ibach, C. M. Schneider, and J. Kirschner, J. Magn. Magn. Mater. **86**, L137 (1990).
[10]See, for example, see: (a) D. T. Pierce, J. Unguris, and R. J. Celotta, MRS Bull. June, 19 (1988); J. L. Robins, R. J. Cellotta, J. Unguris, D. T. Pierce, B. T. Jonker, and G. A. Prinz, Appl. Phys. Lett. **52**, 1918 (1988); (b) R. Allenspach, M. Stampanoni, and A. Bischof, Phys. Rev. Lett. **65**, 3344 (1990); (c) J. Pommier, P. Meyer, G. Penissard, J. Ferré, P. Bruno, and D. Renard, *ibid.* **65**, 2054 (1990); (d) G. Bochi, H. L. Hug, D. I. Paul, S. Stiefel, A. Moser, I. Parashikov, H.-J. Güntherodt, and R. C. O'Handley, *ibid.* **75**, 1839 (1995); (e) M. Speckmann, H. P. Oepen, and H. Ibach, *ibid.* **75**, 2035 (1995).
[11]Q. Jiang, H.-N. Yang, and G.-C. Wang, Phys. Rev. B **52**, 14911 (1995).
[12]J.-P. Qian and G.-C. Wang, J. Vac. Sci. Technol. A **8**, 4117 (1990).
[13]H.-N. Yang, G.-C. Wang, and T.-M. Lu, *Diffraction From Rough Surfaces and Dynamic Growth Fronts* (World Scientific, Singapore, 1993); T.-M. Lu, H.-N. Yang, and G.-C. Wang, Mater. Res. Soc. Proc. **367**, 283 (1995); K. Fang, R. Adame, H.-N. Yang, G.-C. Wang, and T.-M. Lu, Appl. Phys. Lett. **66**, 2077 (1995); Y.-L. He, H.-N. Yang, T.-M. Lu, and G.-C. Wang, Phys. Rev. Lett. **69**, 3770 (1992).
[14]A. K. Schmidt and J. Kirschner, Ultramicroscopy 42–44, 483 (1992); J. J. deMiguel, A. Cebollada, J. M. Gallego, R. Miranda, C. M. Schneider, P. Schuster, and J. Kirschner, J. Magn. Magn. Mater. **73**, 1 (1991); C. M. Schneider, P. Bressler, P. Schuster, J. Kirschner, J. J. deMiguel, and R. Miranda, Phys. Rev. Lett. **64**, 1059 (1990); H. Li and B. P. Tonner, Surf. Sci. **237**, 141 (1990); G. L. Nyberg, M. T. Kief, and W. F. Egelhoff, Jr., Phys. Rev. B **48**, 14509 (1993).
[15]H.-N. Yang, G.-C. Wang, and T.-M. Lu, Phys. Rev. B **51**, 17932 (1995).
[16]R. Altsinger, H. Busch, M. Horn, and M. Henzler, Surf. Sci. **200**, 235 (1995); M. Horn, U. Gotter, and M. Henzler, J. Vac. Sci. Technol. B **6**, 727 (1988).
[17]Q. Jiang, B. Poserow, and G.-C. Wang (unpublished).
[18]J. Wollschläger, J. Falta, and M. Henzler, Appl. Phys. A **50**, 57 (1990).
[19]Y.-L. He and G.-C. Wang, Phys. Rev. Lett. **71**, 3834 (1993), and references therein; Y.-L. He and G.-C. Wang, Mater. Res. Soc. Proc. **237**, 429 (1992).
[20]R. Allenspach, M. Taborelli, M. Laddolt, and H. C. Siegmann, Phys. Rev. Lett. **56**, 953 (1986).
[21]D. S. Chuang, C. A. Ballentine, and R. C. O'Handley, Phys. Rev. B **49**, 15084 (1994).
[22]F. Huang, G. J. Mankey, and R. F. Willis, Surf. Sci. **297**, L79 (1993); S. Z. Wu, G. J. Mankey, F. Huang, and R. F. Willis, J. Appl. Phys. **76**, 6434 (1994).
[23]Q. Jiang, H.-N. Yang, and G.-C. Wang (unpublished).
[24]M. Donath, G. Schönhense, K. Ertl, and V. Dose, Appl. Phys. A **50**, 49 (1990).
[25]Y.-L. He and G.-C. Wang, J. Appl. Phys. **76**, 6446 (1994).
[26]C.-H. Chang and M. H. Kryder, J. Appl. Phys. **75**, 6864 (1994), and references therein.
[27]M. Rao, H. R. Krishnamurthy, and R. Pandhit, Phys. Rev. B **42**, 856 (1990); M. Acharyya and B. K. Chakrabarti, Physica A **192**, 471 (1993); P. Jung, G. Gray, R. Roy, and P. Mandel, Phys. Rev. Lett. **65**, 1873 (1990); **68**, 1437 (1992); C. N. Luse and A. Zangwill, Phys. Rev. E **50**, 224 (1994).
[28]Q. Jiang, H.-N. Yang, and G.-C. Wang, J. Appl. Phys. **79**, 5112 (1996).
[29]C. N. Luse and A. Zangwill, J. Appl. Phys. **79**, 4942 (1996).
[30]Y.-L. He and G.-C. Wang, Phys. Rev. Lett. **70**, 2336 (1993).

Magnetic behavior of $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ pseudomorphic films on Cu(100)*

S. Z. Wu, F. O. Schumann,[a] G. J. Mankey, and R. F. Willis
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 2 October 1995; accepted 18 December 1995)

We set out to compare the ferromagnetic behavior of $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ films of varying stoichiometry epitaxially grown on Cu(100). The thickness chosen was 5 ML over a wide alloy composition range. Using a scaling law for the thickness vs Curie temperature, we extrapolate the bulk fcc film thickness and plot the bulk fcc Curie temperature as a function of composition. The results suggest (a) that the Invar effect is quenched in these ultrathin pseudomorphic fcc films of $Fe_xNi_{(1-x)}$ alloys and (b) the magnetic phase of Fe atoms is the low-spin ferromagnetic phase. The $Co_xNi_{(1-x)}$ films show the expected monotonic decrease in T_c with increasing Ni content. © *1996 American Vacuum Society.*

I. INTRODUCTION

Recently, the study of the electronic and magnetic properties of fcc Fe, Co, and Ni films and their alloys has attracted a lot of attention. The electronic band structures of pure Cu, Ni, Co, and Fe films with (100) surface orientation have been described in many publications.[1–10] Preliminary measurements of the Curie temperatures of ultrathin $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ alloy films deposited on Cu(100) have been reported in an earlier publication.[11] These measurements show that for a few single monolayers film thickness, the Curie temperature increases with increasing Fe or Co content and the thickness dependence can be expressed by an empirical finite-size scaling formula. Meanwhile, these ultrathin alloy films with (100) surface orientation have different magnetic anisotropy behavior. Fe deposited on various substrates has a reorientation of easy axis of magnetization from perpendicular to in-plane as a function of film thickness (a few monolayers), and growth temperature (at a fixed film thickness).[12–20] Ni films have a transition from parallel to perpendicular magnetization at a few monolayers thickness and a transition back to in-plane at greater thickness.[21,22] Pure Co films grown on Cu(100) substrates show an in-plane magnetic anisotropy for all thicknesses.[23–25] For $Fe_xNi_{(1-x)}$ alloys, our measurements show that the anisotropy behavior depends on the alloy composition, with $Fe_{0.75}Ni_{0.25}$ and $Fe_{0.5}Ni_{0.5}$ showing a reorientation of preferred magnetization from perpendicular to parallel at 4 ML, while higher Ni content alloys $Fe_{0.25}Ni_{0.75}$ and $Fe_{0.1}Ni_{0.9}$ have the magnetization lying in-plane for all the film thicknesses in the range that we have investigated (to be published). On the other side, $Co_xNi_{(1-x)}$ alloys with varying composition all show a preferred magnetization in plane.

Face-centered-cubic Fe can exist in three different states: high-spin ferromagnetic, low-spin ferromagnetic, and antiferromagnetic state depending on the lattice constant.[26–31] At the Cu lattice constant, the energy differences between these states become very small, and this results in a structural instability of Fe/Cu(100) films,[32] which in turn explains the different magnetic properties of Fe/Cu(100) reported in the literature. In bulk $Fe_xNi_{(1-x)}$, there is an Invar effect observed at $x=65\%$, where the Curie temperature collapses and the structure is a mixture of fcc and bcc. If we increase the Fe content further, a transition from fcc to bcc structure occurs.[33] As we go from pure Ni to Co or Fe, the electron number is reduced. At 8.4 number of $3d$ electrons we are at the maximum in the magnetic moment curve.[34] However, the basic physics describing this behavior are somewhat clouded by the change in crystal structure as we go from fcc Ni to hcp Co to bcc Fe. The structural change introduces a difference due to the orbital moment. In addition to the crystal structure change, the lattice constant of the bulk material also changes. This has a greater effect because as the lattice constant is reduced the overlap between the orbitals of adjacent atoms becomes greater, and this will result in a reduced moment and ultimately produce a paramagnetic phase. In this article we will show that we have stabilized an fcc low-spin ferromagnetic Fe phase in the $Fe_xNi_{(1-x)}$ alloy films produced on a Cu(100) substrate, and the Invar "collapse" effect was not observed in these metastable ultrathin films, which is confirmed in a separate study.[35]

II. EXPERIMENT

In the ultrathin film limit, the higher surface free energy of the ferromagnetic film verses the Cu substrate prevents the film from wetting the substrate completely. The study of the initial stages of growth of these systems has been the subject of numerous publications. Briefly, for low temperature growth (around 100 K), the mobility of the surface atoms is reduced to eliminate the possibility of Cu atoms diffusing from step edges and partially covering the film surface. This has the advantage of completely eliminating segregation, but the film continuously roughens with increasing thickness, with the roughness increasing according to a scaling law.[36,37] For such a film, the thickness is no longer a well-defined quantity, since areas of the substrate are covered with islands of varying size and thickness. If the growth temperature is below the temperature where bulk diffusion becomes signifi-

*Published without author corrections.

[b]Current address: Department of Physics, Louisiana State University, Baton Rouge, LA 70803.

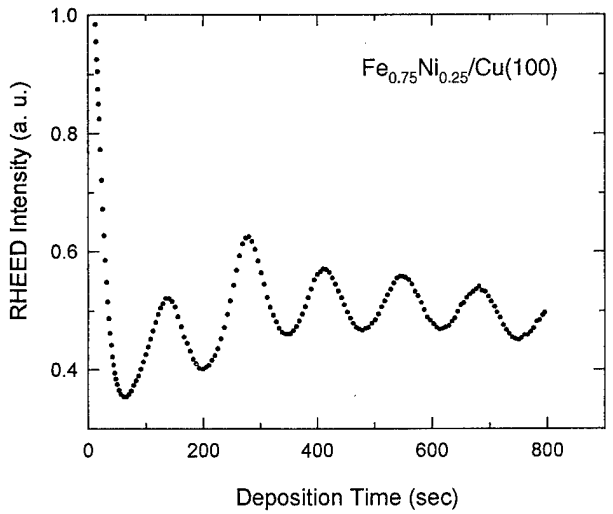


FIG. 1. The oscillations of RHEED intensity of a 5 ML $Fe_{0.75}Ni_{0.25}$ growth with an electron beam energy of 3.0 kV. It reveals a layer-by-layer growth.

cant while still allowing for surface diffusion, Cu will segregate onto the film surface from step edges (defect sites) within a distance defined by the surface diffusion rate and the flux of the evaporation source. If the flux is high enough, the amount of Cu covering the film can be reduced to a few percent of a monolayer. It is our contention that it is more important to have a uniform film than it is to eliminate Cu segregation near defect sites.

Thin films of Fe, Co, Ni, and Fe–Ni, Co–Ni alloys were produced on a Cu substrate with (100) surface orientation in ultrahigh vacuum using molecular beam epitaxy (MBE). The growth temperature was 400 K for $Fe_xNi_{(1-x)}$ and 300 K for $Co_xNi_{(1-x)}$ alloy films, high enough to produce smooth films while low enough to limit interdiffusion at the film/substrate interface to the first few layers. The deposition rate was controlled to be 1 ML/min with the flux of the sources calibrated with reflection high energy electron diffraction (RHEED) oscillations. The films were then cooled to 200 K and the magnetization was measured by a surface magneto-optic Kerr effect (SMOKE) magnetometer and the electron band structure was studied by ultraviolet photoemission spectroscopy (UPS) with 21.2 eV photons. The incident angle is 45° from the surface normal and the energy distribution curve of the photoemitted electrons was collected along the surface normal.

Figure 1 shows the oscillations of RHEED intensity of the (0,0) beam during the growth of 5 ML $Fe_{0.75}Ni_{0.25}$ with electron beam energy 3.0 kV. Similar oscillations were observed for all the $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ alloy films, which confirms a good epitaxial layer-by-layer growth with the relatively low intensity of the first peak showing a tendency of initial bilayer growth. Meanwhile, observation in low energy electron diffraction (LEED) on the grown films has shown that the surface structure is fcc, which further verifies a pseudomorphic growth of these alloy film systems.

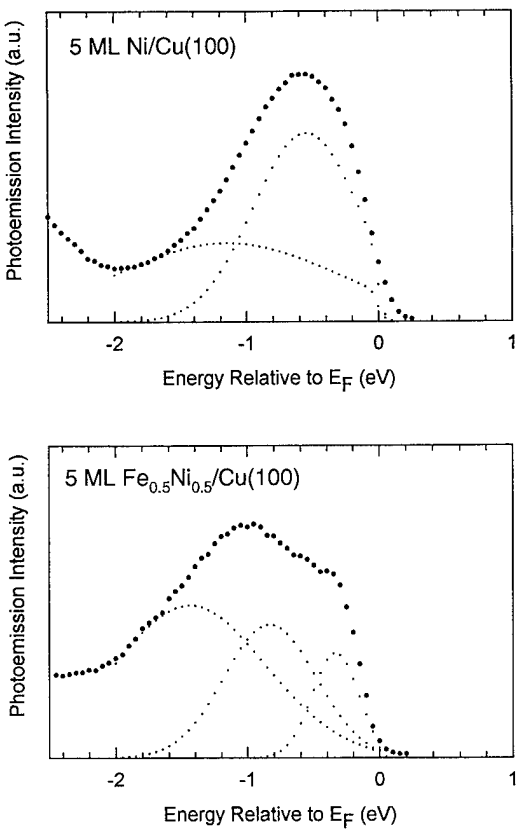


FIG. 2. The Gaussian fittings of the energy distribution curves between −2 and 0 eV (Fermi energy level) with a photon energy of 21.2 eV for both 5 ML Ni and 5 ML $Fe_{0.5}Ni_{0.5}$ on Cu(100).

III. RESULTS AND DISCUSSION

The electronic band structures for occupied states of Ni, Co, and Fe films with (100) surface orientation have previously been measured in detail using spin resolved photoemission and inverse photoemission. The ferromagnetic exchange splittings of the Δ_5-symmetry $3d$ bands are estimated to be ΔE_{ex}=0.3, 1.2, and 1.2 eV for Ni(100), Co(100), and Fe(100), respectively. For $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ alloys, we expect the $3d$ band filling and the exchange splitting to vary continuously from that of pure Ni to Co or Fe.

At 21.2 eV photon energy in normal emission, we are exploring the transitions about halfway from the Brillouin zone center to the zone edge. Only transitions from initial states of Δ_1 or Δ_5 symmetry are allowed with emission from the Δ_5 symmetry band approximately three times as intense in our experimental arrangement. In addition, the energy of the minority-spin band near the Fermi energy is most easily extracted, since peak widths increase linearly for transitions farther from E_F.[38]

Figure 2 gives a sample of the Gaussian fittings of the energy distribution curves between −2 and 0 eV (Fermi energy level) for 5 ML Ni and $Fe_{0.5}Ni_{0.5}$ films after the Shirley background intensity has been subtracted from the raw data. In Fig. 3, UPS measurements of 5-ML-thick alloy films of varying composition are presented. The positions of the minority- (down arrows) and majority- (up arrows) spin

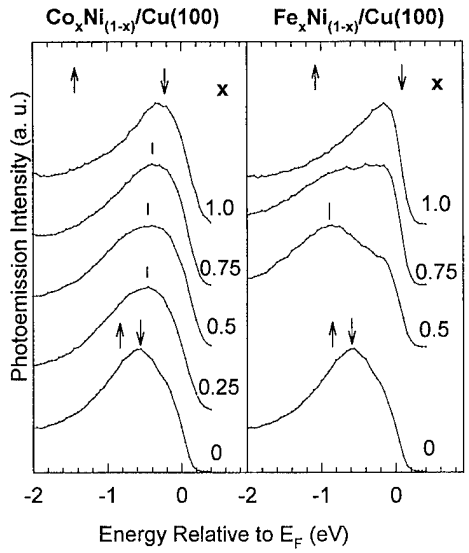


FIG. 3. The energy distribution curves measured with ultraviolet photoemission spectroscopy on 5 ML $Co_xNi_{(1-x)}$ and $Fe_xNi_{(1-x)}$ alloy films as a function of composition x. The up and down arrows indicate the energies of the majority and minority spin Δ_5 bands for the pure films.

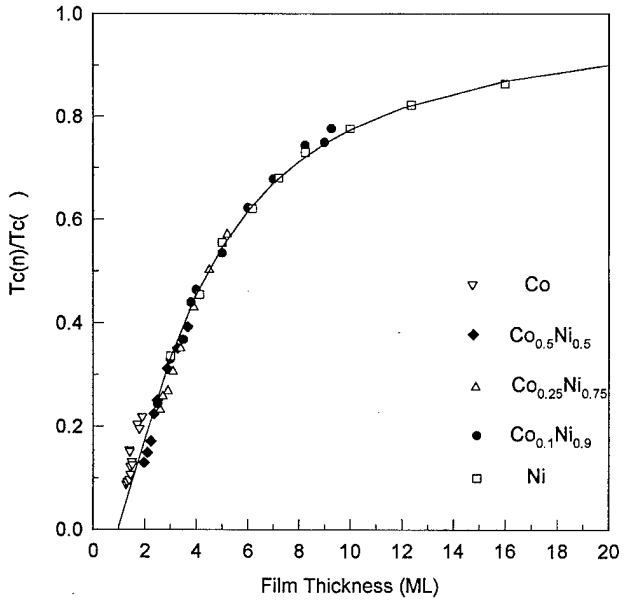


FIG. 4. The normalized Curie temperature $T_c(n)/T_c(\infty)$ as a function of film thickness for $Co_xNi_{(1-x)}$ alloy films with varying composition.

bands in these fcc structures are indicated. The spectra for the pure films are consistent with other work, with the main peak coming from an initial Δ_5 minority-spin state. For $Co_xNi_{(1-x)}$ alloys, the trend is as expected: the peak energy gradually increases as Co concentration increases. For the $Fe_xNi_{(1-x)}$ alloys, a different behavior is observed. At 50% Fe concentration, a peak is observed at -0.9 eV, much lower than the pure Co film which has the same number of valence electrons. However, without spin discrimination we cannot identify either the spin character or the symmetry of this transition. For 75% Fe concentration, there are no clearly resolved peaks, suggesting that the minority-spin band is above E_F as in pure Fe.

Our previous measurements of Curie temperature of thin films have shown that T_c can be expressed by a finite-size empirical scaling law:

$$T_c(n)=\frac{T_c(\infty)}{[1+(n-n'/n_0)]^{-\lambda}},$$

with $T_c(n)$ the Curie temperature for ultrathin films with n ML thickness, and $T_c(\infty)$ represents the bulk Curie temperature. From our previous study of $Co_xNi_{(1-x)}$ alloys, we found that n' is found to be around 1, which means the Curie temperature reaches zero for all the films with 1 ML thickness, and λ is a constant which is found to be 1.3.[39] We have kept these parameters constant and used this scaling law to extrapolate the bulk Curie temperature $T_c(\infty)$ for $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ alloy systems. In Fig. 4 we have plotted the normalized Curie temperature $T_c(n)/T_c(\infty)$ as a function of film thickness for $Co_xNi_{(1-x)}$ alloy films with varying composition. All the data points for these fcc films are indeed following this scaling law curve which intersects the horizontal axis at 1 ML thickness with $T_c=0$ K.

In the mean-field approximation, the Curie temperature is proportional to $J^*s(s+1)$, where J is an effective exchange parameter and s is the spin moment in Bohr magnetons μ_B. The solid lines in Fig. 5 are the plots of bulk Curie temperatures for $Co_xNi_{(1-x)}$ and $Fe_xNi_{(1-x)}$ alloys as a function of Co or Fe atomic concentration x that we calculated by using the method proposed by Vonsovskii of correlating the Curie temperature with the interactions between different types of atoms in alloys.[40] The diamond and round data points are from our experimental measurements of Curie temperature as a function of film thickness and then extrapolating the

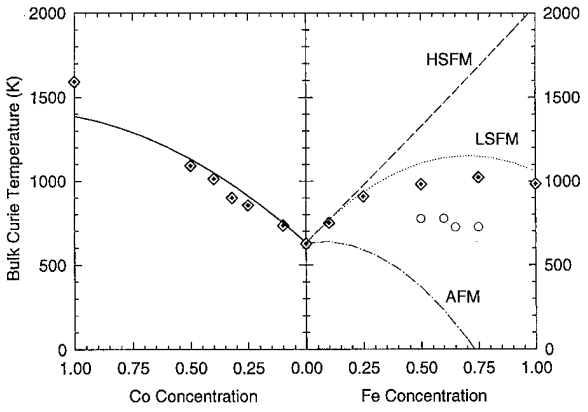


FIG. 5. Bulk Curie temperatures for $Co_xNi_{(1-x)}$ (left) and $Fe_xNi_{(1-x)}$ (right) alloys as a function of Co or Fe atomic concentration x. For $Fe_xNi_{(1-x)}$ alloys, Fe atoms can exist in three phases: high-spin ferromagnetically coupled Fe (HS FM Fe), low-spin ferromagnetically coupled Fe (LS FM Fe), and antiferromagnetically coupled Fe (AFM Fe). The diamond data points are our experimental results extrapolating from the thin film Curie temperature measurements. The round points are from a closer look at concentrations near the Invar effect: $Fe_{0.65}Ni_{0.35}$. Our data show that Fe is in the low-spin ferromagnetic phase in all the alloy films we produced.

bulk value using the above scaling law. The round data are from a recent closer look at the concentration around $Fe_{0.65}Ni_{0.35}$ where Invar effect is popularly observed in bulk material. These measurements have confirmed that the structure of these alloys are fcc at this thickness range, there is no structural transition from fcc to bcc as observed in bulk. Comparison of these results with the theoretical results strongly suggest that the Fe atoms in $Fe_xNi_{(1-x)}$ alloy films exist in the low-spin ferromagnetic phase. Besides, no sudden drop in Curie temperature has been observed at and above 65% Fe concentration, so the Invar effect has been quenched in these thin film systems. Further measurements of thicker $Fe_xNi_{(1-x)}$ films, around 65% Fe content has suggested a structural relaxation or spin reduction occurring later.[35]

IV. SUMMARY

By tuning the elemental composition, we are able to achieve $Fe_xNi_{(1-x)}$ and $Co_xNi_{(1-x)}$ films with varying electron number in random substitutional alloys of $3d$ ferromagnets. Our study of growth, structure, electron band, and magnetic measurements combined has confirmed an fcc pseudomorphic growth of these alloys films in the ultrathin thickness range. The Fe atoms in these metastable $Fe_xNi_{(1-x)}$ alloys appear to exist in the low-spin ferromagnetic phase. Besides, no structural transition from fcc to bcc is observed, and the Invar "collapse" effect appears to have been quenched in these ultrathin films near 65% Fe concentration.

ACKNOWLEDGMENT

Funding for this research was provided by the National Science Foundation under Grant No. NSF-DMR 9121736.

[1]R. Courths and S. Hufner, Phys. Rep. **112**, 53 (1984).
[2]J. A. Knapp, F. J. Himpsel, and D. E. Eastman, Phys. Rev. B **19**, 4952 (1979).
[3]W. Eberhardt and E. W. Plummer, Phys. Rev. B **21**, 3245 (1980).
[4]R. Miranda, D. Chandesris, and J. Lecante, Surf. Sci. **130**, 269 (1983).
[5]C. M. Schneider, P. Schuster, M. Hammond, H. Ebert, J. Noffke, and J. Kirshner, J. Phys. C **3**, 4349 (1991).
[6]W. Clemens, T. Katchel, O. Rader, S. Blugel, C. Carbone, and W. Eberhardt, Solid State Commun. **81**, 739 (1992).
[7]A. A. Hezaveh, G. Jennings, D. Pescia, R. F. Willis, K. Prince, M. Surman, and A. Bradshaw, Solid State Commun. **57**, 329 (1986).
[8]M. F. Onellion, C. L. Fu, M. A. Thompson, J. L. Erskine, and A. J. Freeman, Phys. Rev. B **33**, 7322 (1986).
[9]D. P. Pappas, K.-P. Kamper, B. P. Miller, H. Hospter, D. E. Fowler, C. R. Brundle, A. C. Luntz, and Z.-X. Shen, Phys. Rev. Lett. **66**, 504 (1991).
[10]G. J. Mankey, R. F. Willis, and F. J. Himpsel, Phys. Rev. B **48**, 10 284 (1993).
[11]G. J. Mankey, S. Z. Wu, F. O. Schumann, F. Huang, M. T. Kief, and R. F. Willis, J. Vac. Sci. Technol. A **13**, 1531 (1995).
[12]Z. Q. Qiu, J. Pearson, and S. D. Bader, Phys. Rev. Lett. **70**, 1006 (1993).
[13]D. Li, M. Freitag, J. Pearson, Z. Q. Qiu, and S. D. Bader, Phys. Rev. Lett. **72**, 3112 (1994).
[14]D. P. Pappas, C. R. Brundle, and H. Hopster, Phys. Rev. B **45**, 8169 (1992).
[15]H. Fritzsche, J. Kohlhepp, H. J. Elmers, and U. Gradmann, Phys. Rev. B **49**, 15 665 (1994).
[16]G. Lugert, W. Robl, L. Pfau, M. Brockmann, and G. Bayreuther, J. Magn. Magn. Mater. **121**, 498 (1993).
[17]D. P. Pappas, K.-P. Kamper, and H. Hopster, Phys. Rev. Lett. **64**, 3179 (1990).
[18]R. Allenspach and A. Bischof, Phys. Rev. Lett. **69**, 3385 (1992).
[19]N. C. Koon, B. T. Jonker, F. A. Volkening, J. J. Krebs, and G. A. Prinz, Phys. Rev. Lett. **59**, 2463 (1987).
[20]B. Heinrich, K. B. Urquhart, J. R. Dutcher, S. T. Purcell, J. F. Cochran, and A. S. Arrott, J. Appl. Phys. **63**, 3385 (1988).
[21]S. Z. Wu, G. J. Mankey, F. Huang, and R. F. Willis, J. Appl. Phys. **76**, 6434 (1994).
[22]G. Bochi, C. A. Ballentine, H. E. Inglefield, S. S. Bogomolov, C. V. Thompson, and R. C. O'Handley, J. Appl. Phys. **75**, 6430 (1994).
[23]J. G. Gay and R. Richter, J. Appl. Phys. **61**, 3362 (1987).
[24]D.-S. Wang, R. Wu, and A. J. Freeman, J. Appl. Phys. **73**, 6745 (1993).
[25]P. Krams, F. Lauks, R. L. Stamps, B. Hillebrands, and G. Guntherodt, Phys. Rev. Lett. **69**, 3674 (1992).
[26]V. L. Moruzzi, P. M. Marcus, K. Schwarz, and P. Mohn, Phys. Rev. B **34**, 1784 (1986).
[27]G. L. Krasko, Phys. Rev. B **36**, 8565 (1987).
[28]G. W. Fernando and B. R. Cooper, Phys. Rev. B **38**, 3016 (1988).
[29]D. Bagayoko and J. Callaway, Phys. Rev. B **28**, 5419 (1983).
[30]V. L. Moruzzi, Phys. Rev. Lett. **57**, 2211 (1986).
[31]M. Podgorny, J. Magn. Magn. Mater. **78**, 352 (1989).
[32]I. A. Abrikosov, O. Eriksson, P. Soderlind, H. L. Skriver, and B. Johansson, Phys. Rev. B **51**, 1058 (1995).
[33]E. F. Wasserman, *Ferromagnetic Materials*, edited by K. H. J. Buschow and E. P. Wohlfarth (North-Holland, Amsterdam, 1990), Vol. 5, p. 237.
[34]C. Kittel, *Introduction to Solid State Physics*, 6th ed. (Wiley, New York, 1986).
[35]F. O. Schumann, S. Z. Wu, G. J. Mankey, and R. F. Willis (unpublished).
[36]Q. Jiang, A. Chan, and G.-C. Wang, Phys. Rev. B **50**, 11 116 (1994); Q. Jiang and G.-C. Wang, Surf. Sci. **324**, 357 (1995).
[37]M. T. Kief, and W. F. Egelhoff, Jr., Phys. Rev. B **47**, 10785 (1993).
[38]A. Santoni and F. J. Himpsel, Phys. Rev. B **43**, 1305 (1991).
[39]F. Huang, M. T. Kief, G. J. Mankey, and R. F. Willis, Phys. Rev. B **49**, 3962 (1994). Although this scaling law was initially introduced to fit into the Curie temperature measurement of ultrathin films approaching single monolayer thickness, we found that it works on those thicker films as well, in our case, films up to 20 ML. Besides, it works on different films on substrates with different surface orientation. It does not show any apparent dependency on film morphology or microstructure.
[40]S. V. Vonsovskii, *Magnetism* (Wiley, New York, 1974).

Epitaxial growth, structure, and composition of Fe films on GaAs(001)-2×4

E. Kneedler, P. M. Thibado, B. T. Jonker,[a] B. R. Bennett, B. V. Shanabrook, R. J. Wagner, and L. J. Whitman
Naval Research Laboratory, Washington, DC 20375-5343

(Received 27 November 1995; accepted 1 April 1996)

The structure and composition of Fe films grown on As-terminated GaAs(001)-2×4 surfaces at 175 °C has been studied *in situ* with scanning tunneling microscopy (STM), photoelectron diffraction (PED), and x-ray photoelectron spectroscopy (XPS). The GaAs surfaces were prepared by molecular beam epitaxy (MBE) and exhibited large atomically well-ordered terraces. We find that the 2×4 reconstruction has a significant impact on the Fe nucleation and growth, with initial nucleation occurring at As-dimer sites. STM reveals that the first half-monolayer of Fe forms small two-dimensional islands along the As-dimer rows before growing onto the adjacent Ga-rich rows, with no evidence of substrate disruption. PED indicates that the growth is predominantly layer by layer, with the growth front for the nth deposited layer limited to the $(n+1)$th layer. XPS spectra show that the Fe films include a concentration gradient of Ga and As out-diffused from the interface, with some of the As segregating to the Fe surface, similar to previous results obtained for growth on non-MBE prepared GaAs surfaces. Possible mechanisms for the film growth and the origins of the intermixing are discussed. © *1996 American Vacuum Society.*

I. INTRODUCTION

The behavior of magnetic thin films on semiconductor substrates is of great practical interest due to the potential for integrating such films into novel semiconductor heterostructure devices. For very thin films, the magnetic and electronic properties of the heterostructure may depend strongly on the nature of the metal/semiconductor interface. This dependence can arise from the atomic-scale structure of the interface, the occurrence of interdiffusion, or the nature of the initial film growth. Before magnetic heterostructure devices can be widely realized, the characteristics of the metal/semiconductor interface must be well understood.

One magnetic system of special interest is Fe on GaAs, which is considered to have one of the most well-ordered and abrupt metal/semiconductor interfaces. Fe grows epitaxially (bcc) on GaAs (110) and (001), with the film crystallographic axes coincident with those of the substrate, due in part to the fact that the bcc Fe lattice constant is approximately half that of zinc-blende GaAs ($2a_{Fe}/a_{GaAs}=1.013$).[1,2] Although the films are epitaxial, studies of Fe deposition on both cleaved GaAs(110) and sputter-annealed (001)-$c(8\times2)$ surfaces have found the growth to be three-dimensional (3D) and observed Ga and As intermixed in the first ~50 Å of Fe.[3,4] Magnetic measurements of Fe films on GaAs(001) revealed anomalous behavior, with films as thick as 200 Å exhibiting substantially lower magnetization than expected from bulk properties.[5] Furthermore, such films often have an in-plane uniaxial component to the magnetic anisotropy,[5] although an ideal bcc Fe(001) film should have fourfold symmetry.

One factor that may contribute to the magnetic anisotropy is the substrate surface reconstruction, which is thought to influence the Fe nucleation and growth[6]; no studies have yet focused on this effect, however. A more complete understanding of the magnetic behavior requires a detailed description of both the physical and compositional structure of the Fe films. In this article we describe the structure and composition of Fe films grown on the As-terminated GaAs(001)-2×4 surface, which is the most commonly utilized surface for the growth of compound semiconductor device structures. The samples were prepared by molecular beam epitaxy (MBE) and characterized *in situ* with scanning tunneling microscopy (STM), photoelectron diffraction (PED), and x-ray photoelectron spectroscopy (XPS). We find that the 2×4 surface reconstruction has a significant impact on the Fe nucleation and film growth, with initial nucleation occurring at As-dimer sites.

II. EXPERIMENT

The experiments were performed in a multichamber ultrahigh vacuum (UHV) molecular beam epitaxy and surface analysis facility consisting of two MBE chambers, a PED/XPS chamber, and an STM chamber, all connected via a UHV sample transfer system. The GaAs(001) surface was prepared via homoepitaxial MBE growth on a silicon-doped substrate oriented within 0.1° of (001). After growing a 1-μm-thick buffer layer of Si-doped GaAs ($n\sim10^{17}/cm^3$) at 580 °C, a 300 Å layer of undoped GaAs was grown at a reduced rate combined with growth interrupts to eliminate Si contamination of the surface and produce large well-ordered terraces. The substrate was then cooled to 550 °C under an As flux from a valved cracker source, at which point the As source was both valved and shuttered. The sample was cooled further to 200 °C under no flux prior to transfer to another chamber. This procedure consistently produced an atomically well-ordered, As-dimer terminated (2×4)-reconstructed surface as indicated by reflection high energy electron diffraction (RHEED) and confirmed by STM.[7]

Fe deposition was performed primarily in a second MBE chamber using a high-temperature Knudsen cell source at a

[a]Electronic mail: jonker@anvil.nrl.navy.mil

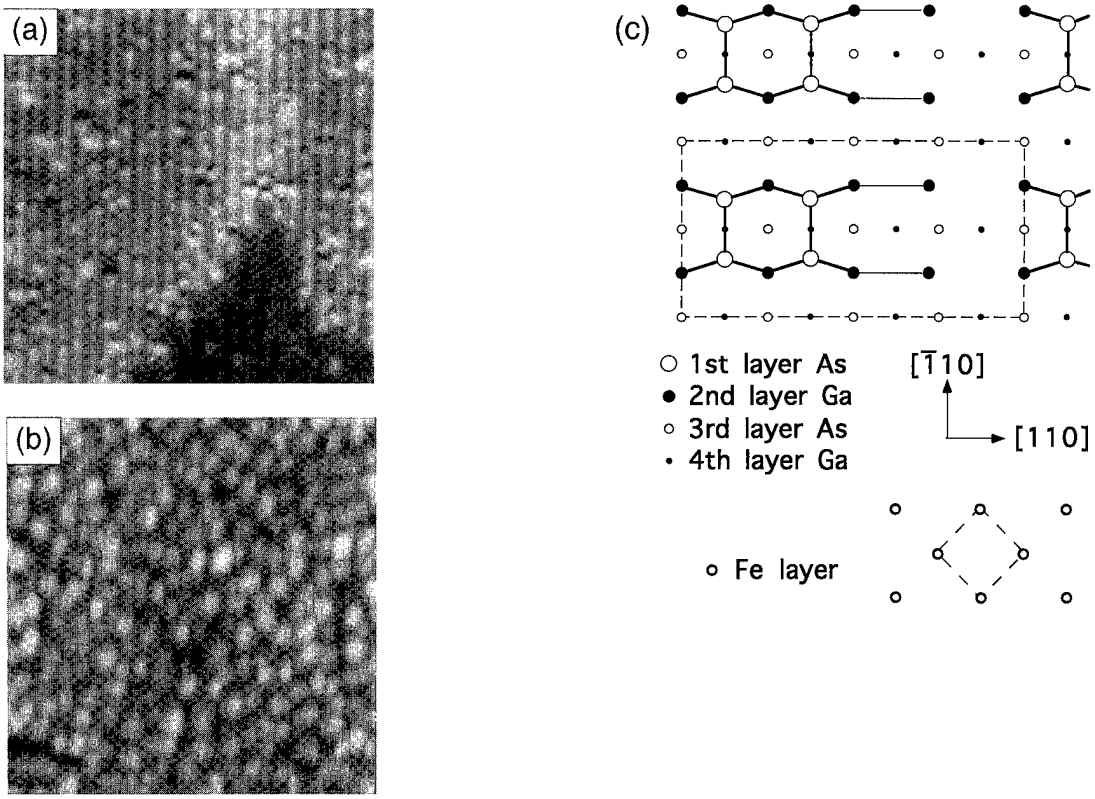


FIG. 1. (a) Filled state STM image (3 V) of 0.1 ML Fe on GaAs(001)-2×4. Image size 380 Å×380 Å. A v-shaped step edge, 2.8 Å high, is visible in the lower part of the image. (b) Filled state image (3 V) of 0.5 ML Fe. The surface corrugation is 1–2 Å. (c) Model of the 2×4 reconstructed surface, and the bcc (001) surface.

rate of 0.05 monolayer/s. We define 1 monolayer (ML)$\equiv 1.22\times10^{15}$ atoms/cm^2, corresponding to the atomic density of the bcc Fe(001) plane, which is approximately twice the atomic density of the bulk terminated GaAs(001) surface (and would produce a film ~1.4 Å thick if deposited uniformly). For the XPS core level measurements, the evaporations were performed in the PED/XPS chamber at a rate of 0.025 ML/s with an electron-beam heated rod evaporator.[8] The substrate temperature was 175 °C for all Fe depositions. The Fe deposition sources were regularly calibrated by x-ray fluorescence measurements performed on thick films (~50 Å). Relative Fe coverages were also confirmed by XPS.

After Fe growth the sample was cooled to room temperature and transferred to one of the other chambers for study. All data were collected with the sample at or slightly above room temperature. PED and XPS spectra were recorded using a monochromatized Al $K\alpha$ source and a 120 mm hemispherical analyzer with an angular resolution of ±3° (PED) and ±15° (XPS). Sample motion was computer-controlled via a stepper motor-driven two-axis goniometer, with the angle between the x-ray source and analyzer fixed at 70.8°. High resolution core level spectra ($\Delta E=0.6$ eV) were recorded at an emission angle of 15° from the surface to maximize surface sensitivity. STM images were acquired with a constant current of 0.1 nA and a sample bias of −3.0 V and are displayed in gray scale.

III. RESULTS AND DISCUSSION

A filled-state STM image of a surface covered with 0.1 ML of Fe, the lowest coverage examined, is displayed in Fig. 1(a). At this coverage most of the surface is unperturbed GaAs(001)-2×4, with bright rows oriented along the [1̄10] direction (spaced 16 Å apart) and a small corrugation visible along each row (every 8 Å). As illustrated in Fig. 1(c), the structure of this surface is believed to consist of rows of As-dimer pairs separated by Ga-terminated rows (where the As-dimer pairs are missing[9]); the bright rows observed in the STM image are associated with the rows of As dimers. The deposited Fe gives rise to the bright "bumps" that occur almost exclusively on the As-dimer rows, typically covering one dimer pair. The association of these features with Fe (and not surface defects, for instance) was confirmed by multibias imaging: whereas the Fe-related features become much more prominent at lower bias voltages, the appearance of typical defects observed on an Fe-free surface has little bias dependence.[7] These results indicate that the Fe adatoms preferentially adsorb on the As-dimer rows, forming 2D islands with an average diameter of 8 Å. It is important to note that the 2×4 surface reconstruction remains intact on the sites adjacent to the Fe islands, demonstrating that any disruption of this surface is limited to atomic-scale reactions between the Fe and As dimers in the immediate 2×4 unit cell.

The STM image obtained after 0.5 ML of Fe deposition is shown in Fig. 1(b). Although the surface is almost completely covered by 2D Fe islands, the rowlike structure of the (2×4) substrate reconstruction is still evident in the general alignment of the islands: the islands are elongated along the $[\bar{1}10]$ direction with an average size of 15 Å by 25 Å as determined from the 2D autocorrelation function. The islands are now typically wider than the As-dimer rows, indicating that Fe adsorption, initially confined to the dimer rows, has now taken place on the missing dimer (Ga-terminated) rows. The fact that the islands appear flat suggests that they are at least two Fe layers thick in these Ga-terminated areas. Although most Fe islands are approximately the same height, a small fraction of them appear noticeably higher, consistent with some additional Fe growth on the predominantly flat islands.

Whereas STM provides an excellent view of the atomic-scale surface topography following Fe deposition, PED and XPS provide complementary information regarding the structure and composition within the Fe overlayers. The integrated Fe $2p$ emission intensity as a function of polar emission angle for a range of Fe coverages between 0.5 and 31 ML is shown in Fig. 2. At this kinetic energy (772 eV) the photoelectrons are in the forward scattering regime, so that intensity maxima are expected along the directions of near-neighbor atoms (i.e., along low-index directions) [Fig. 2(c)]. In particular, the low-index directions within bulklike crystals produce well-defined peaks. Each peak in the PED spectrum for the 31-ML-thick Fe film, for example, can be identified with a low-index direction within a bcc crystal, as labeled in Fig. 2.

Since forward scattering can be observed only when atoms are above the plane of the emitting atom, the information provided by PED can be utilized to determine early growth morphology.[10] If the angle-dependent emission of an adsorbate is measured for a series of coverages, the onset of forward-scattering peaks at angles which characterize the structure signals the occupation of second, third, or fourth layer sites. A single atomically flat monolayer, for example, would produce no forward-scattering peaks, while 2 ML of layer-by-layer bcc growth would result in a single peak in the [111] direction (for the [110] azimuth). If the absolute growth rate is known, multilayer growth or clustering can be identified during deposition of the first few monolayers by the occurrence of forward-scattering peaks corresponding to the occupation of $(n+1)$th layer sites well before completion of the nth monolayer.

The polar scan along the [110] azimuth for 0.5 ML of Fe on GaAs(001)-2×4 clearly shows a maximum in the [111] direction [Fig. 2(a)], indicating that some second layer sites are already occupied. Although the PED spectra do not reveal the identity of the second layer scatterers (e.g., Fe versus As), spectra along the [110] and $[\bar{1}10]$ azimuths are equivalent, demonstrating that the occupied sites have a bcc-like structure. (Note that this would not be the case if the emitting Fe atoms were substitutionally embedded in the zinc-blende substrate.) These data indicate that the Fe adatoms are bonding either above or within the surface plane of the As dimers, with minimal disruption of the substrate surface, concurrent with some occupation of second Fe monolayer sites before completion of the first layer. These results are consistent with the STM results, which indicated that some bilayer Fe growth may occur both between the As-dimer rows and on top of the initial 2D islands. Note that within the resolution of the PED spectra, no evidence for third layer occupation is yet observed.

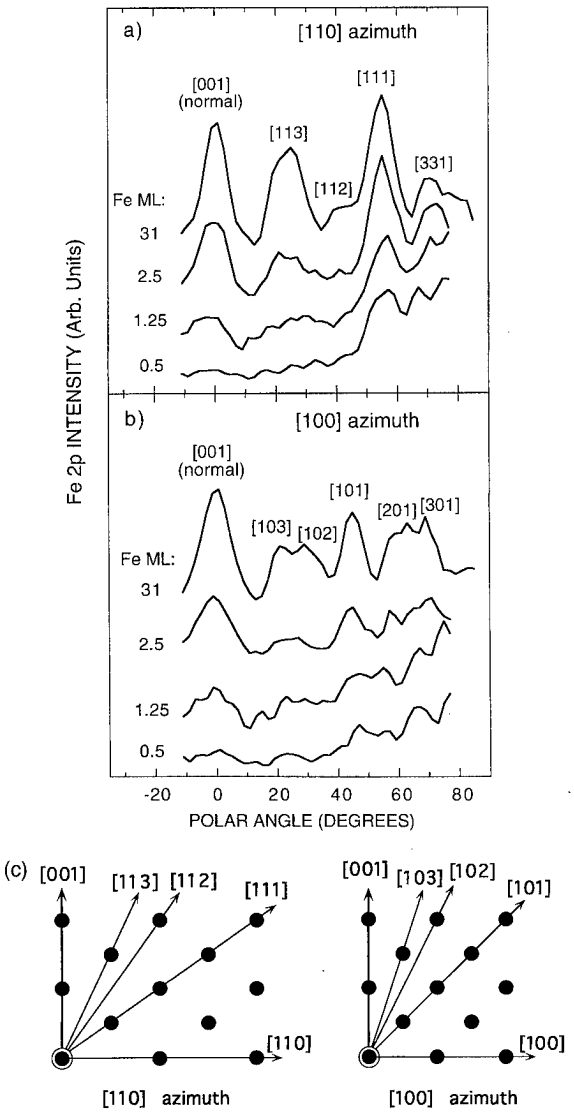


FIG. 2. (a) Fe $2p$ photoelectron polar diffraction spectra in the [110] azimuth (electron kinetic energy=772 eV) as a function of Fe coverage on GaAs(001)-2×4. The spectra are vertically offset for clarity. (b) Spectra in the [100] azimuth. (c) An illustration of forward-scattering conditions in the [110] and [100] projections for a multilayer bcc (001) film. Note that the observation of specific forward-scattering peaks can be correlated with occupation of particular layer sites within the film.

The PED data obtained at a coverage of 1.25 ML exhibit a strong peak along [111] for scans in the [110] azimuth, and weaker peaks along [101] in the [100] azimuth and along [001] for polar scans in both azimuths (Fig. 2). The [111] peak signals the expected occupation of second layer sites,

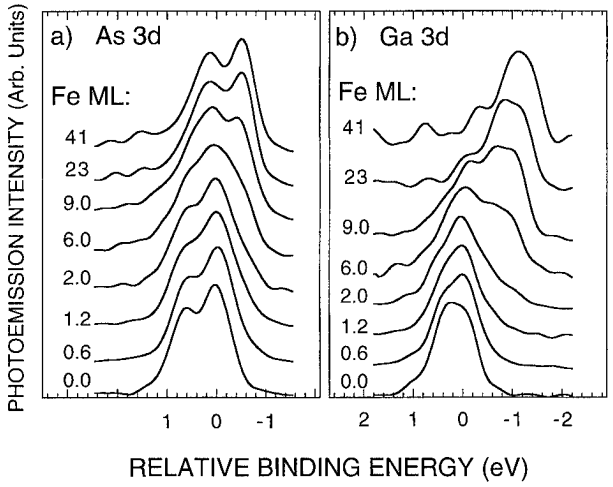


FIG. 3. XPS spectra of As and Ga $3d$ core levels as a function of Fe coverage on GaAs(001)-2×4. Each spectrum has been normalized to its peak intensity and vertically offset for clarity. Core levels shown for clean GaAs have been shifted 0.15 eV to lower binding energy to compensate for band bending.

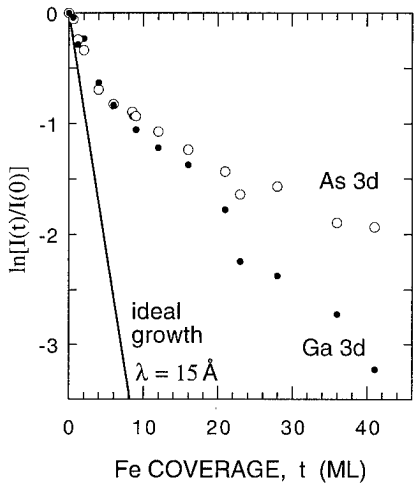


FIG. 4. Normalized Ga and As $3d$ core level intensities as a function of Fe coverage. The attenuation of the substrate levels expected for ideal Fe layer-by-layer growth (without interdiffusion) is shown for comparison as the solid line, calculated for an emission angle of 15° from the surface and an electron mean free path of 15 Å.

whereas the appearance of the [101] peak suggests occupation of some third layer bcc-like sites. Although the weak [001] peak is also indicative of third layer occupation in a simple forward-scattering picture, single-scattering calculations[11] show that this peak has contributions from higher order diffraction effects associated with the first two layers. Since this [001] peak appears together with the [101] peak, we interpret it here as further evidence for slight third layer site occupation.

As the third equivalent monolayer of Fe is deposited, the PED polar scans begin to exhibit most of the features typical of bulklike films, as can be seen by comparing the 2.5 and 31 ML data of Fig. 2. For 2.5 ML, peaks along [001] and [101] are pronounced due to the formation of the third layer, as expected, and some fourth layer site occupation is indicated by the appearance of a peak in the [113] direction. However, the absence of peaks in the [112] and [102] directions implies that there is no significant fifth layer site occupation during growth of the third ML. We thus observe that Fe grows in a predominantly layer-by-layer fashion, with the growth front for the nth deposited layer limited to the $(n+1)$th layer. Although sites in the $(n+1)$th layer are occupied before the nth layer is completed, extreme 3D growth such as that observed on the Fe/GaAs(110) system[12] can be ruled out for Fe/GaAs(100)-2×4 from our PED and STM results. For coverages $\leq$3 ML, the PED and STM data generally show the formation of predominantly flat islands with some multilayer growth occurring, but little or no tendency towards 3D clustering or significant disruption of the substrate.

The elemental composition of the Fe films is elucidated by XPS measurements as a function of Fe coverage. As shown in Figs. 3 and 4, both Ga and As can be detected even for the higher coverages studied (41 ML). Since the STM data show the GaAs surface to be completely covered by a uniform Fe overlayer after the first few monolayers[7] and the mean free path for the As and Ga $3d$ photoelectrons is ~15 Å, we attribute the persistance of these signals to out-diffusion of Ga and As from the interface. The nature of these out-diffused atoms is revealed somewhat by the $3d$ core level spectra, which evolve considerably with increasing Fe coverage (Fig. 3). Two regimes are evident: (i) coverages less than ~2 ML, where the levels show little change; and (ii) higher coverages, where distinctly shifted core levels appear. In the first regime, which we associate with the bonding of Fe to the GaAs surface, the As level broadens slightly with little change in binding energy, and the Ga level shows a slight shift in spectral weight towards lower binding energy. If this behavior is interpreted simply in terms of charge transfer accompanying bonding with the Fe adatoms, it reflects an increase in charge density on the Ga atoms but little change on the As. These features are consistent with interfacial bonding between Fe and As and concomitant debonding between As and Ga. For coverages of a few monolayers, the core levels give no indication of Fe-induced substrate surface disruption, consistent with the STM and PED results.

As the Fe coverage is increased (>2 ML), a second regime of behavior is observed in which the substratelike Ga and As $3d$ features are replaced by peaks at lower binding energy that show a considerable evolution in character with coverage. Within the resolution of our experiment, at least one new peak is observed for both the As and Ga levels, as seen in the spectra for a 6-ML-thick film. While the new As $3d$ feature is relatively constant in binding energy as a function of coverage, the Ga $3d$ spectra exhibit a more complex coverage dependence, with the new peak(s) shifting continuously to lower binding energy. For a 41-ML-thick Fe film, the net shift is 0.45 eV for the new As $3d$ doublet and 1.15 eV for the Ga, and the peaks associated with bulk GaAs are no longer observed. (The signal-to-noise ratio of the spectra

is poor at higher coverages, and the small-amplitude oscillations at higher binding energy are residues of the smoothing and background subtraction and are not considered significant.) The abrupt change in character of the $3d$ levels suggests that both Ga and As atoms are being displaced from the interface after the formation of a few monolayers of Fe, i.e., after the Fe film has assumed a metallic character. Similar core level shifts observed for Fe on GaAs(110) have been associated with Ga and As atoms multiply coordinated with Fe atoms in an increasingly metallic environment.[3]

The persistence of As and Ga $3d$ signals for thicker Fe films is illustrated in Fig. 4. The reduced intensity, defined as $\ln[I(t)/I(0)]$ where t is the Fe film thickness, is plotted for the As $3d$ and Ga $3d$ levels as a function of Fe coverage. The behavior expected for layer-by-layer growth without interdiffusion is included for comparison as the solid line. At coverages up to 2 ML, the experimental data are in reasonable agreement with this ideal growth. At coverages above 2 ML, however, the data depart significantly from this ideal behavior due to As and Ga out-diffusion. Although the Ga intensity continues to decrease monotonically, it does so much more slowly than expected. The As signal also initially decreases, but then plateaus at ~10% of its initial intensity, suggestive of As surface segregation. This behavior is qualitatively similar to that observed previously[3,4] and demonstrates that while Ga out-diffusion and incorporation is limited to the interface region, some As continuously segregates to the Fe surface.

The exact location of the out-diffused As atoms is expected to play an important role in understanding the modified magnetic behavior of the Fe film. Chambers *et al.* proposed an interstitial face-centered As site within the Fe film based on an anomalous feature in electron-stimulated Fe LMM Auger electron diffraction spectra from bcc Fe films grown on a sputter-annealed GaAs(001)-$c(8\times2)$ surface.[13] To determine the sites of the out-diffused As in our samples, we have performed polar scans of the intensity of the photon-stimulated As $L_3M_{45}M_{45}$ Auger level (kinetic energy of 1225 eV) as shown in Fig. 5. It is important to consider contributions to the emission from substrate As atoms, which is significant at near-normal angles. For polar angles >45°, however, we believe that substrate contributions become small relative to near-surface As emission. The features in both azimuths at large polar angles suggest significant As occupation of subsurface bcc sites [see Fig. 2(c)]. The overall increase in intensity at larger polar angles suggests that some of the As is also adsorbed on the surface of the Fe film, as observed for growth on the $c(8\times2)$ surface. We observe no indication for an fcc interstitial As site.

The observed out-diffusion of Ga and As at higher coverages must be reconciled with the fact that the STM images show no evidence of surface disruption during the initial nucleation of the Fe overlayer. Although the STM images provide little insight regarding the composition of the quasi-ordered 2D islands observed along the As-dimer rows, one would expect any exchange reactions with the surface atoms to result in a more disordered surface topography than is

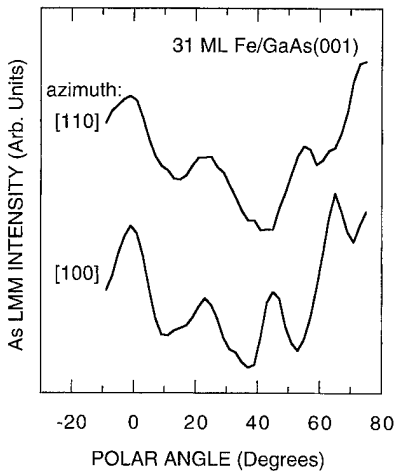


FIG. 5. Photoelectron diffraction spectra of the photon-stimulated As $L_3M_{45}M_{45}$ Auger transition (kinetic energy=1225 eV) for the [110] and [100] azimuths. Note that the spectra have many of the features characteristic of a bcc structure [see Fig. 2(c)].

observed. In addition, the core level data indicate that significant disruption of the substrate surface occurs only after deposition of 2–3 ML of Fe. A possible scenerio which accounts for the origin of the out-diffused Ga and As and the corresponding driving mechanism is proposed as follows: we speculate that the Fe would ideally like to bond to an As-terminated bulklike (1×1) surface. One pathway to such a surface on GaAs(001)-2×4 is for the Fe to initially bond to the As-dimer rows and then, with increasing coverage, to bond to less favorable Ga atoms between the dimer rows (in the second GaAs layer). As the film grows, Fe at some point (~2 ML) displaces the Ga atoms in this substrate layer, including those below the As-dimer rows. This results in an Fe/As interface and frees one layer of Ga and a half layer of As (from the dimer rows) for out-diffusion into the growing Fe film. Given that the onset of out-diffusion appears to be correlated with the onset of metallicity, we further speculate that the displacement of the second layer (substrate) Ga atoms and first layer As is driven by the accompanying change in chemical potential at the interface. The liberated Ga and As atoms are incorporated within the growing film, while a fraction of the As segregates to the surface.

IV. CONCLUSIONS

We have studied the nucleation and growth of Fe on MBE-grown GaAs(001)-2×4 at 175 °C using *in situ* STM, PED, and XPS. STM reveals that Fe initially forms small 2D islands along the As-dimer rows before growing onto the adjacent Ga-rich rows; the images show no evidence of substrate disruption. PED shows that the Fe grows in a predominantly layer-by-layer mode, with the growth front for the nth deposited layer limited to the $(n+1)$th layer. As previously observed for Fe grown on GaAs(001)-$c(8\times2)$ and GaAs(110), the films include a gradient of Ga and As con-

centrations, with some of the As segregating to the Fe surface. We speculate that As in the (2×4) dimer rows and the second substrate layer Ga are displaced by the Fe in order to make an Fe/As interface, and that this disruption is driven by a chemical potential gradient created when the Fe film becomes metallic with increasing thickness.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research. Two of the authors (E. K. and P. M. T.) further acknowledge support by National Research Council Postdoctoral Fellowships at the Naval Research Laboratory.

[1]G. A. Prinz and J. J. Krebs, Appl. Phys. Lett. **39**, 397 (1981).
[2]J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **34**, 630 (1979).
[3]M. W. Ruckman, J. J. Joyce, and J. H. Weaver, Phys. Rev. B **33**, 7029 (1986).
[4]S. A. Chambers, F. Xu, H. W. Chen, I. M. Vitomirov, S. B. Anderson, and J. H. Weaver, Phys. Rev. B **34**, 6605 (1986).
[5]J. J. Krebs, B. T. Jonker, and G. A. Prinz, J. Appl. Phys. **61**, 2596 (1987).
[6]B. T. Jonker, *Epitaxial Growth Processes*, Proc. SPIE **2140**, 118 (1994).
[7]P. M. Thibado, E. Kneedler, B. T. Jonker, B. R. Bennett, B. V. Shanabrook, and L. J. Whitman, Phys. Rev. B R10481 (1996).
[8]B. T. Jonker, J. Vac. Sci. Technol. A **8**, 3883 (1990).
[9]J. Zhou, Q. Xue, H. Chaya, T. Hashizume, and T. Sakurai, Appl. Phys. Lett. **64**, 583 (1993).
[10]S. A. Chambers, Adv. Phys. **40**, 357 (1991).
[11]D. J. Friedman and C. S. Fadley, J. Electron Spectrosc. Relat. Phenom. **51**, 689 (1990).
[12]R. A. Dragoset, P. N. First, Joseph A. Stroscio, D. T. Pierce, and R. J. Celotta, Mater. Res. Soc. Symp. Proc. **151**, 193 (1989).
[13]S. A. Chambers, I. M. Vitomirov, S. B. Anderson, H. W. Chen, T. J. Wagener, and J. H. Weaver, Superlattices Microstructures **3**, 563 (1987).

Mn 3*s* multiplet splitting of pseudomorphic Mn overlayers on Ru(001)

T. K. Sham[a)]
Department of Chemistry, The University of Western Ontario, London N6A 5B7, Canada

M. L. Shek
National Synchrotron Light Source, Brookhaven National Laboratory, Upton, New York 11973

J. Hrbek and D. G. Van Campen
Chemistry Department, Brookhaven National Laboratory, Upton, New York 11973

(Received 2 October 1995; accepted 4 December 1995)

The Mn 3*s* multiplet splitting for a series of Mn overlayers on Ru(001) has been measured with synchrotron radiation photoemission spectroscopy. These overlayers were prepared by vapor deposition of Mn onto a Ru(001) substrate at a monolayer/minute dosing rate at room temperature. Under these conditions, the Mn overlayers grow pseudomorphically to multilayer coverages. It is found that this system exhibits a noticeable increase in the Mn 3*s* multiplet splitting and a positive binding energy shift as the coverage increases from monolayer to multilayer (12 ML). These results are attributed to the increase in the atomic volume, and hence the local magnetic moment at the Mn site on the average, with increasing coverage. © *1996 American Vacuum Society.*

I. INTRODUCTION

Fabrication and characterization of magnetic thin films have attracted considerable attention in recent years partly because of the need to develop better materials for magnetic memory and related technology and partly because of the development of new and better tools, such as ultrahigh vacuum (UHV) surface science techniques and synchrotron radiation. These developments have led to more precise thin film engineering and characterization.

Mn, and its neighboring 3*d* transition metals, continue to be the subject of investigation,[1–8] since these metals possess unpaired electrons in its 3*d* shell and their magnetic thin films can be prepared under well-defined conditions using modern surface science technology. The electronic structure of Mn on Ru(001) has been investigated in several studies.[1–3,7]

It has been generally recognized that at modest deposition rates ($\leq$1/2 ML/min), Mn grows pseudomorphically on Ru(001) at room temperature up to ~3 ML. It exhibits a (1×1) low energy electron diffraction (LEED) pattern indistinguishable from that of clean Ru except for the less intense diffraction spots. Above this coverage, Mn forms a $(\sqrt{3}\times\sqrt{3})R$ 30° LEED pattern that has been associated with the Laves phase.[1] The $\sqrt{3}$ structure can grow to multilayers depending on the condition, and at higher coverage, the Mn overlayer becomes disordered.[3] Both the $(\sqrt{3}\times\sqrt{3})R$ 30° and the disordered structure are metastable; upon annealing, Mn aggregates to form three-dimensional (3D) islands exposing a large area of Ru surface.[4] More recently, we reported that at higher dosing rate (~1 ML/min), the (1×1) pseudomorphic structure persists at multilayer coverages at room temperature.[7]

The pseudomorphic growth of Mn on Ru has been very interesting in that Mn must adopt a structure with Mn–Mn distance of 2.65 Å. This is considerably larger than the closest Mn–Mn distance of 2.24 Å in α-Mn. Thus, pseudomorphic Mn on Ru has an expanded structure in which the atomic volume of Mn increases.[1] This behavior may lead to higher magnetic moment at the Mn site compared to its bulk or close-packed counterparts (such as the Laves phase). Considerable work has been done along this direction.[1,2,5,6] The electronic structure and chemical properties of this system have also been investigated.[3,4]

One of the methods for measuring the magnetic moment at the transition metal site is conventional XPS (x-ray photoelectron spectroscopy) in which the multiplet splitting of the 3*p* and 3*s* levels of the 3*d* transition elements, for example, is investigated.[9,10] The Mn 3*s* multiplet splitting arises from the exchange interaction in the final state involving the 3*s* and the 3*d* electrons (no unpaired 3*d* electron, no multiplet splitting) and the 3*s* spectrum appears as a doublet with predominantly majority (↑) and minority (↓) photoelectron spin components. The magnitude of the splitting in the case of the 3*s* level of 3*d* transition metals, for example, is proportional to the number of unpaired electrons in the 3*d* orbital (multiplicity) and the exchange interaction between the 3*s* and the 3*d* electrons.[9,10] The origin of the splitting in terms of exchange interaction has been confirmed by spin polarized photoemission studies,[11] although quantitative agreement is not always obtained[12] as the result of (i) the exchange interaction calculated based on the atomic model overestimates the splitting and (ii) configuration interactions leading to satellite structures. Despite these complications, the 3*s* multiplet splitting is still very useful for the elucidation of the systematics of the magnetic moment. In general, the larger the splitting, the large the magnetic moment. More recently, the magnetic properties of transition elements have also been investigated with x-ray absorption magnetic circular dichroism measurements at the 3*d* transition metal $L_{3,2}$ edges.[8,13]

In this article, we report photoemission measurements of the Mn 3*s* multiplet splitting for a series of Mn overlayers on Ru(100). In comparison with previous studies, our results have the following new features: (i) the Mn overlayers were prepared with a higher dosing rate (~1 ML/min) that allows

[a)]Electronic mail: sham@uwo.ca

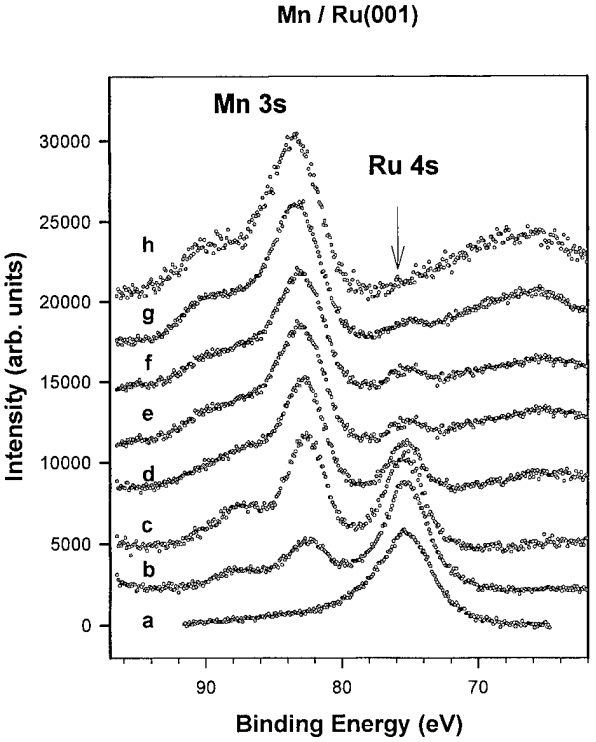


FIG. 1. Mn 3*s* photoelectron spectrum of Mn on Ru(001) with 180 eV photon energy: (a) 0 ML, pure Ru(001); (b) 1; (c) 2; (d) 4; (e) 6; (f) 8; (g) 10; (h) 12 ML.

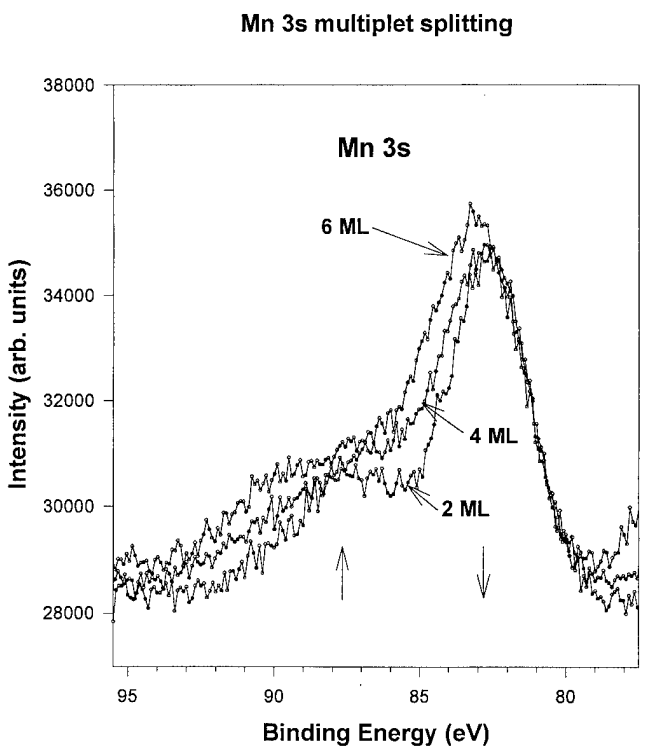


FIG. 2. Close inspection of 2, 4, and 6 ML spectra showing the asymmetric broadening. The positions for the predominantly majority (↑) and minority (↓) photoelectron spins components are noted.

for stepwise preparation of pseudomorphic films up to 12 ML thick; (ii) the photoemission experiments were carried out with optimized photon energy (hence, cross section) and with energy resolution better than conventional XPS. The experimental procedure is described in Sec. II, followed by results and discussion in Sec. III, and summary and conclusions in Sec. IV.

II. EXPERIMENTAL PROCEDURE

Photoemission measurements were carried out at the U7 beamline of the National Synchrotron Light Source (NSLS) using a plane grating monochromator. This beamline is equipped with a UHV chamber that has angle-integrated capabilities in a VSW 100 mm hemispherical electron energy analyzer. Measurements were carried out at 180 eV photon energy with the analyzer operating at 25 eV pass energy. The overall experimental resolution was ~0.4 eV. The Ru(001) crystal was positioned at 45° relative to the incoming beam and the axis of the analyzer. Mn overlayers on Ru(001) were prepared *in situ* with a temperature-controlled doser (ceramic-coated tungsten basket) operating at 700 °C and the Ru crystal is at room temperature. The dosing rate under this condition was ~1 ML/min as calibrated from thermal desorption measurements on the same Ru(001) substrate. The Mn overlayers were prepared sequentially at steps of two ML at a time after the initial depositions of the first two monolayers. The sample preparation procedure has been discussed elsewhere.[3,4]

III. RESULTS AND DISCUSSION

Figure 1 shows the photoemission spectra for a series of coverages from 1 to 12 ML of Mn on Ru(001). Independent experiments carried out in a separate chamber under identical conditions indicate that at 1 ML/min, the Mn films exhibit a (1×1) LEED pattern for the range of coverage studies here.

Several features from Fig. 1 are immediately apparent. First there are three peaks at the monolayer coverage, and at least four, perhaps five, peaks at higher coverages. The three-peak feature at monolayer coverage corresponds to Ru 4*s* (calibrated to 75.0 eV) and the multiplet split Mn 3*s* with the minority (↓) and majority (↑) spin peak at binding energies of ~82.6 and ~87.3 eV, respectively. As expected, the Mn 3*s* peaks become more intense relative to the Ru 4*s* signal as the Mn coverage increases. Second, there appears to be a broad peak at ~66 eV that also becomes more intense with increased Mn coverage. This peak is one of the satellite peaks originating from the photoemission of the Mn 3*p* level;[9,11,13] it was too weak to be detected at low coverage. Third, the Mn 3*s* doublet exhibits a noticeable increase in the apparent splitting as the coverage increases; this is accompanied by a positive binding energy shift.

Close inspection of Fig. 1 reveals the following: (i) the 1 and 2 ML spectra have well-defined doublets with same multiplet splitting; (ii) the 4 and 6 ML spectra exhibit a noticeable asymmetric broadening (to higher binding energy) in both components; the higher energy component is poorly defined; and (iii) the 8, 10, and 12 ML spectra exhibit similar

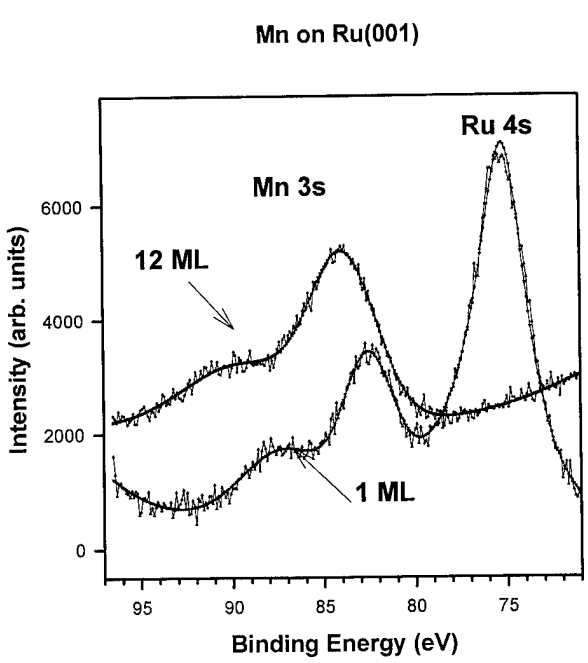


FIG. 3. Comparison of the 1 with the 12 ML spectrum; the solid line is the fit.

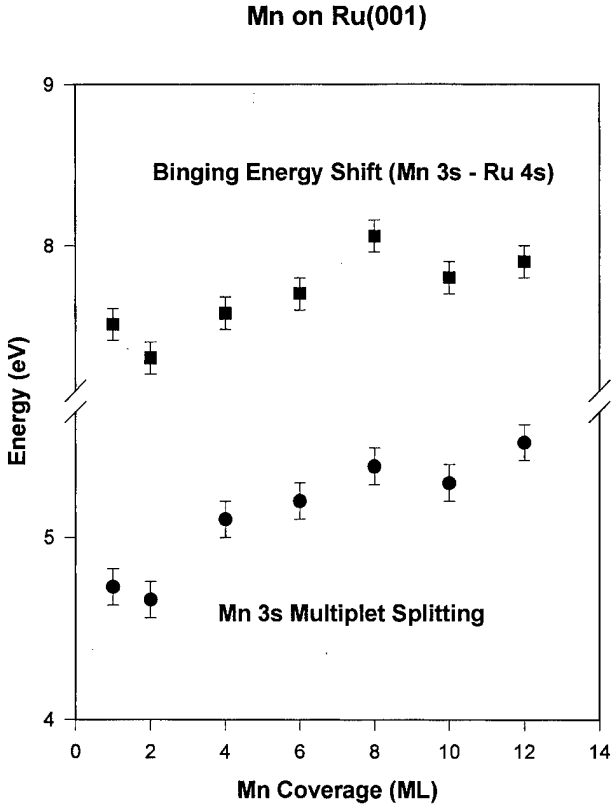


FIG. 4. A plot of multiplet splitting and binding energy shift (Mn 3*s* minority spin–Ru 4*s*) vs. Mn coverage (ML).

doublets with the largest splitting. The components are slightly broader then those for lower coverage and are shifted to higher binding energy. Figure 2 shows the evolution of the Mn 3*s* multiplets from 2 to 6 ML. It can be seen that the doublets broaden and shift progressively towards higher binding energy as coverage increases. The broadening is attributed to surface inhomogeneity due to the presence of Mn atoms with slightly larger magnetic moment. Figure 3 compares the two extreme cases (1 and 12 ML); the increase in splitting and accompanying binding energy shift at high coverages are clearly visible.

The spectra shown in Fig. 1 have also been analyzed by fitting. The main consideration here is that the linewidth for the doublet is not required to be the same, as was assumed previously in some analyses.[10] This consideration comes about from the spin-resolved photoemission studies of the Fe 3*s* multiplets.[11] In the Fe study,[11] it was clearly shown that the majority spin photoelectron peak is considerably broader than that of the minority spin. This linewidth difference between the two components is related to the lifetime of the corehole and manybody effects.[11,12] A couple of configuration interaction satellites related to the majority spin are known to present at higher binding energies.[12] The presence of the satellites reduces the intensity of the majority spin peak.

We fitted all the spectra with a Doniach–Sunjic lineshape for the Ru 4*s*, and Gaussians for the Mn doublets and the Mn 3*p* satellites. The choice of using Gaussian in the latter is motivated by the surface inhomogeneity that is expected. We consistently obtained reasonable fits with the linewidth of the majority spin peak being ~50% broader than the minority spin peak, which has a width of ~3–3.4 eV.[14] The results of multiplet splitting and binding energy shift (minority spin peak referenced to Ru 4*s*) as a function of coverage are summarized in Fig. 4.

It is quite clear from these observations that the multiplet splitting, and hence the average magnetic moment at the Mn site in these overlayers, increases from 4.7 to 5.5 eV. It is interesting to compare our results with some previous work. McFeely *et al.*[15] showed that α-Mn, which has a magnetic moment of $3.5\mu_B$, exhibits a splitting of ~4 eV, while ionic Mn compounds, which have a magnetic moment of $5\mu_B$, exhibit an ~6 eV splitting. Heinrich *et al.*[2] reported splitting of 4.30 and 4.45 eV for the Laves phase and the expanded structure of Mn(2 ML)/Ru(001), respectively. Our 2 ML result is in good agreement with the result of Heinrich *et al.*[2] The much larger splitting observed in our work confirms that the Mn overlayer remains pseudomorphic, consistent with our other measurements.[7] This observation means that the atomic volume of Mn, on the average, increases with coverage in these overlayers. By scaling with the results of McFeely *et al.*,[15] we obtain magnetic moment of ~$3.9\mu_B$ and ~$4.6\mu_B$ at 1 and 12 ML coverage, respectively. It is apparent that at a higher dosing rate (~1 ML/min), the thermally unstable, expanded (1×1) pseudomorphic structure is preserved at room temperature. This observation is in contrast to a recent observation of Mn on Ir, where it was found that the multiplet splitting of the expanded structure (~2 ML Mn, pseudomorphic on Ir) is larger than that of a thicker layer (10 ML on Ir).[8] The discrepancy between the Mn/Ir and our re-

sults may be related to the difference in the morphology of the overlayers.

It should be noted that the observed binding energy shift (Fig. 4) can be attributed to the effect of volume induced rehybridization. It is well known that the condensation of free atoms of transition metals favors $s \rightarrow d$ charge rehybridization. Thus, a lattice expansion would most likely induce a $3d \rightarrow 4s$ charge transfer in Mn. Since the binding energy shift is primarily determined by the coulomb integral, F($3s$, valence), and for Mn, F($3s$,$3d$)>F($3s$,$4s$), a $3d \rightarrow 4s$ electron redistribution would lead to a positive binding energy shift as observed. More work is needed to substantiate these effects.

IV. SUMMARY AND CONCLUSION

We have presented Mn $3s$ multiplet splitting systematics for a series of Mn overlayers adsorbed on Ru(001) at room temperature. The trend of the multiplet splitting indicates an increase in magnetic moment at the Mn site ($3.9–4.6\mu_B$) with coverage from 1 to 12 ML, respectively. We conclude that this trend is associated with an increase of atomic volume of Mn; this observation is consistent with pseudomorphic growth. The binding energy shift with increasing coverage is also noted and interpreted in terms of the volume effect on $4s-3d$ charge rehybridization.

ACKNOWLEDGMENTS

Research carried out at Brookhaven National Laboratory was supported under Contract No. DE-AC02-76CH00016 with the U.S. Department of Energy and supported by its Division of Chemical Sciences, Office of Basic Energy Science. Research at the University of Western Ontario was supported by NSERC of Canada and the Ontario Centre for Materials Research.

[1]A. S. Arrott, B. Heinrich, C. Liu, and S. T. Purcell, J. Vac. Sci. Technol. B **3**, 566 (1985).
[2]B. Heinrich, C. Liu, and A. S. Arrott, J. Vac. Sci. Technol. B **3**, 766 (1985).
[3]J. Hrbek, T. K. Sham, and M.-L. Shek, Surf. Sci. **191**, L772 (1987).
[4]J. Hrbek, J. Phys. Chem. **94**, 1564 (1990).
[5]V. Dupuis, M. Maurer, M. F. Ravet, J. Dekoster, S. Andrieu, J. F. Bobo, F. Baudelet, P. Bauer, and A. Fontaine, Phys. Rev. B **48**, 5585 (1993).
[6]B. Heinrich and J. F. Cochran, Adv. Phys. **42**, 523 (1993).
[7]T. K. Sham, M. L. Shek, M. Kuhn, and J. Hrbek, Physica B **208&209** 469 (1995).
[8]W. L. O'Brien and B. P. Tonner, J. Vac. Sci. Technol. A **13**, 1544 (1995).
[9]C. S. Fadley and D. A. Shirley, Phys. Rev. A **2**, 1109 (1970).
[10]J. F. van Acker, Z. M. Stadnik, J. C. Fuggle, H. J. W. M. Hoekstra, K. H. J. Buschow, and G. Stroink, Phys. Rev. B **37**, 6827 (1988).
[11]F. U. Hillelbrecht, R. Jungblut, and E. Kisker, Phys. Rev. Lett. **65**, 2450 (1990).
[12]B. Hermsmeier, C. S. Fadley, M. O Krause, J. Jimenez-Mier, P. Gerard, and S. T. Manson, Phys. Rev. Lett. **61**, 2592 (1988).
[13]See Proceedings of the XAFS IIIV Conference, Physica B **208&209** (1995).
[14]In the case of 4 and 6 ML where the majority spin peak is better represented with two Gaussians of which the weighted average gives the expected position of the majority spin peak. The energy positions obtained from the fitting are in fact very close (within 0.1 eV in most cases) to those obtained by inspection with differentiation. The inherent lifetime broadening for Mn $3s$ and Ru $4s$ core hole is 2.1 and 3.2 eV, respectively [from *Unoccupied Electronic States*, edited by J. C. Fuggle and J. E. Inglesfield (Springer, Berlin, 1992), Appendix B].
[15]F. R. McFeely, S. P. Kowalszcyk, L. Ley, and D. A. Shirley, Solid State Commun. **15**, 1051 (1974).

Temperature dependent magnetic surface anisotropy in ultrathin Fe films

D. P. Pappas[a)]
Virginia Commonwealth University, Richmond, Virginia 23284-2000

(Received 2 October 1995; accepted 7 May 1996)

The temperature dependence of the uniaxial surface anisotropy, $K_u(T)$, in ultrathin Fe films (4–10 atomic layers) grown on Cu(100) is determined by comparing the dipole and anisotropy energies at the spin reorientation temperature for films of varying thickness. It is observed that the uniaxial anisotropy has a weak temperature dependence compared to the bulk constants for Fe. The measured exponent of $\Gamma_u=2.6(0.5)$ agrees well with the $l(l+1)/2$ law, which is obtained from a spin fluctuation model where $l=2$ at the surface. This shows that the spin reorientation transition can be understood as being driven by thermal spin fluctuations. © *1996 American Vacuum Society.*

Ferromagnetism in films of transition metals a few atomic layers (ALs) thick has become well established since the development of modern vacuum equipment has allowed the preparation of high quality films.[1] The most striking behavior of some of these films, e.g., Fe grown on Cu(100) or Ag(100), is the presence of magnetic anisotropies large enough to overcome the dipole energy of the film and pull the magnetization perpendicular to the plane of the film.[2,3] The presence of magnetic anisotropy is very important to the magnetic ordering of these systems because it has been shown that the critical behavior of the magnetism of very thin films comes very close to that of a true two-dimensional system.[4] Since an isotropic two-dimensional system at finite temperature has no long range order,[5] it is very important to understand how anisotropy enters into the problem. The Hamiltonian for a Heisenberg spin system can be written as a sum of exchange energy, spin—spin (dipole) coupling, and spin–orbit terms (with no external field),

$$\mathcal{H}=\mathcal{H}_{\rm ex}+\mathcal{H}_D+\mathcal{H}_{\rm so}. \tag{1}$$

If the exchange term is taken to be isotropic, e.g., using the Heisenberg model, then the anisotropy originates either in the dipole or the spin–orbit terms.

For a constant saturation magnetization M_s, the classical dipole energy/unit area of the film is obtained from

$$E_D=-\tfrac{1}{2}\mathbf{H}_{\rm loc}\cdot M_s a_0, \tag{2}$$

where a_0 is the lattice constant. The local field $\mathbf{H}_{\rm loc}$ at the point $\mathbf{r}_i$ exerted by a collection of point dipoles $\mathbf{p}_j$ at $\mathbf{r}_{ij}=|\mathbf{r}_i-\mathbf{r}_j|$ is given by

$$\mathbf{H}_{\rm loc}=-\sum_{j\neq i}\left[\frac{\mathbf{p}_j}{r_{ij}^3}-3\frac{\mathbf{r}_{ij}\cdot\mathbf{p}_i\mathbf{r}_{ij}}{r_{ij}^5}\right]\cdots. \tag{3}$$

Evaluating this sum for perpendicular and in-plane magnetizations gives rise to a demagnetizing field $4\pi M_s$ in the plane of the film. Therefore, the total demagnetizing energy of the film (*shape* anisotropy energy) is

$$E_D=2\pi M_s^2, \tag{4}$$

which always favors magnetization in the plane of the film. It has been shown that only the presence of a dipole term is sufficient to cause spontaneous magnetization.[6]

The spin–orbit term originates from the interaction of the electrons with the crystal lattice. In bulk bcc-Fe, for example, the spin orbit term is anisotropic, favoring magnetization in the [100]-like directions.[7] The bulk anisotropy energy/unit volume in Fe is about an order of magnitude smaller than the shape anisotropy energy density in a thin film (10^6 vs 10^7 erg/cm^3) and, hence, is unlikely to be a dominant contribution in Eq. (1). The breaking of symmetry at a surface can also cause the spin–orbit interaction to be anisotropic. In cubic materials the anisotropy vanishes to fourth order, however, at the surface the broken symmetry brings the second order anisotropy term into play. Neél[8] has discussed the surface anisotropy with a phenomenological theory. Using the classical form of the pairwise interactions between neighboring atoms, he predicted that a strong anisotropy perpendicular to the plane of the film is a natural consequence of the reduced symmetry at, e.g., the (100) face of a cubic lattice. The contribution of this term can be included into the Hamiltonian as

$$\mathcal{H}_{\rm so}=-2K_u\sum_i(\mathbf{S}_i\cdot\mathbf{n})^2, \tag{5}$$

where K_u is a uniaxial surface anisotropy energy/unit area and the factor of 2 includes both the top and bottom surfaces. Here, a positive value of K_u will minimize the energy, thus favoring magnetization perpendicular to the film. The energy density associated with this surface anisotropy term then becomes

$$E_{\rm Surf}=\frac{2K_u}{d}, \tag{6}$$

with d being the film thickness. The first efforts to evaluate the direction and strength of K_u were made by Neél using typical elastic and magneto-striction constants. He predicted that K_u at the (100) face would favor magnetization perpendicular to the film and be on the order of 1 erg/cm^2. Thus, for a film of thickness less than about 10 Å the surface anisotropy energy/unit area, $E_{\rm Surf}>10^7$ erg/cm^3, would be large enough to overcome the dipole energy. However, it is neces-

a)Electronic mail: DPAPPAS@CABELL.VCU.EDU

sary to do a fully relativistic bandstructure calculation in order to predict from first principles the strength and sign of the magneto-crystalline anisotropy in transition metals. *Ab initio* electronic structure calculations which include the spin–orbit interaction correctly predicted that the moments of some thin transition metal films orient themselves perpendicular to the film,[9] however it is clear that the difference between perpendicular and in-plane energies is small, and the calculations are pushing to the limits of present computational techniques.[10,11]

These theoretical considerations in combination with experimental studies that confirmed that at low thickness the magnetization is perpendicular to the plane of the film has sparked much work over the past few years on this topic. In particular, it has been observed that films thinner than 5–6 AL of Fe grown on Ag(100)[12] are perpendicularly magnetized, and become magnetized in the plane of the film above this critical thickness. A similar spin reorientation behavior for Fe grown on Cu(100) is observed with a critical thickness of about 6 AL. These observations agree with expectations from ferromagnetic resonance results, in which the uniaxial anisotropy for Fe/Ag(100) was measured to be on the order of 1 erg/cm^2.[13] Contributions to K_u may also arise from epitaxial strain,[14] however it is interesting to note that the spin reorientation transition thickness is very close for Fe grown on Ag(100) and Cu(100), even though the structures are entirely different (bcc and fct, respectively).

The first indications of a temperature dependent spin reorientation from perpendicular at low T to in-plane at high T was observed by Jonker *et al.*,[15] using spin resolved photoemission on Fe films grown on Ag(001), and was then verified by Volkening *et al.*[16] using the conversion electron Mössbauer effect. Subsequent studies of Fe/Cu(100) grown at low temperature showed that the temperature dependent spin reorientation transition is reversible and occurs over a narrow temperature and thickness range[17] with an accompanying loss of magnetic signal as the reorientation occurs for both Fe/Ag(100) and Fe/Cu(100).[18] This loss of signal near has been investigated both from an experimental[19,20] and theoretical perspective,[21,22] and has been shown to be due to the formation of microscopic stripe domains in the region of thickness and temperature that $E_D = E_{\mathrm{Surf}}$. This is important because it shows that the higher order anisotropy terms are sufficient to maintain short range magnetic order.

Observation of a temperature dependent spin reorientation leads to the consideration of the temperature dependence of the anisotropy constants with the conclusion that the surface anisotropy energy, E_{Surf}, decreases faster than the dipole energy, E_D, as the temperature is increased. This behavior can be expected because the origin of K_{Surf} is the same as that of the bulk magneto-crystalline anisotropy, i.e., the spin–orbit interaction. Callen and Callen have discussed the temperature dependence of the bulk anisotropy constants[23] using a spin fluctuation model. By allowing each spin to sample a local temperature independent anisotropy $K_l(0)$ and then allowing the spins to deviate from their equilibrium positions, the following relation is obtained:

$$\frac{K_l(T)}{K_l(0)} = \left[\frac{M_s(T)}{M_s(0)}\right]^{\Gamma}, \tag{7}$$

where l is the order of the anisotropy and $\Gamma = l(l+1)/2$. Good agreement with experiment is obtained using this expression for bulk bcc-Fe with exponent of $\Gamma_B = 10$, where $l=4$ for a cubic system. For the surface, however, an exponent of $\Gamma_u = 3$ is expected because the anisotropy comes in at second order, i.e., $l=2$.[9] Spin fluctuation theory has been used successfully to understand the spin reorientation transition,[21,22] however most emphasis has been placed on the nature of the domain formation at the reorientation transition rather than the cause of the temperature dependence of the uniaxial anisotropy. It has also never been shown experimentally that spin fluctuations can explain this effect without including thermal stress and bandstructure changes into the problem. In the present work, the exponent Γ is obtained for ultrathin Fe films grown on Cu(100) by comparing the saturation magnetization and thickness at which the spin reorientation occurs. This is achieved by finding the normalized magnetization $[M_s(T)]/[M_s(0)]$ at the spin reorientation temperature T_R for a given thickness,[17] and equating the dipole energy density to the surface energy density at that thickness. Good agreement with the spin fluctuation model is found.

In order to find $[M_s(T)]/[M_s(0)]$ it is necessary to fit the dependence of M_s with respect to temperature in the range where the transition occurs because the long range order is reduced at T_R.[18] Figure 1 illustrates the temperature and thickness dependence of the magnetization in thin Fe films which were grown at 125 K and annealed while the secondary electron polarization was monitored.[17] The reduction of signal at T_R is demonstrated in the center panel for the 5.4 AL film and has been shown to be due to the formation of microscopic domains in this temperature range.[20–22] The 4 and 10 AL films show only remanent magnetization perpendicular and in-plane, respectively. These results match with magneto-optical Kerr effect (MOKE) studies on Fe/Cu(100), where the border between perpendicular and in-plane magnetization was observed at $\approx$6 AL.[12,3] The relevant aspect of these curves for this work is the decrease of the magnetization with temperature up to about 300 K. In the spin wave regime, i.e., for $T < T_c/2$, it is well known that the bulk and surface magnetization can be fit using the Bloch law $T^{3/2}$ law.[7,24,25] For a thin film, however, it can be shown[26] that the behavior of the magnetization is more appropriately fit with a function of the form

$$\frac{M_s(T)}{M_s(0)} = 1 - kT \ln T. \tag{8}$$

The low temperature fit of the polarization data of the 5.6 AL film to Eq. (8) is shown in Fig. 2. Since it has been shown that the films are in a single domain state in this temperature region,[12,19] this fit can be used to determine the magnetization at the spin reorientation temperature. A prefactor of $k = 0.000\,68$ is obtained for the 5.6 AL film. Using this fit, it is now possible to determine the relation between $M_s(T)$ and

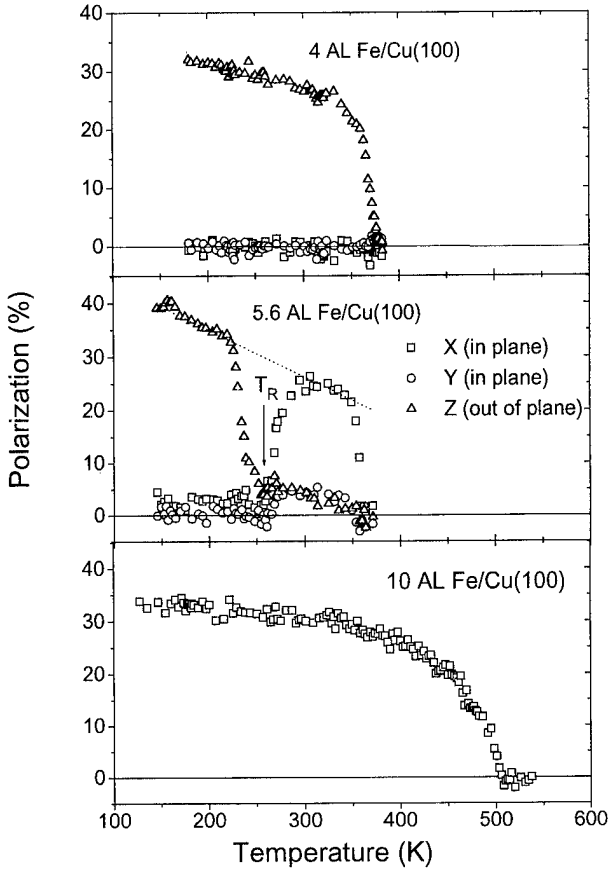


FIG. 1. Temperature dependence of the spin polarization of secondary electrons from a 4 AL (top), 5.6 AL (center), and a 10 AL film (bottom panel).

$K_u(T)$ for the first time from the data of Ref. 17. This can be obtained by observing that at the transition temperature T_R; the dipole energy/unit volume E_D of the film is equal to the energy due to the uniaxial anisotropy $K_u(T_R)$. From Eq. (4) and Eq. (6) for a given thickness d_R, we obtain

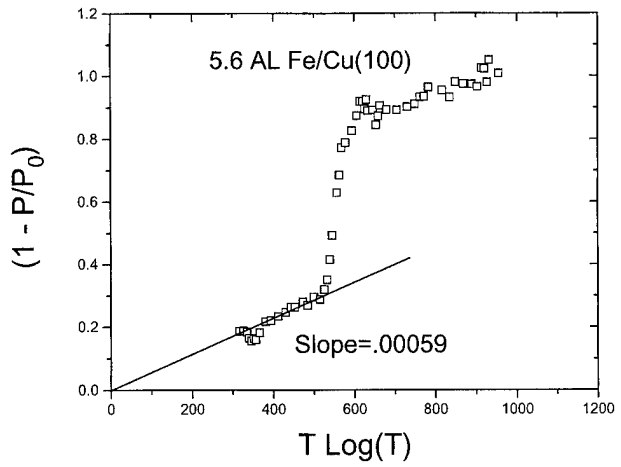


FIG. 2. Fit of the data from Fig. 1 to find the coefficient k of the equation $M_s(T)=M_s(0)[1-k\times T\ln T]$ in the spin wave regime for 5.6 AL film.

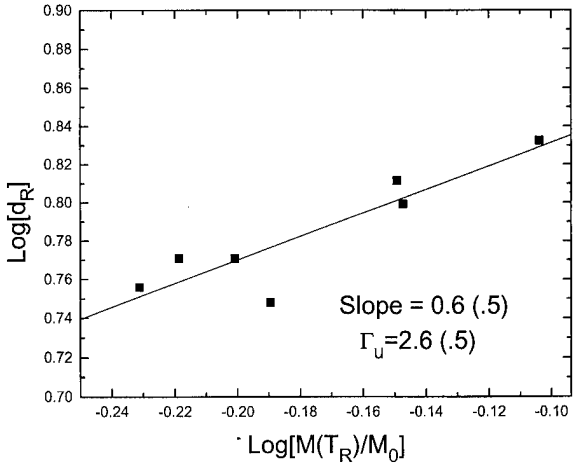


FIG. 3. Log–log plot of the thickness d_R vs magnetization $M_s(T_R)$ at the spin reorientation transition temperature. The data from this plot is obtained from the same data set as Ref. 17, where the corrected thickness calibration (see Ref. 18) is used.

$$\frac{2K_u(T_R)}{d_R}=2\pi M_s^2(T_R). \tag{9}$$

Here, the normalized anisotropy and magnetization can be used, and the relation of Eq. (7) substituted for the value of $[K_u(T)]/[K_u(0)]$ to find

$$\frac{2}{d_R}\left[\frac{M_s(T_R)}{M_s(0)}\right]^{\Gamma_u}=2\pi\left[\frac{M_s(T_R)}{M_s(0)}\right]^2 \tag{10}$$

and thus

$$d_R=\frac{1}{\pi}\left[\frac{M_s(T_R)}{M_s(0)}\right]^{\Gamma_u-2}. \tag{11}$$

A log–log plot of d_R vs $[M_s(T_R)]/[M_s(0)]$ will thus give a slope of Γ_u-2. Therefore, for each of seven films that demonstrated a spin reorientation transition the normalized magnetization is determined at T_R using Eq. (8). The results are plotted in Fig. 3. The linear least squares fit shown gives a slope of 0.6 ± 0.5, and therefore $\Gamma_u=2.6\pm0.5$.

In this work, the magnetic properties of Fe grown on Cu(100) substrates at low temperature are studied. The uniaxial anisotropy is observed to fall off faster than the bulk anisotropy as the magnetization decreases. A reduced exponent relative to the bulk is determined that agrees well with the spin fluctuation model exponent of $l(l+1)/2=3$ where $l=2$ at the surface. This demonstrates that thermal fluctuations have a much smaller effect on the anisotropy at the surface than in the bulk of a cubic material because the next closest energy minimum is 180° away versus only 90° in the bulk. This shows that the magnetic anisotropy can be understood using a thermal spin fluctuation model. However this does not rule out the possibility that thermal stress and/or band structure changes affect the magnetic properties of the films.

Acknowledgments: The authors would like to acknowledge H. Hopster, D. L. Mills, R. P. Erickson, and S. S. P. Parkin for their guidance in obtaining and interpreting the

data. This material is based upon work supported by the National Science Foundation under Grant No. DMR 94-58004. Additional support from Research Corporation under Grant No. CC3778 and the Jeffress Trust No. J338 is also acknowledged.

[1]U. Gradmann and J. Müller, Phys. Status Solidi **27**, 313 (1968); U. Gradmann, Appl. Phys. Lett. **3**, 161 (1974).
[2]D. Pescia, M. Stampanoni, G. L. Bona, A. Vaterlaus, R. F. Willis, and F. Meier, Phys. Rev. Lett. **58**, 2126 (1987).
[3]C. Liu, E. R. Moog, and S. D. Bader, Phys. Rev. Lett. **60**, 2422 (1988).
[4]Z. Q. Qiu, J. Pearson, and S. D. Bader, Phys. Rev. Lett. **67**, 1646 (1991).
[5]N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[6]R. P. Erickson, Phys. Rev. B **46**, 14 194 (1992).
[7]C. Kittel, *Introduction to Solid State Physics*, 5th ed. (Wiley, New York, 1976).
[8]Neél, J. Phys. Radium **15**, 225 (1954).
[9]J. G. Gay and R. Richter, J. Appl. Phys. **61**, 3362 (1987).
[10]G. H. O. Daalderop, P. J. Kelly, and M. F. H. Schuurmans, Phys. Rev. B **41**, 11919 (1990).
[11]J. B. Staunton, Rep. Prog. Phys. **57**, 1289 (1995).
[12]S. D. Bader, Donqi Li, and Z. Q. Qui, J. Appl. Phys. **76**, 6419 (1994).
[13]B. Heinrich, K. B. Urquhart, A. S. Arrott, J. F. Cochran, K. Myrtle, and S. T. Purcell, Phys. Rev. Lett. **59**, 1756 (1987).
[14]R. H. Victora and J. M. Maclaren, Phys. Rev. B **47**, 11 583 (1993).
[15]B. T. Jonker, K.-H. Walker, E. Kisker, G. A. Prinz, and C. Carbone, Phys. Rev. Lett. **57**, 142 (1986).
[16]F. A. Volkening, B. T. Jonker, J. J. Krebs, N. C. Koon, and G. A. Prinz, J. Appl. Phys. **63**, 3869 (1988).
[17]D. P. Pappas, K.-P. Kamper, and H. Hopster, Phys. Rev. Lett. **64**, 3179 (1990).
[18]D. P. Pappas, C. R. Brundle, and H. Hopster, Phys. Rev. B **45**, 8169 (1992).
[19]R. Allenspach and A. Bischof, Phys. Rev. Lett. **69**, 3385 (1992).
[20]Z. Q. Qiu, J. Pearson, and S. D. Bader, Phys. Rev. Lett. **70**, 1006 (1993).
[21]P. J. Jensen and K. H. Bennemann, Phys. Rev. B **42**, 849 (1990).
[22]A. Kasuba and V. L. Pokrovsky, Phys. Rev. Lett. **70**, 3155 (1993); Phys. Rev. B **48**, 10335 (1993).
[23]H. B. Callen and E. R. Callen, J. Phys. Chem. Solids **27**, 1271 (1966); E. R. Callen and H. B. Callen, *ibid.* **16**, 310 (1960).
[24]G. T. Rado, Bull. Am. Phys. Soc. **2**, 127 (1957).
[25]D. L. Mills and A. A. Maradudin, J. Phys. Chem. Solids **28**, 1855 (1967).
[26]R. P. Erickson and D. L. Mills, Phys. Rev. B **46**, 861 (1192).

Distinguishing the close-packed hexagonal and face centered cubic phases of the metallization of diamond by polarization-dependent extended x-ray absorption fine structure

K. M. Kemner,[a] W. T. Elam, V. G. Harris, Y. U. Idzerda, and J. A. Wolf
Naval Research Laboratory, Washington, DC 20375

(Received 2 October 1995; accepted 22 January 1996)

Polarization-dependent extended x-ray absorption fine structure (PD-EXAFS) measurements have been made on a single-crystal 1000-Å-thick Co film deposited on (100) diamond. By comparing the EXAFS signals corresponding to the in-plane and out-of-plane structure relative to the film plane, we clearly determine that the film does establish the face centered cubic (fcc) phase. Distinguishing between the hcp and fcc phases can thus be performed by qualitatively noting the polarization dependence of the Fourier transform peak amplitudes corresponding to the third and fourth coordination shells. This EXAFS study shows that the Co film studied here is unambiguously in the fcc phase. The Fourier transform peaks corresponding to the first nearest neighbors, parallel and perpendicular to the film, have been compared. Results indicate less than a 0.0075 Å difference between in-plane and out-of-plane bondlengths, no significant change in coordination number and less than a 0.001 Å^2 difference in EXAFS Debye–Waller factors. All of these results indicate no tetragonal distortion of the fcc Co crystal lattice. This technique should prove valuable in future studies when trying to determine the crystallinity of thin films. © *1996 American Vacuum Society.*

I. INTRODUCTION

The interest in the theoretical and experimental investigations of metastable phases and their unique characteristics has steadily increased during the past few years.[1] By using lattice-matched substrates, the growth of a wide variety of exotic phases of common materials has been performed. Similarly, the application of these new materials in the field of magnetoelectronics has also provided exciting results.[2]

Among the 3*d* metals, Co has generated renewed interest for after having been stabilized in the body centered cubic (bcc)[3] and the thermodynamically stable close-packed hexagonal (hcp)[4–6] phases, recent results seemed to indicate that it had now been grown in the metastable face centered cubic (fcc) phase.[7] Because of the unique behavior of metastable fcc Co, a new field of applications seems to be very promising.[8]

II. EXPERIMENT

Polarization-dependent extended x-ray absorption fine structure (PD-EXAFS) measurements have been made on a single-crystal 1000-Å-thick Co film deposited on (100) diamond. The growth (molecular beam epitaxy) and characterization (reflection high energy electron diffraction, x-ray diffraction, vibrating sample magnetometry, quantum interference magnetometry, and ferromagnetic resonance) of this film has been documented elsewhere.[7] X-ray absorption spectra were collected in electron yield mode[9] utilizing a low temperature total electron yield (TEY) detector[10] that kept the sample temperature at 90±3 K during the data acquisition. The data were collected near the boiling point of liquid nitrogen to reduce effects of any thermal Debye–Waller terms that dampen the amplitude of the EXAFS signal. A reduction in the thermal Debye–Waller terms increases the range of useful data that can be Fourier transformed, and thus increases the resolution of the Fourier transform peaks. Measurements were made on the National Institute of Standards and Technology EXAFS beamline X23A2 at the National Synchrotron Light Source (Brookhaven National Laboratory, Upton, NY). This beamline is equipped with a fixed exit double crystal monochromator that uses (311) Si crystals. Due to the use of the (311) Si crystals, no harmonic rejection was required. These experiments were performed with stored beam currents of 190–230 mA and beam energy of 2.54 GeV. The initial x-ray intensity was monitored with a 35.4-cm-long ionization chamber filled with free flowing nitrogen gas.

The extended fine structure appearing above the Co *K* absorption edge (7709 eV) was measured with the incident x-ray beam at a glancing angle incidence of less than 10° with respect to the film surface and the x-ray beam's electric field vector parallel and then perpendicular to the film surface. Because the EXAFS signal originates with a $\cos^2$ angular dependence with respect to the electric-field vector of the incident radiation and the bond between the absorbing and backscattering atom,[11] these orientations allow preferential sampling of the local in-plane and out-of-plane structure, respectively. At least two successive energy scans were taken at a minimum of three slightly different azimuthal angles so as to differentiate between the EXAFS oscillations and the Bragg reflections from the diamond substrate, which were later deglitched.[12] A different azimuthal angle was obtained by rotating the sample about an axis normal to the sample by approximately 2°.

The experimentally measured x-ray absorption cross section $\mu_e(E)$, where E is the x-ray photon energy, is propor-

[a]Electronic mail: kemner@anvil.nrl.navy.mil

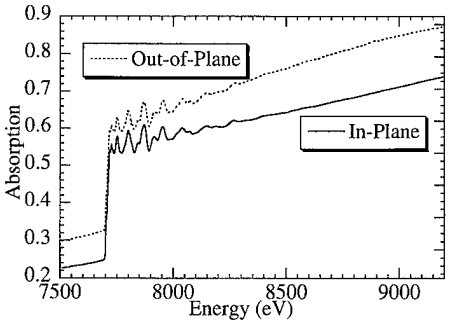


FIG. 1. Raw XAFS data for the average in-plane (solid line) and out-of-plane (dashed line) local structure surrounding the Co atoms in a 1000-Å-thick Co film deposited on diamond.

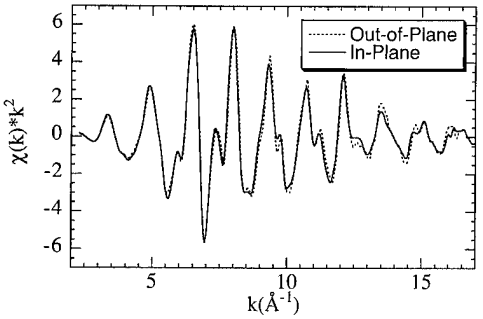


FIG. 2. Normalized $\chi(k)*k^2$ data corresponding to the average in-plane (solid line) and out-of-plane (dashed line) structure surrounding the Co atoms in a 1000-Å-thick Co film deposited on diamond.

tional to I_e/I_o for TEY data, I_o and I_e being the incident x-ray and total electron yield intensities, respectively. The raw TEY data corresponding to the average in-plane (solid line) and out-of-plane (dashed line) structure surrounding the Co atoms in this film are shown in Fig. 1. The pre-edge background absorption μ_p is determined from a quadratic fit to the data roughly 200–50 eV below the edge energy and then extrapolated over the entire range of the spectrum. The K-edge absorption μ is then isolated by subtracting μ_p from μ_e. The smoothly varying atomic absorption μ_a is determined by fitting the postedge data of μ with a cubic-spline function. The absorption spectra are then step normalized by dividing by the amplitude of the absorption edge. This value is determined by making two linear fits to the data: typically -200 to -50 eV and 100 to 300 eV relative to the edge. The two fits are extrapolated to absorption threshold, and their difference is taken to be the amplitude of the absorption edge. The zero energy was assigned to be the first maximum of the first derivative of the TEY data after the onset of absorption. The normalized oscillations $[\chi(E)]$ beginning at about 20–50 eV and extending hundreds of eV (and in this case 1500 eV) are defined to be the extended x-ray absorption fine-structure region of the XAFS. The normalized oscillations are then converted to units of Å^{-1} according to the equation $k=[2m(E-E_{\text{edge}})]^{1/2}$ $(2\pi)/\text{h}$. Four successive normalized $\chi(k)$, data were averaged and the standard deviation of each data point was calculated. The normalized $\chi(k)*k^2$ data corresponding to the average in-plane (solid line) and out-of-plane (dashed line) structure surrounding the Co atoms in this film are shown in Fig. 2. The data in Fig. 2, for k values of 3.0–16.0 Å^{-1}, were Fourier transformed with a 0.5 Å^{-1} modified Hanning window function to reduce truncation ripple.[13] The resulting Fourier transformed data corresponding to the average in-plane (solid line) and out-of-plane (dashed line) structure surrounding the Co atoms in this film are shown in Fig. 3.

Inspection of Fig. 3 shows very little anisotropy between the in-plane and out-of-plane structure. Specifically, there is very little difference between the in-plane and out-of-plane amplitudes for the third and fourth peaks of the Fourier transform (3.5–5.0 Å). The primary source of these peak amplitudes is due to colinear multiple scattering of the outgoing and backscattered photoelectron wave. This focusing effect increases these peak amplitudes above that expected for a single, unfocused outgoing and backscattering event. Amplification of these peaks is typical of a close-packed crystal phase.

To illustrate this, we have used the theoretical codes developed by Rehr and co-workers[14] to simulate the Fourier transforms of Co in bcc, fcc, and hcp structures. Further, we have also utilized the ability of the FEFF6 code to incorporate (1) the orientation of the electric field vector of the incoming x-ray beam relative to the film surface and (2) the temperature of the sample during data acquisition. The results of the theoretical simulations for the average in-plane and out-of-plane local environments of Co in bcc, fcc, and hcp (relative to the basal plane) structures at 90 K are shown in Figs. 4(a)–4(c). These figures resulted from the theoretically generated $\chi(k)*k^2$ data being Fourier transformed in the same manner as the experimental $\chi(k)*k^2$ data described above. Due to the cubic symmetry of the bcc and fcc crystal structures, no polarization dependence is seen in the theoretically generated Fourier transforms. Comparison of polarization-dependent experimental data in Fig. 3 to the theoretical data in Figs. 4(a)–4(c) unequivocally shows that the Co film studied here exists in the fcc structure.

The Fourier transform data representing the in-plane and out-of-plane first coordination shell (1.52–2.84 Å) was fil-

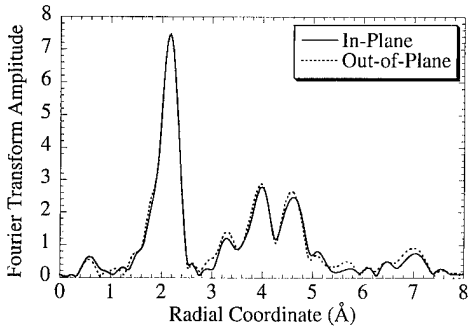


FIG. 3. Fourier transformed experimental data with the x-ray beam's electric field vector in the plane of the film (solid line) and perpendicular to the plane of the film (dashed line).

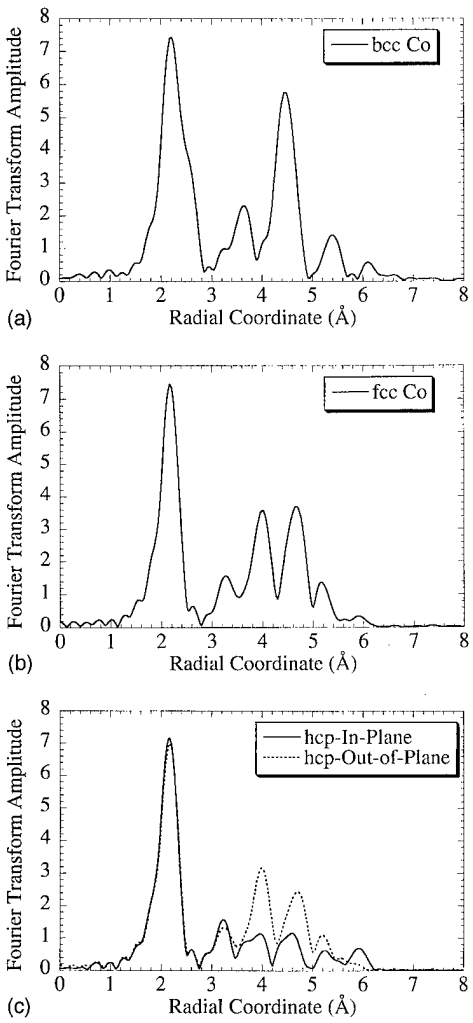


FIG. 4. (a) Fourier transform of the theoretically generated bcc crystal structure Co data; (b) Fourier transform of the theoretically generated fcc crystal structure Co data; (c) Fourier transform of the theoretically generated hcp crystal structure Co data parallel (solid line) and perpendicular (dashed line) to the basal plane.

tered and back-transformed with a modified Hanning window function of 0.25 Å to reduce truncation ripple. These filtered data were compared to each other using the EXAFS ratio analysis technique[13] to determine any relative differences in the coordination number, radial distance, or EXAFS Debye–Waller factors. Results from this analysis indicate less than a 0.0075 Å difference between in plane and out-of-plane bondlengths, less than a difference of 1.3 atoms in coordination number and less than a 0.001 $Å^2$ difference in EXAFS Debye–Waller factors. These differences represent a doubling of the residual during the fitting procedure and are typically much larger than the magnitude of experimental error. These values are much smaller than typical differences from EXAFS experiments because the EXAFS spectra for two different polarizations are being compared *to each other* and not to another standard. This way, the errors that would be introduced by chemical shifts and different background removals are eliminated. Therefore, within the limits of this EXAFS experiment, no tetragonal distortion of the fcc Co crystal lattice is seen. This result indicates that the majority of the 1000-Å-thick film is relaxed and not anisotropically distorted by pseudomorphic growth.

III. SUMMARY

We have employed the polarization-dependent properties of EXAFS to probe the average local environment in a 1000-Å-thick Co film deposited on diamond. By qualitatively noting the polarization dependence of the third and fourth Fourier transform peak amplitudes, this EXAFS study shows that the Co film studied here is unambiguously in the fcc phase. The Fourier transform peaks corresponding to the first nearest neighbors, parallel and perpendicular to the film, have been compared. Results indicate no tetragonal distortion of the fcc Co crystal lattice, therefore indicating that the majority of the film is relaxed and not pseudomorphic. This technique should prove valuable in future studies when trying to determine the crystallinity of thin films.

ACKNOWLEDGMENTS

This research was conducted while K. M. K. and J. A. W. held a National Research Council/Naval Research Lab Fellowship. The X23-A2 beamline at Brookhaven National Laboratory is supported in part by the National Institute of Standards and Technology. The NSLS is supported by the Department of Energy (Division of Materials Sciences and Division of Chemical Sciences of the Office of Basic Energy Sciences) under Contract No. DE-AC02-76 CH00016.

[1]G. A. Prinz, J. Magn. Magn. Mater. **100**, 469 (1991), and references therein.
[2]Physics Today, April (1995).
[3]Y. U. Idzerda, W. T. Elam, B. T. Jonker, and G. A. Prinz, Phys. Rev. Lett. **62**, 2480 (1989).
[4]G. A. Prinz, Phys. Rev. Lett. **54**, 1051 (1985).
[5]P. C. Riedi, T. Dmuelow, M. Rubinstein, G. A. Prinz, and S. B. Qadri, Phys. Rev. B **36**, 4595 (1987).
[6]J. J. de Miguel, A. Cebollada, J. M. Gallego, R. Miranda, C. M. Schneider, P. Schuster, and J. Dirschner, J. Magn. Magn. Mater. **93**, 1 (1991), and references therein.
[7]J. A. Wolf, J. J. Krebs, and G. A. Prinz, Appl. Phys. Lett. **65**, 1057 (1994).
[8]D. Bartlet, F. Tsui, D. Glick, L. Lauhon, T. Mandrekar, C. Uher, and R. Clarke, Phys. Rev. B **49**, 1521 (1994), and references therein.
[9]A. Erbil, G. S. Cargil III, R. Frahm, and R. F. Boehme, Phys. Rev. B **37**, 2450 (1988).
[10]K. M. Kemner, J. Kropf, and B. A. Bunker, Rev. Sci. Instrum. **65**, 3667 (1994).
[11]E. A. Stern, in *X-Ray Absorption: Basic Principles of EXAFS, SEXAFS, and XANES*, edited by D. C. Koningsberger and R. Prins (Wiley, New York, 1988), Chap. 1.
[12]For more information see K. M. Kemner, Z. Wang, R. A. Mayanovic, and B. A. Bunker, Nucl. Instrum. Method B **71**, 345 (1992).
[13]D. E. Sayers and B. A. Bunker, in *X-Ray Absorption: Basic Principles of EXAFS, SEXAFS, and XANES*, edited by D. C. Koningsberger and R. Prins (Wiley, New York, 1988), Chap. 6.
[14]J. J. Rehr, S. I. Zabinsky, and R. C. Albers, Phys. Rev. Lett. **69**, 3397 (1992).

CUMULATIVE AUTHOR INDEX

All authors published to date in Volume 14 are listed alphabetically with the issue (in parentheses) and page numbers following the dash. An (E) after the page number indicates Erratum.

Abe, Takayuki–(4) 2474
Abe, Yoshiaki–(3) 2096
Abelson, J. R.–(4) 2674
Abernathy, C. R.–(1) 118; (4) 2567, 2582
Abraham, D. L.–(2) 1206
Abraham, I. C.–(2) 710
Abstreiter, G.–(2) 698
Achouche, M.–(4) 2555
Ackermann, L.–(2) 832
Actis, R.–(3) 2236
Adachi, Hiroshi–(3) 2093, 2114
Adams, D. P.–(3) 1642; (4) 2438
Adderton, Dennis–(1) 426
Agius, B.–(4) 2543
Agraït, N.–(4) 2772
Agraït, N.–(2) 1289
Ahmed, H.–(2) 662
Ahn, C. C.–(4) 3030
Ahn, Ho Young–(3) 1966
Aikawa, I.–(1) 324
Aiura, Yoshihiro–(2) 1060
Ajmera, P. K.–(2) 727
Akamine, Shinya–(2) 812
Akwani, I.–(2) 929
Aldao, C. M.–(4) 2519
Alexandre, F.–(1) 147
Alexiev, D.–(1) 152
Allegrini, M.–(2) 868, 1509
Allen, Lynn R.–(2) 724
Allen, P. N.–(1) 132
Allen, Ronald L.–(1) 498
Aller, I.–(2) 820
Almeida, J.–(2) 1527
Alperin, J.–(3) 2343
Alstrin, April L.–(4) 2742
Alterovitz, S. A.–(4) 2674
Alvis, Roger–(1) 231, 452
Ambrée, P.–(3) 1712
Amirkhanov, R. N.–(3) 2135
Amman, M.–(1) 54
An, M. H.–(4) 2667
Anand, S.–(4) 2794
Ancona, M. G.–(3) 1918
Andersohn, L.–(2) 1032
Anderson, S. G. H.–(4) 2636
Anselm, K. A.–(3) 2256
Aoki, Susumu–(3) 2110
Aono, M.–(4) 2399
Aoyama, Hisayuki–(2) 849
Arabczyk, W.–(2) 982
Arai, Toyoko–(2) 1378
Arlinghaus, H. F.–(1) 294
Asakawa, K.–(3) 1764
Ascolani, H.–(3) 1655
Ascoli, Cesare–(2) 845
Asonen, H.–(3) 1736
Aspnes, D. E.–(4) 2812, 3040, 3070
Atalar, A.–(4) 2456
Auciello, O.–(2) 602
Aviles, A. F.–(2) 929
Aydil, Eray S.–(2) 738

Bachmann, K. J.–(4) 3019, 3040
Baillargeon, J. N.–(3) 2244
Baiocchi, F. A.–(1) 202
Baker, A. A.–(2) 1438
Bakhtizin, R. Z.–(2) 1000; (3) 2135
Baksheyev, D. G.–(4) 2399
Baldereschi, A.–(4) 2936, 3000
Baldeschwieler, J. D.–(2) 852
Baldeschwieler, John D.–(2) 1302
Balmashnov, A. A.–(1) 471
Bammerlin, M.–(2) 1280, 1285
Banaszak Holl, Mark M.–(4) 2824
Bandić, Z. Z.–(4) 2948
Bando, Hiroshi–(2) 1060
Banerjee, S.–(1) 255
Banerjee, S. K.–(2) 1344
Banoo, K.–(4) 2725
Bänziger, U.–(3) 2000
Baptist, R.–(3) 2119
Barbato, G.–(2) 1531
Bardi, J.–(4) 3000
Baribeau, J.-M.–(3) 1675; (4) 2895
Barney, Bryan–(1) 483
Baró, A. M.–(2) 1145
Barr, A. L.–(2) 1217
Barrett, M.–(1) 447
Baselt, David R.–(2) 789
Baski, A. A.–(2) 992
Basmaji, P.–(3) 2212
Bass, J. M.–(4) 3075
Bassani, F.–(2) 868
Bauer, J.–(4) 2485
Bauer, M. L.–(1) 15
Beckman, E.–(2) 877
Behler, S.–(2) 1176
Bell, A. E.–(3) 2041
Bell, F. H.–(1) 96; (3) 1796; (4) 2493
Benard, William–(4) 2407
Bender, H.–(1) 358
Bender, John E., IV–(4) 2824
Bending, S. J.–(2) 1202
Bennett, B. R.–(4) 3193
Bennett, B.R.–(3) 2195
Benninghoven, A.–(4) 2712
Bensaoula, A.–(3) 2357
Benson, T. E.–(1) 483
Berg, E. W.–(3) 1807
Berndt, R.–(2) 1149
Berry, L. A.–(3) 1853
Berthod, C.–(4) 3000
Bertrams, Th.–(2) 925, 1141
Bessolov, Vasily N.–(4) 2761
Beton, P. H.–(2) 943, 1024; (3) 1596
Bhattacharya, P.–(3) 2312
Bhattacharya, P. K.–(3) 2378
Bhuva, B.–(3) 2068
Bi, W. G.–(3) 2282, 2327; (4) 2918
Biberger, M.–(4) 2627
Bicknell-Tassius, Robert–(3) 2170
Bieth, C.–(3) 2119
Bimberg, D.–(3) 1712; (4) 2961
Bindley, U.–(3) 2374
Binggeli, N.–(4) 3000
Bischoff, L.–(3) 1621
Bitner, Theodore W.–(4) 2850
Bittner, T. C.–(1) 202
Björkqvist, M.–(2) 957
Blanc, N.–(2) 901, 1513
Block, T. R.–(3) 2221
Block, Thomas R.–(3) 2216
Bluhm, H.–(2) 1180
Blume, Richard–(2) 716
Boatner, L.–(2) 1096
Boden, Ted–(3) 1846
Böhm, Christoph–(2) 842
Böhni, H.–(2) 1162
Bonanni, A.–(4) 2961
Boney, C.–(3) 2259
Bonnell, Dawn A.–(3) 1607
Bonnet, Noél–(2) 1552
Booh, K. H.–(2) 1221
Boonstra, Tom–(1) 543
Borngräber, J.–(1) 373
Boswell, E. C.–(3) 1895, 1910
Botchkarev, A. E.–(4) 2674
Botreau, M.–(2) 1492
Böttcher, E. H.–(3) 1712
Bouju, X.–(2) 816
Bousetta, A.–(3) 2357
Boussaad, S.–(2) 1476
Bowers, K. A.–(3) 2349
Boyd, R. D.–(3) 2005
Bradshaw, A. M.–(2) 1117
Braeckelmann, Gregory–(3) 1828
Bratina, G.–(2) 623; (4) 2967
Brault, Pascal–(4) 2767
Brett, M. J.–(2) 679; (4) 2595
Breun, R. A.–(2) 710
Brewer, P. D.–(3) 2335
Briggs, G. A. D.–(2) 966
Brigham, R.–(1) 353
Brillson, L. J.–(4) 2961
Britten, J. A.–(3) 2005
Brodie, I.–(3) 1986
Brouquet, Pierre–(4) 2767
Brown, April S.–(3) 2170
Brown, E. R.–(3) 1745
Brown, S. R.–(2) 1167
Brożek, Tomasz–(1) 571, 577
Brugger, J.–(2) 901
Bruley, J.–(3) 2024
Brunner, J.–(2) 698
Bruno, G.–(2) 691
Brunt, T. A. –(2) 1383
Brutscher, J.–(4) 2701
Bruynseraede, Y.–(2) 1121
Buisset, J.–(2) 1117
Bullemer, B.–(4) 2684
Burger, A.–(2) 1083, 1090, 1096
Burke, Aaron–(3) 1828
Burkett, S. L.–(4) 3131
Burnham, N. A.–(2) 794, 1308
Butcher, K. S. A.–(1) 152

Caballero, J. A.–(3) 1758
Calandra, C.–(4) 2901
Cale, Timothy S.–(4) 2615
Calhoun, L.–(1) 159
Campbell, J. C.–(3) 2256
Campbell, P. A.–(2) 1167, 1373
Campbell, S. A.–(3) 1706
Campbell, Scott–(3) 2147
Campen, D. G. Van–(4) 3199
Camphausen, S. M.–(3) 2024
Campisano, S. U.–(1) 414
Campisano, Salvatore U.–(4) 2693
Cantwell, Gene–(3) 2259
Cao, J. L.–(1) 42
Cao, R.–(4) 2882, 2914
Capano, M. A.–(3) 2278
Capezzuto, P.–(2) 691
Car, Roberto–(4) 2809
Cardona, M.–(1) 187(E)
Carducci, James–(2) 716
Carlsson, S-B.–(4) 2794
Carneiro, K.–(2) 1531
Caron, M.–(4) 2627
Carpick, R. W.–(2) 1289; (4) 2772
Carpick, Robert W.–(2) 1341
Carter-Coman, Carrie–(3) 2170
Cartier, E.–(4) 2855
Casasanta, V.–(3) 2318
Casey, Sean M.–(4) 2742
Castell, M. R.–(2) 966
Castleberry, H.–(4) 3019
Celio, M. R.–(2) 1492
Cescato, Lucila–(3) 1784
Cevc, G.–(2) 1498
Chabal, Y. J.–(4) 3095
Chaban, E. E.–(4) 3095
Chaichimansour, M.–(3) 2263
Chalmers, Scott–(3) 2147
Champaign, H.–(3) 1914
Chan, David–(1) 571
Chan, Yi-Jen–(4) 2550
Chang, J. C. P.–(3) 2225, 2271
Chang, P. C.–(3) 2252
Chang, Ying-Lan–(4) 2914
Chao, K.-J.–(2) 909, 914, 948
Chapek, D. L.–(1) 336
Charych, D. H.–(2) 1481
Chatterjee, T.–(3) 2267
Chau, T. T.–(1) 527
Chen, C. L.–(3) 1745
Chen, Dong–(2) 864, 979
Chen, Dongmin–(2) 974
Chen, G. Y.–(2) 1313
Chen, H.–(4) 2690
Chen, H. C.–(4) 3126
Chen, H. H.–(4) 2531
Chen, K. M.–(3) 2199
Chen, K. T.–(2) 1083, 1090
Chen, K.-T.–(2) 1096
Chen, Mao-Chieh–(1) 167
Chen, P.–(3) 2203
Chen, W. L.–(3) 2225; (4) 2739
Chen, X.–(2) 1136; (3) 1582
Chen, Y. J.–(4) 2753
Cheng, C. C.–(1) 85
Cheng, K. Y.–(3) 2244
Cheng, T. S.–(3) 2346
Cheng, X.–(1) 489
Chernozatonskii, L. A.–(3) 2080
Cheung, Nathan W.–(1) 236, 552
Cheung, R.–(4) 2636
Chi, L. F.–(2) 1503
Childress, J.–(4) 3176
Chin, T. P.–(3) 2225, 2325; (4) 2739
Chizhov, I.–(2) 995
Cho, A. Y.–(3) 2244
Cho, Jin Woo–(3) 1924
Choi, W. B.–(3) 2050
Choi, Y. J.–(2) 1188
Chou, C.–(1) 474
Chou, L. J.–(3) 2327
Chou, L.-J.–(3) 2233
Chow, D. H.–(3) 2163, 2335
Christensen, K. N.–(2) 744
Christman, S. B.–(4) 3095
Chu, S. N. G.–(3) 2244, 2297
Chua, S. J.–(3) 1719
Chung, M.–(3) 2020
Chung, Seoung Woo–(1) 478

Cinti, R. C.–(2) 1005, 1010
Ciotti, M. T.–(2) 1395
Clarke, A. P.–(4) 2636
Clarysse, T.–(1) 358, 380, 390, 404
Clawson, A. R.–(4) 3047, 3089
Cleaver, J. R. A.–(2) 662
Clei, A.–(4) 2555
Coldren, Larry–(3) 2147
Cole, M. W.–(3) 1758
Cole, R. J.–(4) 2424
Coleridge, P. T.–(3) 2290
Collins, R.–(1) 132
Collins, W. E.–(2) 1083, 1090, 1096
Colton, Richard J.–(2) 789
Coluzza, C.–(2) 1527
Constantinidis, G.–(2) 657
Cook, J. W., Jr.–(3) 2259, 2349
Cooke, G. A.–(1) 132, 283, 348
Cooper, James–(2) 716
Couillard, J. Greg–(3) 2381
Cowles, J.–(3) 2221
Craig, M.–(1) 255
Craighead, H. G.–(1) 69; (2) 617
Craighead, Harold G.–(3) 2381
Cramer, H.-G.–(4) 2712
Cramer, L.–(2) 1117
Crean, Gerard–(2) 716
Cricenti, A.–(2) 1395, 1527
Crockett, R. V.–(3) 1752
Croke, E. T.–(4) 3030
Cuberes, M. T.–(3) 1655
Cuomo, J. J.–(3) 2024, 2050
Cuppini, D.–(2) 1531
Currie, J. F.–(4) 2627
Curtin Carter, Margaret M.–(3) 1867
Cutler, P. H.–(3) 2020, 2037

Dąbrowski, J.–(2) 982
Dagenais, M.–(4) 2753
Daghlian, Charles–(1) 440, 457
Dai, Zhenxi–(3) 1591
Dalal, Pranav–(4) 2407
Daleiden, J.–(3) 1780
D'Ambrosio, P.–(3) 2020
Daniels-Race, T.–(4) 2725
Dankowski, S. U.–(3) 2275
Dao, Thuy–(1) 577
D'Asaro, L. A.–(3) 2252
Das Sarma, S.–(4) 2817
Davidson, J. L.–(3) 2068
Davidson, Ross–(3) 1867
Davies, J.–(3) 1582
Davies, M. C.–(2) 1557; (3) 1582
Davies, Martyn C.–(2) 1432
Davis, P. R.–(3) 2090
Davis, Paul R.–(3) 1952
Dawson, Douglas E.–(3) 2170
Dayton, J. A., Jr.–(3) 2046
De Bisschop, P.–(1) 311
de Boeij, P. L.–(4) 3080
de Boer, M.–(2) 597
Debusschere, I.–(1) 265
Deferm, L.–(1) 265
Défourneau, D.–(2) 1070
Deguchi, Osamu–(2) 687
Delain, Etienne–(2) 1552
de Lozanne, A. L.–(2) 1210, 1217
de Lozanne, Alex–(3) 1591
de Lyon, T. J.–(3) 2362
deMarco, John–(1) 63
Deng, C.–(3) 1650
Denlinger, J. D.–(4) 3171
Dennis, M.–(1) 447
De Paoli, Marco-A.–(3) 1784
de Potter, M.–(3) 1773
Deppert, K.–(4) 2794
Dereux, A.–(2) 816
de Rooij, N. F.–(2) 901
DeRose, J. A.–(2) 1476
Derro, D. J.–(2) 1217
Descouts, Pierre–(2) 1422
Deshmukh, Shashank C.–(2) 738
Deshpande, M. R.–(1) 54
Deshpandey, Chandra–(1) 521
Desogus, S.–(2) 897
De Stasio, G.–(2) 1395
Dew, S. K.–(2) 679; (4) 2595
De Wolf, P.–(1) 358, 380; (2) 1513
Dhar, N. K.–(3) 2366
Diebold, A. C.–(1) 196
Diehl, D. J.–(2) 744
Dietler, Giovanni–(2) 1250
Dietz, N.–(4) 3019, 3040
Dietzsch, Ch.–(3) 1621
Dinan, J. H.–(3) 2366
DiNardo, N. J.–(3) 1607
Disch, S. B.–(2) 710
Di Ventra, M.–(4) 2936
Dixson, R.–(2) 1540
Dobbs, Drena–(2) 1405
Dobrowolska, M.–(3) 2374
Döhler, G. H.–(3) 2175, 2275
Dongmo, Samuel–(2) 1552
Donnelly, V. M.–(1) 85
Donohue, J. F.–(1) 466
Doughty, C.–(3) 1853
Douglas, R. A.–(4) 2432
Dowben, P. A.–(4) 2957
Downey, Daniel F.–(1) 329
Downey, S. W.–(1) 301
Dowsett, M. G.–(1) 132, 283, 348
Drabik, Timothy J.–(4) 2445
Dreier, M.–(2) 1418
Drescher, G.–(4) 2485
Driad, R.–(1) 147
Driesel, W.–(3) 1621
Drigo, A. V.–(3) 1739
Droz, Eric–(2) 1422
Du, Q.–(3) 2343
Duane, Michael–(1) 218
Duc, Tran Minh–(2) 1492
Dumania, P.–(2) 856
Dumortier, Chantal–(2) 1427
Dunn, A. W.–(2) 943; (3) 1596
Durbha, A.–(4) 2582
Dürig, U.–(2) 901, 1259
Dutoit, Michel–(3) 1660
Dyer, D. J.–(1) 38

Eby, R. K.–(2) 852
Eckstein, J. N.–(2) 1210
Edmond, J. A.–(3) 2349
Edwards, Hal–(2) 1217
Eftekhari, G.–(1) 181
Egami, Takeshi–(3) 1607
Egan, P. O.–(4) 2510
Egan, R. J.–(1) 152
Eggs, C.–(1) 102
Eguchi, K.–(2) 1353
Ehrlich, S. N.–(4) 3035
Eisele, I.–(4) 2684
Eisele, K. M.–(3) 1780
Eisenmenger, W.–(2) 1079
Elam, W. T.–(4) 3207
Eldridge, T.–(4) 2780
Elliot, D.–(3) 1675
El-Masry, N. A.–(3) 2349
Emerson, A. B.–(1) 301
En, William–(1) 552
Enachescu, M.–(2) 1229, 1336
Endo, Michiaki–(3) 2114
Eng, L.–(2) 1162
Eng, L. M.–(2) 1191, 1386
Engelborghs, Yves–(2) 1427
English, John–(3) 2147
Engstrom, Herbert–(1) 511
Eppell, Steven J.–(4) 2407
Erickson, Andrew–(1) 426
Erickson, Jon W.–(1) 353
Eriksson, N.–(1) 184
Erler, B.–(2) 970
Ermert, Helmut–(3) 2100
Ersland, P.–(1) 437
Eschrich, R.–(2) 1498
Escorcia-Aparicio, Ernesto J.–(4) 3164
Esser, N.–(1) 187(E)
Ester, G. R.–(2) 1373
Esteve, Joan–(4) 2609
Evtukh, A. A.–(3) 2130
Eyink, Kurt G.–(3) 2278

Falk, G.–(2) 1327
Fan, Y.–(2) 623
Fan, Z.–(4) 2674
Fang, X. M.–(3) 2267
Faricelli, J.–(1) 224
Farrow, R. F. C.–(4) 3148
Feinerman, A. D.–(4) 2524
Felch, S. B.–(1) 336
Felch, Susan B.–(1) 397
Feldman, L. C.–(4) 3095
Felter, T.–(3) 1958
Feng, D.–(4) 2690
Fernandez, A.–(3) 2005
Ferry, D. K.–(4) 2780, 2840
Ferry, David K.–(1) 75
Fetters, L. J.–(2) 1276
Fijol, J. J.–(1) 159
Fimland, B. O.–(1) 187(E)
Finch, Robert–(1) 440
Fine, J.–(4) 2417
Firon, M.–(4) 2543
Fischer, Stefan–(2) 961
Fischer, U. C.–(2) 877
Fitting, H.-J.–(3) 2087
Fleming, J. G.–(3) 1958
Flipse, C. F. J.–(2) 585
Folch, Albert–(4) 2609
Forgotson, N.–(2) 732
Fournier, Françoise–(3) 2170
Fousek, J.–(2) 1191
Foxon, C. T.–(3) 2346
Franciosi, A.–(2) 623; (4) 2961, 2967, 2980, 2994
Frank, E. R.–(2) 1136
Franz, Gerhard–(1) 126
Fratello, V. J.–(3) 2297
Frazier, D. O.–(2) 1083
Frediani, Carlo–(2) 845
Freeouf, J. L.–(4) 3089
Frei, Th.–(1) 278
Freudenberg, J. S.–(1) 483
Frey, L.–(3) 1973
Friedrich, M.–(2) 1191
Fritzsche, W.–(2) 1399
Fritzsche, Wolfgang–(2) 1405
Frost, Michael R.–(1) 329
Fu, Elain S.–(4) 2799
Fu, Shelton–(3) 1607
Fuchs, H.–(2) 877, 1503
Fujihira, Masamichi–(2) 1272, 1378
Fujii, M.–(4) 2731
Fujikura, Hajime–(4) 2888
Fujita, K.–(4) 2731
Fujita, Shizuo–(3) 2349
Fukatsu, S.–(3) 2387
Fukuzawa, Kenji–(2) 812
Funatsu, Jun.–(1) 179
Furdyna, J. K.–(3) 2374
Furuta, Tomofumi–(4) 2500

Gagnon, G.–(4) 2627
Gailhanou, Marc–(3) 1660
Gaillard, Frédéric–(4) 2767
Gaisch, R.–(2) 1149
Gallagher, M. C.–(2) 1126
Gallo, P.-J.–(2) 794, 1308
Gamble, R. C.–(2) 852
Ganière, Jean-Daniel–(3) 1660
Gao, H. J.–(2) 1349
Gao, Hua–(1) 521; (2) 716
García, R.–(4) 2403
Garnaes, J.–(2) 1531
Gauthier, S.–(2) 1131
Gaworzewski, P.–(1) 341, 373
Geenen, L.–(1) 311
Geerts, Wim–(4) 3176
Geib, E.–(3) 2087
Geis, M. W.–(3) 2060
Geller, J. D.–(3) 2318
Gemma, Nobuhiro–(2) 1234
Generosi, R.–(2) 1395, 1527
George, M. A.–(2) 1090, 1096
Gersten, J. I.–(1) 30
Geva, M.–(1) 202; (3) 2301
Ghots, S. S.–(3) 2135
Gibbons, R.–(1) 348
Gibson, M. A.–(1) 489
Gill, David M.–(3) 2166, 2192
Gill, W. N.–(2) 772
Gilmer, D. C.–(3) 1706
Gimzewski, J. K.–(2) 827
Girard, C.–(2) 816
Givargizov, E. I.–(3) 2030, 2034
Gladfelter, W.–(2) 623
Gladfelter, W. L.–(3) 1706
Glaser, E.R.–(3) 2195
Gledhill, H. C.–(2) 1167
Gluck, Robert–(1) 440
Gobrecht, J.–(2) 1255
Goddard, D. T.–(2) 966
Göhl, A.–(3) 2056
Göken, M.–(2) 1157
Goldman, R. S.–(4) 3035
Golubov, A. A.–(2) 1224
Gölz, A.–(4) 3107
Gomez, J.–(1) 311
Gómez-Herrero, J.–(2) 1145
Gómez-Rodríguez, J. M.–(2) 1005
Gómez-Rodríguez, J.-M.–(2) 1010
Gompf, B.–(2) 1079
Gondermann, J.–(2) 698
Gondran, Carolyn–(3) 1846
González-Borrero, P. P.–(3) 2212
Goodhue, W. D.–(3) 2318
Goodman, K. W.–(4) 3171
Goorsky, M. S.–(4) 2922
Göpel, Wolfgang–(2) 961
Goplen, B.–(3) 1942, 1947, 1990
Gorbatkin, S. M.–(3) 1853
Gori, G.–(2) 1531
Gossard, Arthur–(3) 2147
Göthelid, M.–(2) 957
Gotoh, Yasuhito–(3) 1970
Gotszalk, T.–(2) 856
Graham, M. J.–(4) 2882
Gramstorff, B. F.–(3) 1745
Gray, A.–(3) 2366
Gray, H. F.–(3) 2072
Gray, Struan M.–(2) 1015, 1043
Green, M. L.–(1) 301
Gregory, D. A. C.–(4) 2424
Gremaud, G.–(2) 794, 1308
Greve, D. W.–(1) 489

Griffin, Peter–(1) 231
Grizzle, J. W.–(1) 483; (4) 2531
Groeneveld, R. H. M.–(2) 861, 1206
Grundmann, M.–(4) 2961
Gruverman, A.–(2) 602
Guckenberger, R.–(2) 1498
Gueorguiev, Y. M.–(4) 2462
Guggisberg, M.–(2) 1255, 1285
Guido, L. J.–(1) 54
Guinn, K. V.–(1) 85; (4) 2510
Gujrathi, S. C.–(4) 2627
Gulden, K. H.–(3) 2175
Gulyaev, Y. V.–(3) 1982
Gulyaev, Yu. V.–(3) 2080
Günter, P.–(2) 1191, 1386
Günther, E.–(2) 1327
Güntherodt, H. -J.–(2) 1255
Güntherodt, H. J.–(2) 1418
Güntherodt, H.-J.–(2) 970, 1176, 1264, 1280, 1285
Günzel, R.–(4) 2701
Guo, X.–(4) 3152
Güthner, Peter–(4) 2428
Gyalog, T.–(2) 1280

Haas, T. W.–(3) 2278
Häberle, W.–(2) 1449
Habermann, T.–(3) 2056
Hadam, B.–(2) 698
Haddad, G. I.–(3) 2225; (4) 2739
Hadjiiski, L.–(2) 1563
Haefke, H.–(2) 1280
Hagberg, M.–(1) 184
Hagmann, Mark J.–(2) 838
Halas, N. J.–(2) 593
Hale, G. D.–(2) 593
Halfpenny, P. J.–(2) 1373
Hall, Anselm P.–(4) 2850
Hallett, P.–(2) 1444
Hamaguchi, S.–(3) 1819; (4) 2603
Hamann, D. R.–(4) 3095
Hamers, R. J.–(2) 1136, 1360
Hamers, Robert J.–(2) 1038
Hammiche, A.–(2) 1486
Han, P.–(4) 2690
Han, P. G.–(2) 824
Hane, Masami–(1) 560
Hanish, C. K.–(1) 483; (4) 2531
Hanish, P. D.–(1) 483
Hänni, W.–(3) 2056
Hansma, Helen G.–(2) 1390
Hanson, C. M.–(4) 3047, 3089
Hanzawa, Satoru–(3) 2096
Hara, Shigeta–(2) 1369
Harmand, J. C.–(1) 147; (4) 2555
Harmon, E. S.–(3) 2271
Harp, G. R.–(4) 3148
Harrington, William L.–(1) 329
Harriott, Lloyd–(1) 63
Harris, J. S., Jr.–(3) 2208
Harris, James S., Jr.–(2) 642; (3) 1725
Harris, V. G.–(4) 3207
Harris, W. C., Jr.–(1) 305
Harrison, Walter A.–(2) 1243
Harsch, William C.–(3) 2259
Hartman, Robert L.–(3) 1952
Hartmann, E.–(2) 1229, 1336
Hartmann, U.–(2) 1224
Hartnagel, H. L.–(2) 820
Haruyama, Yuichi–(2) 1060
Hasegawa, Hideki–(4) 2872, 2888
Hasenack, Claus Martin–(1) 538
Hashimoto, Keiichi–(4) 2645
Hashizume, T.–(2) 1000, 1522
Hashizume, Tamotsu–(4) 2872, 2888
Hauenstein, R. J.–(4) 2948
Hawryluk, A. M.–(3) 2005
He, Y. W.–(3) 2349
Hegde, R. I.–(3) 1697
Hegner, M.–(2) 1418
Hegner, Martin–(2) 1466
Heike, S.–(2) 1522
Heim, M.–(2) 1498
Heimbrook, L. A.–(1) 202
Heinemann, B.–(1) 287
Heinrich, F.–(3) 2000
Hellemans, L.–(1) 380; (2) 1513
Hellemans, Louis–(2) 1427
Henderson, D. O.–(2) 1083, 1090
Henderson, Eric–(2) 1405, 1413
Henderson, R. H.–(3) 2309
Hendriks, H. T.–(3) 2233
Henini, M.–(2) 943, 1024, 1202
Henniger, U.–(1) 373
Henning, Albert K.–(1) 440, 457
Herman, G. S.–(2) 1126
Herzel, F.–(1) 287
Hess, P.–(3) 1600
Hesse, E.–(3) 1621
Heuberger, Manfred–(2) 1250
Heun, S.–(4) 2980
Heyd, A. R.–(4) 2674
Heyvaert, I.–(2) 1121
Hieda, Hiroyuki–(2) 1234
Hietschold, M.–(2) 1461
Higashi, Yasuhiro–(2) 763
Higgins, S. R.–(2) 1360
Highberg, Ivar–(1) 533
Hill, C.–(1) 348
Hill, Rod–(1) 547
Hillebrand, A.–(2) 1498
Hinch, B. J.–(2) 1032
Hinds, B. J.–(4) 2812
Hingst, Th.–(3) 2087
Hinode, Kenji–(2) 687
Hinrich, S.–(2) 982
Hirano, Takayuki–(3) 1885
Hiraoka, Akiko–(4) 2707
Hirashita, N.–(1) 324
Ho, I. H.–(4) 3013
Ho, P. S.–(4) 2636
Hobart, K. D.–(3) 2229
Hobler, G.–(1) 272
Hobson, W. S.–(1) 118; (3) 1758; (4) 2567
Hochwitz, Todd–(1) 440, 457
Hodul, D.–(2) 679
Hoke, W. E.–(3) 2233
Holland, C. E.–(3) 1986
Holland, M. C.–(3) 2305
Holloway, P. H.–(1) 159; (4) 2582
Homma, T.–(2) 1184
Homma, Yoshikazu–(2) 763
Hommel, D.–(4) 2757
Hong, J.–(4) 2567
Hong, M.–(3) 2297
Hooper, S. E.–(3) 2346
Hoppe, Matthew M.–(3) 2216
Höppner, K.–(1) 373
Höppner, W.–(1) 373
Hopwood, J.–(2) 732
Hörber, J. K. H.–(2) 1449
Horn, J.–(2) 820
Hosokawa, Akihiro–(1) 494
Hosono, A.–(3) 1973
Hotta, H.–(3) 1764
Hou, S. Y.–(3) 2297
Hou, Xiao-yuan–(4) 2973
Houghton, D. C.–(3) 1675
Houriet, Raymond–(3) 1660
Hourston, D. J.–(2) 1486
Howald, L.–(2) 1255, 1264, 1280, 1285
Howell, M.–(3) 2068
Hoyle, P. C.–(2) 662
Hoyler, Charlotte–(1) 126
Hrbek, J.–(4) 3199
Hren, J. J.–(3) 2024, 2050
Hsieh, K. C.–(3) 2233, 2327
Hsu, C. C.–(2) 1105
Hsu, Y.–(3) 2286
Hu, B. H.–(3) 2374
Hu, C.–(3) 2256
Hu, C.-W.–(2) 938
Hu, Evelyn–(4) 2914
Hu, Jun–(2) 1341
Hu, L. Q.–(4) 2690
Hu, Y. Z.–(2) 744; (3) 1687; (4) 2543
Huang, C.-L.–(1) 224
Huang, Jin-Lin–(2) 1064
Huang, L. J.–(4) 2895
Huang, M.–(3) 1895, 1910
Huang, X. D.–(4) 2690
Huang, Y.–(1) 433
Huang, Z.-H.–(3) 2020, 2037
Hubler, U.–(2) 1176
Hughes, G.–(2) 1531
Hughes, P. J.–(2) 632
Hughes, W. C.–(3) 2259, 2349
Hugon, M. C.–(4) 2543
Hundhausen, M.–(2) 1268; (4) 2845
Hunter, A. T.–(4) 3030
Huq, S. E.–(3) 1910
Hussey, R. J.–(4) 2882
Huth, C.–(3) 2000
Hwang, Seong-Don–(4) 2957
Hwang, W.-Y.–(3) 2301
Hybertsen, Mark S.–(4) 2809
Hyuga, Fumiaki–(2) 652

Iacona, Fabio–(4) 2693
Ibbotson, D. E.–(4) 2510
Idzerda, Y. U.–(4) 3207
Ikai, Atsushi–(2) 1410
Ikemiya, Norihito–(2) 1369
Il'in, E. M.–(3) 1982
Illig, M.–(4) 2485
Ingrey, S.–(4) 2895
Ino, M.–(2) 751
Inoue, Kazunori–(3) 1970
Inoue, Takahito–(3) 2105
Irene, E. A.–(2) 744; (3) 1687, 1697; (4) 2543
Ishida, E.–(1) 255, 336
Ishida, Emi–(1) 397
Ishihara, O.–(1) 112
Ishii, Masami–(2) 1331
Ishikawa, H.–(3) 1812
Ishikawa, Junzo–(3) 1970
Ishikawa, Yuuichi–(3) 2110
Ishitani, A.–(2) 1109
Ismail, K.–(4) 2776
Itasaka, M.–(2) 1365
Itoga, Toshihiko–(4) 2707
Itoh, Junji–(3) 1885, 1902, 2105
Itoh, Shigeo–(3) 1977
Itoh, T.–(3) 1577
Ivanov, D. I.–(4) 2462
Ivey, D. G.–(1) 1
Iwadate, Kazumi–(4) 2500
Iwakuro, Hiroaki–(2) 707
Iwasaki, T.–(2) 1153
Iwata, Futoshi–(2) 849
Iwatsuki, Masashi–(3) 2110

Jackson, Andrew–(3) 2147
Jackson, D. E.–(2) 1557
Jackson, M. D.–(4) 2424
Jacobi, S.–(2) 1503
Jacobs, B.–(4) 2537
Jahnel, F.–(1) 272
Jamba, D. M.–(3) 2362
Jansen, R.–(2) 1173
Jawarani, D.–(4) 2636
Jenkins, L. C.–(3) 2346
Jensen, C. P.–(2) 1531
Jensen, K. L.–(3) 1873, 1942, 1947, 1990, 1994, 2072
Jentzsch, A.–(3) 2000
Jeon, Il Cheol–(2) 1318
Jeong, I. C.–(2) 1188
Jernigan, G. G.–(3) 2229
Jess, P.–(2) 1176
Jessing, J. R.–(3) 1899
Jesson, D. E.–(3) 2199
Jia, A. W.–(3) 2371
Jiang, Q.–(4) 3180
Jiang, X. R.–(3) 1635
Jiménez, C.–(4) 2623
Jin, C. G.–(4) 3052
Jin, Cai-xia–(4) 2973
Jin, S.–(3) 2011
Jing, Z.–(4) 2832
Jobst, B.–(4) 2757
Johansson, Lars S. O.–(2) 1015, 1043
Johansson, Mikael K.-J.–(2) 1015, 1043
Johnson, F. G.–(3) 2180; (4) 2753
Johnson, M. A. L.–(3) 2349
Johnson, M. D.–(4) 2799
Jokerst, Nan Marie–(3) 2170
Jondle, Daniel–(2) 1405
Jones, Erin C.–(1) 236
Jonker, B. T.–(4) 3193
Jørgensen, J. F.–(2) 1531
Joubert, O.–(1) 96; (3) 1796; (4) 2493
Jouneau, Pierre-Henri–(3) 1660
Jovin, T. M.–(2) 1399
Joyce, S. A.–(2) 1126
Joyner, C. F.–(1) 294
Juhel, M.–(1) 147
Jung, T. M.–(3) 2072
Jusko, O.–(1) 48; (2) 1531

Kaczer, B.–(4) 2864
Kaindl, Josef–(1) 126
Kalburge, A.–(3) 2203
Kaldis, E.–(2) 970
Kaloyeros, Alain E.–(3) 1828
Kamath, K.–(3) 2312
Kamenitsa, D.–(1) 260
Kamlet, L. I.–(1) 483
Kanehori, Keiichi–(4) 2707
Kaneko, Akira–(2) 606
Kanemaru, Seigo–(3) 1885, 1902, 2105
Kang, E. T.–(3) 1611
Kang, Jie Hun–(3) 1924
Kang, W. P.–(3) 1889, 2068
Kania, D. R.–(3) 2005
Kanski, J.–(2) 957
Kao, K. C.–(1) 527
Karakashian, A. S.–(3) 2318
Karlsson, U. O.–(2) 957
Karp, A.–(2) 1438
Karpov, I.–(2) 623
Kässer, T.–(2) 1296
Kassing, R.–(2) 1563
Kasuya, A.–(2) 938
Kaszuba, Phil–(1) 440
Katayama, Itsuo–(2) 988
Katayama, T.–(4) 3176
Katsumata, Akira–(2) 849
Kaufmann, L. M. F.–(2) 861
Kava, J.–(1) 474
Kavanagh, K. L.–(4) 2952, 3035

Kavanagh, Karen L.–(2) 1238
Kawaguchi, Makoto–(2) 849
Kawakami, R. K.–(4) 3164
Kawakatsu, Hideki–(1) 11; (2) 872
Ke, Kuang-Han–(1) 521; (2) 716
Ke, Mao-long–(4) 2786
Keil, U. D.–(3) 2275
Kelly, K. F.–(2) 593
Kemner, K. M.–(4) 3207
Kendrick, R.–(2) 1197
Kent, M.–(1) 466
Kernen, Peter–(2) 1466
Kerns, D. V.–(3) 2068
Khalil, N.–(1) 224
Khang, Y.–(2) 1221, 1536
Kharas, D.–(1) 437
Khargonekar, P. P.–(1) 483
Khemka, V.–(2) 732
Khor, K. E.–(4) 2817
Khoury, Maroun–(1) 75
Kiesel, P.–(3) 2175, 2275
Kim, Donghwan–(3) 1906
Kim, E.–(3) 2357
Kim, Jae Jeong–(1) 478
Kim, Jae Wan–(3) 1924
Kim, Jong Min–(3) 1924
Kim, Keunjoo–(4) 2667
Kim, Sam-Dong–(2) 642
Kim, Sanghee–(2) 1318
Kim, Woo Shik–(1) 478
Kimura, K.–(4) 3052
King, Hubert W.–(3) 1867
King, O.–(4) 2753
King, R. L.–(4) 2417
Kinoshita, Takashi–(1) 560
Kinser, D. L.–(3) 2068
Kirmse, K. H. R.–(2) 710
Kishimoto, J.–(2) 1000
Kishimoto, T.–(3) 1973
Kisliuk, A.–(3) 1864
Klebanoff, L. E.–(4) 3140
Klein, J.–(2) 1070, 1131
Kleinke, M. U.–(4) 2432
Klemens, F. P.–(4) 2510
Kliese, R.–(2) 1051
Klimecky, P.–(1) 483
Klyachko, D.–(2) 974
Klyui, N. I.–(3) 2130
Knapp, H. F.–(2) 1498
Kneedler, E.–(4) 3193
Kneissl, M.–(3) 2175, 2275
Knight, T. J.–(1) 489
Ko, P. K.–(2) 824
Kobayashi, Katsuyoshi–(2) 1075
Kobayashi, Kiyoshi–(2) 804
Kobayashi, M.–(3) 2371
Kobayashi, N. P.–(3) 2203
Kobzev, A. P.–(2) 657
Koch, F.–(2) 1229, 1336
Koch, M. W.–(3) 2184, 2322
Kochanski, G. P.–(2) 1032; (3) 2011
Kodama, Satoshi–(4) 2888
Kodis, M. A.–(3) 1942, 1947, 1990, 1994
Kodiyalam, S.–(4) 2817
Koenders, L.–(1) 48
Kohmoto, S.–(3) 1764
Koide, Y.–(3) 1812
Kojima, Hisao–(4) 2707
Kollakowski, St.–(3) 1712
Komura, T.–(2) 906
Kondo, Seiichi–(2) 687
Konenkova, Elena V.–(4) 2761
Köning, R.–(1) 48
Kopanski, J. J.–(1) 196, 242
Kopanski, Joseph J.–(1) 426
Kopka, Peter–(3) 2100
Kopp, W. F.–(3) 2236
Kosakovskaya, Z. Ya.–(3) 2080
Koshita, Gen–(3) 2096
Köster, T.–(2) 698
Kougami, Hiroshi–(1) 11
Kowaka, Masahiko–(1) 494
Koyanagi, Satoshi–(4) 2872
Kraft, Robert–(1) 543
Kragler, K.–(2) 1327
Krainsky, I. L.–(3) 2046
Krausch, G.–(2) 1296
Krause, H.-P.–(1) 91
Krogh, B. H.–(1) 489
Kronik, L.–(4) 3089
Krüger, D.–(1) 287, 341
Kruit, P.–(3) 1635
Kruse, N.–(2) 1264
Kryuchenko, Yu. V.–(3) 1934
Kuan, T. S.–(3) 2286
Kub, F. J.–(3) 2229
Kubalek, Erich–(2) 842
Kuhl, D.–(3) 1712
Kuk, Y.–(2) 1188, 1221
Kuk, Young–(2) 1536
Kulakov, M. A.–(4) 2684
Kulik, A. J.–(2) 794, 1308
Kump, Kenneth S.–(4) 2407
Kump, M. R.–(1) 196
Kuo, H. C.–(3) 2252
Kuo, J. M.–(3) 1758, 2252
Kuo, L. H.–(4) 3052
Kurabayashi, Toru–(1) 136
Kurihara, Kenji–(4) 2500
Kuroda, R.–(2) 1353
Kuroda, S.–(1) 324
Kuroda, Tsukasa–(2) 707
Kurogi, H.–(2) 751
Kurokawa, Y.–(2) 1184
Kurps, R.–(1) 287, 341
Kurz, H.–(2) 698; (4) 3107
Kusumi, Y.–(2) 638
Kuwano, Hiroki–(2) 812
Kuzmicz, W.–(1) 408
Kuznetsov, A. V.–(3) 2030
Kwo, J.–(3) 2297
Kwon, Sang Jik–(3) 1906
Kyogaku, M.–(2) 1353

Labardi, M.–(2) 868, 1296, 1509
Labelle, C. B.–(4) 2574
LaBrosse, J.–(1) 489
Lacaze, E.–(2) 1131
Lacklison, D. E.–(3) 2346
Lafond, D.–(1) 22
Lafontaine, H.–(3) 1675
Laframboise, S.–(3) 1675
Lahiri, I.–(3) 2271
Lalinský, T.–(2) 657
Lambers, E. S.–(1) 118; (4) 2567
Lammert, Michael D.–(3) 2216
Lamouri, A.–(3) 2046
Landheer, D.–(4) 2895
Lang, H. P.–(2) 970, 1176
Lányi, Š.–(2) 892, 897
La Rocca, G. C.–(2) 868
Larson, L.–(1) 255
Larsson, A.–(1) 184
Lau, W. M.–(4) 2895
Lau, Y. Y.–(3) 2126
Launay, P.–(1) 147
La Via, F.–(1) 414
Law, Mark E.–(1) 213
Layadi, N.–(4) 2510
Lebedev, Mikhail V.–(4) 2761
Leblanc, R. M.–(2) 1476
Lebreton, C.–(2) 1356
Lecoeur, J.–(2) 1131
Lee, C.–(1) 437
Lee, Chun Gyoo–(3) 1966
Lee, D. R.–(4) 2832
Lee, Eric J.–(4) 2850
Lee, Evans–(2) 716
Lee, G. S.–(2) 727
Lee, Gil U–(2) 789
Lee, H. J.–(4) 2667
Lee, H. P.–(3) 2151
Lee, J. T. C.–(4) 2510, 2574
Lee, J. W.–(1) 118; (3) 1752; (4) 2567, 2582
Lee, Jong Duk–(3) 1906, 1966
Lee, K. B.–(4) 2667
Lee, S. H.–(4) 3013
Lee, S. M.–(2) 1536
Lee, Sunghee–(4) 2824
Lee, Y. H.–(4) 2667
Legay, P.–(1) 147
Lehmann, T.–(2) 1280
Leibovitch, M.–(4) 3089
Leibsle, F. M.–(4) 2424
Lemonias, P. J.–(3) 2233
Lencová, Bohumila–(3) 1860
Lennard, W. N.–(4) 2895
Leo, G.–(3) 1739
Leone, Stephen R.–(4) 2742
Leong, J-K.–(4) 3113
Leta, D. P.–(2) 1276
Leuschner, R.–(2) 1327
Levey, Chris–(1) 440, 457
Levine, Jules D.–(3) 2008
Lew, A. Y.–(4) 2940
Lewis, Jenny–(1) 521
Ley, L.–(2) 1268; (4) 2845, 3008
Li, B.–(1) 91
Li, D.–(4) 2674
Li, J. T.–(2) 1149
Li, L. K.–(3) 2354
Li, Min-Qian–(2) 1410
Li, N. Y.–(3) 2271; (4) 2952
Li, Q.–(3) 1889, 2068
Li, S. F. Y.–(3) 1611
Li, Shaozhong–(3) 2381
Li, Xiaodong–(4) 2660
Li, Xiao-Yu–(1) 571
Li, Y.–(1) 447; (3) 2151
Liao, Hung–(4) 2615
Lieber, Charles M.–(2) 1064
Lim, S. C.–(2) 929
Lin, H. T.–(4) 2922
Lin, Mou-Shiung–(1) 167
Lin, X. F.–(2) 995
Lin, Zhangda–(2) 707
Linder, Barry P.–(1) 552
Linker, F.–(3) 1870
Linnemann, R.–(2) 856
Lio, A.–(2) 1481
Lipp, S.–(3) 1973
Litovchenko, V. G.–(3) 1934, 2130
Litvak, Herbert E.–(1) 516
Liu, Da-Jiang–(4) 2799
Liu, G. L.–(1) 324
Liu, G. Y.–(3) 1963
Liu, J.–(3) 2024, 2050
Liu, J. S.–(3) 1963
Liu, Jie–(2) 1064
Liu, N.–(2) 1349
Liu, Q.–(2) 744; (3) 1697
Liu, W. K.–(2) 647; (3) 2267, 2339
Liu, X. F.–(3) 1889
Liu, Y.–(2) 1276
Liu, Youfan–(3) 2126
Livi, S.–(2) 1531
Lo, W.–(3) 1587
Loan, James F.–(3) 1828
Loh, F. C.–(3) 1611
Löhndorf, M.–(2) 1214
Lombardi, D. R.–(1) 54
Lombardo, Salvatore–(4) 2693
Longo, M.–(3) 1739
Looser, H.–(2) 1386
Lopata, J.–(3) 1758
Loppacher, C.–(2) 1280
Lord, Susan M.–(2) 642
Lorincik, J.–(4) 2417
Losurdo, M.–(2) 691
Lothian, J. R.–(3) 1758
Lovergine, N.–(3) 1739
Lowell, John–(1) 248
Lowney, J. R.–(1) 242
Lü, Z. G.–(4) 3107
Lu, Z. H.–(4) 2882
Lubyshev, D. I.–(3) 2212
Lucovsky, G.–(4) 2812, 2832, 3107
Ludeke, R.–(4) 2855
Ludwig, C.–(2) 1079
Luftman, H. S.–(1) 202
Luning, Scott–(1) 231
Lüpke, G.–(4) 3107
Luscher, Paul–(2) 716
Lüthi, R.–(2) 1255, 1264, 1280, 1285
Lux-Steiner, M.–(2) 1296
Lyszczarz, T. M.–(3) 2060

Ma, Shawming–(1) 566
Ma, Weiluan–(3) 2339
Ma, Y.–(1) 301
Ma, Y.-R.–(2) 943
Ma, Y. Y.–(1) 42
MacDonald, N. C.–(3) 1630
Machajdík, D.–(2) 657
Mackie, W. A.–(3) 2041, 2090
Mackie, William A.–(3) 1952
Madhukar, A.–(3) 2203
Magee, C.–(1) 260
Magee, C. W.–(1) 336
Magno, R.–(3) 2195
Magrì, A.–(1) 369
Magri, R.–(4) 2901
Maher, D. M.–(1) 421; (2) 744
Mahner, E.–(3) 2056
Mahoney, L. J.–(3) 1745
Mahr, K.–(2) 1268; (4) 2845
Mai, H. A.–(2) 995
Maillot, P.–(1) 336
Maiti, C. K.–(3) 1682
Makino, Takahiro–(4) 2500
Malik, S. M.–(1) 336
Malikova, L.–(4) 3089
Malinowski, M.–(3) 1958
Mallet, P.–(2) 1070
Malzer, S.–(3) 2175, 2275
Mamedov, S.–(3) 1864
Manalis, S. R.–(4) 2456
Mancini, A. M.–(3) 1739
Mändl, S.–(4) 2701
Manfra, M. J.–(3) 1745
Mang, K. M.–(2) 1221, 1536
Manger, Dirk–(3) 1828
Manghi, F.–(4) 2901
Mango, F.–(2) 868
Mankey, G. J.–(4) 3171, 3189
Mannaerts, J. P.–(3) 2297
Mansano, Ronaldo Domingues–(1) 538
Mantese, L.–(4) 3070
Mantiply, Bryant–(1) 452
Maranowski, Kevin–(3) 2147
Marchant, Roger E.–(4) 2407
Marchenko, R. I.–(3) 2130
Marchetti, E.–(2) 1509
Marchiando, J. F.–(1) 242

Marder, B. M.–(4) 2438
Marega, E., Jr.–(3) 2212
Margaritondo, G.–(2) 1527
Mariani, Tullio–(2) 845
Mariolle, D.–(1) 22
Markert, J. T.–(2) 1217
Marks, R. F.–(4) 3148
Marsh, Thomas–(2) 1413
Marshall, A. F.–(3) 2208
Mårtensson, Per–(2) 933
Marti, O.–(2) 1296
Martin, Linda–(2) 1405
Martínez-Duart, J. M.–(4) 2623
Marton, D.–(4) 2417
Maruo, Tetsuya–(2) 763
Matsubara, Atsuko–(4) 2707
Matsuda, H.–(2) 1353
Matsui, Y.–(2) 638
Matsumoto, Kazuhiko–(2) 1331
Matsumoto, Yuji–(2) 1114
Matsuo, Naoto–(4) 2770
Matsuoka, H.–(1) 112
Matthai, C. C.–(4) 2786, 3075
Maya, L.–(1) 15
Mayer, T. M.–(3) 1642; (4) 2438
Maynard, H. L.–(4) 2510, 2574
McCann, P. J.–(3) 2267
McCarthy, J. M.–(3) 1650
McClellan, James H.–(4) 2445
McClure, M. T.–(3) 2050
McFeely, F. R.–(4) 2824
McGill, T. C.–(4) 2790, 2948, 3030
McIlroy, D. N.–(4) 2957
McIntyre, N. Stewart–(3) 1867
McMaster, T. J.–(2) 1438
McMurray, J.–(4) 3113
McNabb, J. W.–(2) 617
McQuoid, H.–(2) 1531
McVittie, James P.–(1) 560, 566
McWaid, T. H.–(2) 1540
Mearini, G. T.–(3) 2046
Meassick, S.–(3) 1914
Mei, X. B.–(3) 2327
Meindl, James D.–(1) 192
Meinen, F.–(2) 877
Melloch, M. R.–(2) 861; (3) 2271
Melngailis, John–(1) 179
Memmert, U.–(2) 1224
Men, F. K.–(2) 909, 914
Méndez, J.–(2) 1145
Mercanti, D.–(2) 1395
Mercer, T. W.–(3) 1607
Merz, James–(4) 2914
Mesa, G.–(4) 2403
Meshcheryakova, A. L.–(3) 2034
Meshkinpour, M.–(4) 2922
Meyer, C.–(4) 3107
Meyer, E.–(2) 1255, 1264, 1280, 1285
Meyer, J. A.–(2) 710
Meyer, K.-P.–(2) 1180
Mi, Jian–(3) 1660
Michels, A.–(2) 877
Micovic, M.–(3) 2301
Mielke, F.–(2) 1224
Miethe, K.–(1) 132
Mileham, J. R.–(1) 118
Miles, M. J.–(2) 809, 1438, 1444, 1472
Miles, R. H.–(3) 2335
Miller, D. L.–(3) 2301
Miller, Richard–(2) 1405, 1413
Millo, O.–(1) 30
Mil'shtein, S.–(1) 437
Minne, S. C.–(4) 2456
Mishima, S.–(2) 953
Mishori, B.–(4) 3089
Miskovsky, N. M.–(3) 2020, 2037
Mitchell, I. V.–(4) 2895
Mitsui, Yasuhiro–(4) 2707
Mitterauer, J.–(3) 2083
Miwa, S.–(4) 3052
Miyake, Masayasu–(2) 763
Miyano, J.–(2) 751
Miyazaki, S.–(2) 1268; (4) 2845, 3008
Miyoshi, Tadaki–(4) 2770
Mladenov, G. M.–(4) 2462
Mlynek, J.–(2) 1296
Moers, M. H. P.–(2) 597
Mohammad, S. N.–(4) 2674
Möller, W.–(4) 2701
Möllers, R.–(4) 2712
Molvar, K. M.–(3) 1745
Mönch, Winfried–(4) 2985
Montelius, L.–(4) 2794
Moore, Randy–(1) 498
Morand, Y.–(1) 22
Moreno, M.–(3) 1655
Morgan, Brent A.–(2) 1238
Morgenstern, G.–(1) 102
Mori, Y.–(2) 1109
Moriarty, P.–(2) 943, 1024; (3) 1596
Morimoto, H.–(3) 1973
Morita, S.–(2) 953
Morita, Seizo–(2) 1055
Morita, Y.–(2) 1109
Moriya, N.–(3) 2297
Morkoç, H.–(4) 2674
Morris, S. J.–(1) 260
Mosbacher, J.–(2) 1449
Mozolová, Ž.–(2) 657
Mu, R.–(2) 1083, 1090
Mueller, W.–(3) 2046
Muggelberg, C.–(2) 966
Mukhopadhyay, M.–(3) 1682
Müller, G.–(3) 2056
Müller, T.–(2) 1296
Mundschau, M.–(4) 3126
Münster, S.–(2) 1563
Munz, Andreas W.–(2) 961
Murakami, Masanori–(3) 1812
Murase, Katsumi–(4) 2500
Murata, H.–(4) 3013
Murdfield, Th.–(2) 877
Murison, R.–(3) 1736
Murphy, R. A.–(3) 1745
Murtaza, S. S.–(3) 2256
Murthy, C. S.–(1) 278
Müssig, H.-J.–(2) 982
Myers, A. F.–(3) 2024, 2050

Nagahara, L. A.–(2) 800
Nagase, Masao–(4) 2500
Nagatomo, Y.–(2) 751
Nagy, J. O.–(2) 1481
Nakagiri, Nobuyuki–(2) 887
Nakagiri, T.–(2) 1353
Nakahara, S.–(1) 202
Nakamura, M.–(2) 1109
Nakamura, T.–(2) 1184
Nakane, Hideaki–(3) 2114
Nakasuji, Mamoru–(4) 2467
Namatsu, Hideo–(4) 2500
Namba, Y.–(1) 42
Narita, Tadashi–(4) 2645
Nason, D.–(2) 1096
Nazuka, Yutaro–(3) 2105
Neddermeyer, H.–(2) 925, 1051, 1141
Nefedov, I. S.–(3) 1982
Nejo, H.–(4) 2399
Nelep, C.–(3) 2130
Neo, Yoichiro–(3) 2096
Nerino, R.–(2) 897
Neubauer, Gabi–(1) 426
Neudeck, Gerold W.–(3) 1670
Neumann, G.–(3) 2000
Neumann, P.–(2) 1157
Never, James–(1) 440
Nichols, C.–(3) 1819
Nichols, K. B.–(3) 1745, 2236
Nicolaescu, Dan–(3) 1930
Niedermann, Ph.–(3) 2056
Niehuis, E.–(4) 2712
Nielsen, L.–(2) 1531
Niimi, H.–(4) 2812
Niiyama, Takahiro–(3) 1977
Nishihara, Yoshikazu–(2) 1060
Nishikawa, Osamu–(3) 2110
Nishimura, Kazumi–(2) 652
Nishina, Y.–(2) 938
Nishizawa, Jun-ichi–(1) 136
Nitishin, P. M.–(3) 1745
Noh, H. P.–(2) 1188
Nojiri, Kazuo–(3) 1791
Nolte, D. D.–(3) 2271
Nordström, Lars–(4) 3160
Norman, J. A. T.–(2) 1032
Novikov, S. V.–(3) 2346
Noya, Atsushi–(2) 674
Nunan, Peter–(1) 218
Nxumalo, J. N.–(1) 386

Oasa, Takahiko–(2) 1055
O'Connor, S. D.–(2) 852
O'Connor, Stephen D.–(2) 1302
Odaki, T.–(2) 1365
Oden, P. I.–(2) 1313
Oesterschulze, E.–(2) 832, 856, 1563
Offer, G.–(2) 1444
Ogletree, D. F.–(2) 1289; (4) 2772
Ogo, I.–(3) 1630
Ohashi, T.–(3) 1577
Ohba, Yasuyuki–(2) 988
Ohkouchi, S.–(2) 953
Ohlberg, D. A. A.–(3) 1958
Ohnishi, Hideaki–(2) 988
Ohta, Akira–(2) 674
Ohta, M.–(2) 953
Oka, Kunihiko–(2) 1060
Okada, Yoshitaka–(3) 1725
Oki, A. K.–(3) 2221
Oki, Aaron K.–(3) 2216
Okuda, S.–(3) 1973
Okusako, Takahiro–(2) 1055
Okuyama, F.–(2) 768
Olsson, L. Ö.–(2) 957
Olthoff, S.–(2) 1019
Omstead, Thomas R.–(3) 1828
Onoda, Hiroshi–(4) 2645
Oral, A.–(2) 1202
Orloff, Jon–(1) 80
Orton, J. W.–(3) 2346
Osaka, F.–(2) 953
Osaka, T.–(2) 1184
O'Shea, S. J.–(2) 1383
O'Steen, M. L.–(4) 2948
Osvald, J.–(2) 657
Otsuka, I.–(2) 1153, 1184
Ott, Adina K.–(4) 2742
Otterbeck, Markus–(2) 842
Ouellet, L.–(4) 2627
Oulevey, F.–(2) 794, 1308
Oura, Kenjiro–(2) 988
Overbury, S. H.–(2) 918
Overney, R. M.–(2) 1276
Owman, Fredrik–(2) 933
Ozzello, T.–(4) 2967

Padmore, H. A.–(4) 3119
Paggel, J. J.–(4) 2980, 2994
Palmetshofer, L.–(1) 272
Pang, S. J.–(2) 1349
Pang, S. W.–(2) 612; (3) 1807; (4) 2531
Pao, Y. C.–(3) 2240
Pappas, D. P.–(4) 3203
Parab, K. B.–(1) 260
Paranthaman, M.–(1) 15
Park, Byung Gook–(3) 1966
Park, Chan–(2) 1318
Park, D. G.–(4) 2674
Park, J. Y.–(2) 1188
Park, Jae Hyun–(1) 478
Park, R. M.–(1) 159
Park, Sang-Kyu–(2) 1318
Park, W.–(3) 2263
Park, Y. J.–(2) 1536
Park, Young-Bae–(4) 2660
Parker, D. L.–(3) 1899
Parker, M. R.–(4) 3131
Parker, Marie-Claire–(2) 1432
Partridge, A.–(2) 585
Pascual, J. I.–(2) 1145
Paskova, T.–(3) 1729
Pasquarello, Alfredo–(4) 2809
Passlack, M.–(3) 2297
Patrin, J.–(3) 1870
Patterson, George A.–(3) 2187
Paulson, W. M.–(3) 1697
Pawlik, M.–(1) 390
Pearson, P.–(1) 348
Pearton, S. J.–(1) 118; (3) 1752, 1758; (4) 2567, 2582
Peden, C. H. F.–(2) 1126
Pei, S. S.–(3) 2252
Pekarik, John–(1) 440
Peluso, T.–(3) 1739
Pelz, J. P.–(4) 2864
Pennycook, S. J.–(3) 2199
Pereira, R. G.–(1) 106; (3) 1773
Peressi, M.–(4) 2936
Perez, J. M.–(2) 929
Pérez-Casero, R.–(4) 2623
Pérez-Rigueiro, J.–(4) 2623
Perfetti, P.–(2) 1395, 1527
Perry, M. D.–(3) 2005
Petersen, J.–(2) 1079
Peterson, Gregory G.–(3) 1828
Petitprez, E.–(3) 2212
Pettersson, P. O.–(4) 3030
Phillips, J.–(3) 2312
Phillips, P.–(1) 283
Pianetta, P.–(4) 2914
Pickett, Warren E.–(4) 3137
Picotto, G. B.–(2) 897, 1531
Piechotka, M.–(2) 970
Piel, H.–(3) 2056
Pierrat, Christophe–(1) 63
Pillai, D.–(4) 3131
Pinsukanjana, Paul–(3) 2147
Plekhanov, P. S.–(3) 2030
Plumlee, Jason E.–(3) 2041
Poirier, G. E.–(2) 1453
Poker, D. B.–(3) 1853
Pollak, Fred H.–(4) 3089
Pollock, H. M.–(2) 1486
Poon, M. C.–(2) 824
Poon, Vincent M. C.–(3) 1938
Popova, T. B.–(3) 2346
Porath, D.–(1) 30
Posselt, M.–(1) 278
Pourmir, F.–(2) 1131
Prato, Stefano–(2) 593
Prengle, Scott–(1) 543
Preuninger, Fred–(1) 571
Prewett, P. D.–(3) 1910
Prins, M. W. J.–(2) 1206
Prins, R.–(2) 1264

Privitera, V.–(1) 369
Pu, Bryan–(2) 716
Pu, Bryan Y.–(1) 521
Pupeter, N.–(3) 2056
Py, C.–(3) 2119

Qiu, J. Y.–(4) 2505
Qiu, Z. Q.–(4) 3164
Quate, C. F.–(4) 2456

Radojkovic, P.–(2) 1229, 1336
Rafailovich, M. H.–(2) 1276
Raineri, V.–(1) 414
Raisanen, A. D.–(4) 2961
Rajavel, R. D.–(3) 2362
Rakhshandehroo, M. R.–(2) 612
Ralston, J.–(3) 1780
Ram, P.–(4) 3089
Ramachandran, T. R.–(3) 2203
Ramadan, A. R.–(2) 1383
Ramam, A.–(3) 1719
Ramaswami, S.–(4) 2636
Rammohan, K.–(4) 2922
Rangelow, I. W.–(2) 856
Rashap, B. A.–(1) 483
Rasing, Th.–(2) 861
Rau, Richard–(4) 2445
Ray, S. K.–(3) 1682
Rayment, T.–(2) 1383
Reading, M.–(2) 1486
Reddy, K.–(1) 255
Reed, M. A.–(1) 54
Řehůřek, P.–(2) 892
Reichert, A.–(2) 1481
Reidsema, Cindy–(3) 1828
Reiner, J.–(3) 2318
Remmes, N. B.–(4) 2957
Ren, F.–(1) 118; (3) 1758; (4) 2567, 2582
Ren, J.–(3) 2349
Resch-Esser, U.–(1) 187(E)
Resh, J. S.–(2) 593
Rhee, Shi-Woo–(4) 2660
Rhoades, R. L.–(3) 1853
Rich, D. H.–(4) 2922
Richardson, B. E.–(4) 2786
Richardson, R. A.–(4) 2510
Richter, H. H.–(1) 102
Richter, W.–(1) 187(E); (4) 2757, 3058
Rietman, Edward A.–(1) 504
Ring, Ken M.–(2) 1238
Ringel, S. A.–(3) 2157
Ristein, J.–(4) 3008
Ritter, G.–(1) 102
Rivera, M.–(2) 1472
Roberts, C. J.–(2) 1557; (3) 1582
Robinson, G. Y.–(3) 2248
Rockett, A.–(4) 2674
Rodgers, Mark–(1) 426
Roditchev, D.–(2) 1070
Rogers, T. J.–(3) 2236
Rolfe, S. J.–(3) 1675
Romanato, F.–(3) 1739
Romanowicz, J.–(4) 3126
Roos, B.–(1) 373
Roskosand, H. G.–(2) 698
Rossberg, A.–(2) 970
Rossnagel, S. M.–(3) 1819, 1853; (4) 2603
Rossow, U.–(4) 3019, 3040, 3070
Rotenberg, E.–(4) 3171
Röttger, B.–(2) 925, 1051
Rousset, S.–(2) 1131
Rowell, N. L.–(3) 1675
Rowland, W. H., Jr.–(3) 2259, 2349
Rubini, S.–(4) 2980
Rubloff, G. W.–(2) 772
Rudman, D. A.–(2) 1210
Rueckwald, E. R.–(3) 2184
Ruegsegger, S. M.–(1) 483
Ruff, M.–(3) 2275
Rugar, D.–(2) 1197
Ruiter, A. G. T.–(2) 597
Ruskell, Todd G.–(2) 864
Rust, H.-P.–(2) 1117
Ruzic, D.–(3) 1819
Ryssel, H.–(3) 1973

Sacedón, J. L.–(3) 1655
Sacks, R. N.–(3) 2157, 2187
Sacks, W.–(2) 1070
Sadakane, Shinji–(3) 1970
Saemann-Ischenko, G.–(2) 1327
Sáenz, J. J.–(2) 1145; (4) 2403
Saggio, M. G.–(1) 369
Sailor, Michael J.–(4) 2850
Saimoto, S.–(4) 2636
Saito, Takashi–(2) 872
Sakata, Y.–(3) 1764
Sakuraba, Hiroshi–(1) 136
Sakurai, T.–(2) 1000
Salamanca-Riba, L.–(3) 2366
Salling, Craig T.–(2) 1322
Salmeron, M.–(2) 1289, 1481; (4) 2772
Salmeron, Miquel–(2) 1341
Samara, D.–(2) 1344
Sampson, A.–(1) 466
Samuelson, L.–(4) 2794
Santos, M. B.–(2) 647; (3) 2267, 2339
Saravanan, K.–(1) 474
Sarid, Dror–(2) 864, 979
Sarkar, Dipankar–(2) 593
Sasaki, Akira–(2) 849
Sasaki, Katsutaka–(2) 674
Sase, Touko–(2) 674
Sassaki, R. Massami–(4) 2432
Sato, D. L.–(3) 2151
Sato, K.–(1) 112
Savolainen, P.–(3) 1736
Scandella, L.–(2) 1255, 1264
Scarselli, M. A.–(2) 1395, 1527
Schad, R.–(2) 1206
Schade, U.–(3) 1712
Schäfer, J.–(4) 3008
Schetzina, J. F.–(3) 2259, 2349
Schierbaum, Klaus-Dieter–(2) 961
Schlachetzki, A.–(1) 48
Schlammadinger, J.–(2) 1399
Schlapbach, Louis–(2) 1250
Schlittler, R. R.–(2) 827
Schlueter, J.–(3) 1837
Schlueter, James–(3) 1846
Schmid, P. G.–(3) 1706
Schmidt, D.–(4) 2812
Schneider, W.-D.–(2) 1149
Schneir, J.–(2) 1540
Schnieders, A.–(4) 2712
Scholz, W.–(2) 832
Schreiber, E.–(3) 2087
Schreyer, A.–(2) 1162
Schulman, J.–(3) 2163
Schumacher, A.–(2) 1264
Schumann, F. O.–(4) 3189
Schütze, M.–(2) 698
Schwartzkopff, M.–(2) 1229, 1336
Schweizer, E. K.–(2) 1117
Schwentner, N.–(1) 91
Schwoebel, P. R.–(3) 1986
Seco, Miquel–(4) 2609
Segawa, Kazuhito–(2) 1331
Segawa, Yusaburo–(2) 1238
Seibles, L.–(3) 2011
Seiferth, F.–(4) 2753
Seiler, D. G.–(1) 196
Selberherr, S.–(1) 224
Semenovich, V.–(3) 2130
Semenza, G.–(2) 1418
Semenza, Giorgio–(2) 1466
Seo, A.–(2) 1365
Serra, A.–(2) 1527
Servat, Jordi–(4) 2609
Setiagung, C.–(3) 2371
Seuret, Ch.–(2) 1386
Sgarzi, P.–(2) 1509
Shakesheff, K. M.–(2) 1557; (3) 1582
Sham, T. K.–(4) 3199
Shan, Hongching–(1) 521; (2) 716
Shan, Jun–(2) 1038
Shan, W.–(3) 2267
Shanabrook, B. V.–(4) 3193
Shanabrook, B.V.–(3) 2195
Shapira, Yoram–(4) 3089
Shareef, I. A.–(2) 772
Shaw, J. L.–(3) 2072
Sheergar, M. K.–(4) 2595
Shek, M. L.–(4) 3199
Shen, Dian-Hong–(2) 707
Shi, Song–(4) 2914
Shido, S.–(2) 1353
Shih, C. K.–(1) 447; (2) 909, 914, 948, 1344
Shimada, Y.–(2) 1353
Shimitzu, Tetsushi–(2) 1060
Shimizu, D. T.–(1) 386
Shimizu, Hiroyasu–(4) 2467
Shimizu, Keizo–(3) 2105
Shin, Y. G.–(4) 2667
Shinabe, S.–(2) 1029
Shindo, H.–(2) 1365
Shiojima, Kenji–(2) 652
Shiraki, Y.–(3) 2387
Short, S. W.–(3) 2374
Shuh, David K.–(4) 2909
Shul, R. J.–(1) 118
Siciliano, P.–(2) 1527
Sieg, R. M.–(3) 2157
Siegel, M.–(1) 474
Siegrist, Theo–(1) 63
Sigmon, T. W.–(3) 1650
Sik, H.–(1) 147
Simionescu, A.–(1) 272
Simonton, R.–(1) 260
Simpson, William C.–(4) 2909
Sin, J. K. O.–(2) 824
Sin, Johnny K. O.–(3) 1938
Sinclair, Robert–(1) 231
Singh, David J.–(4) 3137, 3160
Singh, J.–(3) 2312, 2378
Sinitsyn, N. I.–(3) 1982, 2080
Šišolák, J.–(2) 657
Sleight, J. W.–(1) 54
Slinkman, James–(1) 440, 457
Smalbrugge, E.–(2) 861
Smith, A. R.–(2) 909, 914, 948
Smith, F. W.–(3) 2236
Smith, G. D. W.–(3) 1895
Smith, Gene–(1) 533
Smith, H.–(1) 433
Smith, H. E.–(1) 305, 404
Smith, J. S.–(3) 2175
Smith, N. V.–(4) 3119
Smithe, D. N.–(3) 1942, 1947, 1990
Smy, T.–(2) 679; (4) 2595
Snauwaert, J.–(1) 380; (2) 1513
Sobiesierski, Z.–(4) 3058, 3065
Sokolov, J.–(2) 1276
Soler, J. M.–(2) 1145
Solletti, J. M.–(2) 1492
Solomon, G. S.–(3) 2208
Soltz, David–(3) 1784
Sommer, F.–(2) 1492
Song, H. B.–(3) 1889
Song, J. J.–(3) 2267
Song, Juho–(2) 727
Song, M.–(2) 1486
Sorba, L.–(2) 623; (4) 2994
Sosso, A.–(2) 897
Sotto, M.–(2) 1131
Sou, I. K.–(4) 2973
Spallas, J. P.–(3) 2005
Spangenberg, B.–(2) 698
Spence, J. C. H.–(3) 1587
Spicer, W. E.–(4) 2914
Spindt, C. A.–(3) 1986
Spinella, C.–(1) 414
Sprengepiel, Jörg–(2) 842
Springer, C.–(1) 187(E)
Sproule, G. I.–(3) 1675
St. John, Pamela M.–(1) 69
Stahrenberg, K.–(4) 2757
Stair, Kathleen A.–(3) 2187
Stalder, A.–(2) 1259
Stallcup II, R. E.–(2) 929
Stanley, C. R.–(3) 2305
Stefanov, E.–(2) 1229
Stehle, M.–(2) 820
Steimetz, E.–(4) 3058
Stevens, F.–(1) 38
Stingeder, G.–(1) 272
Stone, D. R.–(4) 2753
Stopka, M.–(2) 832
Strachan, W. J.–(3) 2318
Streb, D.–(3) 2275
Streetman, B. G.–(3) 2256
Streetman, Ben G.–(3) 2278
Streit, D. C.–(3) 2221
Streit, Dwight C.–(3) 2216
Streller, U.–(1) 91
Stringfellow, G. B.–(4) 3013, 3113
Strohmaier, R.–(2) 1079
Su, Chao-Shin–(4) 2550
Subrahmanyan, Ravi–(1) 218
Suga, T.–(3) 1577
Sugawara, Y.–(2) 953
Sugawara, Yasuhiro–(2) 1055
Sugimoto, Y.–(3) 1764
Sugimura, Hiroyuki–(2) 887
Suh, M. S.–(4) 2667
Suh, Y. D.–(2) 1188
Sukidi, N.–(4) 3019
Sullivan, John J.–(3) 1828
Sultan, A.–(1) 255
Sumita, Isao–(2) 606
Sumitani, K.–(1) 112
Summers, C. J.–(3) 2263
Sun, D. L.–(4) 3126
Sung, Kuo-Tung–(4) 2550
Suto, S.–(2) 938
Suzuki, A.–(3) 1812
Suzuki, M.–(2) 953
Suzuki, Y.–(4) 3176
Suzuki, Yoshihiko–(2) 887
Svensson, Stefan P.–(3) 2166, 2192
Swartzentruber, B. S.–(3) 1642
Swope, Richard–(3) 1702
Szabadi, M.–(3) 1600

Taborelli, Mauro–(2) 1422
Taferner, W. T.–(3) 2357
Tagare, M. V.–(3) 2325
Tait, R. N.–(2) 679
Takács, L.–(2) 1399
Takahashi, E.–(2) 1048
Takai, M.–(3) 1973
Takano, H.–(1) 112

Takano, Hajime–(2) 1272
Takata, Keiji–(2) 882
Takebe, T.–(4) 2731
Takeuchi, T.–(3) 1764
Takeyama, Mayumi–(2) 674
Takimoto, K.–(2) 1353
Talin, A. Alec–(2) 1238
Tamura, H.–(2) 751
Tan, K. L.–(3) 1611
Tanaka, K.–(2) 1365; (4) 3052
Tanaka, Ken-ichi–(2) 1114
Tanaka, Kuniyoshi–(2) 1234
Tang, H. T.–(4) 2895
Tang, S. W.–(3) 1963
Tang, Tiantong–(1) 80
Tang, Y.–(4) 2922
Taniguchi, Masateru–(3) 1977
Tanimoto, Masafumi–(2) 1547
Tanoue, Hisao–(3) 1885
Tansley, T. L.–(1) 152
Tao, M.–(4) 2674
Tapfer, L.–(3) 1739
Tasch, A. F.–(1) 260
Tay, S. P.–(4) 2882
Tazi, A.–(2) 1476
Teichert, J.–(3) 1621
Teitsworth, S. W.–(4) 2725
Tejada, Javier–(4) 2609
Temst, K.–(2) 1121
Tendler, S. J. B.–(2) 1557; (3) 1582
Tendler, Saul J. B.–(2) 1432
Tepermeister, I.–(4) 2510
Tepore, A.–(2) 1527
Teraguchi, N.–(3) 1812
ter Beek, Marcel–(1) 218
Terhorst, M.–(4) 2712
Terry, F. L., Jr.–(1) 483
Teschke, O.–(4) 2432
Teuschler, T.–(2) 1268; (4) 2845
Thanikasalam, P.–(4) 2840
Thiagarajan, P.–(3) 2248
Thibado, P. M.–(3) 1870; (4) 3193
Thibado, P.M.–(3) 2195
Thibado, Paul M.–(3) 1607
Thomas, H.–(2) 1280
Thomas, S., III–(3) 1807; (4) 2531
Thommen-Geiser, V.–(2) 1176
Thompson, Carl V.–(1) 179
Thompson, Liliana–(1) 231
Thompson, P. E.–(3) 2229
Thomson, D. J.–(1) 386
Thundat, T.–(1) 15; (2) 1313; (3) 2199
Tian, C.–(1) 272
Tian, S.–(1) 260
Tiffin, D.–(1) 447
Tillack, B.–(1) 102
Ting, D. Z.-Y.–(4) 2790
Tkachenko, V. A.–(4) 2399
Tobin, J. G.–(4) 3152, 3171
Tobin, P. J.–(3) 1697
Tofte, Jan–(3) 2147
Toivonen, M.–(3) 1736
Tokumoto, H.–(2) 602, 800, 1109
Tokumoto, Hiroshi–(2) 812
Toma, Yasushi–(3) 1902
Tomich, D. H.–(4) 2952
Tomomura, Y.–(3) 1812
Tong, Linsu–(3) 1938
Tong, S. Y.–(4) 3152
Torgashov, G. V.–(3) 1982, 2080
Török, J.–(2) 892
Tortonese, M.–(2) 1210
Touchi, Kenshin–(4) 2645
Toussaint, S. L. G.–(2) 585
Towe, E.–(3) 2309
Towner, F. J.–(3) 2315
Tran, L.–(3) 2221
Tremblay, Y.–(4) 2627
Tret'yakov, V. V.–(3) 2346
Trexler, J. T.–(1) 159
Trezza, J. A.–(3) 2208
Trinick, J.–(2) 1444
Troyon, Michel–(2) 1552
Tsai, V. W.–(2) 1540
Tsai, W.–(2) 679
Tskhovrebova, L.–(2) 1444
Tsubouchi, N.–(1) 112
Tsuji, Hiroshi–(3) 1970
Tsukamoto, Kazuyoshi–(4) 2770
Tsukui, K.–(3) 1812
Tsunokuni, Kazuyuki–(3) 1791
Tu, C. W.–(3) 2271, 2282, 2327, 2331; (4) 2918, 2952
Tung, Y. S.–(2) 1083, 1090
Turkot, R.–(3) 1819
Turner, I. G.–(2) 1167
Twichell, J. C.–(3) 2060
Twigg, M.E.–(3) 2195

Uchida, H.–(1) 324
Uchihashi, Takayuki–(2) 1055
Ueda, A.–(2) 1083
Ueyama, H.–(2) 953
Umemoto, Donald K.–(3) 2216
Unger, Marc A.–(2) 1302
Uppal, Parvez N.–(3) 2192
Usami, N.–(3) 2387

Vaidya, Sheila–(1) 63
Valcheva, E.–(3) 1729
Vallier, L.–(1) 96; (3) 1796
van den Oetelaar, L. C. A.–(2) 585
Vandervorst, W.–(1) 265, 311, 358, 380, 390, 404; (2) 1513
Van Haesendonck, C.–(2) 1121
Van Hove, M.–(1) 106; (3) 1773
van Hulst, N. F.–(2) 597
van IJzendoorn, L. J.–(2) 585
van Kempen, H.–(2) 861, 1173, 1206
van Kesteren, H. W.–(2) 1206, 1214
Van Rossum, M.–(1) 106; (3) 1773
Vanzetti, L.–(4) 2961
Vasanelli, L.–(3) 1739
Vasileska, D.–(4) 2780
Vatel, Olivier–(2) 1547
Vautrot, Philippe–(2) 1552
Vehoff, H.–(2) 1157
Venable, D.–(2) 744
Venables, D.–(1) 421
Verdonck, Patrick–(1) 538
Vereb, G.–(2) 1399
Vesenka, James–(2) 1405, 1413
Veuillen, J.-Y.–(2) 1005, 1010
Vigneron, J. P.–(2) 816
Villarreal, L. M.–(2) 929
Villarrubia, J. S.–(2) 1518, 1540
Vinckier, Anja–(2) 1427
Viswanathan, Chand R.–(1) 571, 577
Vogt, A.–(2) 820
Vogt, T. J.–(3) 2248
Volk, R.–(2) 877
von Criegern, R.–(1) 272
von Seggern, H.–(2) 1327
Voronin, A. B.–(3) 2034
Vu, Bao–(4) 2588

Wada, Y.–(2) 1522
Wadas, A.–(2) 1214
Waddill, G. D.–(4) 3152
Wagner, B. K.–(3) 2263
Wagner, P.–(2) 1418
Wagner, Peter–(2) 1466
Wagner, R. J.–(4) 3193
Wago, K.–(2) 1197
Walba, D. M.–(1) 38
Walck, S. D.–(3) 2278
Wallis, C. R.–(4) 2725
Walpole, James N.–(1) 179
Walther, Steven R.–(1) 329
Walzer, K.–(2) 1461
Wamsley, C. C.–(3) 2322
Wandel, K.–(3) 1712
Wang, Baoping–(3) 1938
Wang, C. L.–(2) 744
Wang, D. W.–(2) 1349
Wang, Dawen–(3) 2381
Wang, G.-C.–(4) 3180
Wang, J. Y.–(4) 3126
Wang, Jie–(4) 2973
Wang, Kang L.–(3) 2381
Wang, Li–(1) 80
Wang, R. H.–(4) 2690
Wang, Rynn–(2) 716
Wang, Shi-Qing–(1) 174; (3) 1837, 1846; (4) 2656
Wang, W. I.–(3) 2286, 2343, 2354
Wang, Y.–(3) 1687; (4) 2543
Wang, Yajun–(2) 1038
Wang, Yaxin–(3) 2046
Wang, Z.–(4) 2674
Wang, Z. L.–(3) 2263
Wang, Z. Z.–(2) 1356
Warmack, R. J.–(2) 1313; (3) 2199
Warren, Patricia–(3) 1660
Warwick, A.–(4) 3171
Wasilewski, Z.–(3) 2290
Watanabe, T.–(4) 2731
Watanabe, Teruo–(3) 1977
Watanuki, Osaaki–(2) 804
Watts, Josef S.–(3) 1670
Wawro, A.–(2) 938
Weaver, J. H.–(4) 2519
Weeks, John D.–(4) 2799
Wei, H. Y.–(3) 2366
Weichold, M. H.–(3) 1899
Weierstall, U.–(3) 1587
Weightman, P.–(4) 2424
Weir, D. G.–(3) 2233
Weiss, B. L.–(2) 632
Welch, Michael–(1) 521; (2) 716
Weldon, M. K.–(4) 3095
Welland, M. E.–(2) 1019, 1383
Wells, Randy–(1) 440
Wells, Timothy N. C.–(2) 1422
Welser, R. E.–(1) 54
Wen, H. J.–(4) 2855
Wendelken, J. F.–(2) 918
Wendt, A. E.–(2) 710
Weng, Y. M.–(4) 2505
Wenisch, H.–(4) 2757
Werlen, Raymond C.–(2) 1422
Werner, S.–(2) 832
Westwood, D. I.–(4) 2786, 3058, 3065
Whaley, R. D.–(4) 2753
Whelan, Mike–(1) 498
Whidden, T. K.–(4) 2840
Whitman, L. J.–(2) 992; (3) 1870; (4) 3193
Whitman, L.J.–(3) 2195
Wicks, G. W.–(3) 2184, 2322
Wieder, H. H.–(3) 2293; (4) 2944, 3035
Wiesendanger, R.–(2) 1180, 1214
Wijers, C. M. J.–(4) 3080
Wilkening, G.–(2) 1531
Williams, C. C.–(1) 433; (2) 1536; (4) 3113
Williams, Clayton C.–(1) 426
Williams, Ellen D.–(4) 2799
Williams, F. J.–(4) 2519
Williams, P. M.–(2) 1557; (3) 1582
Williams, R. H.–(4) 2786
Williams, R. Stanley–(2) 1238
Williamson, J. R.–(2) 1344
Williamson, R. L.–(2) 809, 1472
Willis, R. F.–(2) 995; (4) 3171, 3189
Willke, T. L.–(4) 2524
Wilshaw, P. R.–(3) 1895, 1910
Wilson, David L.–(4) 2407
Wilson, I. H.–(2) 1105
Winesett, J.–(3) 2339
Winfield, M. O.–(2) 1438
Wise, M. L.–(1) 301
Wojtowicz, M.–(3) 2221
Wojtowicz, Michael–(3) 2216
Wolf, J. A.–(4) 3207
Wolf, R. M.–(2) 1173
Wolff, A.–(1) 102
Wolff, G.–(2) 982
Wolter, J. H.–(2) 861
Wong, G. K. L.–(4) 2973
Wong, H.–(2) 824
Wong, W. S.–(4) 2952
Wood, C. E. C.–(3) 2180, 2366
Woodall, J. M.–(3) 2225, 2271, 2325; (4) 2739
Woods, R. Claude–(2) 710
Woolf, D. A.–(2) 943, 1024; (4) 2424, 3058, 3065
Workman, Richard K.–(2) 864, 979
Wu, J. Z.–(2) 710
Wu, O. K.–(3) 2362
Wu, Robert–(2) 716
Wu, S. Z.–(4) 3189
Wu, X. X.–(3) 2175

Xiang, Qi–(3) 2381
Xiao, Xu-dong–(2) 1341
Xie, Qianghua–(3) 2203
Xie, Tianbao–(3) 2090
Xin, S. H.–(3) 2374
Xu, J. B.–(2) 1105
Xu, J. F.–(3) 1889, 2068
Xu, Jun–(2) 918
Xu, Lei–(2) 1410
Xue, Ping–(4) 2407
Xue, Z. Q.–(2) 1349

Yakimova, R.–(3) 1729
Yamada, Hirofumi–(2) 812
Yamada, Hiroshi–(2) 757
Yamada, Koji–(2) 1369
Yamada, T.–(3) 2371
Yamada, Toshishige–(2) 1243
Yamamoto, M.–(3) 1764
Yamamoto, T.–(4) 2731
Yamamoto, Takuma–(2) 887
Yamamoto, Yasuji–(2) 988
Yamamoto, Yoshihisa–(2) 1243
Yamanishi, Yoshiki–(2) 1055
Yamashita, M.–(3) 1973
Yamazaki, Kazuo–(3) 1791
Yan, C. H.–(3) 2331
Yan, J.–(3) 1706
Yang, H.-N.–(4) 3180
Yang, J.–(4) 3131
Yang, S.-H.–(1) 260
Yang, T.–(3) 2263
Yang, X.–(4) 2961
Yang, Z.–(3) 2354; (4) 2973
Yannoni, C. S.–(2) 1197
Yano, Fumiko–(4) 2707
Yano, K.–(2) 1353
Yao, T.–(2) 906, 1029, 1048; (4) 3052
Yarmoff, Jory A.–(4) 2909
Yasuda, T.–(4) 3052

Yasuda, Takashi–(2) 1238
Yasui, Takanari–(2) 1238
Yeh, Wen-Kuan–(1) 167
Yeo, I.-S.–(4) 2636
Yeung, Y. H.–(4) 2973
Yin, A.–(3) 2374
Yokoo, Kuniyoshi–(3) 2096
Yokoyama, Hiroshi–(3) 2105
Yoo, Minsun–(3) 1591
Yoo, Woo Sik–(3) 1702
Yoshikawa, A.–(3) 2371
Yoshikawa, T.–(3) 1764
Yoshimaru, M.–(2) 751
Yoshimura, M.–(2) 906, 1029, 1048
Youn, C. J.–(4) 2667
Young, A. P.–(3) 2293; (4) 2944
Young, A. T.–(4) 3119
Young, Michael–(1) 452
Yu, E. T.–(4) 2940
Yu, J.–(1) 42
Yu, Z.–(3) 2259
Yuan, C. W.–(2) 1210
Yura, S.–(3) 1973

Zaidman, and E. G.–(3) 1990
Zaidman, E. G.–(3) 1873, 1942, 1947, 1994
Zakharchenko, Y. F.–(3) 1982
Zakharchenko, Yu. F.–(3) 2080
Zaugg, Frank–(2) 1466
Zavracky, Paul M.–(4) 2588
Zawadzki, P.–(3) 2290
Zegenhagen, J.–(1) 187(E)
Zeindl, H. P.–(1) 287, 341
Zengerle, R.–(4) 2537
Zettler, J. -T.–(4) 2757
Zettler, J.-T.–(4) 3058
Zhang, F. C.–(3) 2378
Zhang, J.–(3) 1600
Zhang, K. Z.–(4) 2824
Zhang, R.–(1) 1
Zhang, Xuemei–(3) 2339
Zhang, Y. H.–(4) 2940
Zhang, Z.–(4) 2684
Zhang, Zhenyu–(2) 909
Zhao, Bin–(1) 174; (4) 2656
Zhao, C. Y.–(2) 744
Zhbanov, A. I.–(3) 1982
Zheng, Y. D.–(4) 2690
Zheng, Z.–(2) 1210
Zhirnov, V. V.–(3) 2030, 2034, 2050
Zhou, J. J.–(3) 2151
Zhou, Y. H.–(4) 2505
Zhu, C. C.–(3) 1963
Zhu, M. H.–(3) 1963
Zhu, S. M.–(4) 2690
Zhu, W.–(3) 2011
Zinck, J. J.–(3) 2163
Zoethout, E.–(4) 3080
Zong, X. F.–(4) 2505
Zotov, A. V.–(4) 2684
Züger, O.–(2) 1197

Get the Technical Training You Need Now!

For more than 20 years, the AVS has been providing practical job training by recognized professionals in the areas of:

- *Applied Vacuum Technology*
- *Surface Analysis and Materials Characterization*
- *Materials, Thin Films, and Coatings: Processing and Properties*

Courses are offered all year long across the U.S. Call the 800 number below to find out when courses will be offered in your area.

Ask us about our On-Site Training Program too!

Call AVS Short Courses today at 1-800-888-1021.

For Convenient Training at Your Facility ...

Try AVS Industrial Training Videos

We bring vacuum technology training directly to your workplace—on video—for convenient employee training, review, and self-paced learning. Training programs currently available are:

- *Properties of Vacuum System Materials*
- *Sputter Deposition and Ion Beam Processes*
- *Fundamentals of Capture Pumping*

Each program provides 6-8 hours of training and includes 20 illustrated workbooks.

Order your tapes today! Call the AVS at 212-248-0200.

Learn Vacuum Technology from the Experts ...

Buy an *AVS Classic Series* reference book—out-of-print references by industry-founding experts:

- *Vacuum Technology and Space Simulation*—Santeler, Holkeboer, Jones, and Pagano
- *Field Emission and Field Ionization*—Gomer
- *Handbook of Electron Tube and Vacuum Techniques*—Rosebury
- *The Physical Basis of Ultrahigh Vacuum*—Redhead, Hobson, and Kornelsen
- *Ionized Gases*—von Engel
- *Vacuum Sealing Techniques*—Roth
- *Basic Data of Plasma Physics: The Fundamental Data on Electrical Discharges in Gases*—Brown
- *Handbook of Materials and Techniques for Vacuum Devices*—Kohl
- *Plasma Diagnostics*—Lochte-Holtgreven
- *Quadrupole Mass Spectrometry*—Dawson

Order your Classic today! Call 1-800-488-BOOK to order.

Vacuum Technology Resources for Only $15!

The AVS offers a variety of monographs, providing detailed information on vacuum safety, systems, operation, fundamentals, and related topics. These educational references give an in-depth tutorial on a single topical area. They range from 33–168 pages and cost $15 each.

Call the AVS at 212-248-0200 for a list of monographs.

Advertisers Direct Response

Journal of Vacuum Science & Technology B

Microelectronics and Nanometer Structures, Processing, Measurement, and Phenomena

◊JULY/AUGUST 1996 ◊

Listed below are the companies that have advertised in this issue. If you would like to know more about a product or service please complete the information below and fax this form to the appropriate company.

ADVERTISERS:

C4.	Adhesion Int'l. A Quad Group Fax: 509-458-4555	C3.	Leybold Vacuum Fax: 412-325-3557	A7.	Stanford Research Fax: 408-744-9049
A11.	EPI Fax: 612-653-0725	A3.	MDC Vacuum Parts Fax: 510-887-0626	A4.	Surface Interface Fax: 415-965-8207
A4.	Hiden Analytical Fax: 44 1925 416 518	A1. A9.	MKS Instruments Fax: 508-975-0093	A15.	Thermionics Northwest Fax: 360-385-6617
A15.	Instruments SA Fax: 908-549-5125	A13.	Omicron Fax: 412-831-9828	A13.	Vacuum Gauge & Test Fax: 916-273-1371
C2.	Leybold Inficon Fax: 315-437-3803	A15.	Princeton Research Fax: 609-924-4970		

FROM:

_____
Your Name

_____
Title

_____
Institution/Dept.

_____
Address

_____
City/State/Zip

_____
Phone / Fax

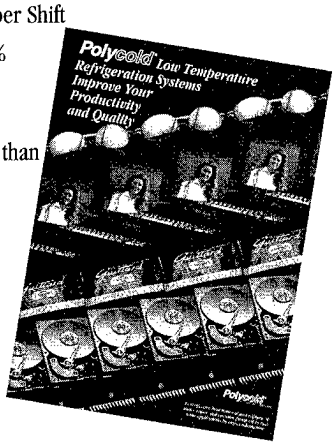
1. **Specialty:**
   a. □ Applied Surface Science
   b. □ Electronic Mat'l. & Processing
   c. □ Nanotechnology
   d. □ Plasma Science & Technology
   e. □ Surface Science
   f. □ Thin Film
   g. □ Vacuum Metallurgy
   h. □ Vacuum Technology
   i.□ Other: _____
2. **Need for information:**
   j. □ Immediate k. □ Future
3. **Purchasing Authority**
   l. □ Specialty m.□ Recommend

Achieve Faster Pumpdowns and Better Deposition With Polycold Cryogenics

Pump water vapor up to 100,000 l/s directly in the chamber to reach Base Pressure Faster

- ▶ Achieve More Vacuum Cycles per Shift
- ▶ Increase throughput up to 75%
- ▶ Improve Deposition Quality & Reproducibility
- ▶ Faster Response & Lower Cost than Liquid Nitrogen Meissners
- ▶ Easy to Install in Existing Vacuum Systems
- ▶ High Reliability
- ▶ CFC-FREE
- ▶ Available for Lease

Call to find out how to easily retrofit your system for improved performance and for our FREE Product Selection Guide



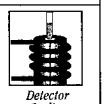
Semiconductor & Optical Deposition

Magnetic Disc/Flat Panel Display Deposition

Gas Chillers used for Testing materials

Cryotrapping Volatiles

Low Temperature Plasma Etch/Magnetic Disc Cooling Stations


Detector Cooling

Polycold® SYSTEMS INTERNATIONAL

67 Mark Dr. • San Rafael, CA 94903 • Fax: (415) 499-0927 • Phone: (415) 479-0577

INDEX TO ADVERTISERS

Adhesion International, A Quad Group Co. . . . . Cover 4
EPI A11
Hiden Analytical Ltd. A4
Instruments SA A15
Leybold Inficon, Inc. . . . . Cover 2
Leybold Vacuum Products, Inc. . . . . Cover 3
MDC Vacuum Products A3
MKS Instruments A9
Omicron A13
Polycold Systems A20
Princeton Research Instruments, Inc. A15
Stanford Research Systems A7
Surface Interface A4
Thermionics Northwest Inc. A15
Vacuum Gauge & Test A13

Advertising Sales Office

American Institute of Physics
500 Sunnyside Boulevard
Woodbury, NY 11797-2999
Telephone (516) 576-2440
Fax (516) 576-2481

Advertising Manager: Richard T. Kobel

Advtsg. Sales Representative: Robert G. Finnegan

Production Manager: Betty Aroesty

For Scanning Electron Microscopes, Surface Analysis, Research and Industrial Processing Industries

Design Flexibility With Application Proven Reliability

Low vibration lubricant-free magnetic bearing TURBOVAC Turbomolecular Pumps deliver reliable vacuum service. They provide trouble-free vacuum service while operating in any position. Convection cooling, ISO-K and CF inlet flanges are standard. The permanent magnetic bearings resist wear and require no maintenance. Converter and power supply require no battery backup. Accessories are available for standard and corrosive applications.

Leybold's Worldwide service assures you fast response.

Call Us For Specifics! 1-800-433-4021

ISO 9001

Innovative Vacuum Technology

LEYBOLD
VACUUM PRODUCTS, INC.

5700 Mellon Road • Export, PA 15632-8900

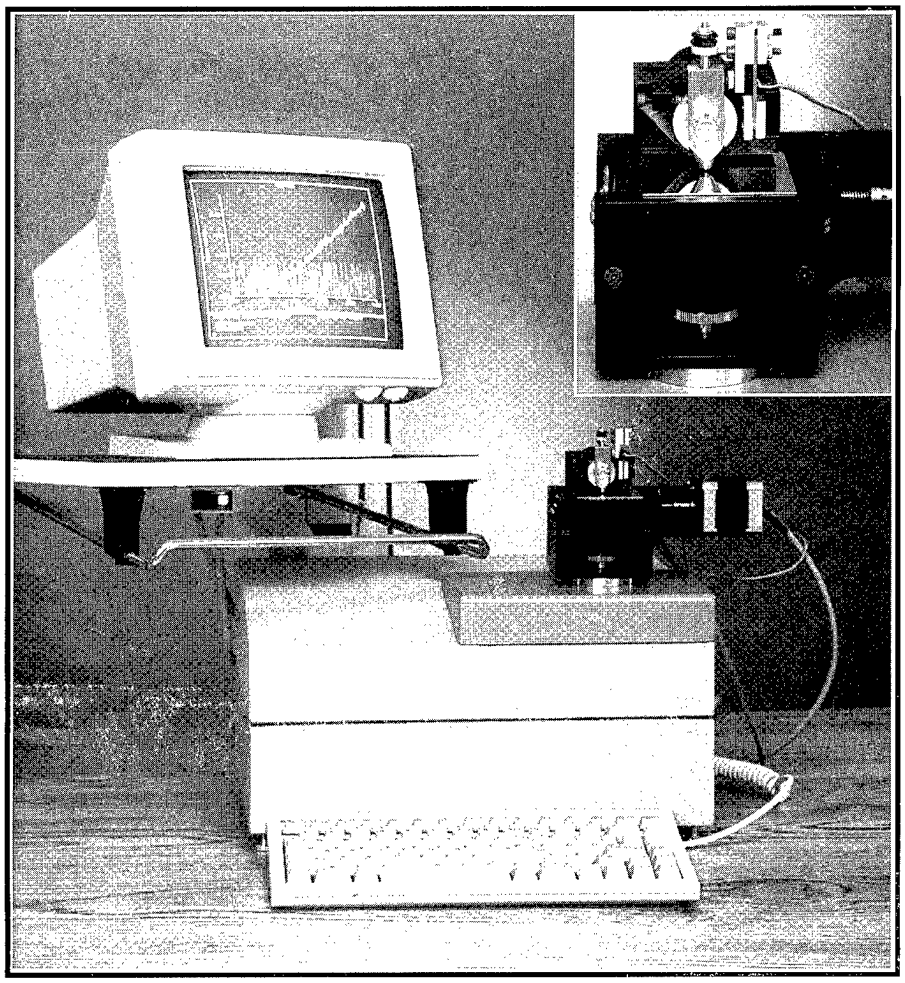
LEYBOLD

SCRATCH COATING ADHERENCE TESTER

The Adhesion International Stylometer is a software driven diamond scratch coating adherence tester. It is intended for the testing of thin films and is especially valuable in the testing of tribological or wear type coatings whose adherence normally exceed the measurement range of all other known adherence test instruments.

A spherical diamond is pressed against a coating surface at a constant rate of loading as the sample moves at a constant rate of travel. Force is increased until the coating catastrophically tears. The force at which tearing occurs, up to the substrate crushing force, is a measure of relative adherence, for samples of similar character and thickness. Typically, highly adherent coatings display micro-cracks normal to the diamond travel direction. These occur when the diamond drag force exceeds the tensile strength of the coating. As the force increases, the cracks become more tightly spaced until the coating tears or the substrate mechanically fails.

It is believed that the relative value of coating strength and elasticity will eventually be deduced for coatings of similar friction coefficient.



TOOLING

A breadbox sized enclosure contains an IBM compatible computer, high speed analog to digital converter, power supplies, signal processing, force application and control systems.

The scratch assembly, shown in the photo insert, contains the movable diamond stylus and the constant rate of travel sample stage.

An acoustic transducer, mounted on the stylus, is used to identify the point of initiation of micro-cracking and senses the coating or substrate failure. Acoustic energy, applied diamond force, the force required to move the sample and effective friction between the diamond and the coating are plotted as a function of travel distance.

ADHESION™ INTERNATIONAL
Test Equipment • Testing • Consulting

A Quad Group Company
1815 S. Lewis
Spokane, WA 99204
(509) 458-4556
FAX (509) 458-4555
Toll Free (800) 883-3481