Journal of Vacuum Science & Technology B

JVST B

Second Series
Volume 17, Number 4
July/August 1999

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

Papers from the 26th Conference on the Physics and Chemistry of Semiconductor Interfaces

20000201 042

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



A revolutionary ceramic diaphragm gives our new CDG superior performance—at lower cost.



Old-fashioned metal-diaphragm CDGs are about to get a wake-up call. Our new CDG uses a remarkable aluminum oxide ceramic diaphragm to produce measurements with superior accuracy and repeatability.

Even small temperature changes affect metal diaphragms, creating zero drift that can dramatically impact a CDG's accuracy. Alumina ceramic cuts temperature-induced zero drift by almost 90%.

Ceramic is also more rigid than metal, so it's not deformed by repeated pressure cycling or over-pressure. This further decreases zero drift and shortens the time needed for readings to stabilize during each process cycle, as well as eliminating the need for isolation valves. In addition, ceramic has nearly perfect corrosion resistance—so our CDG maintains its superior repeatability through every process. It is also fully compatible with existing installations.

You'd expect a CDG with such dramatically improved performance to cost more, but it actually costs *less*. And because it needs so little maintenance, it keeps on costing less.

If a CDG that delivers better repeatability, lower zero drift and lower cost of ownership sounds like your cup of tea, call us today.

Now with DeviceNet™!

LEYBOLD LEYBOLD INFICON

The Instrumental Difference™

United States
Phone: 315 434-1100
Fax: 315 437-3803

France
Phone: +33 01 69 824800
Fax: +33 01 69 075738

Netherlands
Phone: +31 0348 577411
Fax: +31 0347 420489

Korea
Phone: +82 02 270 1330
Fax: +82 02 277 5322

Germany
Phone: +49 221 347-0
Fax: +49 221 347-1250

United Kingdom
Phone: +44 0181 9717000
Fax: +44 0181 9717005

Japan
Phone: +81 045 471 3311
Fax: +81 045 471 3323

Taiwan
Phone: +886 3 5615156
Fax: +886 3 5618112

Visit our website for other sales offices worldwide. http://www.leyboldinficon.com

E-mail: reachus@inficon.com

™DeviceNet is a trademark of the Open DeviceNet Vendor Association, Inc.

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
	27 January 2000	Final 01 Oct 98 through 30 Sep 99

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

5. FUNDING NUMBERS
G: N00014-99-1-0145

6. AUTHOR(S)
C. R. Schulte

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, AZ 85252-0036

8. PERFORMING ORGANIZATION REPORT NUMBER
F1-1999

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
ATTN: ONR 251: Elizabeth L. Ford
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

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Twenty-sixth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces was held in San Diego, California, and the enclosed report contains abstracts of papers presented.

14. SUBJECT TERMS
Semiconductor Interfaces

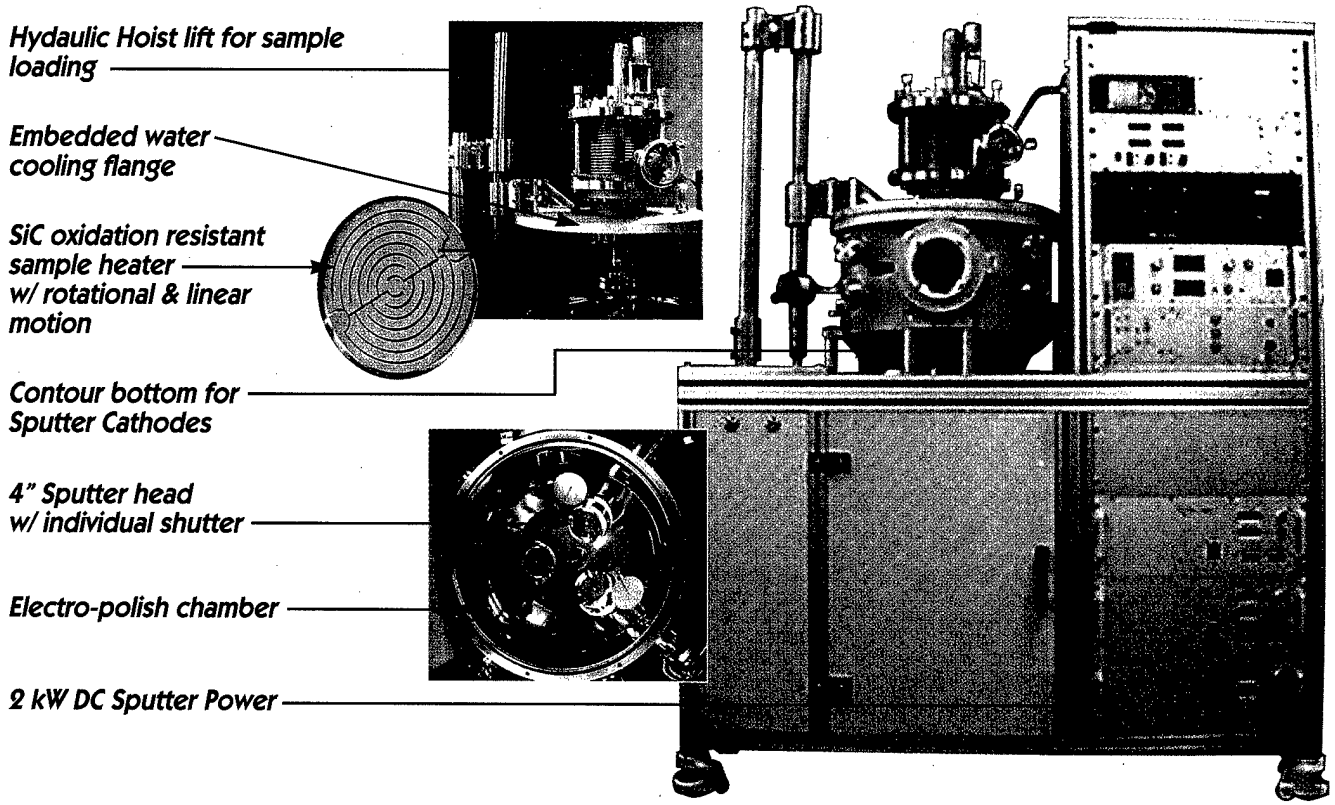
15. NUMBER OF PAGES
182

16. PRICE CODE

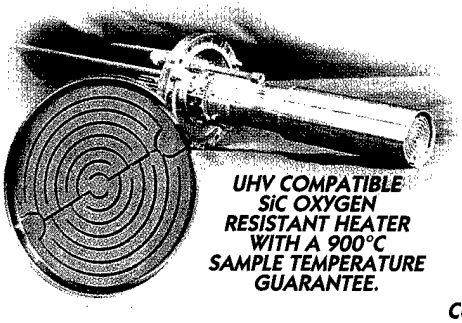
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

REPRODUCED AT GOVERNMENT EXPENSE

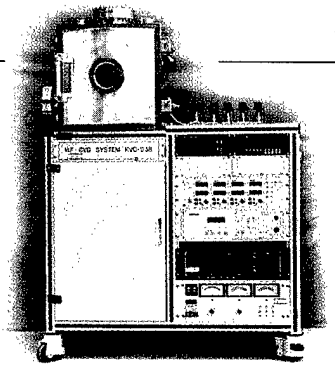
THE BEST NAME FOR TOP QUALITY VACUUM TECHNOLOGY AT THE BEST BOTTOM LINE COST.



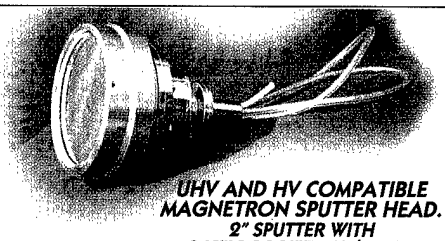
INNOVAC SPUTTER SYSTEM – Under $100k



UHV COMPATIBLE SiC OXYGEN RESISTANT HEATER WITH A 900°C SAMPLE TEMPERATURE GUARANTEE.



CUSTOM PVD & CVD DEPOSITION SYSTEMS.



UHV AND HV COMPATIBLE MAGNETRON SPUTTER HEAD. 2" SPUTTER WITH 2 KW DC POWER IS $5,500.



INNOVAC

INNOVATION IN VACUUM COMPONENTS & SYSTEMS

Call toll free
877-877-1556
Visit our web site
www.innovaccum.com

NORTH AMERICAN HEADQUARTERS: 50 Harrison St., Hoboken, NJ 07030, USA • Tel: 201-963-5450 • Fax: 201-963-5449

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,
MCNC, 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, Caller Box 13994, 10 Park Plaza, Ste. 4A,
Research Triangle Park, NC 27709, Telephone: (919) 361-2787 and (919) 361-2342;
FAX: (919) 361-1378; E-mail: jvst@jvst.org

Associate Editors:
Eric Kay, Review Articles
D. Bonnell (2001) Univ. of Pennsylvania
S. Felch (2000) Varian Associates
J. Givens (99) Micron Technology, Inc.
J. P. McVittie (2000) Stanford University
G. Owen (2000) Hewlett Packard Labs.
Y. Yasuda (2001) Nagoya University

JVST Publication Committee:
Dorota Temple, Chair, MCNC
John E. Crowell, UC San Diego
Joseph D. Geller, Geller Microanalytical Lab.
J. William Rogers, Jr., Univ. of Washington
Pete Sheldon, NREL

JVST Editorial Board:
J. C. Bean (2001) Univ. of Virginia
S. A. Chambers (99) Pacific NW National Lab.
A. Czanderna (2000) NREL
M. Kushner (2000) Univ. of Illinois
T. Mayer Sandia Natl. Labs
D. Monroe (2000) Lucent Bell Labs
D. Ruzic (99) Univ. of Illinois
P. A. Thiel (2001) Iowa State Univ.

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

The Journal of Vacuum Science & Technology B is published six times annually (Jan/Feb, Mar/Apr, May/Jun, Jul/Aug, Sep/Oct, Nov/Dec) by the American Vacuum Society (AVS) through the American Institute of Physics (AIP). 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 and 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 & Technology*, 10 Park Plaza, Caller Box 13994, Research Triangle Park, North Carolina 27709-3994; e-mail jvst@jvst.org. 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 Contrbutors," 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 & 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.

Electronic Physics Auxiliary Publication Service (EPAPS): For a nominal fee, authors may submit material that is supplemental to a paper. EPAPS deposits must be in electronic media, and can include text, figures, movies, computer programs, etc. Retrieval instructions are footnoted in the related published paper. Direct requests to the Editor; for additional information see http://www.aip.org/pubservs/paps.html

Advertising Rates will be supplied on request from AIP's Advertising Division, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502. 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 $15 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/99 $15.00.

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, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Telephone: (516) 576-2268; Fax: (516) 576-2327; Internet: rights@aip.org.

Copyright © 1999 American Vacuum Society. All rights reserved.

American Vacuum Society

Officers

Stephen Rossnagel, *President*
IBM T. J. Watson Research Center

Paula Grunthaner, *President-Elect*
Jet Propulsion Lab., Caltech

Jerry M. Woodall, *Immediate Past-President*
Yale University

Joseph E. Green, *Secretary*
University of Illinois

John W. Coburn, *Treasurer*
Univ. of CA, Berkeley

Directors

Cammy R. Abernathy
University of Florida

Yip-Wah Chung
Northwestern University

Elizabeth Dobisz
Naval Research Laboratory

Gregory J. Exarhos
Pacific Northwest National Lab.

Peter Sheldon
National Renewable Energy Laboratory

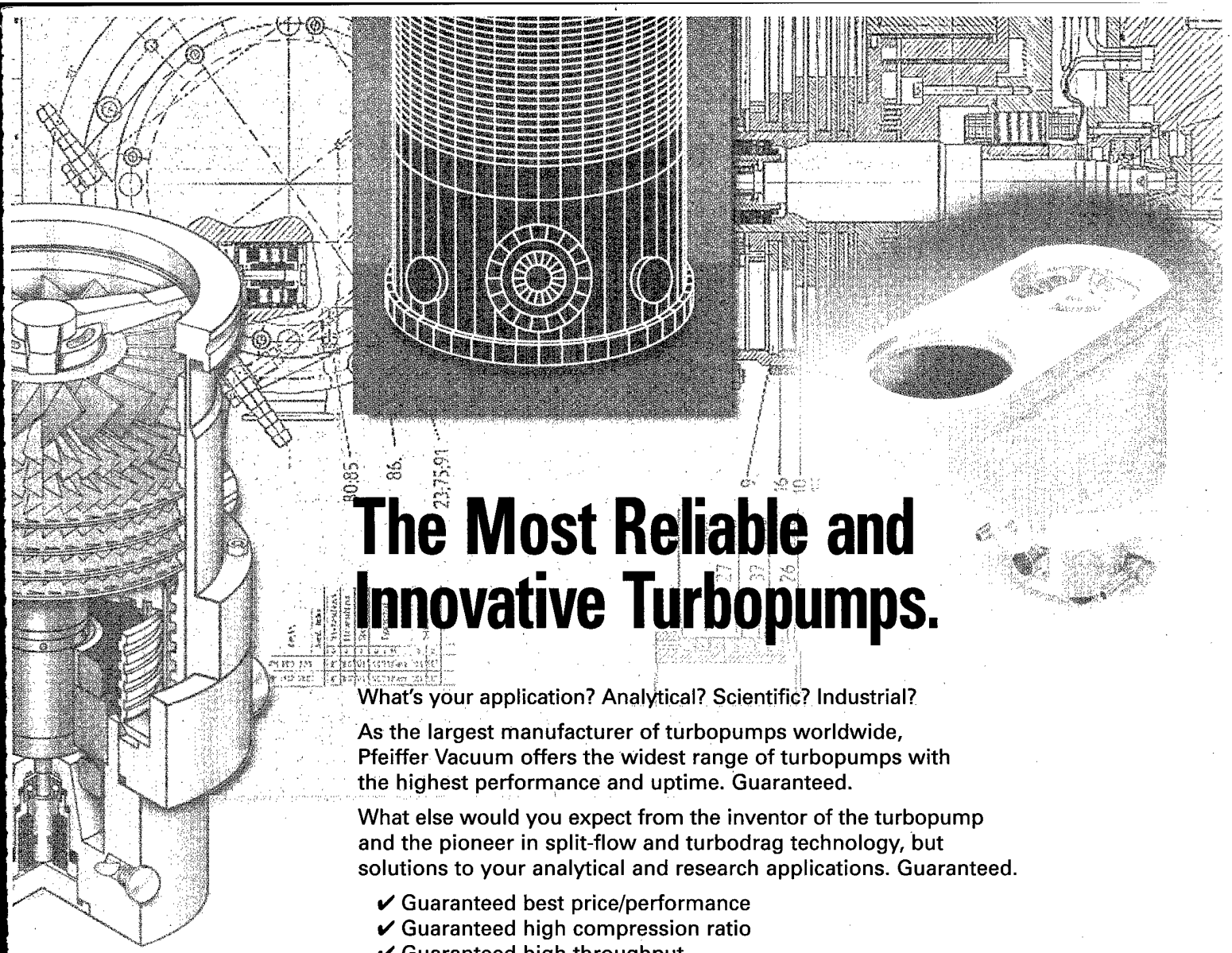
Peter M. A. Sherwood
Kansas State University

JVST

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

AVS Membership Information may be obtained from

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



The Most Reliable and Innovative Turbopumps.

What's your application? Analytical? Scientific? Industrial?

As the largest manufacturer of turbopumps worldwide, Pfeiffer Vacuum offers the widest range of turbopumps with the highest performance and uptime. Guaranteed.

What else would you expect from the inventor of the turbopump and the pioneer in split-flow and turbodrag technology, but solutions to your analytical and research applications. Guaranteed.

- ✔ Guaranteed best price/performance
- ✔ Guaranteed high compression ratio
- ✔ Guaranteed high throughput
- ✔ Guaranteed high tolerable foreline pressure
- ✔ Guaranteed lowest cost of ownership
- ✔ Guaranteed worldwide service
- ✔ Guaranteed application support

For total vacuum solutions, Pfeiffer Vacuum offers a complete range of rotary vane pumps as well as gauges, leak detectors and RGAs. Call Pfeiffer Vacuum for the turbodrag pump that's right for your application. Guaranteed.



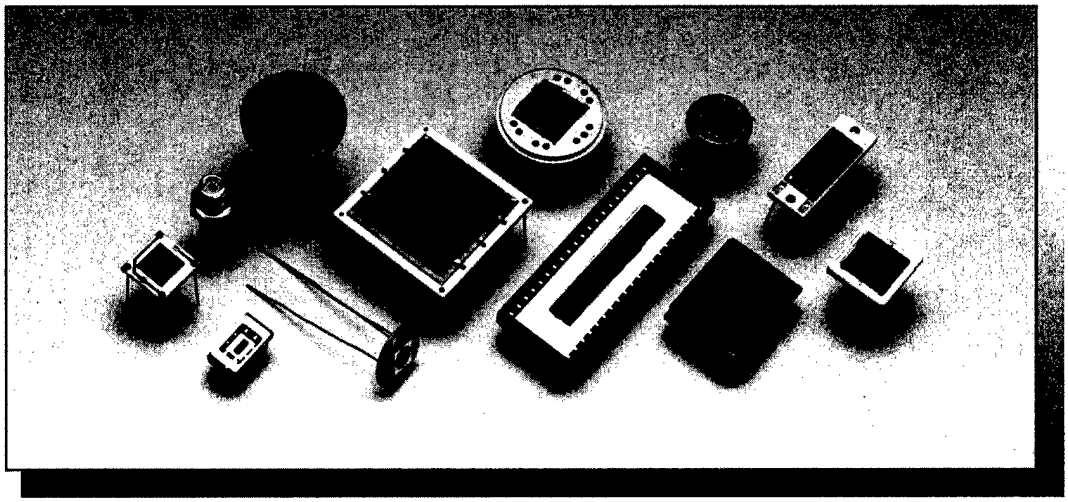
Circle to receive literature.
Circle to have a sales rep call.

PFEIFFER VACUUM
TECHNOLOGY INC.

24 Trafalgar Square, Nashua, NH 03063-1988
Tel: 603-578-6500 • Fax: 603-578-6550
http://www.pfeiffer-vacuum.com

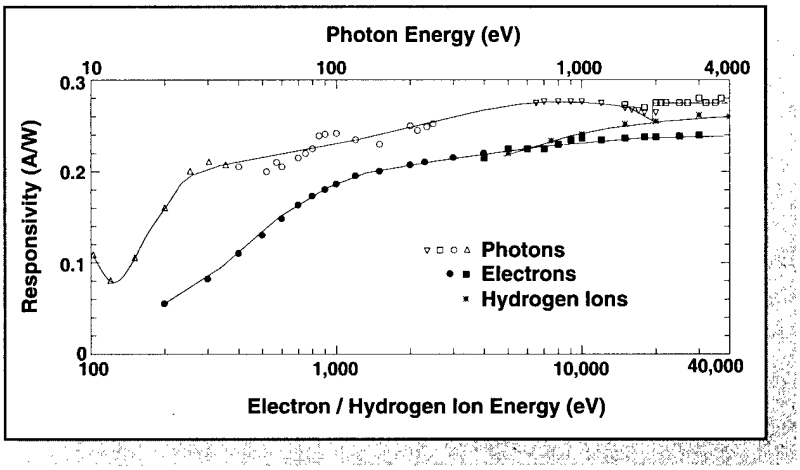
ABSOLUTE

XUV SILICON PHOTODIODES



Features

- 100% Internal Collection Efficiency
- 8 Decades of Linearity
- No Change in QE with 100 Mrad (Si) 124 eV Photons
- May be Operated Without Bias
- Cryogenically and UHV Compatible
- With Integrated Bandwidth Limiting Filters



TYPICAL RESPONSIVITY OF AXUV PHOTODIODES TO PHOTONS, ELECTRONS AND HYDROGEN IONS

SXUV Series Photodiodes are Available for Infinite Hardness to XUV Photons

INTERNATIONAL RADIATION DETECTORS INC.
2545 West 237th Street, Unit I ◆ Torrance, California 90505-5229
(310) 534-3661 ◆ FAX (310) 534-3665 ◆ email: irdinc@earthlink.net
http://www.ird-inc.com

IRD

DEVELOPED IN COLLABORATION WITH NIST, NIH, LLNL, LANL, NCAR

Journal of Vacuum Science & Technology B

JVST B

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

Second Series
Volume 17, Number 4
Jul/Aug 1999

Regular Articles

Application and calibration of a quartz needle sensor for high resolution scanning force microscopy
W. Clauss, J. Zhang, D. J. Bergeron, and A. T. Johnson **1309**

Surface structure characterization of DNA oligomer on Cu(111) surface using low temperature scanning tunneling microscopy
Chiho Hamai, Hiroyuki Tanaka, and Tomoji Kawai **1313**

Microstructure and electrical properties of Sb nanocrystals formed in thin, thermally grown SiO_2 layers by low-energy ion implantation
Anri Nakajima, Hiroshi Nakao, Hiroaki Ueno, Toshiro Futatsugi, and Naoki Yokoyama **1317**

Fabrication of a metal nanostructure on the Si(111) surface
D. Rogers and H. Nejoh **1323**

Mechanical properties, stress evolution and high-temperature thermal stability of nanolayered Mo–Si–N/SiC thin films
P. Torri, J.-P. Hirvonen, H. Kung, Y.-C. Lu, M. Nastasi, and P. N. Gibson **1329**

Characterization of bending in single crystal Si beams and resonators
J. W. Weigold, W. H. Juan, S. W. Pang, and J. T. Borenstein **1336**

(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) by the American Vacuum Society through the American Institute of Physics, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502. 1999 subscription rates are: US$763. POSTMASTER: Send address changes to *Journal of Vacuum Science & Technology B,* SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086. Periodicals postage paid at Thorofare, NJ 08086, and at additional mailing offices.

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.*

Subscription Prices (1999)

	U.S.A. and Poss.	Can., Mex., Central & S. America & Caribbean	Europe, Asia, Africa & Oceania*
JVST B†	$763	$793	$814
JVST B‡	$877	$905	$929
JVST A and B†	$877	$935	$977
JVST A and B#	$877	$905	$905
JVST A and B‡	$992	$1050	$1092

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

Back-number Prices: 1999 single copy: $135. Prior to 1999 single copies: $105.

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. Missing issue requests will be honored only if received within six months of publication date (nine months for Australia and Asia).

Single-copy orders (current and back issues) should be addressed to American Institute of Physics, Circulation and Fulfillment Division, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Telephone 800-344-6902 (or 516-576-2270 outside the U.S.A.), Fax at 516-349-9704, or E-mail at subs@aip.org.

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 and Fulfillment/ Reprints, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Fax: 516-349-9704; Telephone: 800-344-6909 (U.S. and Canada) or 516-576-2234.

Document Delivery: Copies of journal articles can be ordered for online delivery from the Articles in Physics online document delivery service (http://ojps.aip.org/jhtml/artinphys/).

Microform: *Journal of Vacuum Society & 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 and Fulfillment/ Single Copy Sales, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Fax: 516-349-9704; Telephone: 800-344-6908 (U.S. and Canada) or 516-576-2277.

Online Access: The *Journal of Vacuum Science and Technology B* is available online to AVS members at no additional charge; for details, please see http://www.aip.org/ojs. Abstracts of journal articles are available from AIP's SPIN Database, via AIP's Online Journal Publishing Service (OJPS) (http://ojps.aip.org).

In situ **ellipsometric study of the formation process of metalorganic vapor-phase epitaxy-grown quantum dots**
Jeong-Sik Lee, Shigeo Sugou, Hong-Wen Ren, and Yasuaki Masumoto **1341**

High temperature reaction of nitric oxide with Si surfaces: Formation of Si nanopillars through nitride masking and oxygen etching
K. Prabhakaran and T. Ogino **1346**

Low energy electron beam decomposition of metalorganic precursors with a scanning tunneling microscope at ambient atmosphere
H. Brückl, J. Kretz, H. W. Koops, and G. Reiss **1350**

Two-dimensional dopant profiling of patterned Si wafers using phase imaging tapping mode atomic force microscopy with applied biases
M. W. Nelson, P. G. Schroeder, R. Schlaf, and B. A. Parkinson **1354**

Fabrication of a microcavity structure with a polyimide thin film prepared by vacuum deposition polymerization
Y. Sakakibara and T. Tani **1361**

Low voltage electron beam lithography in PMMA
A. Olkhovets and H. G. Craighead **1366**

Surface roughness development during photoresist dissolution
Lewis W. Flanagin, Vivek K. Singh, and C. Grant Willson **1371**

Relations between the solubility speed and the electrical conductivity of phenol novolak polymer solutions
T. Takeda and M. Saka **1380**

Low-stress W/Cr films for SCALPEL® mask scattering layers
David L. Windt **1385**

Microelectron gun integrating a point-source cathode
Zoubida Hammadi, Michel Gauch, and Roger Morin **1390**

Low brightness and high emittance electron gun for a reducing image projection system
Mamoru Nakasuji and Hiroyasu Shimizu **1395**

Generalization of electrostatic lens characteristics using the Picht ray trajectories
Yoshimi Kawanami and Tohru Ishitani **1400**

X-ray photoelectron spectroscopy analyses of oxide-masked organic polymers etched in high density plasmas using SO_2/O_2 gas mixtures
C. Monget and O. Joubert **1406**

Fabrication and characterization of chromium based single-electron transistors with evaporated chromium oxide barrier tunnel junctions
Yu. A. Pashkin, J. P. Pekola, and L. S. Kuzmin **1413**

Process sensing and metrology in gate oxide growth by rapid thermal chemical vapor deposition from SiH_4 and N_2O
Guangquan Lu, Laura L. Tedder, and Gary W. Rubloff **1417**

Relaxation of strained Si layers grown on SiGe buffers
S. B. Samavedam, W. J. Taylor, J. M. Grant, J. A. Smith, P. J. Tobin, A. Dip, A. M. Phillips, and R. Liu **1424**

High density plasma deposition of device quality silicon nitride. II. Effects of thickness on the electrical properties
M. C. Hugon, F. Delmotte, B. Agius, and E. A. Irene **1430**

Characterization of Al, Cu, and TiN surface cleaning following a low-*K* dielectric etch
P. J. Matsuo, T. E. F. M. Standaert, S. D. Allen, G. S. Oehrlein, and T. J. Dalton **1435**

Reaction and thermal stability of cobalt disilicide on polysilicon resulting from a Si/Ti/Co multilayer system
A. Alberti, F. La Via, and F. Rimini **1448**

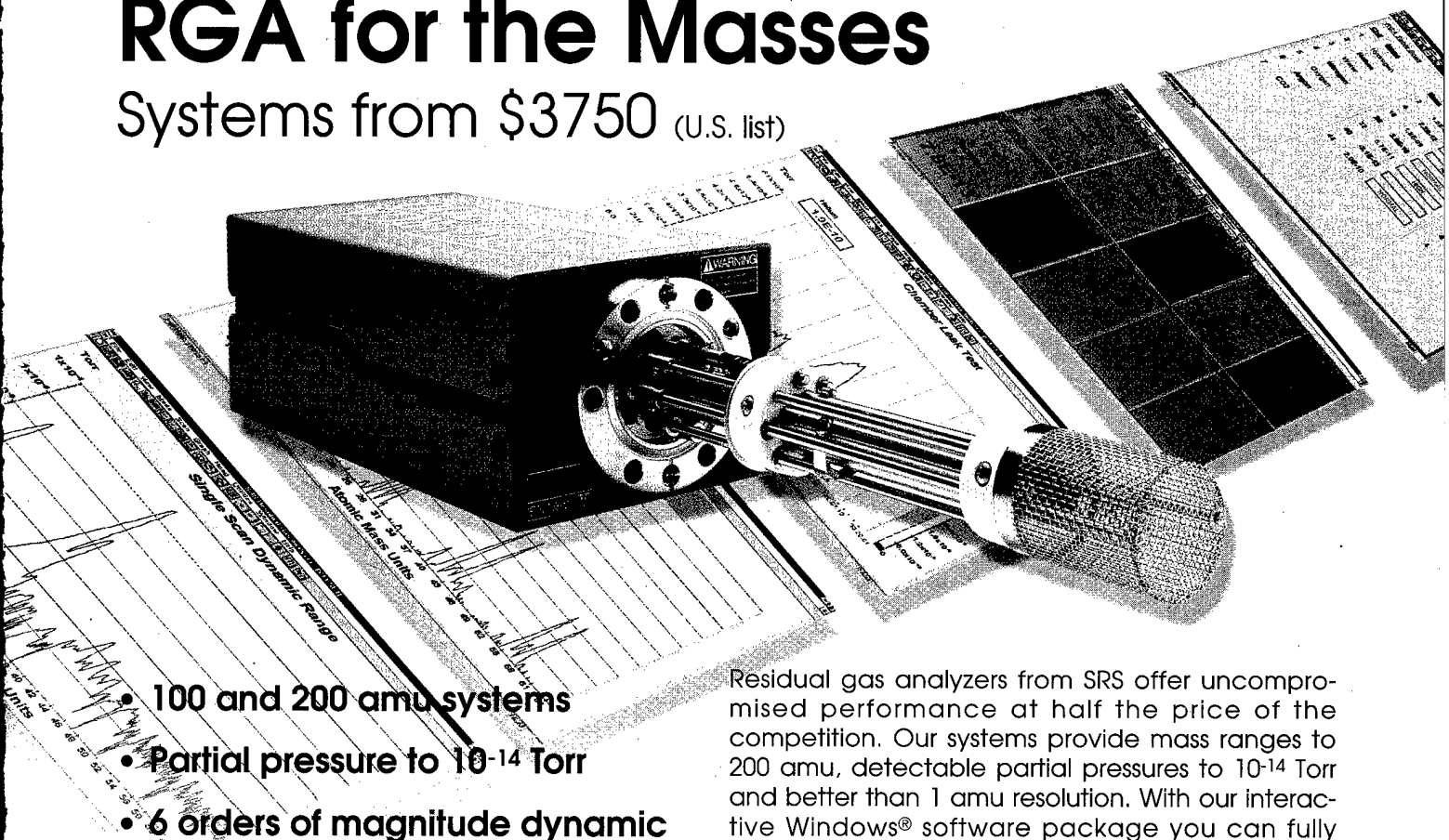
Impact of plasma-enhanced chemical vapor deposited oxide characteristics on interconnect via resistance and device performance of four-transistor static random access memory with polysilicon load resistors
C.-F. Lin, W. T. Tseng, M. S. Feng, and Y. L. Wang **1456**

Study of crystal orientation in Cu film on TiN layered structures
Kazuhide Abe, Yusuke Harada, and Hiroshi Onoda **1464**

(Continued)

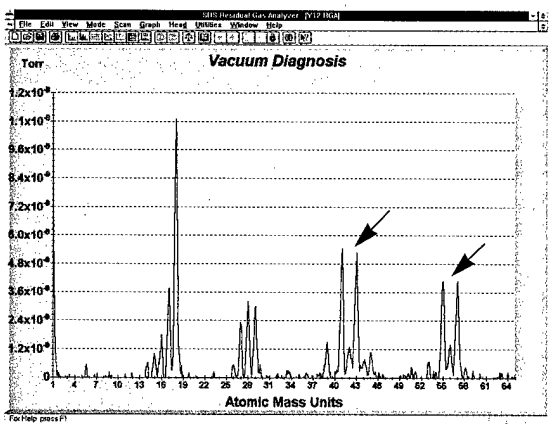
RGA for the Masses

Systems from $3750 (U.S. list)

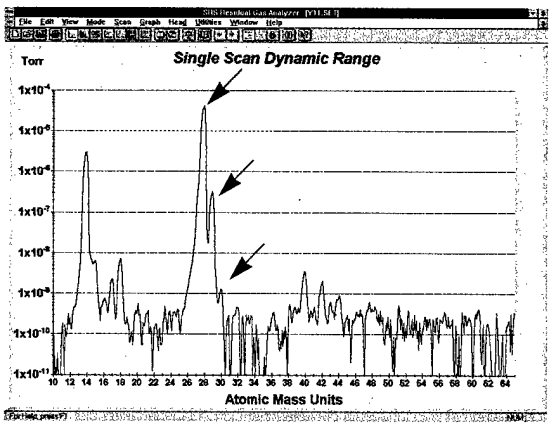


- **100 and 200 amu systems**
- **Partial pressure to 10^{-14} Torr**
- **6 orders of magnitude dynamic range in a single scan**
- **Real-time Windows® software**
- **Mass spectra, leak detection and pressure vs. time modes**
- **Multi-head operation**
- **Optional Electron Multiplier**

Residual gas analyzers from SRS offer uncompromised performance at half the price of the competition. Our systems provide mass ranges to 200 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@thinkSRS.com • WWW: www.thinkSRS.com

Windows is a registered trademark of Microsoft Corporation

Barrier properties of Ta–RuO_2 diffusion barrier for dynamic random access memory capacitor bottom electrodes
Dong-Soo Yoon, Hong Koo Baik, Sung-Man Lee, and Sang-In Lee 1470

Superlatticed negative differential-resistance heterojunction bipolar transistor
Shiou-Ying Cheng, Jung-Hui Tsai, Po-Hung Lin, and Wen-Chau Liu 1477

High-quality InAlAs layers grown on (411)A-oriented InP substrates by molecular beam epitaxy
Takahiro Kitada, Keisuke Nii, Tetsuya Hiraoka, Satoshi Shimomura, and Satoshi Hiyamizu 1482

Evolution of microstructure and dislocation dynamics in $In_xGa_{1-x}P$ graded buffers grown on GaP by metalorganic vapor phase epitaxy: Engineering device-quality substrate materials
Andrew Y. Kim, Wendy S. McCullough, and Eugene A. Fitzgerald 1485

Strain in AlGaN layer studied by Rutherford backscattering/channeling and x-ray diffraction
M. F. Wu, Shude Yao, A. Vantomme, S. M. Hogg, G. Langouche, J. Li, and G. Y. Zhang 1502

Blueshift of $In_{0.2}Ga_{0.8}N$/GaN single quantum well band gap energy by rapid thermal annealing
G. Li, S. J. Chua, J. H. Teng, W. Wang, Z. C. Feng, Y. H. Huang, and T. Osipowicz 1507

New self-aligned processes for III–V electronic high speed devices
J. Etrillard, C. Besombes, L. Bricard, and A. M. Duchenois 1510

Effect of short-time helicon-wave excited N_2–Ar plasma treatment on the interface characteristic of GaAs
Satoshi Wada, Keisuke Kanazawa, Nariaki Okamoto, and Hideaki Ikoma 1516

Nitridation of GaAs(001) surface: Auger electron spectroscopy and reflection high-energy electron diffraction
Igor Aksenov, Hidekazu Iwai, Yoshinobu Nakada, and Hajime Okumura 1525

Inductively coupled plasma damage in GaN Schottky diodes
X. A. Cao, A. P. Zhang, G. T. Dang, H. Cho, F. Ren, S. J. Pearton, R. J. Shul, L. Zhang, R. Hickman, and J. M. Van Hove 1540

Effect of rapid thermal annealing on radio-frequency magnetron-sputtered GaN thin films and Au/GaN Schottky diodes
Ching-Wu Wang, Jin-Yuan Liao, Chih-Liang Chen, Wen-Kuan Lin, Yan-Kuin Su, and Meiso Yokoyama 1545

Plasma assisted two stage selenization process for the preparation of low resistivity ZnSe films
R. Jeyakumar, G. K. Chadda, S. T. Lakshmikumar, and A. C. Rastogi 1549

Arrays of ungated GaAs field emitters fabricated by wet or dry etching
F. Ducroquet, P. Kropfeld, O. Yaradou, and A. Vanoverschelde 1553

Novel lateral field emission device fabricated on silicon-on-insulator material
Minhee Yun, Anthony Turner, Ronald J. Roedel, and Michael N. Kozicki 1561

Field emission properties of diamondlike carbon films deposited by ion beam assisted deposition
Jing Wang, Wen-Zhi Li, and Heng-De Li 1567

High aspect ratio all diamond tips formed by focused ion beam for conducting atomic force microscopy
A. Olbrich, B. Ebersberger, C. Boit, Ph. Niedermann, W. Hänni, J. Vancea, and H. Hoffmann 1570

Sensitivity analysis of the field emitter
Lei Wei, Wang Baoping, Yin Hanchun, and Guohui Li 1575

High field characteristics of dielectric spacers in thin-film electrode vacuum gaps
Xianyun Ma and T. S. Sudarshan 1580

Brief Reports and Comments

Mechanism of highly preferred (002) texture of Ti films sputter deposited on water-absorbed borophosphosilicate glass films
Tomoyuki Yoshida, Koyu Aoki, and Yasuichi Mitsushima 1585

Microfabrication and testing of suspended structures compatible with silicon-on-insulator technology
A. A. Ayón, K. Ishihara, R. A. Braff, H. H. Sawin, and M. A. Schmidt 1589

Rapid Communications

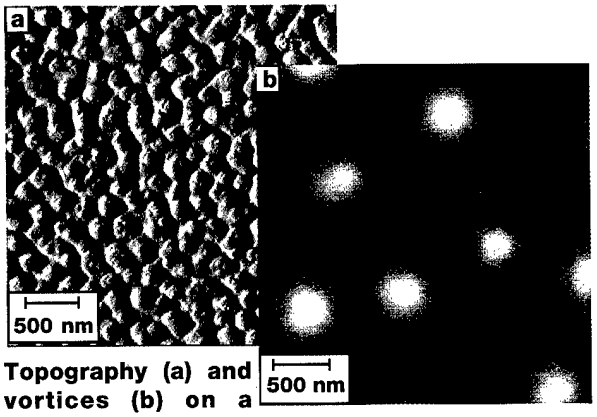
Low-energy electron-beam lithography using calixarene
A. Tilke, M. Vogel, F. Simmel, A. Kriele, R. H. Blick, H. Lorenz, D. A. Wharam, and J. P. Kotthaus 1594

New method to prepare W–B^+–N ternary barrier to Cu diffusion by implanting BF_2^+ ions into W–N thin film
Dong Joon Kim, Yong Tae Kim, and Jong-Wan Park 1598

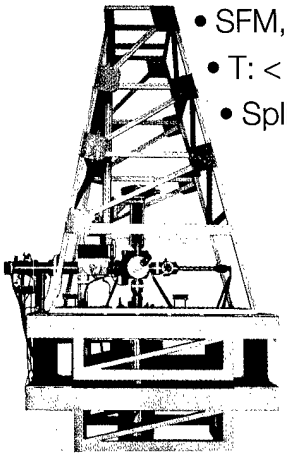
(Continued)

CRYOGENIC SFM

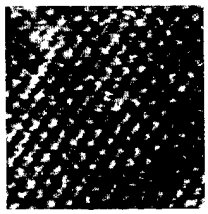
Low Temperature Scanning Force Microscopy in Magnetic Fields



Topography (a) and vortices (b) on a superconducting $YBa_2Cu_3O_{7-x}$ film at T = 7.6 K.



- SFM, MFM, STM in UHV
- T: < 9 K - 300 K
- Split-Coil 8-Tesla Magnet



Contact mode AFM on NaCl at 7.8 K in UHV,
(3.9 nm x 3.9 nm).

H. J. Hug, B. Stiefel, A. Moser, P. J. A. van Schendel, O. Knauff, S. Martin and H.-J. Güntherodt, University of Basel

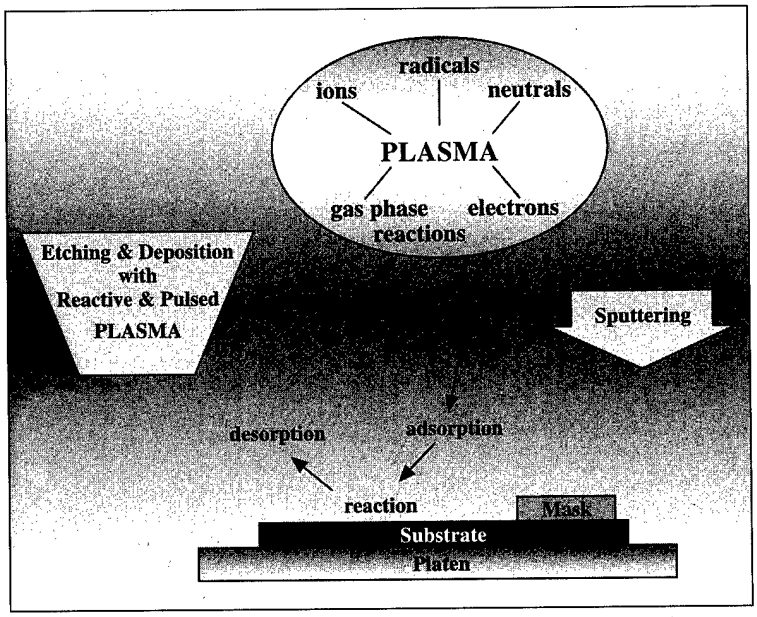
OMICRON
INSTRUMENTS FOR SURFACE SCIENCE

Headquarters:
OMICRON VAKUUMPHYSIK GmbH
Idsteiner Str. 78 • D-65232 Taunusstein • Germany
Tel. +49 (0) 6128 9 87 - 0 • Fax +49 (0) 6128 9 87 - 185

USA:
OMICRON Associates
1226 Stoltz Road • Bethel Park, PA 15102
Tel. (412) 831-2262 • Fax (412) 831-9828

www.omicron.de or www.omicron-instruments.com

For Complete Characterisation of Plasma Processes ...



... Hiden's EQP

In today's demanding world of plasma research, process transfer and process optimisation, our desire for improved understanding of plasma chemistry has never been greater.
Hiden Analytical's universally acclaimed **EQP** provides the answers. Complete gas, ion and radical characterisation at the wafer level and throughout the reactor.
For high performance **mass** and **energy** analysis of:

- **Positive Ions**
- **Neutrals**
- **Negative Ions**
- **Radicals**

in **deposition** and **etching** with **reactive** and **pulsed plasma,** contact **Hiden Analytical.**

Hiden Analytical manufacture an extensive range of quadrupole mass spectrometers for vacuum diagnostics, process gas analysis and end point detection. All operate under the control of Hiden's Windows™ MASsoft PC for fast, flexible and user-friendly data acquisition and management.
For further details, visit our Web site at: **http://www.hiden.co.uk**
or e-mail: **info@hiden.demon.co.uk**

HIDEN ANALYTICAL
Excellence in Plasma Diagnostics

Corporate Office and Service Center: Hiden Analytical Inc., 75 Hancock Road, Suite D, Peterborough, NH 03458-1100. Tel:1-888 96 HIDEN
European Headquarters: Hiden Analytical Ltd., 420 Europa Boulevard, Warrington WA5 5UN, England. Tel: +44 (0)1925 445225 Fax: +44 (0)1925 416518

Growth of a near-atomic protrusion on molybdenum field emitter tips under argon ion bombardment
F. Okuyama, H. Sugie, and M. Sato **1602**

Shop Notes

Multilayer resist films applicable to nanopatterning of insulating substrates based on current-injecting scanning probe lithography
Hiroyuki Sugimura, Osamu Takai, and Nobuyuki Nakagiri **1605**

Papers from the 26th Conference on the Physics and Chemistry of Semiconductor Interfaces

Preface **1611**

Dielectrics, Quantum Dots, and Optoelectronic Interfaces

Deliberately designed interfaces for monolithic integration in optoelectronics
Tairan Wang, N. Moll, Kyeongjae Cho, and J. D. Joannopoulos **1612**

Pyroelectronics: Novel device concepts based on nitride interfaces
G. Zandler, J. A. Majewski, and P. Vogl **1617**

Selective area chemical vapor deposition of titanium oxide films: Characterization of $Ti(OC_3H_7)_4$ as an electron beam resist
William J. Mitchell and Evelyn L. Hu **1622**

Reduction and creation of paramagnetic centers on surfaces of three different polytypes of SiC
P. J. Macfarlane and M. E. Zvanut **1627**

Finite linewidth observed in photoluminescence spectra of individual $In_{0.4}Ga_{0.6}As$ quantum dots
J. L. Spithoven, J. Lorbacher, I. Manke, F. Heinrichsdorff, A. Krost, D. Bimberg, and M. Dähne-Prietsch **1632**

Atomic structure of stacked InAs quantum dots grown by metal-organic chemical vapor deposition
O. Flebbe, H. Eisele, T. Kalka, F. Heinrichsdorff, A. Krost, D. Bimberg, and M. Dähne-Prietsch **1639**

Quantum dot-like behavior of GaInNAs in GaInNAs/GaAs quantum wells grown by gas-source molecular-beam epitaxy
H. P. Xin, K. L. Kavanagh, Z. Q. Zhu, and C. W. Tu **1649**

Nitride Epitaxy, Photonic Probes, and Band Offsets

Influence of active nitrogen species on high temperature limitations for (000$\bar{1}$) GaN growth by rf plasma-assisted molecular beam epitaxy
T. H. Myers, M. R. Millecchia, A. J. Ptak, K. S. Ziemer, and C. D. Stinespring **1654**

Nitridation of the GaAs(001) surface: Thermal behavior of the (3×3) reconstruction and its evolution
J. Lu, L. Haworth, P. Hill, D. I. Westwood, and J. E. Macdonald **1659**

Use of ultrathin ZnSe dipole layers for band offset engineering at Ge and Si homo/heterojunctions
S. P. Wilks, R. H. Williams, M. Pan, P. R. Dunstan, and B. C. C. Cowie **1666**

AlN and GaN epitaxial heterojunctions on 6H–SiC(0001): Valence band offsets and polarization fields
Angela Rizzi, Roberta Lantier, Fulvia Monti, Hans Lüth, Fabio Della Sala, Aldo Di Carlo, and Paolo Lugli **1674**

Epitaxial growth and electronic structure of lanthanide silicides on *n*-type Si(111)
S. Vandré, T. Kalka, C. Preinesberger, and M. Dähne-Prietsch **1682**

GaP(001) and InP(001): Reflectance anisotropy and surface geometry
N. Esser, W. G. Schmidt, J. Bernholc, A. M. Frisch, P. Vogt, M. Zorn, M. Pristovsek, W. Richter, F. Bechstedt, Th. Hannappel, and S. Visbeck **1691**

***In situ* reflectance difference spectroscopy of II–VI compounds: A real time study of N plasma doping during molecular beam epitaxy**
D. Stifter, M. Schmid, K. Hingerl, A. Bonanni, M. Garcia-Rocha, and H. Sitter **1697**

Magnetic Interfaces and Photon Scattering

Specular electron scattering in metallic thin films
W. F. Egelhoff, Jr., P. J. Chen, C. J. Powell, D. Parks, G. Serpa, R. D. McMichael, D. Martien, and A. E. Berkowitz **1702**

Spin relaxation of conduction electrons
J. Fabian and S. Das Sarma **1708**

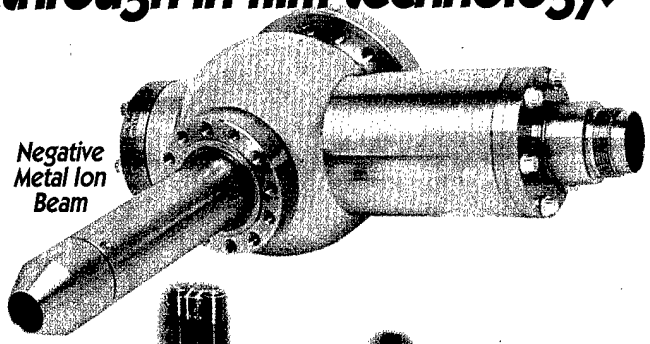
Kinetics of MnAs growth on GaAs(001) and interface structure
F. Schippan, A. Trampert, L. Däweritz, and K. H. Ploog **1716**

(Continued)

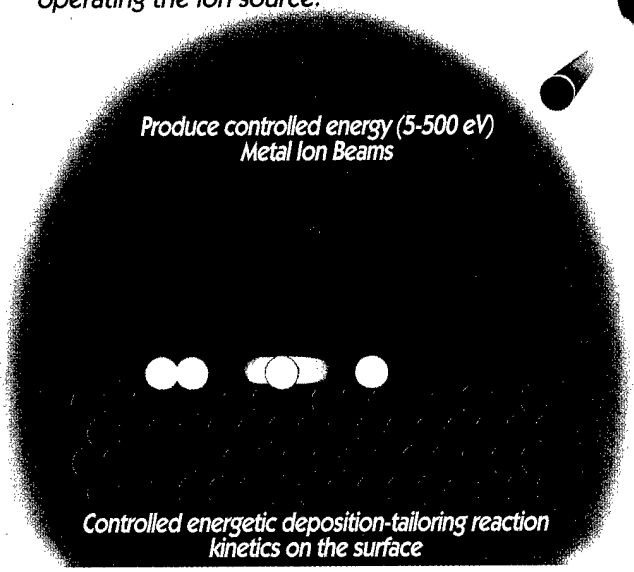
Discover the newest breakthrough in film technology.

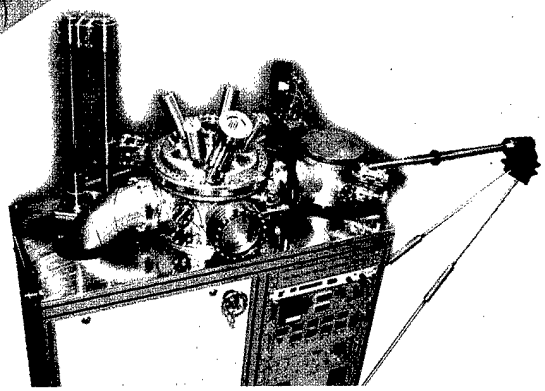
A REVOLUTION IN THIN FILM

The Ion Source produces a variety of metal ion beams (C, Si, Au, Cr, Cu, W, B, etc.) by sputtering (1" diameter target is used).
Metal Ion Beam current of 10-100 μA depending on the target materials.
UHV compatible: maintain ≤ 10^{9} Torr while operating the ion source.



Negative Metal Ion Beam

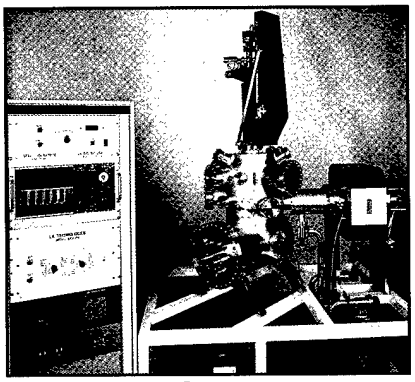




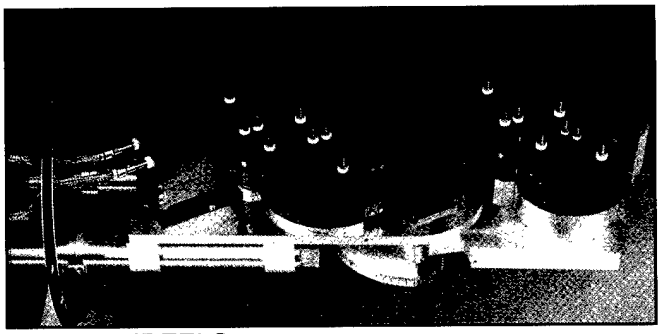
Metal Ion Beam Epitaxy System MIBE-600
The system starts at $250,000

SKION CORPORATION
50 Harrison St., Hoboken, NJ 07030
Tel: 201-963-5450 Fax: 201-963-5449
E-mail: info@skion.com

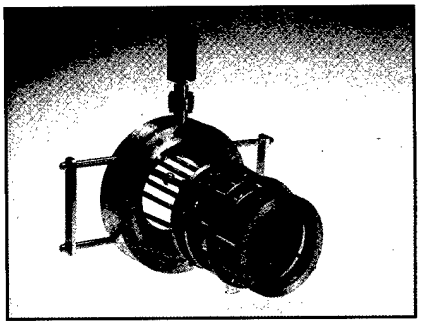
Surface Analysis Components and Systems Since 1985



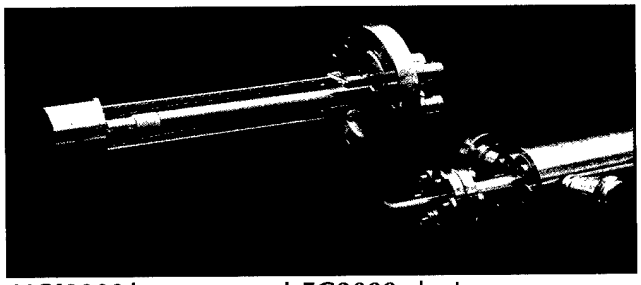
Custom UHV Systems



ELS3000 HREELS



LEED/Auger



NGI3000 ion guns and ***EG3000*** electron guns

LK TECHNOLOGIES

LK Technologies, Inc.
3910 Roll Avenue, Bloomington, IN 47403
Tel (812) 332-4449
Fax (812) 332-4493
http: //www.lktech.com
e-mail: lktech@lktech.com

Reflectance difference spectroscopy and magneto-optical analysis of digital magnetic heterostructures
A. Bonanni, G. Prechtl, W. Heiss, F. Schinagl, S. Holl, H. Krenn, H. Sitter, D. Stifter, and K. Hingerl **1722**

Coherent soft x-ray scattering from InP islands on a semiconductor substrate
M. Adamcyk, C. Nicoll, T. Pinnington, T. Tiedje, S. Eisebitt, A. Karl, R. Scherer, and W. Eberhardt..... **1728**

Infrared study of Si surfaces and buried interfaces
A. Milekhin, M. Friedrich, K. Hiller, M. Wiemer, T. Gessner, and D. R. T. Zahn **1733**

Infrared and Raman studies of confined and interface optical phonons in short-period GaAs/AlAs superlattices with a grating coupler
A. Milekhin, M. Rösch, E. Batke, D. R. T. Zahn, K. Köhler, P. Ganser, V. Preobrazhenskii, and B. Semyagin..... **1738**

Piezoelectric Effects, Minigaps, Defect Centers, and Devices

Spontaneous and piezoelectric polarization effects in III–V nitride heterostructures
E. T. Yu, X. Z. Dang, P. M. Asbeck, S. S. Lau, and G. J. Sullivan..... **1742**

Electric force microscopy of induced charges and surface potentials in GaN modified by light and strain
P. M. Bridger, Z. Z. Bandić, E. C. Piquette, and T. C. McGill..... **1750**

Piezoelectric fields in nitride devices
R. A. Beach and T. C. McGill **1753**

Minigaps in strained silicon quantum wells on tilted substrates
T. J. Thornton, F. Ge, A. Andresen, D. Pivin, J. Bird, and D. K. Ferry..... **1757**

***DX* centers in $Al_{0.37}Ga_{0.63}As/GaAs$ and $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ heterostructures**
H. H. Wieder and Huseyin Sari..... **1761**

Behavior of a new ordered structural dopant source in InAs/(001) GaP heterostructures
V. Gopal, E.-H. Chen, E. P. Kvam, and J. M. Woodall **1767**

Nanoelectronic device applications of a chemically stable GaAs structure
D. B. Janes, V. R. Kolagunta, M. Batistuta, B. L. Walsh, R. P. Andres, Jia Liu, J. Dicke, J. Lauterbach, T. Pletcher, E. H. Chen, M. R. Melloch, E. L. Peckham, H. J. Ueng, J. M. Woodall, Takhee Lee, R. Reifenberger, C. P. Kubiak, and B. Kasibhatla **1773**

Superlattices, III–V Growth, and Characterization

Role of As_4 in Ga diffusion on the GaAs(001)-(2×4) surface: A molecular beam epitaxy-scanning tunneling microscopy study
H. Yang, V. P. LaBella, D. W. Bullock, and P. M. Thibado..... **1778**

Cross-sectional scanning tunneling microscopy of InAsSb/InAsP superlattices
S. L. Zuo, E. T. Yu, A. A. Allerman, and R. M. Biefeld **1781**

Characterization of AlSb/InAs surfaces and resonant tunneling devices
B. Z. Nosho, W. H. Weinberg, W. Barvosa-Carter, A. S. Bracker, R. Magno, B. R. Bennett, J. C. Culbertson, B. V. Shanabrook, and L. J. Whitman..... **1786**

Analysis of buried (Al,Ga)As interfaces after molecular-beam epitaxy overgrowth
M. Wassermeier, R. Hey, M. Höricke, E. Wiebicke, and H. Kostial..... **1791**

Si/SiO_2 and $Si–Si_3N_4$ Interfaces: Defect Properties and Novel Structures

Mechanistic studies of silicon oxidation
M. K. Weldon, K. T. Queeney, Y. J. Chabal, B. B. Stefanov, and K. Raghavachari **1795**

Constraint theory and defect densities at (nanometer SiO_2-based dielectric)/Si interfaces
J. C. Phillips **1803**

Bonding constraint-induced defect formation at Si-dielectric interfaces and internal interfaces in dual-layer gate dielectrics
G. Lucovsky, Y. Wu, H. Niimi, V. Misra, and J. C. Phillips **1806**

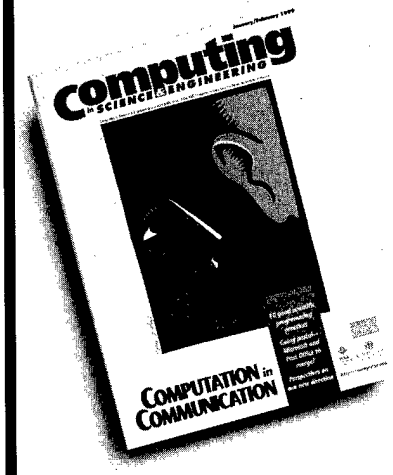
Suppression of boron transport out of p^+ polycrystalline silicon at polycrystalline silicon dielectric interfaces
Y. Wu, H. Niimi, H. Yang, G. Lucovsky, and R. B. Fair..... **1813**

Energy-dependent conduction band mass of SiO_2 determined by ballistic electron emission microscopy
R. Ludeke and Andreas Schenk..... **1823**

Band offsets for ultrathin SiO_2 and Si_3N_4 films on Si(111) and Si(100) from photoemission spectroscopy
J. W. Keister, J. E. Rowe, J. J. Kolodziej, H. Niimi, T. E. Madey, and G. Lucovsky **1831**

(Continued)

A New Joint Publication from AIP and IEEE Computer Society!



An exciting new bimonthly publication, ***Computing in Science and Engineering,*** has grown out of a merger of AIP's ***Computers in Physics*** and ***IEEE Computational Science & Engineering***, published by the IEEE Computer Society. This full-color magazine continues the best of ***CIP***: your favorite in-depth departments—*Scientific Programming, Computer Simulations, Computers in Education*, and more. To these, it adds the best of ***CS&E's*** contents: peer-reviewed theme issues on environmental modeling, computational cosmology, and other vital topics.

Every issue includes theme articles, departments, news reports, and editorial comment. Collateral materials such as source code will be made available electronically over the Internet.

IEEE COMPUTER SOCIETY

AMERICAN INSTITUTE OF PHYSICS

To subscribe, call 800-344-6902 or 516-576-2270. You can also e-mail subs@aip.org.

9 out of 10 Perfectionists Insist On VLSI's Resistivity Standards.



You're a professional. You want your metrology to be right on the money.

And VLSI can help. With its new family of larger and more convenient resistivity standards...

They're now on 200mm wafers to run more easily in today's equipment.



Available resistivity values span six orders of magnitude. And uncertainty is less than 1%!

To eliminate confusion, both the nominal value and measurement area are clearly marked.

The standards are NIST-traceable, of course, and are usable on 4-point probes and non-contact instruments.

For full information, just call. And also get a free "Good Enough ISN'T" button to flaunt your perfectionism.

VLSI Standards: (800) 228-8574. Or on the Internet: www.vlsistd.com

VLSI Standards Incorporated

The Measurement Standards for the Industry.

Interfacial properties of ultrathin pure silicon nitride formed by remote plasma enhanced chemical vapor deposition
V. Misra, H. Lazar, Z. Wang, Y. Wu, H. Niimi, G. Lucovsky, J. J. Wortman, and J. R. Hauser **1836**

Remote plasma enhanced chemical vapor deposition SiO_2 in silicon based nanostructures
M. J. Rack, L. L. Hilt, D. Vasileska, and D. K. Ferry **1840**

Self-organization in $Si/CoSi_2(111)$ heteroepitaxy
T. Meyer, M. Klemenc, T. Graf, and H. von Känel **1848**

Novel Growth, Nano-Machining, and Characterization

Comparison of nanomachined III–V semiconductor substrates
L. Grazulis, D. L. Kelly, D. E. Walker, Jr., D. H. Tomich, K. G. Eyink, and W. V. Lampert **1852**

Properties of nanometer-sized metal–semiconductor interfaces of GaAs and InP formed by an *in situ* electrochemical process
Hideki Hasegawa, Taketomo Sato, and Chinami Kaneshiro..... **1856**

Barrier heights of real Schottky contacts explained by metal-induced gap states and lateral inhomogeneities
Winfried Mönch **1867**

***In situ* electrical determination of reaction kinetics and interface properties at molecular beam epitaxy grown metal/semiconductor interfaces**
L. C. Chen and C. J. Palmstrøm **1877**

Photoemission spectromicroscopy studies on epitaxial lateral overgrowth GaN surfaces
Y. Yang, S. Mishra, F. Cerrina, S. H. Xu, H. Cruguel, G. J. Lapeyre, and J. F. Schetzina **1884**

CUMULATIVE AUTHOR INDEX **1891**

CERAC incorporated

advanced specialty inorganics

evaporation materials

sputtering targets

Whatever you need in inorganics, we'll work with you to make it happen.

All New Catalog
All New Website...
All From Your One Reliable Source

For years, CERAC has been your one source for the world's largest selection of inorganic chemicals, evaporation materials and sputtering targets in nearly any size, shape and composition. And now, we're expanding your resources even more.

First there's our brand new catalog with more new products, quicker and easier references and expanded technical information.

Then there's our newly revised web site with everything that's in our catalog and more, including a complete listing of stock items, in-depth technical product information and the latest industry news. Check in often to see what's new.

Of course, your greatest resources are the helpful people at CERAC. They're ready to answer your questions personally, or to take your order for fast, on-time shipment.

So many convenient resources from a single reliable source... working together just keeps getting easier with CERAC.

Call Today to Order Your FREE Personal Copy of the 1999 CERAC Catalog

414-289-9800

Or Visit Our Web Site to Browse the Catalog On-line

www.cerac.com

CERAC incorporated

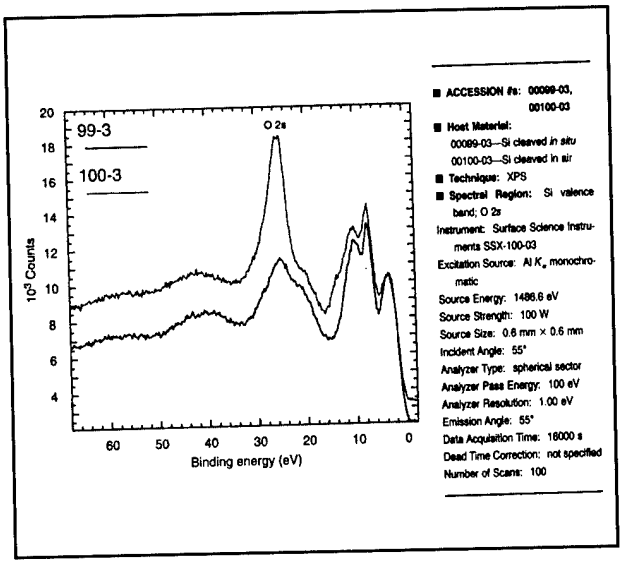
P.O. Box 1178
Milwaukee, WI 53201-1178
marketing@cerac.com

YOUR SINGLE SOURCE FOR INORGANICS FROM R&D TO FULL PRODUCTION

Surface Science Spectra

An International Journal Devoted to Archiving Surface Science Spectra of Technological and Scientific Interest

Add a cost-effective tool to your lab or library by subscribing to ***Surface Science Spectra*** - the definitive international reference guide to spectral data. Applications range from routine identification work to current research in industrial, government, and university labs. Surface Science Spectra offers scientists, engineers, and analysts easy and convenient access to high-quality spectra and related data with every quarterly issue.



Surface Science Spectra is the most comprehensive source available of complete spectral data. An official journal of the American Vacuum Society, *Surface Science Spectra* currently publishes XPS, Auger, UPS, SIMS and EELS/HREELS spectra from a wide range of fields. Features include:

- Extensive Experimental Details Provided
- The Only Peer Reviewed Spectral Archive
- More Spectra Than in Other Journals
- Focused-Topic Issues and Consolidated References on Specific Materials
- Reference, Comparison, and Technical Spectra
- Spectra in Hard Copy and Electronic Format
- Electronic Database Access with Subscription
- Spectral Data Records from Over 60 Labs Worldwide

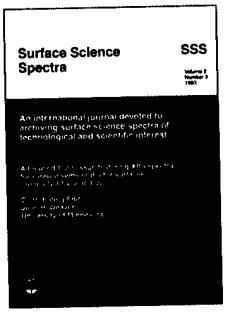
* * * ***Become a part of*** * * *

this

evolving

standard

today!



Surface Science Spectra

To Subscribe or Contribute

Fax or mail this form to SSS, Caller Box 13994, Research Triangle Park, NC 27709, Fax (919)361-1378
Phone(919)361-2498. E-Mail: sss@jvst.org or located at: http://www.vacuum.org/sss.html

Name _____
Organization _____
Address/Mail Stop _____
City/State/Zip _____
Country _____
E-Mail _____

Learn Vacuum Technology from the Experts ...

Buy an *AVS Classic Series* reference book—out-of-print references by industry-founding experts:

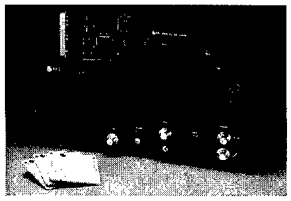
- *Vacuum Technology and Space Simulation—Santeler, Holkeboer, Jones, and Pagano*
- *Field Emission and Field Ionization—Gomer*
- *Handbook of Electron Tube and Vacuum Techniques—Rosebury*
- *The Physical Basis of Ultrahigh Vacuum—Redhead, Hobson, and Kornelsen*
- *Ionized Gases—von Engel*
- *Vacuum 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 1-800-777-4643 to order.

Does Your Scanning System Need an Upgrade?

PC Upgrades for PHI® 590, 595, & 600 Scanning Auger Systems

The RBD ScanMap1 PC Upgrade Package provides Windows® 95, 98 & NT control of data acquisition and analysis on PHI® 590, 595, 600 Scanning Systems.



PC Upgrade Benefits:

- System reliability & performance are greatly improved
- Replaces 18-075/18-175 keyboard and 18-070/18-170 scanning electronics
- Reduced system footprint means more lab space
- RBD Y2K compliant software used to control the system
- Can be used with most PCs

RBD Enterprises also provides the following products and services for PHI® and other surface analysis systems and components:

- Repair of optics and electronics
- Refurbished components and systems
- Field service
- PC-based upgrades for many PHI® systems

RBD ENTERPRISES

563 SW 13th St. Suite 201• Bend, OR • 97702
Phone: 541.330.0723 • Fax: 541.330.0991
E-Mail: info@rbdenter.com
Web Site: http://www.rbdenter.com

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.

Advertisers Direct Response

Journal of Vacuum Science & Technology B

Microelectronics and Nanometer Structures, Processing, Measurement, and Phenomena

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.

◊ JULY/AUGUST 1999 ◊

ADVERTISERS:

A15	CERAC Inc. Fax: (414) 289-9805	Cov 2	Leybold Inficon Inc Fax: (315) 437-3803	A17	RBD Enterprises Fax: (541) 330-0991
A9	Hiden Analytical Inc. Fax: (603) 924-5009	A11	LK Technologies Fax: (812) 332-4493	A11	Skion Corp. Fax: (201) 963-5449
A1	Innovac Corp. Fax: (201) 963-5449	A9	Omicron Associates Fax: (412) 831-9828	A7	Stanford Research Fax: (408) 744-9049
A4	International Radiation Detectors Inc. Fax: (310) 534-3665	A3	Pfeiffer Vacuum Fax: (603) 578-6650	A16	Surface Science Spectra Fax: (919) 361-2498
Cov 3	Kurt J. Lesker Co. Fax: (412) 233-4275	Cov 4	Quad Group, Inc. Fax: (509) 458-4555	A13	VLSI Standards, Inc. Fax: (408) 428-9555

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

Application and calibration of a quartz needle sensor for high resolution scanning force microscopy

W. Clauss,[a] J. Zhang, D. J. Bergeron, and A. T. Johnson
Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 15 January 1999; accepted 21 May 1999)

We have investigated a force sensing technique for high-resolution scanning force microscopy which uses a piezoelectric quartz resonator oscillating perpendicular to the direction of the sample surface at a frequency of 1 MHz. The achievement of true atomic resolution on single-wall carbon nanotube surfaces illustrates the promises of this new technique which is especially well suited for ultrahigh vacuum or low-temperature conditions. The force sensitivity and the mechanical amplitude of the sensor oscillation were characterized semiquantitatively. © *1999 American Vacuum Society.* [S0734-211X(99)08504-2]

I. INTRODUCTION

The use of piezoelectric resonators as distance sensors was introduced into the field of scanning probe microscopy some years ago,[1] mainly driven by applications connected to near-field scanning optical microscopy.[2] For dedicated scanning force microscopy (SFM), this technique was not widely taken into consideration, mainly because of low operating frequencies and the fact that dynamic noncontact or tapping mode techniques with standard cantilevers are already well established. Recently, a scanning force microscope based on a quartz "needle" resonator operating at a frequency of 1 MHz[3] has become commercially available.[4] A sketch of this force detector is shown in Fig. 1. The high operation frequency greatly increases the maximum possible scan speed which is the most severe limitation of this method. Moreover, as shown below, with appropriate tips this sensor even allows the acquisition of true atomic resolution images. We suggest that this surprisingly high resolution results from the low oscillation amplitude ($\leqslant$1 nm) made possible by the high mechanical stiffness of the quartz needle, especially in comparison to usual Si-based cantilevers.

Until now, there exists no quantitative characterization of the frequency response as a function of the tip-sample interaction force. Here, we present the first calibration procedure of the needle sensor response allowing at least an estimate of the actual tip-sample forces. Finally, we describe a simple procedure to determine the amplitude of the needle oscillation.

II. EXPERIMENTAL SETUP

We used a BEETLE[5] type ultrahigh vacuum (UHV) compatible microscope head with a scan range of 5 μm which was modified[6] for the use of the 1 MHz needle sensor. The instrument is capable of operation at temperatures below 4 K, but for the measurements reported here it was operated only under ambient or HV conditions. Electronic control of the microscope is provided by a SCALA system[4] allowing the operator to switch between different modes of tunneling, dynamic force, and conductive-tip force microscopy by the appropriate choice of the feedback signal and the use of either a standard metal tip or a needle sensor. In SFM mode, the feedback uses the phase shift ϕ between the needle output signal and the excitation frequency, which is adjusted by the software to match the piezo resonance. In order to attach the tips of standard[7] or diamond-coated conductive SFM cantilevers[8] to the end of the needle sensors,[9] we developed a micromanipulation stage. A cantilever is positioned on top of the face surface of a needle with an accuracy of 5 μm and connected by gluing with a conductive resin. After the epoxy has cured, the cantilever holder is broken away from the beam by carefully retracting it, leaving the needle with a slight weight increase due to the combined mass of the resin and the part of the cantilever carrying the tip. This results in a shift of the undistorted resonance frequency at 1 MHz by several kHz. A change from ambient conditions to moderate vacuum (10^{-2} T) introduces an increase of the resonance frequency of about 40 Hz. Further lowering of the pressure does not affect the frequency significantly. Cooling the sensor to liquid helium temperatures shifts the frequency by $\sim$500 Hz. The quality factor of the oscillation also varies from $\sim$8000 at ambient to 20 000 under HV conditions. We note that due to the dependence on the environmental conditions, it is not possible to make a calibration as described below that is valid for all needles even if they were identical. Therefore, the reported values should only be taken as a quantitative estimate.

III. HIGH RESOLUTION IMAGES

In order to demonstrate the needle sensor capabilities, we show images of single-wall carbon nanotubes (SWCNs). Figures 2(a) and 2(b) are the topography and phase images, respectively, of a rope containing several SWCNs lying on a flat Au(111) surface. The diameter of individual tubes cannot be determined from this image because the resolution is limited by the diameter of the diamond-coated tip used in this measurement. However, in some cases it was possible to achieve a much higher resolution with the same kind of tips. An example is given in Fig. 3 which shows an area with few

[a]Permanent address: University of Tuebingen, Inst. of Applied Physics, Auf der Morgenstelle 10, D-72076 Tuebingen, Germany; electronic mail: wilfried.clauss@uni-tuebingen.de

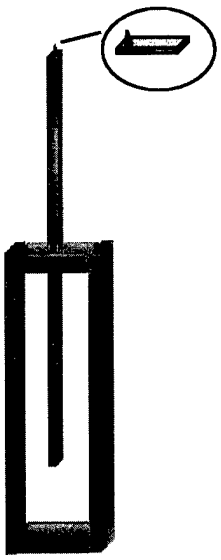


FIG. 1. Sketch of a needle sensor with a cantilever beam attached to the top surface.

individual SWCNs. In addition to the standard topographic image, the local conductivity [Fig. 3(b)] could be measured simultaneously by application of a voltage of several hundred mV between sample and conductive tip. In all signals, the chiral structure of the SWCN surface resulting from the orientation of the graphitic lattice with respect to the tube axis is clearly visible. In addition to the expected primitive periodicity along the spiral lines, a second periodicity compatible with a $\sqrt{3}$-based superstructure can be identified. A detailed interpretation of the image contrast is given elsewhere.[10] This image clearly illustrates that the spatial

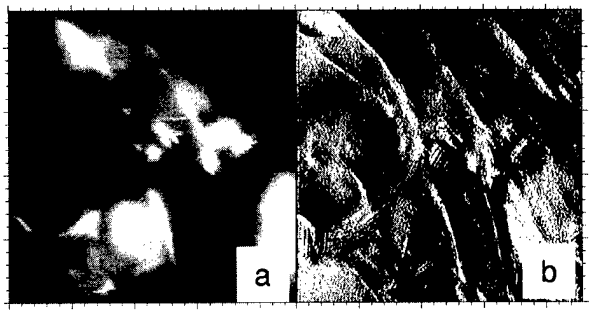


FIG. 2. Overview SFM image (230×230 nm) of nanotubes on a Au(111) surface. (a) Topography, (b) error signal (phase deviation).

resolution of the needle sensor technique is only limited by the actual sharpness of the tip. To account for the extremely high resolution we suggest that the tip fortuitously picked up an individual nanotube serving as very sharp force sensor in subsequent images. We also note that in the high resolution image series taken with phase setpoint values below 2°, no compression of the tube surface was observed. Theoretical and experimental values for the bending strength of SWCN suggest that the forces necessary to perturb the surface geometry are of the order of a few nN.[11,12] This is consistent with the calibration results discussed below.

IV. CALIBRATION OF THE NEEDLE PHASE RESPONSE

In order to get a better estimate about the relation between the phase shift of the needle signal and the applied force we arranged a setup where a soft standard SiN cantilever[7] was imaged by the SFM tip. Using the standard spectroscopy

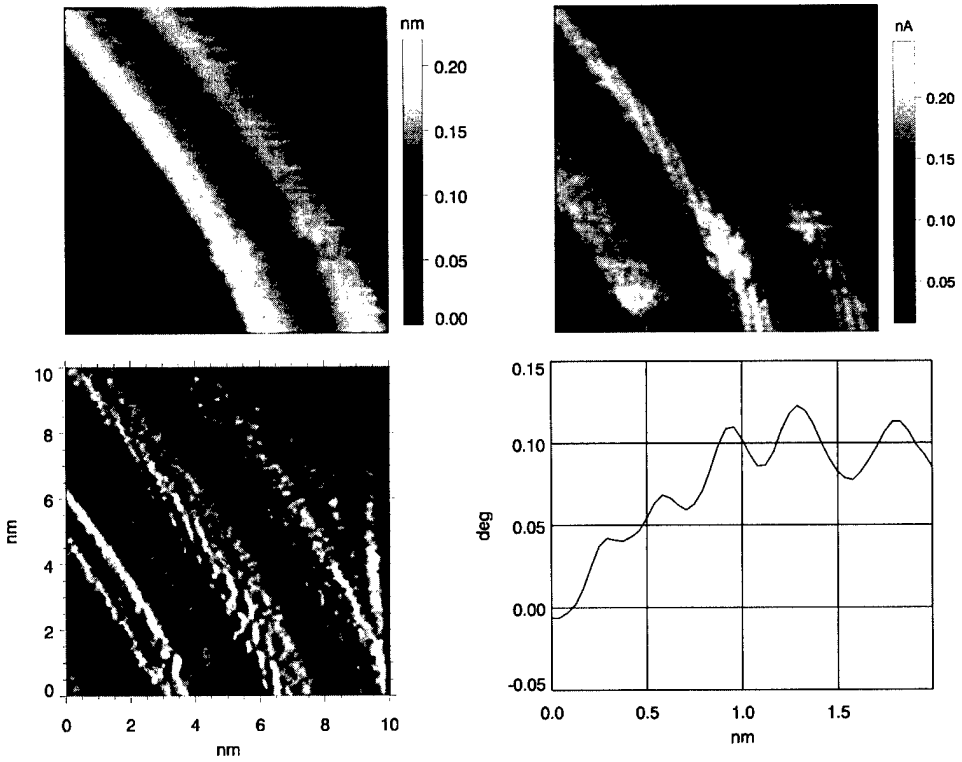


FIG. 3. High resolution SFM image (10×10 nm) of nanotubes with phase setpoint of 1.5°. (a) Topography, (b) current, (c) phase (error signal). The periodicity visible in the line scan perpendicular to the spiral lines results from a $\sqrt{3}\times 1$ superstructure superimposed to the primitive lattice which can be identified along the spiral lines. The two identically looking tubes are probably imaged by a double tip.

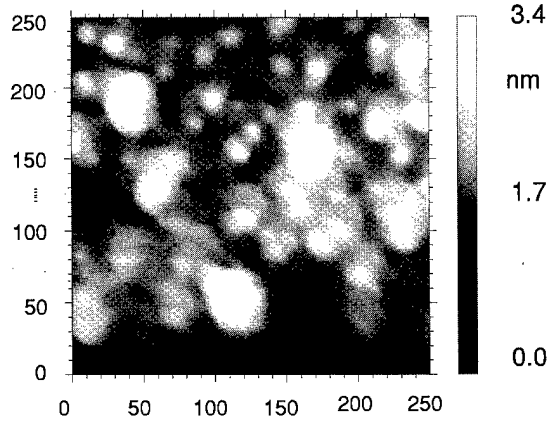


FIG. 4. Scanning force image of the cantilever carrier surface taken before the force-distance spectra on the cantilever itself were taken. The phase setpoint value is 5°.

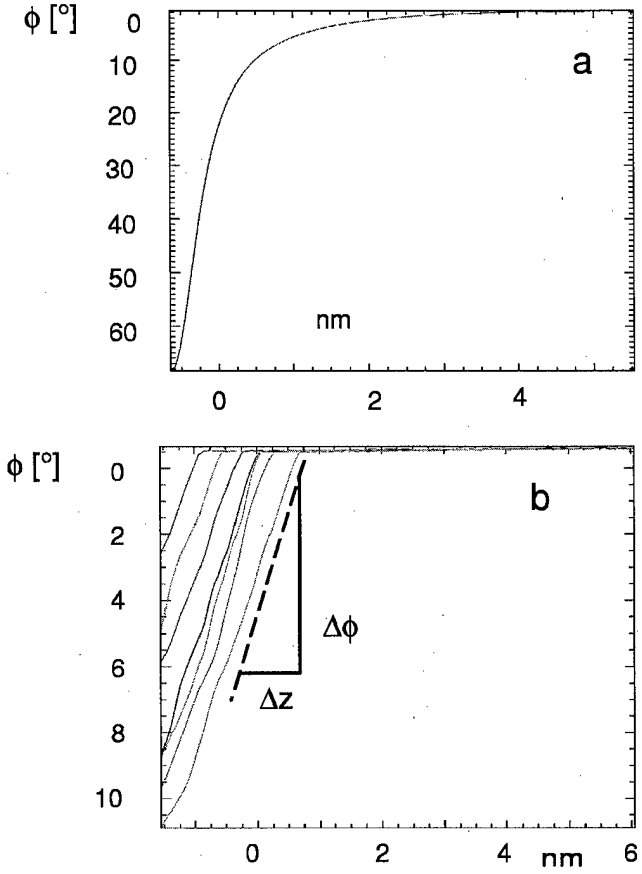


FIG. 5. Force distance curves taken (a) on the cantilever carrier, (b) on cantilever with $k=0.5$ N/nm at relative position $l_i=0.4l$.

software of the microscope controller, the phase response of the needle was measured under variation of the z position of the tip at a fixed lateral position l_i along the cantilever. From standard continuum elasticity, the spring constant of a straight beam with length l, thickness d, and width w is given by $k=cwd^3/l^3$, where c is a constant accounting for the elastic material properties. Therefore, the local spring constant k_i can be described by $k_i=k(l/l_i)^3$. Assuming a linear cantilever response, we expect a phase-distance characteristic with slope $m=\Delta\phi/\Delta z$ which is related to the sensitivity S of the needle by $S=m/k_i$.

Four cantilevers of known geometry with different spring constants k have been used. Additionally, the thickness of the cantilevers was measured by scanning electron microscopy to correct for possible deviations with respect to the nominal value. With this correction taken into account, their k values were 0.095, 0.19, 0.50, and 0.92 N/m, respectively. The calibration measurements were performed under high vacuum to avoid capillary force effects in the following way: First, a coarse approach onto the surface of the cantilever carrier was performed. In order to check the quality of the cantilever tip, an image of the carrier surface was taken immediately after this approach. An example is shown in Fig. 4. Subsequently, the tip was gradually moved towards the end of the cantilever while monitoring the feedback stability. The phase setpoint was chosen as low as possible in the range between 10° and 15°, in order to minimize the static bend of the cantilever beam. We observed a strong trend towards instability of the cantilever if the effective spring constant k_i was lowered by increasing the distance from the carrier edge. On the cantilever with the smallest k value, stable feedback operation could not be established at all, most likely because adhesive or capillary forces were too large. On the other three cantilevers, phase-distance spectra were taken at various positions on the cantilevers between 40% and 100% of the full beam length. Vertical variations were extended over a range of 8-15 nm around the setpoint position. An example of a phase-distance curve is shown in Fig. 5(b). For comparison, a standard phase-distance spectrum taken on top of the SiN carrier is shown in Fig. 5(a). As expected, the force-distance relation on the cantilever is dominated by the bend of the beam, leading to a straight force-distance curve. In contrast, the phase shift on the stiff cantilever surface is strongly nonlinear and reaches very high values. Table I lists the measured sensitivity values from various positions, taking into account the different values of k_i of different tip positions. We suggest that the variation between the different measurements is mostly resulting from the uncertainty of the actual tip position, leading to a corresponding error in the effective spring constant. From the fact that the measurement at the full length of a cantilever (giving the most reliable length estimate) yields the highest sensitivity, we conclude that a realistic calibration factor is around 1°/nN. We note that the sensitivity of an individual sensor also depends on its actual quality factor.

TABLE I. Measured values of the sensitivity S on three different cantilevers.

k_{eff} [N/m]	l_i	sensitivity S [°/nN]
0.19	0.8	0.72
	1.0	1.81
0.5	0.4	0.9
	0.7	0.95
0.92	0.75	0.5
	0.8	0.8

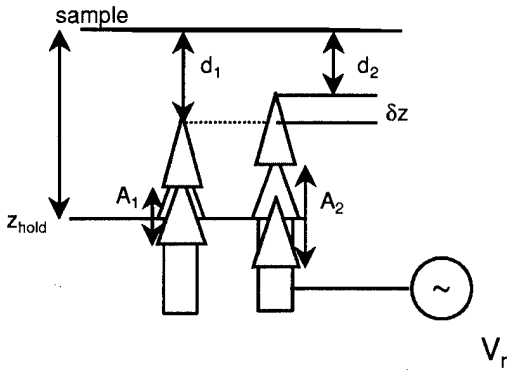


FIG. 6. Sketch of the setup for the measurement of the oscillation amplitude. During stable tip-sample contact the distance is fixed to a value z_{hold}. Afterwards, the minimum tip-sample distance is decreased from d_1 to d_2 by increasing the amplitude from A_1 to A_2.

V. AMPLITUDE OF THE APPLIED OSCILLATION

Along with an estimate for the strength of the interaction force, it is also important to determine the amplitude of the mechanical needle oscillation. This can be done by the following procedure: As shown schematically in the left part of Fig. 6, we assume that the amplitude A_1 is of the same order as the closest distance d_1 at the turning point of the cycle. We further assume that the interaction leading to the phase shift will mostly take place during the part of the cycle with the closest approach (tip near d_1). During the first step of the procedure, a stable tip-sample contact with a phase setpoint in the range of usual imaging parameters is established. Next, the z position of the piezo is fixed to a value z_{hold} (see Fig. 6) and the feedback circuit is disabled. Then, the needle excitation is raised from A_1 to $A_2 \approx 1.5A_1$. The increased mechanical amplitude reduces the minimum tip sample distance at the turning point by δz, causing an increased phase shift. With the feedback still disabled, the tip is backed away manually until the original phase setpoint is restored. Neglecting mechanical drift in the tip-sample distance, this distance correction δz and the value of the amplitude variation $A_1 - A_2$ caused by the change of the excitation signal V_r can be assumed to be identical. Therefore, we get the expression for the amplitude conversion factor D between excitation voltage and mechanical needle oscillation as

$$\delta z = 2(A_2 - A_1) = D(V_{r2} - V_{r1}).$$

The application of this procedure with a variety of amplitudes yields values for D around 2 nm/mV. This can vary for individual needles because of the dependence of the oscillation amplitude on the Q factor. Common excitation values for the sensor operation are in the range of 0.3–0.5 mV, resulting in amplitudes below or around 1 nm, small enough to explain the observed high spatial resolution. Moreover, a 1 nm oscillation amplitude will keep the tip in a distance range where electron tunneling is possible during a large fraction of the cycle. Therefore, this sensor is well suited for simultaneous measurement of mechanical and electronic surface properties and for imaging samples with inhomogeneous electrical conductivity. This is illustrated by current images acquired simultaneously with the phase and topography signal as shown in Fig. 3(b).

VI. SUMMARY

We have shown that dynamic force microscopy with a 1 MHz needle sensor can be successfully used to obtain near-atomic resolution as illustrated by experiments on single-wall carbon nanotubes. A semiquantitative calibration measurement of the force sensitivity shows that this technique is comparable to more sophisticated setups for high resolution dynamic force microscopy used previously.[13,14] Moreover, the applicability for scanning force microscopy under UHV and low-temperature conditions is greatly improved by this setup because no optical components are used.

ACKNOWLEDGMENTS

We thank A. Rinzler and R. S. Smalley for providing SWCN samples and Micro Crystal for needle sensors.

[1]K. Karrai and R. D. Gruber, Appl. Phys. Lett. **66**, 1842 (1995).
[2]E. Betzig, P. L. Finn, and J. S. Weiner, Appl. Phys. Lett. **60**, 2484 (1992).
[3]U. Grunewald, K. Bartzke, and T. Antrack, Thin Solid Films **264**, 169 (1995).
[4]Omicron Instruments, Taunusstein, Germany.
[5]K. Besocke, Surf. Sci. **181**, 145 (1987).
[6]Zeiss Jena GmbH, Jena, Germany.
[7]Digital Instruments, Santa Barbara, CA.
[8]Nanosensors, Wetzlar, Germany.
[9]Micro Crystal/ETA SA, Grenchen, Switzerland.
[10]W. Clauss, M. Freitag, D. J. Bergeron, G. Mele, C. Kane, and A. T. Johnson, Europhys. Lett. (submitted).
[11]R. S. Ruoff and D. C. Lorents, in *Carbon Nanotubes*, edited by M. Endo, S. Iijima, and M. S. Dresselhaus (Pergamon, Oxford, 1996).
[12]H. Edwards, L. Taylor, and W. Duncan, J. Appl. Phys. **82**, 980 (1997).
[13]E. Luethi *et al.*, Z. Phys. B **100**, 165 (1996).
[14]P. Guethner, J. Vac. Sci. Technol. B **14**, 2428 (1996).

Surface structure characterization of DNA oligomer on Cu(111) surface using low temperature scanning tunneling microscopy

Chiho Hamai, Hiroyuki Tanaka, and Tomoji Kawai[a)]
The Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

(Received 7 January 1999; accepted 21 May 1999)

The surface structures of DNA oligomers, *pAAAAAAATTTTTTT*, deposited on Cu(111) surface have been characterized at liquid nitrogen temperature using a scanning tunneling microscope. Four different types of adsorbed structures have been observed in DNA oligomers; (i) an isolated whole molecule, (ii) a shortened molecule, (iii) a cluster, and (iv) a double helix. The internal structures of the oligomers also have been resolved. © *1999 American Vacuum Society.* [S0734-211X(99)08104-4]

I. INTRODUCTION

The scanning tunneling microscope (STM) is an important tool for the study of biomolecules.[1] A large number of reports on the observations of DNA in air or water using STM have been published.[2–6] The invention of STM aroused our interest in DNA base sequencing using STM, but sequencing using STM remains to be achieved.[1] The resolution of the images obtained in air or water seems not to be enough to image the internal structure of DNA.[2–6] For the improvement in resolution, we have performed the observation at liquid nitrogen temperature in ultrahigh vacuum (UHV).

Here, we report the STM observation of a DNA oligomer adsorbed on Cu(111). By performing the observations at liquid nitrogen temperature in UHV, high resolution images have been obtained. We have found that the oligomer forms various structures on Cu(111) and the molecular internal structures are also successfully resolved.

The observed DNA oligomer is *pAAAAAAATTTTTTT* [abbreviated as *A7T7*]. The 14-mer has 14 nucleotide units consisting of seven adenines and seven thymines attached to its phosphate sugar backbone. Since the Watson–Crick base-pairing forces DNA into the double helix structure, the 14-mer is expected to form some interesting structures.

II. EXPERIMENT

A clean Cu(111) surface was prepared by several cycles of annealing at 550 °C and sputtering with Ar^+.[7,8] The *A7T7* oligomer was deposited on the clean Cu(111) surface using the pulse injection technique[9] in order to prevent decomposition of the DNA oligomer because it was well known that large molecules, such as DNA deposited using thermal evaporation[10] can be easily decomposed.[11,12] A pulse valve filled with a DNA aqueous solution was attached to the preparation chamber in vacuum ($\sim 10^{-7}$ Torr). A clean Cu(111) substrate was located just under the pulse valve. The distance from the pulse valve to the substrate was 100 mm. The solution was injected toward the substrate at room temperature when the valve was opened for 1.5 ms. After deposition of the DNA oligomer on the Cu surface, this sample was subsequently transferred to the STM UHV chamber ($\sim 10^{-10}$ Torr). And then surface structures of the DNA were observed using STM at liquid nitrogen temperature (~80 K). All images were obtained in constant current mode at the bias voltage of 2 or 3 V. The atomic resolution of Cu atoms[13] cannot be obtained under the tunneling condition which is suitable for the imaging of DNA oligomers.

III. RESULTS AND DISCUSSION

Figure 1 shows a STM image of an *A7T7* oligomer on the clean Cu(111) surface. The size of the encircled bright object is in agreement with that of one molecule in lateral dimensions. The various observed adsorbed structures can be classified into four different types: (i) an isolated whole molecule [Fig. 2(a)], (ii) a shortened molecule [Fig. 2(b)], (iii) a cluster which is constructed from two or more molecules [Figs. 2(c)–2(e)], and (iv) a double helix [Figs. 2(f) and 2(g)].

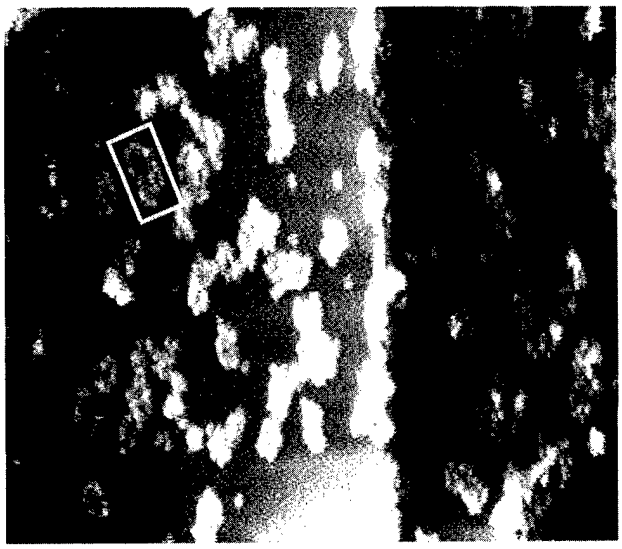


FIG. 1. Image of *A7T7* DNA oligomers on Cu(111) surface. The image was taken in the constant current mode with $I=4$ pA and $V_t=2.0$ V for the area of ~500 Å × ~430 Å. The bright object encircled with a rectangular is an isolated molecular image.

[a)]Corresponding author; electronic mail: kawai@sanken.osaka-u.ac.jp

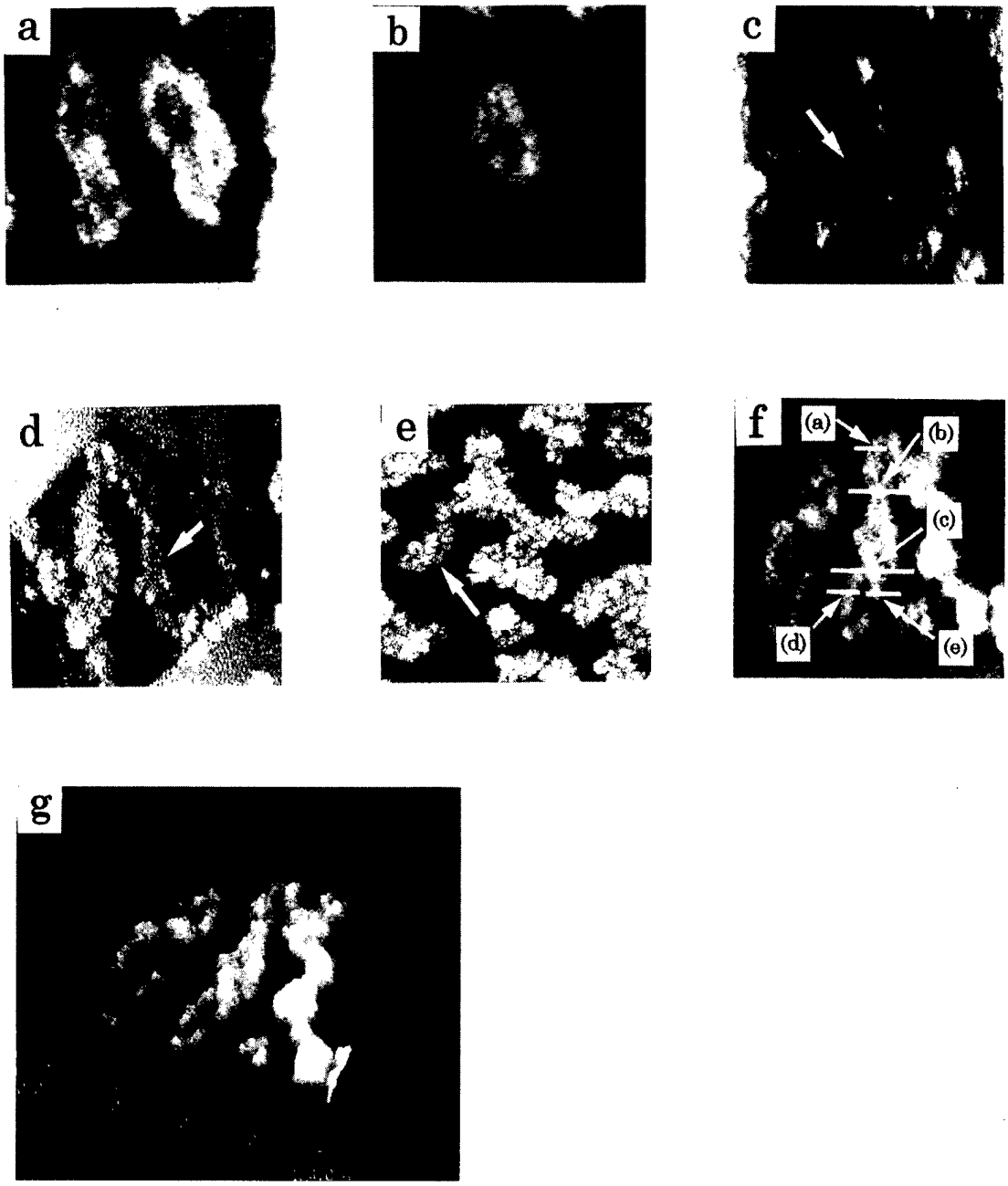


FIG. 2. STM images of *A7T7* DNA oligomers adsorbed on Cu(111) surface. (a) An image of two isolated whole molecules (~90 Å×90 Å). The topographic heights of these molecular images are about 2Å. (b) An image of a shortened molecule (~90 Å×90 Å). The chain length is about half of those measured for the isolated molecules in (a). (c) An image of a cluster (~200 Å×200 Å). A pair of parallel two chains can be seen (arrow). (d) An image of a cluster (300 Å×300 Å). A pair of parallel two chains is observed (arrow). (e) An image of clusters (220 Å×220 Å). A part of twisted chains can be seen (arrow). (f) An image of double helix (~130 Å×130 Å). White lines indicated by arrows show positions where the height of the molecule was measured. The maximum topographic heights of cross sections (a), (b), (c), (d), and (e) are 1.6, 2.9, 3.5, 2.6, and 2.2 Å, respectively. (g) The perspective projection of (f). (a), (b), (c), and (f) The extracted images from Fig. 1. (e) Extracted from Fig. 3. (d) A region picked out in an image which is not presented here. (d) A constant current image taken with $I=5$ pA and $V_t=3.0$ V.

The topographic heights of isolated molecules [Fig. 2(a)] are about 2 Å. Figure 2(b) is a STM image of the shortened molecule, whose chain length is approximately half of that of the isolated molecule shown in Fig. 2(a). Figures 2(c)–2(e) show STM images of "clusters." Parallel chains [Figs. 2(c) and 2(d)] and a twisted chain [Fig. 2(e)] are seen. A "double helix" structure has been found as shown in Figs. 2(f) and 2(g). The value of lateral dimension indicates that the "double helix" is formed from two oligomers. White lines indicated by arrows in Fig. 2(f) show positions where the height of molecule was measured. The maximum heights of cross sections (a), (b), (c), (d), and (e) of Fig. 2(f) are 1.6, 2.9, 3.5, 2.6, and 2.2 Å, respectively. Since the heights of cross sections (b) and (c) of Fig. 2(f) are larger than those of cross sections (a), (d), and (e), (b) and (c) are the positions where the two strands cross. The "double helix" is therefore

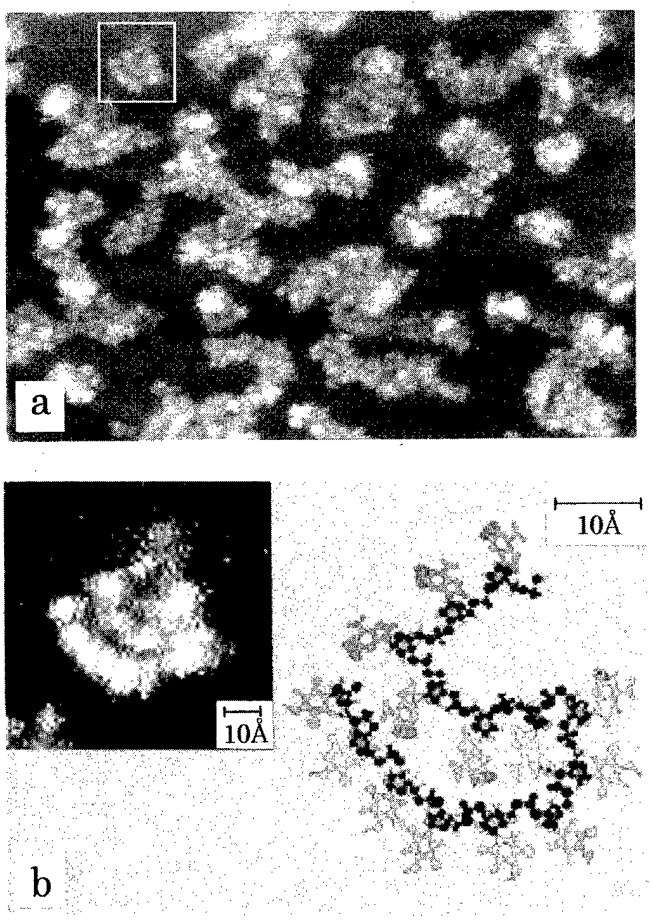


FIG. 3. (a) Image of $A7T7$ DNA oligomers on Cu(111) surface. The internal structures of the oligomers are resolved in this image. The image was taken in the constant current mode with $I=5$ pA and $V_t=3.0$ V for the region of 480 Å×320 Å. The bright object enclosed with a square is an isolated molecule. (b) The inset is the expanded image of the isolated molecular image enclosed with a square in (a). The molecular model is for the isolated molecule of the inset. The model and the molecular image are in agreement in lateral dimensions.

one and a half turned double helix. A double strand DNA in a crystal or solution has about ten base pairs per turn.[14] If the double helix is constructed from two $A7T7$ oligomers, then it ideally has one and a half turns. The number of turns present in the "double helix" image is the same as that of the helix structure existing in a crystal or solution. However the topographic heights of the helix image are considerably smaller than the diameter of a DNA double helix in a solid or a solution (~20 Å).[14] This difference may be attributed to the imaging mechanism or the molecule–surface interaction which transforms the conformation without changing the number of helix turns, but only shortening the height.

In addition to the imaging of four adsorbed structural types, the internal structure of $A7T7$ has been resolved [Fig. 3(a)]. The bright object enclosed with a square in Fig. 3(a) corresponds to an isolated molecule. Its topographic height is about 2 Å. All molecular images consist of small bright maxima that are about 7–10 Å in diameter. The length of one oligomer in STM images is approximately equal to 14 times of the bright maximum diameter (7–10 Å), therefore each bright maximum is considered to be one nucleotide unit. A molecular model whose base planes were parallel to the Cu surface was constructed for an isolated molecular image as shown in Fig. 3(b). The inset of Fig. 3(b) is the expanded image of the isolated molecule enclosed with a square in Fig. 3(a). The model is in agreement with the isolated molecular image shown in the inset in lateral dimensions.

The height of the isolated molecular image (2 Å) in Figs. 2(a) and 3(a), and the diameter of the bright maximum (7–10 Å) in Fig. 3(a) suggest that the small bright maximum observed in Fig. 3(a) corresponds to one base. It has been shown that DNA base molecules adsorbed flat on Cu(111) surfaces are imaged with the height of 1–2 Å.[15] The value of 1–2 Å is relatively consistent with the height of the isolated oligomers (2 Å). And the diameter of the bright maximum (7–10 Å) is a typical value for the lateral size of DNA base molecules lying flat on metal surfaces.[15–17] The discrimination between A and T was not achieved in this experiment. Improving the resolution and optimizing the tunneling condition (such as bias voltage or tunneling current) can lead to identification of DNA bases. However, we cannot assert that only bases are imaged, because we have not measured anything but constant current images.

The STM images depend on two factors, namely local density of state (LDOS) and barrier height. If the LDOS contributes more to the imaging of DNA oligomers compared with the barrier height, then the small bright maximum observed in this study corresponds to one base. This is because the π orbitals of the base exists at the highest energy level of all orbitals of DNA and the flat plane of the base could result in a sufficient interaction with the Cu surface. On the other hand, if the barrier height contributes more to the imaging compared with the LDOS, then it is not possible to infer which part of DNA is imaged most brightly. In order to clarify the imaging mechanism, we would need further investigations.

IV. CONCLUSIONS

DNA oligomers ($A7T7$) adsorbed on Cu(111) surface have been observed with a low temperature STM in UHV. We have found that $A7T7$ oligomers form various adsorbed structures on Cu(111) surfaces. These adsorbed structures are classified into four types: isolated molecule type, cut molecule type, cluster type, and double helix type. The internal structures of $A7T7$ oligomers have also been resolved.

[1] *Scanning Tunneling Microscopy and its Application*, edited by C. Bai (Springer, Berlin, 1992), pp. 251–276.
[2] S. M. Lindsay, T. Thundat, L. Nagahara, U. Knipping, and R. L. Rill, Science **244**, 1063 (1989).
[3] A. Cricenti, S. Selci, A. C. Felici, R. Generosi, E. Gori, W. Djaczenko, and G. Chiarotti, Science **245**, 1226 (1989).
[4] D. P. Allison and T. Thundat, J. Vac. Sci. Technol. A **11**, 816 (1993).
[5] T. W. Jing *et al.*, Proc. Natl. Acad. Sci. USA **90**, 8934 (1993).
[6] R. Guckenberger, M. Heim, G. Cevc, H. F. Knapp, W. Wiegräbe, and A. Hillebrand, Science **266**, 1538 (1994).
[7] I. Bönicke, W. Kirstein, S. Spinzig, and F. Thieme, Surf. Sci. **313**, 231 (1994).

[8]Y. L. Chan, P. Chuang, T. J. Chuang, and R. Klauser, Surf. Sci. **402–404**, 125 (1998).
[9]H. Tanaka and T. Kawai, J. Vac. Sci. Technol. B **15**, 602 (1997).
[10]T. A. Jung, R. R. Schlittler, and J. K. Gimzewski, Nature (London) **386**, 696 (1997).
[11]R. W. Nelson, M. J. Rainbow, D. E. Lohr, and P. Williams, Science **246**, 1585 (1989).
[12]R. J. Levis, Annu. Rev. Phys. Chem. **45**, 483 (1994).
[13]A. R. H. Clarke, J. B. Pethica, J. A. Nieminen, F. Besenbacher, E. Laegsgaard, and I. Stensgaard, Phys. Rev. Lett. **76**, 1276 (1996).
[14]W. Saenger, *Principles of Nucleic Acid Structure* (Springer, New York, 1984).
[15]M. Furukawa, H. Tanaka, and T. Kawai, Surf. Sci. **392**, L33 (1997).
[16]H. Tanaka, J. Yoshinobu, M. Kawai, and T. Kawai, Jpn. J. Appl. Phys., Part 2 **35**, L244 (1996).
[17]H. Tanaka and T. Kawai, Mater. Sci. Eng. **C3**, 143 (1995).

Microstructure and electrical properties of Sb nanocrystals formed in thin, thermally grown SiO_2 layers by low-energy ion implantation

Anri Nakajima,[a] Hiroshi Nakao, Hiroaki Ueno,[b] Toshiro Futatsugi, and Naoki Yokoyama
Fujitsu Laboratories Ltd., 10-1 Morinosato-Wakamiya, Atsugi 243-0197, Japan

(Received 15 January 1999; accepted 16 April 1999)

We have formed Sb nanocrystals in thin, thermally grown SiO_2 layers using low-energy ion implantation which was followed by thermal annealing. These Sb nanocrystals showed good uniformity with respect to size and position. Both the narrow as-implanted profile and the compressive strain that exists near the SiO_2/Si interface are considered to contribute to the uniformity. We found that it was possible to control the size of the Sb nanocrystal by changing the dose of the implanted Sb. The $I-V$ characteristics of the diodes which had Sb nanocrystals with an average diameter of 3.3 nm showed good reproducibility. Moreover, these diodes showed a Coulomb blockade region around 0 V and a Coulomb staircase at 4.2 K. Identical $I-V$ characteristics were obtained for all the samples measured, indicating an excellent reproducibility. Almost the same Coulomb blockade region as that at 4.2 K was observed up to a temperature as high as 100 K even for a diode which had larger Sb nanocrystals (with an average diameter of 5.6 nm). The technique introduced here offers the possibility of developing practical Si-based single-electron devices. © *1999 American Vacuum Society.* [S0734-211X(99)01504-8]

I. INTRODUCTION

Single-electron devices might be used as elements of future solid-state circuits that require high integration and low-power consumption. Many fabrication processes have already been proposed for creating single-electron devices that would be operable at high temperatures. One of the key technologies in this connection is fabrication of nanoscale structures. Electron beam (EB) lithography is the most popular method for nanofabrication.[1–3] However, while EB lithography has many advantages, the resulting low productivity with respect to nanofabrication is a serious drawback in the development of commercial devices. Contrary to this, the self-organizing fabrication process of nanoscale structures is a relatively simple, inexpensive and highly productive method. Another demand that has to be fulfilled by the fabrication technique is its usability in connection with the fabrication processes of conventional large-scale integrated (LSI) devices. When this condition is fulfilled, it becomes possible to use proven, mature Si techniques, which is an advantage for developing commercial devices.

Various techniques have been proposed to date for fulfilling the above two requirements. Nanoscale grain islands in a thin polysilicon film were used for floating dot gates to create a single electron memory that operated at room temperature.[4] The formation of Si nanocrystals has also been discussed based on the early stage of Si film growth for conventional low-pressure chemical vapor deposition (LPCVD) on a SiO_2 film.[5,6] However, these nanocrystals are not sufficient for device applications due to the insufficient uniformity of their size and position. Therefore, we have reported a fabrication technique for improving the uniformity of nanocrystals in earlier works.[7–9] Following the technique, metal nanocrystals in thin, thermally grown SiO_2 layers were formed by using the low-energy ion implantation of Sn^+ ions followed by thermal annealing. With this method, we were able to form Sn nanocrystals that provided excellent size and depth position uniformity. This technique provides high fabrication throughput of nanoscale structure and is extensively used in LSI device fabrication technology.

Metal or semiconductor nanocrystals in barrier materials with low dielectric constants have the advantage that the single-electron effect occurs as well at high temperatures, since the total capacitance of the dots decreases in these materials. The merit of using metal instead of semiconductor nanocrystals as Coulomb islands in single-electron devices is that the energy separation of the quantum levels is negligible in metal nanocrystals. This results in a constant addition energy and makes designing single-electron devices simple.

However, the method[7–9] of using Sn^+ ions for ion implantation is not completely compatible with general LSI fabrication processes, since Sn^+ ions are not typically used in conventional LSI device fabrication. We discussed in an early study,[10] how to overcome this problem by forming Sb nanocrystals in SiO_2 layers by ion implantation and subsequent thermal annealing. Sb is one of the typical doping materials in conventional LSI fabrication. However, the fluctuation in the size and position of the formed Sb nanocrystals was rather large, since low-energy ion implantation was not carried out for the method discussed in that paper.[10] In a later work, we discussed preliminary results for low-energy ion implantation of Sb^+ ions in thin, thermally grown SiO_2 layers followed by thermal annealing and also reported the electrical characteristics of the diodes that employed the fabri-

[a]Present address: Research Center for Nanodevices and Systems, Hiroshima University, 1-4-2 Kagamiyama, Higashi-Hiroshima 739-8527, Japan; electronic mail: nakajima@sxsys.hiroshima-u.ac.jp
[b]Permanent address: Department of Electronic Engineering, Faculty of Engineering, Osaka University, 2-1 Yamada-oka, Suita 565-0871, Japan.

cated structures.[11] In this article, we will discuss the microstructure of these Sb nanocrystals and the electrical characteristics of the diodes in detail.

II. EXPERIMENT

The low-energy implantation of Sb^+ ions was carried out at an energy of 10 keV and a dose of 1×10^{15} and 5×10^{15} ions/cm^2 into thin SiO_2 layers which were thermally grown on Si substrates with varying thickness. The thickness of the SiO_2 layers was 15 and 20 nm, respectively. The oxidation was carried out at 900 °C in dry conditions. Heavily doped n^+-type Si substrates were used. The calculated projected range was 7.1 nm from the SiO_2 surface, and the projected standard deviation was 2.0 nm. Subsequent thermal annealing in an N_2 ambient at 900 °C for 10 min was carried out with an AG Heatpulse 210 system. During annealing, the samples were positioned face down on a virgin Si substrate which was placed on the system's graphite susceptor.

We studied the size and position of Sb nanocrystals by transmission electron microscopy (TEM). We used a Hitachi H-9000NAR (for cross-sectional views) and an H-9000UHR (for planar views) that operated at 300 kV.

To obtain current–voltage ($I-V$) characteristics, we created diode structures. An Al gate with circular areas of 3.1×10^{-4} cm^2 was formed on the above SiO_2 layer structure using conventional photolithography. The back contact was also an Al electrode. $I-V$ measurements were made with an HP 4156A semiconductor parameter analyzer.

III. RESULTS AND DISCUSSION

A. Microstructure of Sb nanocrystals

Figure 1 shows a cross-sectional TEM micrograph of the samples after ion implantation and subsequent thermal annealing. The thickness of the thermally grown SiO_2 layers shown in Fig. 1 is about 15 nm for (a) and (b) and 20 nm for (c). In all cases, Sb precipitated in the SiO_2 layer and nanoscale Sb dots were formed.

Figure 1(a) shows the results of using a dose of 1×10^{15} ions/cm^2. It can be seen that Sb dots with a diameter of about 2 nm were formed. Interestingly, the Sb dots were incorporated at the same depth in the SiO_2 layer (about 5 nm from the SiO_2/Si interface).

In the sample with a higher dose (5×10^{15} ions/cm^2), larger spherical Sb dots with good size uniformity can clearly be seen in the SiO_2 layer [Fig. 1(b)]. The diameter of the dots is around 3–4 nm. The formation of larger dots when a higher dose was used suggests that it might be possible to control the dot size by changing the dose of implantation ions. Moreover, it can be observed that most Sb dots exist at the same depth in the SiO_2 layer (about 10 nm from the SiO_2 surface). The depth of the dot center is almost the same as for the Sb dots in Fig. 1(a). This depth is also close to the calculated peak position of the as-implanted Sb distribution (7.1 nm from the SiO_2 surface). It should be noted that there are a few Sb regions in the SiO_2 layer at the SiO_2/Si interface (marked by arrows). It is assumed that

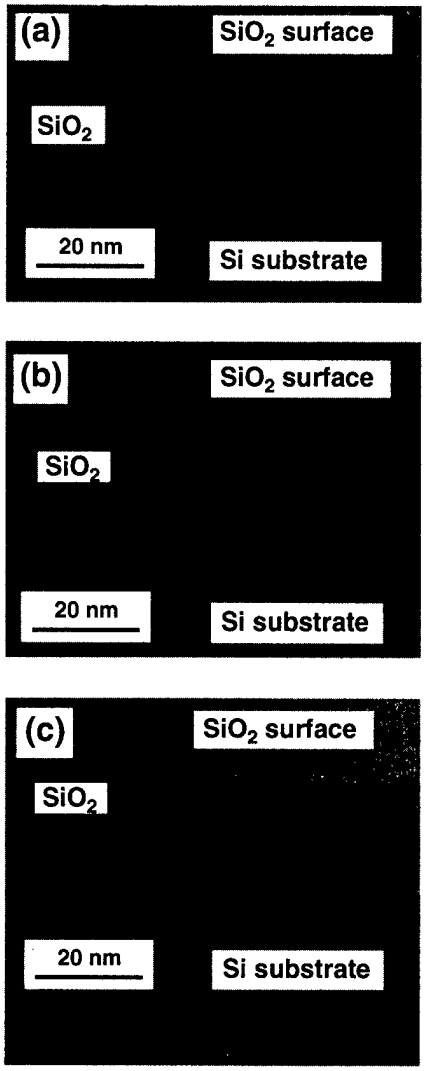


FIG. 1. Cross-sectional TEM micrograph of formed Sb nanocrystals in a thin, thermally grown SiO_2 layer. The thickness of the SiO_2 layer and the dose of implanted Sb are (a) 15 nm and 1×10^{15} ions/cm^2, (b) 15 nm and 5×10^{15} ions/cm^2, and (c) 20 nm and 5×10^{15} ions/cm^2, respectively. In (b), dots grown at the SiO_2/Si interface are marked by arrows. In (c), some of the smaller nanocrystals are marked by arrows.

these Sb atoms came from the Si substrate during thermal annealing. The existence of Sb in the substrate can probably be attributed to the tail of the as-implanted profile. A clear lattice image shows that these Sb regions are crystalline, which suggests that the epitaxial growth occurred on the Si substrate. The dark contrast region in the Si substrate near the SiO_2/Si interface may have been a result of the stress caused by the lattice mismatch between the Sb crystal region and the substrate.

After we changed the SiO_2 thickness (20 nm), two types of dots with different sizes and depth positions [Fig. 1(c)] could be observed even though the dose is the same as in the case of Fig. 1(b). One type of dots was larger (around 6 nm in diameter) and existed around the center of the SiO_2 layer (about 10 nm from the SiO_2 surface). The other type of dots was smaller (about 2 nm in diameter) and existed at a depth

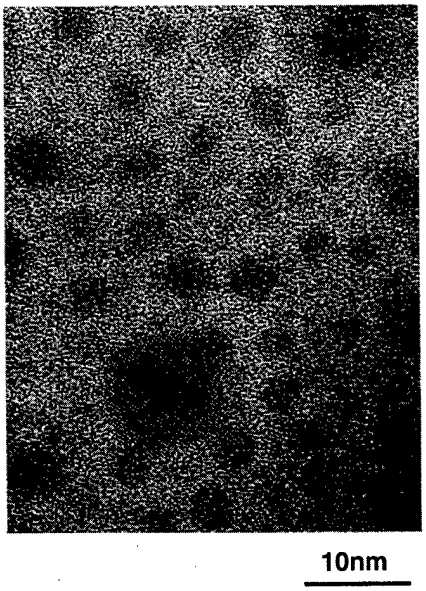


FIG. 2. Plan-view TEM micrograph of the Sb nanocrystals from Fig. 1(b). The comparatively large dot marked by an arrow corresponds to the crystal grown at the SiO_2/Si interface that was shown in Fig. 1(b).

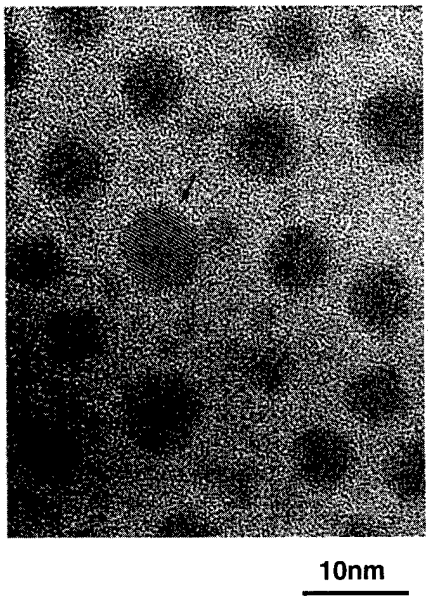


FIG. 3. Plan-view TEM micrograph of the Sb nanocrystals shown in Fig. 1(c). Clear lattice images in the dot (marked by an arrow) indicate that the Sb dots are crystalline.

of about 5 nm from the SiO_2/Si interface. It also seemed that there were more larger dots than the smaller dots.

Figure 2 shows a plan-view TEM micrograph of the Sb dots from Fig. 1(b). A clear contrast between the formed Sb dots and the SiO_2 layer can be observed. Most of the Sb dots are isolated from their neighbors and their diameter is around 3–4 nm. The density of the Sb dots in the plane is 1.3×10^{12} dots/cm^2. The dots are somewhat randomly distributed throughout the plane, yet their size is very uniform. The coverage of the entire Sb region is about 12%. The dots grown at the SiO_2/Si interface, which will be described in the following, were excluded from this coverage estimation. Besides these dots, a comparatively large dot (marked by an arrow) was observed. The shape of this dot was nearly square. It had a side length of about 10 nm, which is close to the length of the left dot which was grown at the SiO_2/Si interface shown in Fig. 1(b). Figure 2 shows a clear lattice image in the dot region. Considering the large size and the low density of the dot, we assume that this large dot corresponds to the dot which was grown at the SiO_2/Si interface shown in Fig. 1(b).

The size distribution of the Sb dots from Fig. 2 is shown in Fig. 4. The average diameter of 142 measured dots was 3.3 nm, with a standard deviation of 1.0 nm. The dots grown at the SiO_2/Si interface are excluded from the statistics. This size uniformity is excellent and comparable with that of Sn nanocrystals in a thin thermally grown SiO_2 layer (average diameter of 4.2 nm with a standard deviation of 1.0 nm), which were formed with the same technique as described in this article, by low-energy ion implantation.[7–9] The excellent uniformity appears for example in comparison to Ag nanocrystals[12] that were formed in a thick glass via ion implantation at rather high energy (60 keV) with a dose of 4×10^{16} ion/cm^2. These nanocrystals showed a much larger standard deviation, 3.7 nm (for an average diameter of 4.2 nm). Moreover, even Si dots that were artificially fabricated by using EB lithography and subsequent dry and wet etching do not have such an excellent size uniformity (a standard deviation of 3.4 nm for an average diameter of 9.8 nm).[1]

Figure 3 shows a plan-view TEM micrograph of the Sb dots from Fig. 1(c). Clear lattice images in the dot (marked by an arrow) indicate that the Sb dots are crystalline. The image reveals a lattice spacing of 3.09 Å. Other TEM micrographs showed lattice spacings of 3.74 and 2.24 Å. These results suggest that the formed nanocrystals consist of metal Sb since the lattice spacings of 3.09, 3.74, and 2.24 Å are almost identical to that of the (012), (003), and (104) plane of metal Sb, respectively. Moreover, most of the Sb nanocrystals are isolated from their neighbors. The density of the Sb dots in the plane is 1.4×10^{12} dots/cm^2. The dots are also randomly distributed throughout the plane. The coverage of the entire Sb region is about 26%. The difference between the coverage for this sample (26%) and that for the sample from Fig. 1(b) (12%) occurred probably since the dots grown at the SiO_2/Si interface were excluded in the coverage estimation for the latter sample.

The size distribution of the Sb dots from Fig. 3 is shown in Fig. 5. The total number of measured dots was 160. Apparently there are two kinds of Sb nanocrystals with differing diameter. One of these types of Sb nanocrystals has a diameter around 2 nm (an average diameter of 2.0 nm) and the other a diameter around 5–6 nm (an average diameter of 5.6 nm). Both types of nanocrystals have good uniformity (a standard deviation of 0.5 nm for the former type and of 1.0 nm for the latter type). There are more nanocrystals of the latter type than of the former type. These facts are consistent with those described in Fig. 1(c).

Summarizing the above results, it can be said with respect

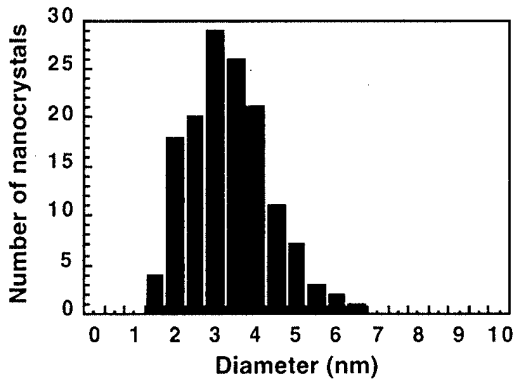


FIG. 4. Size distribution of the Sb nanocrystals shown in Fig. 2. The dots grown at the SiO_2/Si interface are excluded from the statistics.

to the thinner SiO_2 layer (15 nm) that the formed dots have excellent size uniformity (a standard deviation of 1.0 nm for an average diameter of 3.3 nm, as shown in Fig. 4). The depth of the dots is also uniform [Fig. 1(b)]. For the thicker SiO_2 layer (20 nm), two types of dots with differing diameter (Fig. 5) and depth [Fig. 1(c)] are formed. The larger dots have diameters around 6 nm and show good uniformity (a standard deviation of 1.0 nm). They were also formed in the right depth (about 10 nm from the SiO_2 surface). The smaller dots have diameters around 2 nm and good uniformity (standard deviation of 0.5 nm). Their formation depth is also uniform (about 5 nm from the SiO_2/Si interface).

One possible reason for the size and depth uniformity of Sb nanocrystals in thin SiO_2 layers is the fact that a narrow as-implanted Sb distribution is achieved with low-energy ion implantation, which leads to an almost constant concentration of Sb only in the narrow region. This narrow Sb distribution can be sufficiently preserved during thermal annealing due to the slow diffusion of Sb in the SiO_2 region. Since there is a strong correlation between nanocrystal size and as-implanted concentration,[13] the narrow profile will lead to excellent size and depth position uniformity. Another possible reason for uniformity is that a compressive strain exists in the SiO_2 layer near the SiO_2/Si interface (transition region) due to a density mismatch.[14] The formation of Sb nanocrystals near the transition region might have the effect

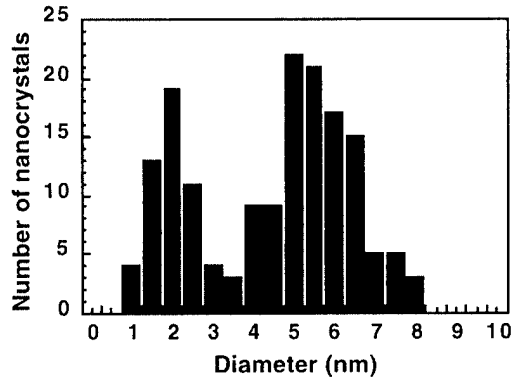


FIG. 5. Size distribution of the Sb nanocrystals shown in Fig. 3.

that the compressive strain is reduced. This would lead to a nanocrystal size and depth appropriate for minimizing the strain energy.

The individual contribution of the above two factors to the uniformity of size and depth can be examined by changing the thickness of the SiO_2 layer while keeping the ion implantation conditions identical. If the peak position of the as-implanted Sb profile was crucial for the Sb precipitation, the depth of Sb nanocrystals from the SiO_2 surface would remain constant. It would be close to the peak position of the as-implanted profile despite a changed thickness of the SiO_2 layer. A comparison of the cross-sectional TEMs for Sb nanocrystals in Fig. 1(b) and in Fig. 1(c) shows that the depth of nanocrystal formation shown in Fig. 1(b) from the SiO_2 surface is almost the same as that for the larger Sb nanocrystals shown in Fig. 1(c). This is observed in spite of the change in the thickness of the SiO_2 layer. These depths are also close to the peak position of the as-implanted profile. Since Sb implantation conditions are identical for both SiO_2 layers, the above results clearly indicate that the peak position of the as-implanted Sb profile is crucial for the Sb precipitation. This suggests that the effect of the narrow as-implanted profile contributes significantly to the size and depth uniformity of Sb nanocrystals in the SiO_2 layers.

Contrary to this, if the compressive strain contributed to the size and depth uniformities, it could be expected that the depth of Sb nanocrystals from the SiO_2/Si interface remained constant regardless of a change of the SiO_2 thickness. This is because it could be expected that Sb always precipitated near transition regions. The depth of the central part of the smaller Sb nanocrystals in Fig. 1(c), which is well separated from the as-implanted peak position, is close to that of the nanocrystals shown in Fig. 1(b). The depths of the two types of Sb nanocrystals are constant with respect to the SiO_2/Si interface. This suggests that the compressive strain also contributes to the size and depth uniformity of Sb nanocrystals.

This means that both the narrow as-implanted profile and the compressive strain near the SiO_2/Si interface contribute to the size and depth uniformity of Sb nanocrystals. Single peaks in the size and depth distribution, which show good uniformity, as shown in Fig. 4 and Fig. 1(b), are considered to be due to the fact that the peak of the narrow as-implanted profile is close to the transition region.

The above result differs from the result for Sn nanocrystals that were formed with the same technique as explained in this study.[9] The main reason for the size and depth uniformity of Sn nanocrystals in a thin thermally grown SiO_2 layer formed via low-energy ion implantation appears to be the compressive strain near the SiO_2/Si interface. There seems to be no contribution from the narrow as-implanted profile. The reason for the difference in the nanocrystal formation mechanism for Sb and for Sn has not been clarified yet. However, we assume that the difference of their diffusion constant in SiO_2 or their association energy may be the reason.

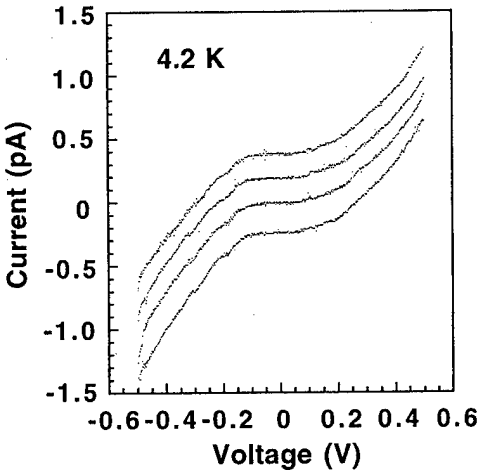


FIG. 6. $I-V$ characteristics of diodes using the structure shown in Fig. 1(b) at 4.2 K. The curves are offset by 2×10^{-13} A to provide clarity.

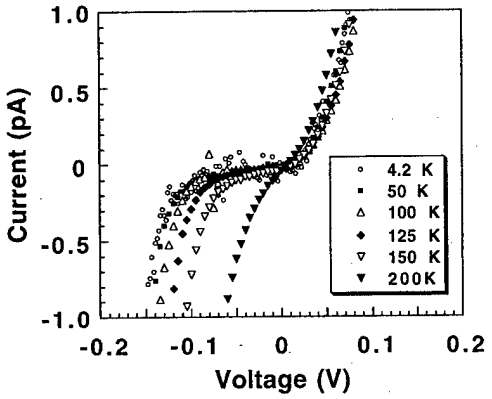


FIG. 7. Temperature dependence of $I-V$ characteristics for the diode shown in Fig. 1(c).

B. Electrical properties

Figure 6 shows the $I-V$ characteristics of diodes using the sample structures shown in Fig. 1(b) at 4.2 K. To confirm the reproducibility of results, $I-V$ characteristics were measured for multiple samples. The curves of the samples were offset by 2×10^{-13} A to provide clarity. It can be observed that almost identical $I-V$ characteristics were obtained for all samples. A Coulomb blockade region of about 0.15 V around 0 V can be seen as well. Clear steplike $I-V$ characteristics can also be noticed, which indicates the presence of a Coulomb staircase. The Coulomb staircases have a periodicity of about 0.15 eV, which coincides well with the Coulomb blockade region. We estimated the charging energy of the nanocrystals as explained in the following. Although the thickness of the SiO_2 layer between the Al gate and the Sb nanocrystals is slightly larger than that between the Si substrate and the Sb nanocrystals [Fig. 1(b)], we assume in a first approximation that the thickness is equal. Under this assumption, the values of the corresponding capacitance (C_1 and C_2) are the same: $C_1=C_2=C$. According to the Coulomb blockade theory, the value of the Coulomb blockade region (0.15 V) equals e/C, where e is the electron charge. Since the total capacitance C_Σ of the Sb nanocrystal is approximately $C_1+C_2=2C$, the charging energy of an Sb nanocrystal ($e^2/2C_\Sigma=e^2/4C$) is 0.04 eV. A charging energy of 0.11 eV is estimated based on the self-capacitance of the Sb nanocrystals shown in Fig. 1(b) (diameter of 3.3 nm). The difference between the experimentally obtained charging energy and that estimated based on the self-capacitance is probably due to the finite distances from the Sb nanocrystal to the Si substrate, to the Al gate, and to the other Sb nanocrystals.

On the other hand, Fig. 7 shows a Coulomb blockade region of about 0.1 V around 0 V for the sample shown in Fig. 1(c) at 4.2 K. As discussed previously, there are more larger nanocrystals (with a diameter of around 5–6 nm) than smaller nanocrystals (with a diameter of around 2 nm) in the sample shown in Fig. 1(c). Therefore, we estimated the charging energy of the larger nanocrystals. A charging energy of 0.07 eV is estimated based on the self-capacitance of the larger Sb nanocrystal that is shown in Fig. 1(c) (with a diameter of 5.6 nm). Since the thickness of the SiO_2 layer between the Al gate and the larger Sb nanocrystals and that between the Si substrate and the larger Sb nanocrystals are almost equal [Fig. 1(c)], the values of the corresponding capacitance (C_1 and C_2) are the same: $C_1=C_2=C$. Following the same analysis as explained above, it can be calculated that the charging energy of the Sb nanocrystal ($e^2/2C_\Sigma$) is 0.025 eV. The difference between the experimentally obtained charging energy in this case and that estimated based on the self-capacitance is also considered to be due to the finite distances from the Sb nanocrystal to the Si substrate, to the Al gate, and to the other Sb nanocrystals.

Since the size of the Sb nanocrystals (with a diameter of 3.3 nm) shown in Fig. 1(b) is smaller than for the nanocrystals in Fig. 1(c) (with a diameter of 5.6 nm), the charging energies of the Sb nanocrystal in Fig. 1(b) are larger than for the nanocrystals in Fig. 1(c), which is concluded from the capacitance estimation of the Sb nanocrystals. This is consistent with the experimentally obtained results reported above.

The overall current levels are consistent with those predicted by an electron direct tunneling model for an SiO_2 barrier based on parameters derived from the TEM images (Fig. 1). The current density J for direct tunneling can be evaluated quantitatively using the equation[15] $J=J_0\,[\varphi\exp(-A\varphi^{1/2})]$, where $J_0=e^2/(2\pi ht_{ox}^2)$, $A=4\pi t_{ox}(2em^*)^{1/2}/h$, $\varphi=\phi_B-V_{ox}/2$. Here, t_{ox} is the oxide thickness, h the Plank constant, V_{ox} the voltage across the oxide, $e\phi_B$ the barrier height at the Al/SiO_2 interface, and m^* is the effective mass of the tunneling electron. We assumed, based on the TEM images (Fig. 1), that $t_{ox}=5$ nm and we also assumed that $m^*=0.35\,m_0$.[15] We concluded therefore that $J=3.1\times10^{-5}$ A/m^2 when $e\varphi=1.5$ eV. Here, m_0 is the electron rest mass. Considering the electrode area $S=3.1\times10^{-8}$ m^2 and the areal coverage of Sb nanocrystals (about 12%), the current I for direct tunneling is obviously 2.0×10^{-13} A. This agrees well with the experimental result in Fig. 6 for the sample in Fig. 1(b) at the voltage of about 0.2 V. Since the voltage of 0.2 V corresponds to $V_{ox}=0.1$ V, the condition

$e\varphi = 1.5$ eV implies that $e\phi_B = 1.55$ eV. The current level for the sample in Fig. 1(c) also matches the calculated level when assuming a slightly lower SiO_2 barrier height. The possible reason for these low SiO_2 barrier heights is the defect states that were caused by ion implantation or the effect of Sb inclusion in the barrier region.

Figure 7 also shows the temperature dependence of the $I-V$ characteristics for the sample in Fig. 1(c). Almost the same Coulomb blockade regions (0.1 V) around 0 V as for 4.2 K are observed up to 100 K. From about 125 K upwards, the Coulomb blockade region becomes smaller. However, at 150 K the Coulomb blockade region remains even. Unfortunately, at 200 K, no Coulomb blockade region can be observed.

IV. CONCLUSION

We have formed Sb nanocrystals in thin, thermally grown SiO_2 layers using low-energy ion implantation which was followed by thermal annealing. The favorable size and depth uniformity of the formed Sb nanocrystals are considered to be a result of both the narrow as-implanted Sb distribution and the compressive strain near the SiO_2/Si interface. This reason differs from the reason for the size and depth uniformity of the Sn nanocrystals in thin, thermally grown SiO_2 layers that occurs when using the same technique as outlined in this study. For Sn nanocrystals, only the compressive strain near the SiO_2/Si interface contributes to the uniformity. We found that it was possible to control the size of the Sb nanocrystal by changing the dose of the implanted Sb. The $I-V$ characteristics of the diodes which had Sb nanocrystals with an average diameter of 3.3 nm show a clear Coulomb blockade region around 0 V and a Coulomb staircase at 4.2 K. Identical $I-V$ characteristics were obtained for all the samples measured, indicating an excellent reproducibility. Almost the same Coulomb blockade region as that at 4.2 K was observed up to a temperature of 100 K even for a diode which had larger Sb nanocrystals (with an average diameter of 5.6 nm). The experimentally obtained charging energies for Sb nanocrystals were consistent with the results obtained from the capacitance estimation of the Sb nanocrystals. The measured overall current levels are consistent with the values predicted by an electron direct tunneling model for an SiO_2 barrier. These results show that the proposed formation technique for metal nanocrystals is favorable with respect to the high temperature operation of Si-based single-electron devices. A high fabrication throughput of nanoscale structures is another benefit of using this technique.

ACKNOWLEDGMENT

The authors thank Dr. Takashi Ito for his encouragement.

[1]A. Nakajima, H. Aoyama, and K. Kawamura, Jpn. J. Appl. Phys., Part 2 **33**, L1796 (1994).
[2]A. Nakajima, T. Futatsugi, K. Kosemura, T. Fukano, and N. Yokoyama, Appl. Phys. Lett. **70**, 1742 (1997); IEEE International Electron Devices Meeting, San Francisco, 8–11 December 1996, pp. 952–954.
[3]A. Nakajima, T. Futatsugi, K. Kosemura, T. Fukano, and N. Yokoyama, Appl. Phys. Lett. **71**, 353 (1997); IEEE Silicon Nanoelectronics Workshop 1997, Kyoto, 8–9 June 1997, pp. 26–27.
[4]K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai, and K. Seki, IEEE International Electron Devices Meeting, Washington, DC, 5–8 December 1993, pp. 541–544.
[5]A. Nakajima, Y. Sugita, K. Kawamura, H. Tomita, and N. Yokoyama, Jpn. J. Appl. Phys., Part 2 **35**, L189 (1996).
[6]A. Nakajima, Y. Sugita, K. Kawamura, H. Tomita, and N. Yokoyama, J. Appl. Phys. **80**, 4006 (1996).
[7]A. Nakajima, T. Futatsugi, N. Horiguchi, and N. Yokoyama, Appl. Phys. Lett. **71**, 3652 (1997).
[8]A. Nakajima, T. Futatsugi, N. Horiguchi, H. Nakao, and N. Yokoyama, IEEE International Electron Devices Meeting, Washington, DC, 7–10 December 1997, pp. 159–162.
[9]A. Nakajima, T. Futatsugi, H. Nakao, T. Usuki, N. Horiguchi, and N. Yokoyama, J. Appl. Phys. **84**, 1316 (1998).
[10]A. Nakajima, T. Futatsugi, N. Horiguchi, and N. Yokoyama, Jpn. J. Appl. Phys., Part 2 **36**, L1552 (1997).
[11]A. Nakajima, H. Nakao, H. Ueno, T. Futatsugi, and N. Yokoyama, Appl. Phys. Lett. **73**, 1071 (1998); IEEE Silicon Nanoelectronics Workshop 1998, Honolulu, 7–8 June 1998, pp. 11–12.
[12]L. C. Nistor, J. van Landuyt, J. D. Barton, D. E. Hole, N. D. Skelland, and P. D. Townsend, J. Non-Cryst. Solids **162**, 217 (1993).
[13]H. Hosono, H. Fukushima, Y. Abe, R. A. Weeks, and R. A. Zuhr, J. Non-Cryst. Solids **143**, 157 (1992).
[14]F. J. Grunthaner, P. J. Grunthaner, R. P. Vasquez, B. F. Lewis, and J. Maserjian, Phys. Rev. Lett. **43**, 1683 (1979).
[15]M. Hiroshima, T. Yasaka, S. Miyazaki, and M. Hirose, Jpn. J. Appl. Phys., Part 1 **33**, 395 (1994).

Fabrication of a metal nanostructure on the Si(111) surface

D. Rogers[a] and H. Nejoh
National Research Institute for Metals, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

(Received 14 January 1999; accepted 21 May 1999)

A metal nanostructure was fabricated on a Si(111) surface in ultrahigh vacuum. The stages in the fabrication were H passivation of the surface, Pb evaporation, depassivation of selected areas with the tip of a scanning tunneling microscope, and heating. Using this procedure, a nanowire and two contact pads to connect it to external probes were fabricated. At each stage of the process, the surface was imaged with scanning tunneling microscopy. © *1999 American Vacuum Society.* [S0734-211X(99)08604-7]

I. INTRODUCTION

Electronic devices with dimensions at the nanometer scale exhibit novel properties based on quantum effects such as Coulomb blockade,[1] quantum conductance,[2] and quantum interference.[3] The fabrication of devices with these dimensions is problematic when the traditional lithographic techniques are utilized, but the capability of scanning tunneling microscopy (STM) to microscopically position a sharpened metal tip over a surface provides a means by which structures at the nanometer scale can be not only imaged but also fabricated. Examples of STM nanofabrication are deposition of material from a STM tip to a surface by field evaporation[4,5] or the application of laser radiation to the tunneling gap,[6] dissociation of gas phase[7–9] or adsorbed molecules[10] to form patterns of molecular fragments, and exposure of a surface resist by the electrons emitted from a STM tip.[11] A monolayer of H atoms chemisorbed on a Si surface is a promising resist for STM nanofabrication.[12,13] Lines of bare Si atoms on a hydrogen passivated Si(100) substrate have been successfully fabricated by STM,[12] and the patterns of depassivation have been used as templates for oxidation[14] and chemical vapor deposition.[15,16]

We have developed a procedure, based on STM exposure of a hydrogen atom resist, to fabricate metal nanostructures on semiconductor surfaces, and in this article we show how our procedure is applied to the fabrication of a nanowire on a Si(111) surface in ultrahigh vacuum (UHV). The procedure which we followed was to first hydrogen passivate the Si(111) surface, then to evaporate Pb, followed by depassivation of selected areas to form the pattern of the device, and finally a heat treatment. At each stage of the procedure, the surface was imaged with STM. Our method differs from other methods based on H depassivation and metal adsorption[17] in that we depassivate the surface after, not before, metal evaporation. In this way we can easily create large contact pads connecting to the nanowire without having to rely on thermal diffusion of metal atoms across long distances. The contact pads are to be used to connect the nanoscale device to external probes.

[a]Author to whom correspondence should be addressed; electronic mail: drogers@chem.utoronto.ca; present address: Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada.

II. EXPERIMENT

The experimental apparatus consisted of an ultrahigh vacuum chamber that contained a JEOL scanning tunneling microscope (JAFM-4500XT), a leak valve and W filament for dissociating hydrogen gas, and provisions for sample heating and Pb evaporation. The base pressure of the chamber was 6.0×10^{-11} Torr. The STM was capable of operation at sample temperatures from room temperature up to 900 °C. All STM measurements reported in this article were made in the constant current mode. STM tips were fabricated by electrochemically etching a 0.3-mm-diam W wire. Inside the UHV chamber each STM tip was cleaned by bringing it close to a Si surface that was heated to 1250 °C for 30 s. For additional tip cleaning, the tip was scanned over a Si surface at a tip to sample tunneling current of 10 nA and a bias voltage of 8 V on the tip. The tip prepared in this manner was clean enough that after scanning it over a Si(111) 7 $\times$7 surface at a tunneling current of 10 nA and bias voltage of -8 V on the tip [the conditions we used to desorb H atoms from the hydrogen passivated Si(111) surface], no contamination on the surface could be seen by STM.

Pb was evaporated from a resitively heated basket which was formed from a 0.3-mm-diam W wire. The basket had a diameter of about 2.5 mm, and its two W leads were spot welded to a current feedthrough. Pieces of 99.999% pure Pb wire of 0.5 mm diam were loaded into the basket which was subsequently outgassed in UHV by passing a 1.8 A current through the basket for 12 h. A current of 2.9 A through the basket was sufficient to evaporate Pb onto the sample surface. To ensure a uniform evaporation rate, the basket was heated until it glowed red, whereupon the sample was turned to face the evaporator for the required time interval.

Si(111) samples were cleaved from an *n*-type wafer (P doped) with a resistivity of $<0.01\ \Omega$ cm. Each sample had dimensions of 1 mm$\times$7 mm$\times$0.25 mm, and was rinsed in acetone and methanol before being loaded into the UHV chamber through a load lock. Inside the chamber each sample was cleaned by resistive heating: it was first outgassed at 550 °C for 12 h, then the oxide was removed by heating at 850 °C for 10 min, and the carbon was removed by heating at 1250 °C for 2 min. In the final step the sample was slowly cooled from 900 °C to room temperature at a rate of at most

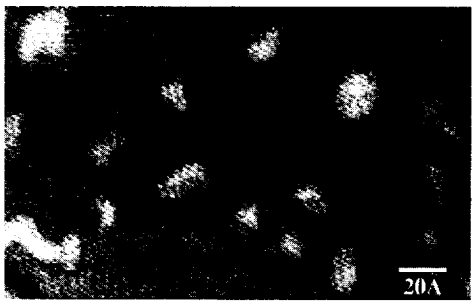


FIG. 1. STM image of the hydrogen passivated Si(111) surface. The image was made with a tip bias voltage of 2.1 V and a tunneling current of 1.0 nA. The horizontal range is 200 Å.

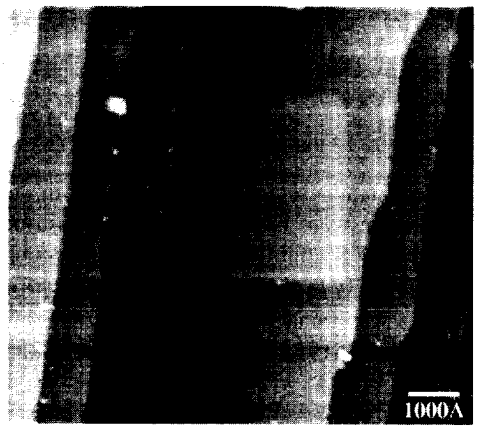


FIG. 2. STM image of a square area of depassivation on the hydrogen passivated Si(111) surface. The depassivated area is brighter than the surrounding surface and has dimensions of 330 nm×330 nm. The image was made with a tip bias voltage of −3.0 V and a tunneling current of 0.20 nA. The horizontal range is 890 nm.

1 °C/s. After this procedure, STM measurements showed that the surface consisted of (111) oriented terraces, which contained the 7×7 reconstruction, separated by atomic steps of height equal to 3.14 Å, the double atomic layer distance of Si(111). The experiments described in this article were undertaken on two different samples cleaved from the same wafer; on one of the samples the average terrace width was 98 nm, and on the other sample the average terrace width was 280 nm, as determined by STM measurements.

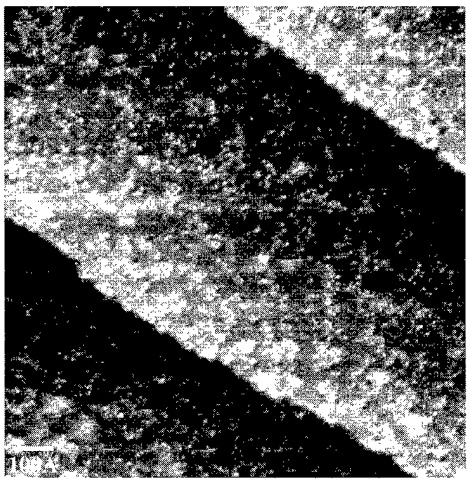
Sample temperatures above 600 °C were measured with an optical pyrometer. Below 600 °C, the sample temperature was inferred from a logarithmic plot of the temperature versus heating current.[18] The plot was based on a measurement of the heating currents at which the two peaks for the desorption of H from a hydrogen passivated Si(111) surface were observed on an ion gauge. The two desorption peaks correspond to desorption from the SiH_2 and SiH_3 species at 420 °C and from the SiH species at 540 °C.[19]

III. HYDROGEN PASSIVATION AND DESORPTION

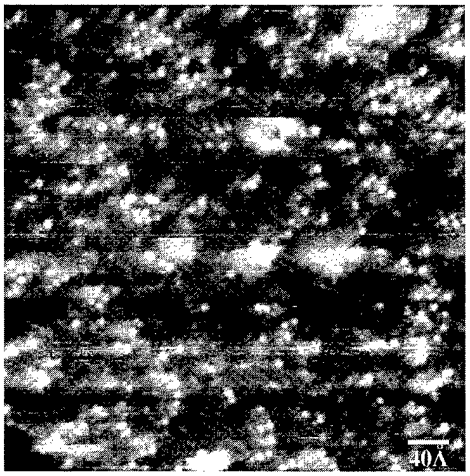
The hydrogen passivated Si(111) surface was formed by heating the 7×7 surface to 300 °C during exposure of the surface to a flux of hydrogen atoms formed by the dissociation of hydrogen gas on a tungsten filament heated to 1900 °C. A hydrogen gas pressure of 4×10^{-5} Torr and time interval of 100 s was sufficient to remove all traces of the 7×7 pattern of Si adatoms, yielding a surface that consisted of a flat substrate upon which clusters of atoms were arrayed as seen in Fig. 1. As shown previously by Owman and Mårtensson,[20] the flat substrate is the underlying rest atom layer of the 7×7 surface after removal of the Si adatoms and stacking faults by atomic H. Each dangling bond of the Si atoms of the substrate is passivated by bonding to a hydrogen atom. The Si adatoms that are liberated in this process are collected into the islands that cover the H passivated substrate.

Hydrogen atoms were desorbed from selected areas of the hydrogen passivated surface by applying a high voltage between the STM tip and the sample (−8.0 V on the tip) while scanning the STM tip over the surface in a desired pattern at a demanded tunneling current of 10 nA. The pattern of depassivation thus created could be observed by scanning the STM tip over the same region of the surface with the current and voltage values typical for imaging (e.g., −2 V on the tip, 1.0 nA). Figure 2 is an STM image of the surface after depassivation. The bright square area of dimensions 330 nm×330 nm is the region that was depassivated by the STM tip. Double atomic layer high steps run from the top to bottom of this image, and the image shows that there was no difficulty in depassivating the surface across these steps. Similar square regions of depassivation were later used in the construction of contact pads for the connection of the nanowire to external probes. To form the depassivated square, the STM tip made 1024 forward and reverse scan lines in a 330 nm×330 nm area. The spacing between traces was 3.3 Å. The injected charge density over the scan area was 31 pC/nm^2. If each trace of the tip is considered separately, the injected charge density along a scan line was 10 pC/nm and the writing speed was 980 nm/s. The depassivated area appears brighter than the rest of the surface since the removal of hydrogen atoms from the Si atoms creates excess surface state density over the depassivated atoms, allowing the required tunneling current between the tip and sample to be met by a tip which is farther from the surface when over the depassivated region than when over the passivated region.

Higher magnification STM images of the depassivated area are shown in Fig. 3. The depassivated surface consisted of small domains of parallel ridges oriented along one of the three $\{1\bar{1}0\}$ directions on the surface, single adatoms which bonded on top of the ridges, and atoms that were clustered into islands. The ridges within the domains were separated by 6.7 Å. This structure is consistent with the 2×1 reconstruction of the Si(111) surface formed by cleavage in UHV.[21,22] According to the π bonded chain model proposed by Pandey,[23] the formation of the 2×1 surface involves a drastic change in the bonding topology of the surface atoms, moving the dangling bonds to nearest neighbor positions where they significantly interact by forming chains of π bonds. The ridges observed in Fig. 3 are the π bonded chains. Since the number of atoms in the top bilayer and adatom layer of the 7×7 surface (2.08/1×1 unit cell) is



(a)



(b)

FIG. 3. High magnification STM images of the depassivated surface. The images show the 2×1 reconstruction with additional Si atoms bonding on top of the 2×1 ridges. The images were made with a tip bias voltage of −3.0 V and a tunneling current of 0.20 nA. The horizontal ranges are 110 nm (a) and 45 nm (b).

greater than the number on the 1×1 and 2×1 surfaces (2/1×1 unit cell),[24] those additional surface atoms which bonded on top of the 2×1 structure or formed surface islands may include Si adatoms of the 7×7 surface that were not removed from the surface during H passivation. These are the Si adatoms which formed the islands seen in Fig. 1.

The 2×1 reconstruction was also formed by Becker *et al.* by STM depassivation of the hydrogen passivated Si(111) 1×1 surface.[25] In that work the STM tip was pulled back about 100 nm from the surface and 60–70 V was applied between the tip and sample resulting in a region of 2×1 reconstruction extending over an 800-nm-wide region of the surface. In our experiment, by moving the tip closer to the surface and using lower bias voltages, we were able to more precisely control the location of the 2×1 reconstruction. By scanning the tip over just a few scan lines, instead of the complete 330 nm×330 nm area, we were able to draw lines of depassivation similar to the lines formed on the H-passivated Si(100) surface.[12] These results will be discussed further in Sec. V.

In one stage of the fabrication of the nanowire, the surface that contained both passivated and depassivated regions was heated to 350 °C. To test the stability of the pattern of depassivation at this temperature, a passivated surface with a square area of depassivation of dimensions 330 nm×330 nm underwent a series of 30 s heat treatments at 300, 350, and 375 °C. The pressure in the UHV chamber did not increase at any time during the heating, and thus we conclude that H did not desorb from the surface. In order to image the same depassivated region before and after heating, the sample was kept in the STM stage during the heat treatments. After heating, the depassivated square could be found within the scan range of the STM piezo tube; no changes in the coarse x and y positions of the sample were required. Even for the highest temperature of 375 °C, no change in the shape of the perimeter of the square could be seen by STM, leading us to conclude that H did not diffuse from the passivated to the depassivated region. Although the perimeter of the square did not change during heating, the structure of the surface within the square did undergo a change.

A STM image of the surface within the depassivated region after heating for 30 s at 300 °C is shown in Fig. 4(a), and details of the surface structures found are shown in the higher magnification images in Figs. 4(b) and 4(c). In these images the dark regions containing parallel ridges are the 2×1 reconstructed regions and the bright regions contain Si adatom structures. Most of the depassivated surface maintained the 2×1 reconstruction after heating, but the isolated Si adatoms and Si islands on the depassivated surface before heating (Fig. 3) combined during heating to form more ordered surface structures. Corner holes of the dimer-adatom-stacking fault (DAS) structures[26] can be clearly seen within the bright regions of Si adatoms. One of the corner holes is indicated in Fig. 4(b). We also observed incomplete DAS 5×5 unit cells, and 2×2 and $c2\times4$ adatom structures within the depassivated regions. These structures are outlined in Figs. 4(b) and 4(c). Holes in the 2×1 layer were observed to be preferentially located adjacent to the adatom structures. The proximity suggests that atoms are taken from the 2×1 layer in order to build the adatom structures.

When Feenstra and Lutz heated a Si(111) 2×1 surface that was formed by cleavage of a Si wafer in UHV, they observed 200-nm-wide domains of 2×1 reconstruction coexisting with 200-nm-wide domains of 5×5 reconstruction.[24] After heating our surface we observed the adatom structures beginning to form throughout the 2×1 surface, and did not observe such wide domains of the 2×1 or adatom structures. Our observations are accounted for by the closely spaced Si adatoms and Si clusters scattered over the surface before heating (Fig. 3) which acted as nucleation sites for the adatom structures.

IV. Pb EVAPORATION

To evaporate Pb onto the hydrogen passivated surface, the sample was removed from the STM stage and positioned in

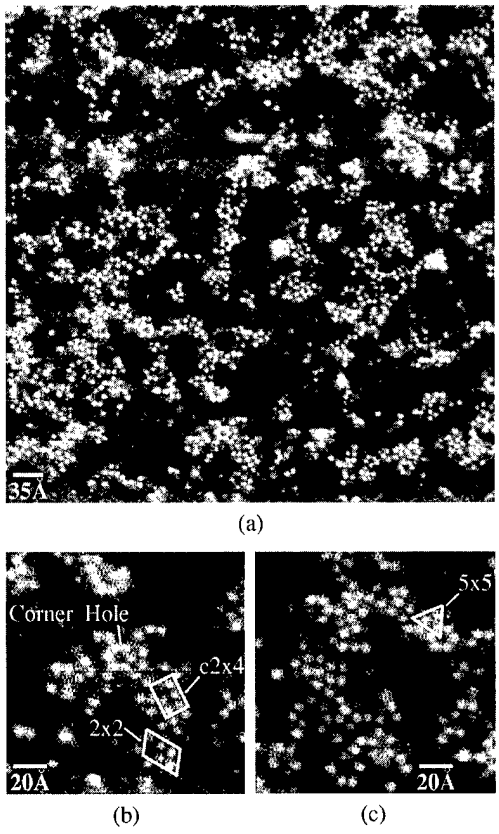


FIG. 4. STM images of the depassivated surface after heating at 300 °C for 30 s. The dark areas contain the parallel rows of the 2×1 reconstruction and the bright areas contain the Si adatom structures. Some of the adatom structures are outlined and a corner hole is indicated in the high magnification images in (b) and (c). The images were made with a tip bias voltage of −3.0 V and a tunneling current of 0.20 nA. The horizontal ranges are 570 (a), 140 (b), and 140 Å (c).

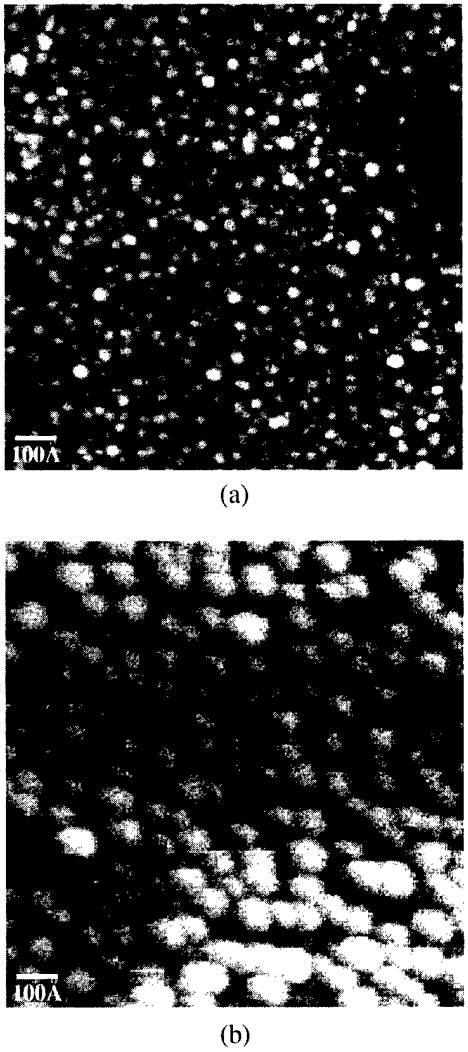


FIG. 5. STM images of the H passivated surface with evaporated Pb before (a) and after (b) heating at 350 °C for 47 s. The tip bias voltages and tunneling currents were −3.0 V and 0.20 nA for (a) and −2.0 V and 0.30 nA for (b). The horizontal ranges are 110 nm.

front of the heated Pb evaporator for 5 s. Figure 5(a) is a STM image of the surface after Pb evaporation. The STM image shows clusters of Pb covering the surface. Heating the surface after evaporation was found to create larger and hence more widely separated Pb clusters on the surface. This can be seen in the STM image of Fig. 5(b) which was recorded at room temperature after a 47 s heat treatment at 350 °C. The Pb cluster width distribution measured from the room-temperature STM images is shown in Fig. 6 as a function of annealing temperature after evaporation. The average cluster width increased from 32 Å on the surface before any heat treatments to 65 Å on the surface that underwent a 30 s heat treatment at 250 °C, and to 89 Å on the surface that underwent a 47 s heat treatment at 350 °C. Heating the surface is the final step in the fabrication of the nanowire.

V. FABRICATION OF THE NANOWIRE

The nanowire and two contact pads connecting to it were formed by desorption of hydrogen from selected areas of the H passivated surface after evaporation of Pb. The H desorption was accomplished by scanning the STM tip over the surface with a bias voltage of −8 V on the tip and a demanded tunneling current of 10 nA, the same procedure used to remove H from the surface before Pb evaporation, as discussed earlier. Figure 7(a) is a STM image of the nanowire and its two connecting contact pads. The contact pads were formed by scanning the STM tip over two 330 nm×330 nm areas and the nanowire between the two contact pads was formed by tracing the tip over four forward and reverse scan lines with a spacing of 3.3 Å between the traces. The nanowire is 20 nm wide and 770 nm long. The contact pads have dimensions of approximately 350 nm×420 nm and 370 nm×260 nm. In the fabrication of this structure, the STM tip was moved 400 nm to the left and scanned over the surface to form the contact pad on the left, and later it was moved 800 nm to the right and scanned over the surface to form the contact pad on the right. The large lateral movement of the STM tip produced piezo creep which we observe as the imperfect square shape of the two contact pads. The pits which are seen covering the surface were formed by etching of the surface by atomic hydrogen during H passivation. The depassivated region appears darker than the passivated region

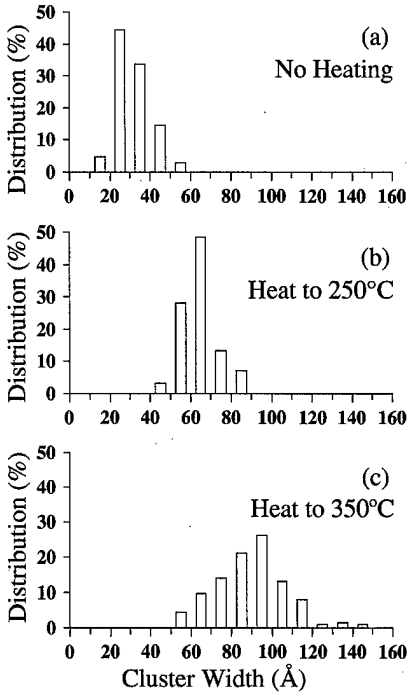


FIG. 6. Distribution of Pb clusters on the H passivated surface over cluster width for the surface before heating (a), after heating at 250 °C for 30 s (b), and after heating at 350 °C for 47 s (c).

since the removal of H allows the Pb atoms in the clusters to bond to the silicon atoms on the bare silicon surface, decreasing the height of the clusters and spreading a film of Pb over the depassivated region.

Depending on the position of the STM tip during depassivation, the electrons emitted by the tip will be either directly incident on the Pb clusters or will pass between neighboring clusters and be directly incident on the H passivated surface. For the electrons which are directly incident on the H passivated surface, the H desorption mechanism is direct electronic excitation of the Si–H bond leading to a repulsive

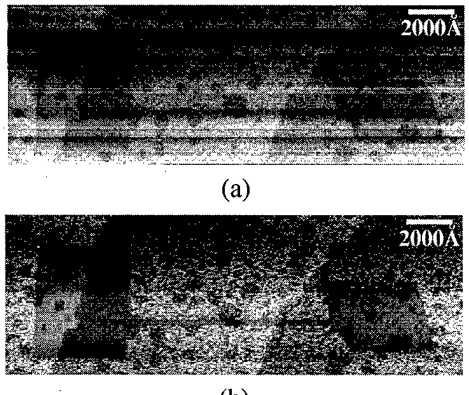


FIG. 7. Nanowire and the two contact pads before (a) and after (b) heating at 350 °C for 47 s. The horizontal lines in image (a) are the result of noise during the STM measurement. The tip bias voltages and tunneling currents were −3.0 V and 0.20 nA for (a) and −2.0 V and 0.30 nA for (b). The horizontal ranges are 1980 nm.

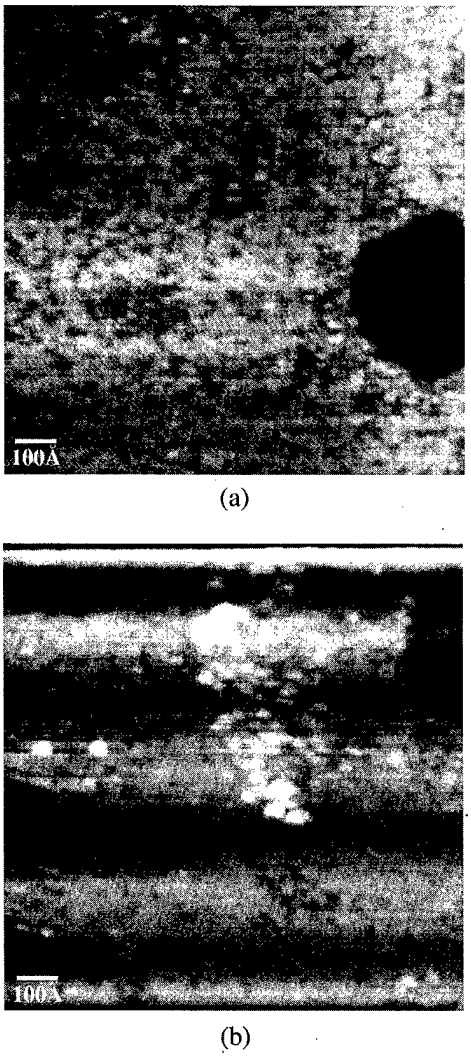
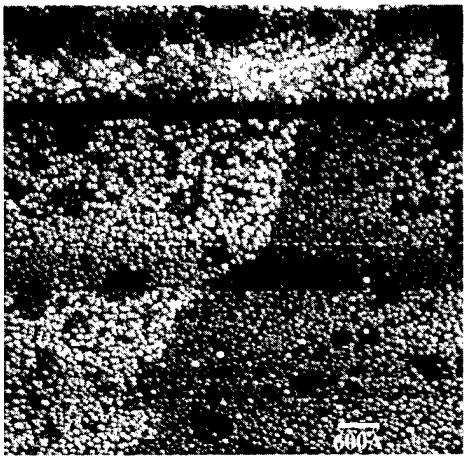


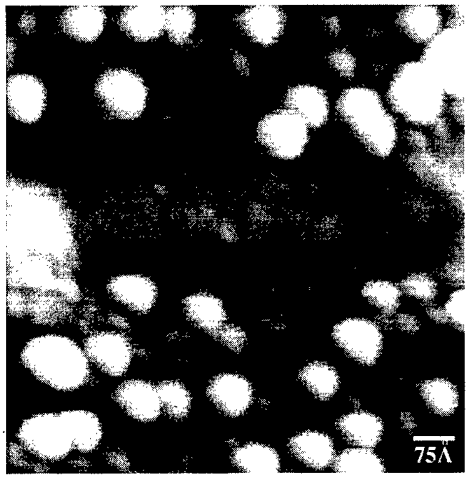
FIG. 8. STM images of the Pb film within the depassivated region before (a) and after (b) heating at 160 °C for 30 s. The same etch pit is visible at the right of each image. The oscillation of the brightness from the top to bottom of image (b) is an artifact of the image processing. The images were made with a tip bias voltage of −3.0 V and a tunneling current of 0.20 nA. The horizontal ranges are 110 nm.

excited state followed by desorption.[27] For the electrons which are directly incident on the Pb clusters, we propose a H desorption mechanism mediated by the Pb clusters. In this mechanism, electron-induced excitation of the Pb cluster is followed by energy transfer to the Si–H bond. This finally results in H desorption by multiple vibrational excitation of the Si–H bond, as proposed for H desorption induced by a low bias voltage ($|V|<5$ V) between a STM tip and a H passivated Si surface.[28]

The structure of the Pb film within the depassivated region can be seen in the STM image of Fig. 8(a). The film was continuous but it contained many vacancies which are seen as the dark areas in the image. We could not observe any atomic structure within the smooth parts of the Pb film. To make the film more continuous the surface was heated to 160 °C for 30 s. After cooling to room temperature the same depassivated region was imaged with STM and found to con-



(a)



(b)

FIG. 9. STM images of the nanowire after heating at 350 °C for 47 s. An atomic step runs from the top to the bottom of the image in (a). Clusters of Pb atoms can be seen on the H passivated surface which surrounds the nanowire. The images were made with a tip bias voltage of −2.0 V and a tunneling current of 0.30 nA. The horizontal ranges are 670 (a) and 84 nm (b).

tain fewer vacancies [Fig. 8(b)]. Heating the surface creates a more continuous Pb film within the depassivated region, possibly increasing its conductivity.

After heating the surface containing the nanowire at 350 °C for 47 s and cooling to room temperature, the nanowire and surrounding regions were imaged by STM. The STM image after heating in Fig. 7(b) shows that the dimensions of the nanowire and contact pads were not changed by the heat treatment. The high magnification STM images of the nanowire in Fig. 9 show that the Pb clusters which surround the wire are larger and more widely spaced after the heat treatment while no change can be seen in the structure of the nanowire itself. The increased spacing between the Pb clusters may decrease the conductivity of the region surrounding the nanostructure.

VI. CONCLUSION

A nanowire and two contact pads to connect it to external probes have been fabricated on a Si(111) surface. The stages in the fabrication were H passivation, Pb evaporation, H depassivation, and heating. The final heat treatment creates a more uniform Pb film inside the nanowire and, in the region surrounding the nanowire, it increases the size of the Pb clusters and the spacing between them. Future work in which the amount of Pb evaporated onto the H passivated surface is decreased should lead to even more widely spaced Pb clusters after the heat treatment, possibly resulting in a decreased conductivity in the region surrounding the nanowire.

ACKNOWLEDGMENTS

The authors thank Dr. S. Odasso, Dr. T. Ohgi, and Dr. D. Fujita for generous assistance. The Science and Technology Agency of Japan is gratefully acknowledged for providing financial support.

[1] *Single Charge Tunneling: Coulomb Blockade Phenomena in Nanostructures*, edited by H. Grabert and M. H. Devoret (Plenum, New York, 1992).
[2] B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon, Phys. Rev. Lett. **60**, 848 (1988).
[3] R. A. Webb, S. Washburn, C. D. Umbach, and R. B. Laibowitz, Phys. Rev. Lett. **54**, 2696 (1985).
[4] R. S. Becker, J. A. Golovchenko, and B. S. Swartzentruber, Nature (London) **325**, 419 (1987).
[5] D. Fujita, Q. Jiang, and H. Nejoh, J. Vac. Sci. Technol. B **14**, 3413 (1996).
[6] J. Jersch and K. Dickmann, Appl. Phys. Lett. **68**, 868 (1995).
[7] M. A. McCord, D. P. Kern, and T. H. P. Chang, J. Vac. Sci. Technol. B **6**, 1877 (1988).
[8] E. E. Ehrichs, S. Yoon, and A. L. de Lozanne, Appl. Phys. Lett. **53**, 2287 (1988).
[9] F. Thibaudau, J. R. Roche, and F. Salvan, Appl. Phys. Lett. **64**, 523 (1994).
[10] G. Dujardin, R. E. Walkup, and Ph. Avouris, in *Computations for the Nano-Scale*, edited by P. E. Blöchl, C. Joachim, and A. J. Fisher (Kluwer, Boston, 1993), p. 223.
[11] M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **6**, 293 (1988).
[12] J. W. Lyding, T.-C. Shen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, Appl. Phys. Lett. **64**, 2010 (1994).
[13] M. Schwartzkopff, P. Radojkovic, M. Enachescu, E. Hartmann, and F. Koch, J. Vac. Sci. Technol. B **14**, 1336 (1996).
[14] T.-C. Shen, C. Wang, J. W. Lyding, and J. R. Tucker, Appl. Phys. Lett. **66**, 976 (1995).
[15] D. P. Adams, T. M. Mayer, and B. S. Swartzentruber, Appl. Phys. Lett. **68**, 2210 (1996).
[16] G. C. Abeln, M. C. Hersam, D. S. Thompson, S.-T. Hwang, H. Choi, J. S. Moore, and J. W. Lyding, J. Vac. Sci. Technol. B **16**, 3874 (1998).
[17] T.-C. Shen, C. Wang, and J. R. Tucker, Phys. Rev. Lett. **78**, 1271 (1997).
[18] D. Rogers and T. Tiedje, Phys. Rev. B **53**, R13227 (1996).
[19] G. Schulze and M. Henzler, Surf. Sci. **124**, 336 (1983).
[20] F. Owman and P. Mårtensson, Surf. Sci. **324**, 211 (1995).
[21] R. E. Schlier and H. E. Farnsworth, J. Chem. Phys. **30**, 917 (1959).
[22] R. N. Feenstra, A. P. Fein, and W. A. Thompson, Phys. Rev. Lett. **56**, 608 (1986).
[23] K. C. Pandey, Phys. Rev. Lett. **47**, 1913 (1981); **49**, 223 (1982).
[24] R. M. Feenstra and A. M. Lutz, Phys. Rev. B **42**, 5391 (1990).
[25] R. S. Becker, Y. J. Chabal, G. S. Higashi, and A. J. Becker, Phys. Rev. Lett. **65**, 1917 (1990).
[26] K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi, J. Vac. Sci. Technol. A **3**, 1502 (1985).
[27] T.-C. Shen, C. Wang, G. C. Abeln, J. R. Tucker, J. W. Lyding, Ph. Avouris, and R. E. Walkup, Science **268**, 1590 (1995).
[28] R. E. Walkup, D. M. Newns, and Ph. Avouris, in *Atomic and Nanometer Scale Modification of Materials*, edited by Ph. Avouris (Kluwer, Dordrecht, 1993).

Mechanical properties, stress evolution and high-temperature thermal stability of nanolayered Mo–Si–N/SiC thin films

P. Torri[a)]
Department of Physics, Accelerator Laboratory, P.O. Box 9, FIN-00014 University of Helsinki, Finland

J.-P. Hirvonen
European Commission, Joint Research Centre, Institute for Advanced Materials, 1755 ZG Petten, The Netherlands

H. Kung, Y.-C. Lu, and M. Nastasi
Los Alamos National Laboratory, Los Alamos, New Mexico 87545

P. N. Gibson
European Commission, Joint Research Centre, Institute for Health and Consumer Protection, I-21020 Ispra, Italy

(Received 18 March 1999; accepted 28 May 1999)

A study of the microstructure, thermal stability, nanoindentation mechanical properties, and residual stress evolution of nanolayered Mo–Si–N/SiC thin films as a function of vacuum annealing time and temperature is reported. Multilayers of Mo–Si–N ($MoSi_{2.2}N_{2.5}$) and SiC were deposited by magnetron sputtering from planar $MoSi_2$ and SiC targets onto single crystal silicon wafers. The relative amount of both components was varied (12.5–50 vol. % of SiC) while keeping the bilayer thickness constant (12 nm), or the bilayer thickness was varied (6–24 nm) with constant Mo–Si–N to SiC ratio (25 vol. % of SiC). Mechanical properties were measured by nanoindentation on as-deposited films and films annealed in vacuum at 500 and 900 °C. Microstructure and thermal stability were examined by cross-sectional transmission electron microscopy, glancing angle x-ray diffraction and nuclear resonance broadening. Stress evolution induced by thermal annealing was determined by measuring optically the change in curvature of coated silicon beams. In the as-deposited state, all films exhibited an amorphous microstructure. At 900 °C SiC still remained amorphous, but Mo–Si–N had developed a microstructure where nanocrystals of Mo_5Si_3 were embedded in an amorphous matrix. The interface between Mo–Si–N and SiC was indirectly shown to be stable at least up to 41 h annealing at 1075 °C in vacuum. The potential of Mo–Si–N as a barrier layer against intermixing between nanolayered $MoSi_2$ and SiC at 900 °C has been demonstrated. Hardness, modulus and residual stress followed the volume fraction rule of mixture of both constituents of the nanolayered Mo–Si–N/SiC structure. Consequently, by optimizing the volume fraction of the constituents, zero residual stress on a silicon substrate is possible after annealing. © *1999 American Vacuum Society.* [S0734-211X(99)09404-4]

I. INTRODUCTION

Molybdenum disilicide is a promising material for high temperature (1200–1600 °C) use since it has excellent high temperature oxidation properties, high melting point (2020 °C), moderate density (6.24 g/cm^2), a brittle-to-ductile transition in a favorable temperature range, and it is reasonably electrically conductive, which allows relatively easy and inexpensive electrodischarge machining. It can be melted and densified without densification aids, so it allows more versatility in processing compared to silicon-based structural ceramics such as SiC and Si_3N_4. The thermal expansion coefficient of $MoSi_2$ is also more closely matched to metals, thus it is easier to join to metals. Possible applications for $MoSi_2$-based materials as a structural high temperature material, high temperature joining material for ceramic parts and as oxidation resistant coatings for refractory metals, Ni-based superalloys, carbon-based alloys and carbon–carbon composites, can be found in many energy production, aerospace, and other industrial areas.[1–3]

In spite of its many promising properties, there are still some limitations to large scale use of $MoSi_2$ in applications. In particular, at intermediate temperatures (typically centered at 500 °C), it suffers from accelerated oxidation or even catastrophic structural disintegration, known as "pest disintegration." Pest disintegration has been associated with the formation of solid MoO_3 into suitable nucleation sites such as preexisting cracks and pores in a material. Due to the large volume expansion this can lead to extensive cracking and total disintegration of the material into powders.[4,5] In addition, the low temperature fracture toughness and high temperature creep resistance and strength of $MoSi_2$ are not optimal and should also be improved. In an attempt to improve the overall performance of $MoSi_2$, several $MoSi_2$-based composites have been synthesized and studied during the last 2 decades. This approach includes the addition of a ductile second phase like metallic Nb or Ta or rigid particles or

a)Electronic mail: Pauli.Torri@helsinki.fi

whiskers of SiC, Al_2O_3, TiB_2, and ZrO_2, and alloying with substitutional species such as WSi_2. A comprehensive review of the mechanical properties of $MoSi_2$ and $MoSi_2$-based composites has been given by Petrovic.[6]

Improvement of mechanical properties due to nanoscale microstructures has been predicted and observed in a number of multilayer structures, especially metal–metal systems. High strength, high temperature coating materials may be utilized as surface modifiers which can extend the operation of bulk structural materials into high temperature and chemically aggressive environments. For such applications, knowledge of the phase evolution, the thermal stability of the layered structure, and the structure–properties relationship of nanolayered coatings as a function of thermal annealing are particularly important. In earlier work, we have studied the structure–mechanical properties relationship and oxidation behavior of nanolayered $MoSi_2$/SiC thin films.[7,8] In these studies, a partial breakdown of the layered structure was observed after 1 h vacuum anneal at 800 °C. At 900 °C 1 h anneal, this layer breakdown had further advanced and during prolonged annealing, the layered structure was totally destroyed, which was associated with a decrease of both hardness and the elastic modulus. The layer breakdown was tentatively explained by a displacive diffusion reaction at the $MoSi_2$ grain boundaries. Thus, diffusion between $MoSi_2$ and SiC should be reduced in order to achieve temperatures which are of practical importance.

One approach to reduce diffusion in a multilayered structure is to use a diffusion barrier. We have previously studied the thermal stability of $MoSi_2$/Mo–Si–N, Mo/Mo–Si–N and Mo/Mo–Si–N/$MoSi_2$ multilayers with the composition of the Mo–Si–N close to $MoSi_2N_x$ ($x \sim 3-4$).[9] These coatings showed excellent thermal stability and the layered structure remained intact without interfacial reaction during 1 h vacuum annealing at least up to 900 °C (highest temperature tested). This kind of strongly nitrogen alloyed molybdenum disilicide has also shown good elevated temperature oxidation resistance and no signs of pest disintegration.[10] For these reasons, the multilayered Mo–Si–N/SiC appears to be a suitable candidate system to study high temperature behavior of nanolayered structures of two constituents which separately have demonstrated good mechanical and chemical performance at elevated temperatures. The aim of this work is to investigate the microstructural evolution and thermal stability of nanolayered Mo–Si–N/SiC thin films and to find the correlation between mechanical properties and microstructure. In addition, we have examined the mechanical properties as well as the stress state of the film as a function of the ratio of constituents and bilayer thickness, and also the influence of thermal annealing on the stress state of the films. The possibility of using Mo–Si–N as a diffusion barrier between $MoSi_2$ and SiC in a nanolayered structure is tested.

II. EXPERIMENTAL PROCEDURES

Magnetron sputtering techniques were used for deposition of both single layer films and multilayers of Mo–Si–N and SiC. The deposition chamber was evacuated to less than 1.3×10^{-6} Pa vacuum prior to the deposition by use of a turbopump. Mo–Si–N was deposited by dc sputtering from a planar $MoSi_2$ target (4 in. diameter, 99.5% purity from Plasmaterials, Inc.) in a gas mixture of 50% N_2 and 50% Ar. Gas flow (70 sccm) and composition was controlled with mass-flow controllers and a multigas analyzer.

SiC was sputtered in pulsed dc mode in an argon plasma by use of a SparcLe unit, which reverses the target voltage for 10 μs at a 2 kHz rate. This pulsing technique is for decharging the surface of the SiC planar sputtering target (4 in. diameter, 99.5% purity, Plasmaterials, Inc.) during the sputtering process. Mo–Si–N was deposited at 0.2 Pa and SiC at 0.4 Pa working gas pressure. In both cases, 300 W dc power was used, which gave deposition rates of 32 nm/min for Mo–Si–N and 8 nm/min for SiC. These rates were measured with a Dektak II A surface profilometer from premasked samples.

Pieces of single crystal (100) silicon wafers were used as substrates. For stress measurements, 0.39 mm×5 mm×25 mm beams were cut from silicon using a diamond saw. Prior to the deposition, the substrates were sputter etched for 30 min in 0.5 Pa Ar with 110 W dc power using a pulsed 25 kHz plasma driver (EXB2000) to remove the native oxide layer. Magnetron units were situated vertically at opposite sides of the deposition chamber. Samples were situated 20 cm from the targets at the sample holder which was rotated with a stepper motor. No extra heating or substrate bias voltage was applied. The entire deposition process for multilayer deposition, including vacuum pumping, gas control, target cleaning, sample moving and deposition, was controlled by a computer.

Single layer films (Mo–Si–N and SiC) and Mo–Si–N/SiC multilayers with the nominal structure 50×(18+6) nm, 100×(9+3) nm, 200×(4.5+1.5) nm, 100×(10.5+1.5) nm, and 100×(6+6) nm were deposited with total coating thickness of 1.2 μm for all samples. In this designation, $N \times (B + C)$, N is the number of bilayers, B is the thickness of the Mo–Si–N layer and C is the thickness of the SiC layer. A four-layer system of $MoSi_2$/Mo–Si–N/SiC/Mo–Si–N with nominal structure of 50×(10+4+6+4) nm was deposited to study if Mo–Si–N can act as a diffusion barrier and prevent layer breakdown and mixing between $MoSi_2$ and SiC which was seen in our earlier work on $MoSi_2$SiC nanolayered structures.[7] Samples consisting of three layers, SiC/Mo–Si–^{15}N/SiC, were also deposited for nitrogen diffusion measurements, since these would give indirect information on the thermal stability of layered structure. In these samples, isotopically enriched $^{15}N_2$ gas (99.98% isotopical purity) was used in the Mo–Si–N layer since it allows for fast ion beam profiling of nitrogen by use of the sharp 429 keV resonance of the $^{15}N(p,\alpha\gamma)^{12}C$ reaction.[11]

All thermal annealings at temperatures up to 900 °C were made in a vacuum quartz-tube furnace in a vacuum better than 10^{-6} Pa. Higher temperature annealings (1050–1075 °C) were done in a quartz-tube furnace with a vacuum better than 5×10^{-4} Pa. Rutherford backscattering spectrometry (RBS) with a 2 MeV He^+ beam and backscattering

(BS) of 2.4 MeV protons (non-RBS from light elements) were used to determine elemental composition of the films using a standard RBS setup with a 170° scattering angle and a 100 mm^2 ion-implanted silicon detector. For the depth profiling of nitrogen, 4.43 MeV γ-rays from the sharp resonance (Γ=124 eV[12]) of the reaction $^{15}N(p,\alpha\gamma)^{12}C$ at 429 keV were detected as a function of probing proton beam energy with a 12.7 cm (diameter)×10.2 cm NaI(Tl) scintillator detector. The energy resolution of the proton beam was about 500 eV. Beams for both BS and nuclear resonance broadening measurements were obtained from the single-ended Van de Graaff accelerator. For determination of major impurities in as-deposited films, time-of-flight elastic recoil detection analysis (TOF-ERDA) was utilized using a 37 MeV Au^{7+} beam obtained from a 5 MV EGP-10-II tandem accelerator.

Surface hardness and elastic modulus of the films were determined by use of a Nanoindenter II from Nano Instruments. This instrument directly measures the load on a triangular pyramid diamond indenter tip as a function of displacement from the surface. Indents were done in displacement-controlled mode and the continuous stiffness method was used. In this method a small modulation signal is added to the loading (vibrating tip). Elastic modulus and hardness are determined from the contact stiffness at each point of the loading curve. Total depth of the indents was 500 nm. Nine indents were made on each sample and the data averaged. Fused silica was used as a calibration standard for indentation tip area function and as a control sample. A comprehensive review of the determination of elastic and plastic properties of thin films on substrates using nanoindentation techniques has been recently given by Nix.[13]

The residual stress in thin films on single crystal silicon substrates was determined using the Stoney equation:[14]

$$\sigma = \frac{E_s t_s^2}{6Rt_f}.$$

Here, σ is the stress in the film, E_s is the biaxial elastic modulus of the (100) Si substrate, t_s is the thickness of the substrate, R is the radius of curvature of the film/substrate system (negative for compressive stress state in the film) and t_f is the thickness of the film. E_s is equal to $E_0/(1-\nu_0)$, where E_0 is Young's modulus and ν_0 is the Poisson ratio. The radius of curvature of silicon beams was measured optically by scanning a helium–neon laser spot over the center line of the silicon beam and recording its reflection with a position sensitive detector. This was done prior to and after film deposition, and after thermal annealings in vacuum. The final stress of the film was calculated from the change of curvature with respect to the uncoated beam.

Cross-sectional transmission electron microscopy (XTEM) specimens were made by mechanical thinning and ion beam milling from as-deposited and vacuum-annealed films. Examinations were made using conventional and high-resolution transmission electron microscopy on a Philips CM30ST microscope operating at 300 kV. Glancing angle x-ray diffraction scans were taken from single layer Mo–Si–N film after 900 °C 3 h vacuum anneal using an unmonochromated Cu source at an incident angle of 0.5°. To control the beam dimensions, a high precision variable slit was placed in front of the sample. For angular resolution, a Soller slit collimator was mounted between the sample and detector. A solid state detector was used to isolate the Cu K_α doublet and to maximize the signal-to-noise ratio.

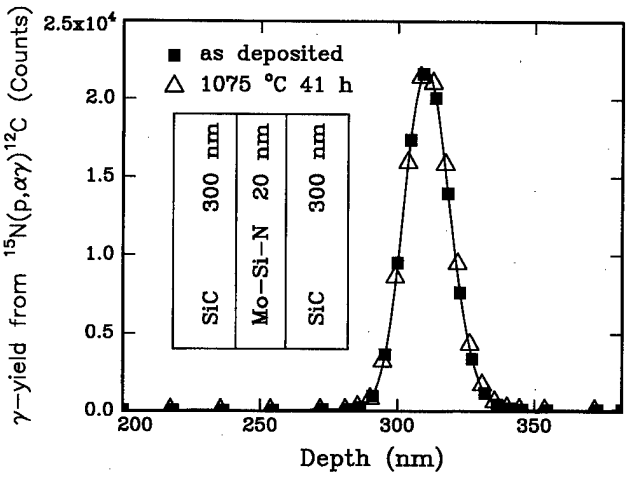


FIG. 1. Nitrogen profiles from (300+20+300) nm SiC/Mo–Si–N/SiC samples measured by nuclear resonance broadening.

III. RESULTS AND DISCUSSION

A. Ion beam analysis

Backscattering measurements of a single layer Mo–Si–N film revealed the elemental ratio of Mo:Si:N to be 1:2.2:2.5. In addition, less than 2 at. % of oxygen, 1% of carbon, and 1.5% of argon were detected in the film by TOF elastic recoil detection (TOF-ERD). Nuclear resonance profiling of nitrogen shows that nitrogen content in the films is constant for anneals of 900 °C for 3 h in vacuum. At 1050 °C, an annealed single layer Mo–Si–N film exhibited some loss (about 30%) of nitrogen very close to the surface (about 30 nm). Deeper in the film, no change in the Mo:Si:N ratio takes place at this temperature. SiC films were observed to be slightly carbon rich (Si:C~0.93) with about 1.5 at. % oxygen and less than 1% argon impurities. By combining areal densities obtained from backscattering data and thickness values measured by surface profilometer from premasked samples, mass densities of 4.4 and 2.9 g/cm^3 were estimated, respectively, for Mo–Si–N and SiC. The density of as-deposited SiC film was about 90% of the full density of bulk silicon carbide (3.2 g/cm^2).

In Fig. 1, nitrogen profiles from SiC/Mo–Si–^{15}N/SiC samples (300 nm+20 nm+300 nm) in as-deposited form and after 41 h of vacuum annealing at 1075 °C are shown. No diffusional broadening of the nitrogen profile can be seen, which indicates good thermal stability of the Mo–Si–N/SiC interface at this temperature.

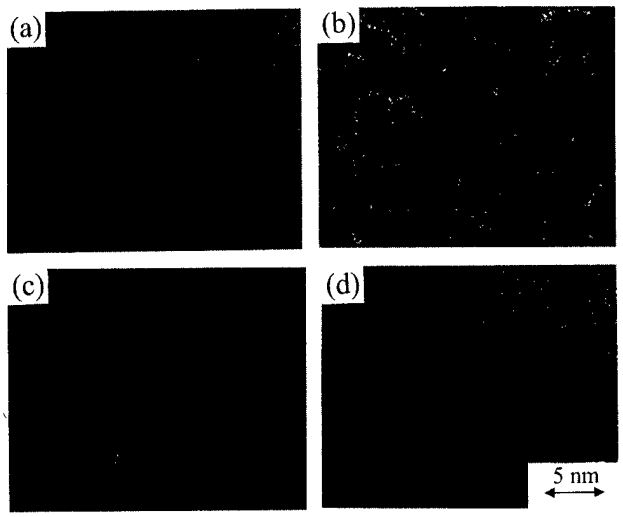


FIG. 2. Transmission electron microscopy (TEM) images taken from Mo–Si–N films in the as-deposited state (a), after 900 °C/3 h vacuum annealing (b), from SiC films in the as-deposited state (c), and after similar annealing (d).

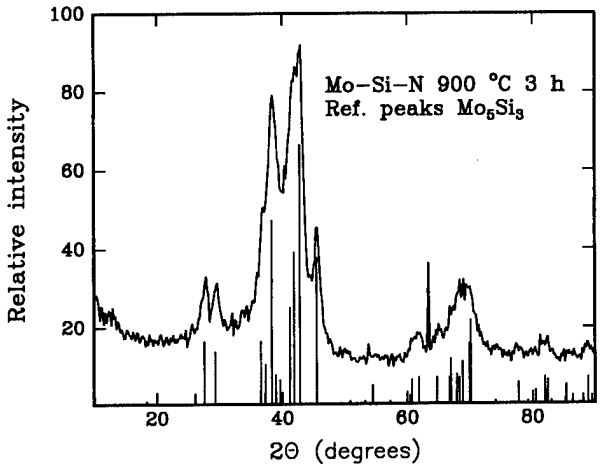


FIG. 4. Glancing angle x-ray diffraction scan from a 900 °C/3 h annealed single layer Mo–Si–N film. Plotted reference lines are from powder diffraction data of tetragonal Mo_5Si_3.

B. Microstructure

All films exhibited an amorphous microstructure in their as-deposited state. No crystallization was detected after a 3 h anneal at 900 °C in SiC single layer films [Fig. 2(d)] or in SiC sublayers in multilayers [Fig. 3(b)]. Similar annealing of Mo–Si–N, either as a single layer [Fig. 2(b)] or a sublayer in nanolayered composite [Fig. 3(b)] did result in partial crystallization, where 3.5 nm diam equaxial grains were observed embedded in an amorphous matrix. After 500 °C annealing for 3 h, no crystallization was detected [Fig. 3(a)]. Glancing angle x-ray diffraction scan taken from 900 °C 3 h annealed single-layer Mo–Si–N film is presented in Fig. 4. The nanocrystalline phase identified by x-ray peak fitting is tetragonal Mo_5Si_3. The formation of this silicide compound is consistent with the Mo–Si–N phase diagram,[15] which indicates that the Mo–Si–N equilibrium state should consist of mainly Mo_5Si_3 and Si_3N_4.

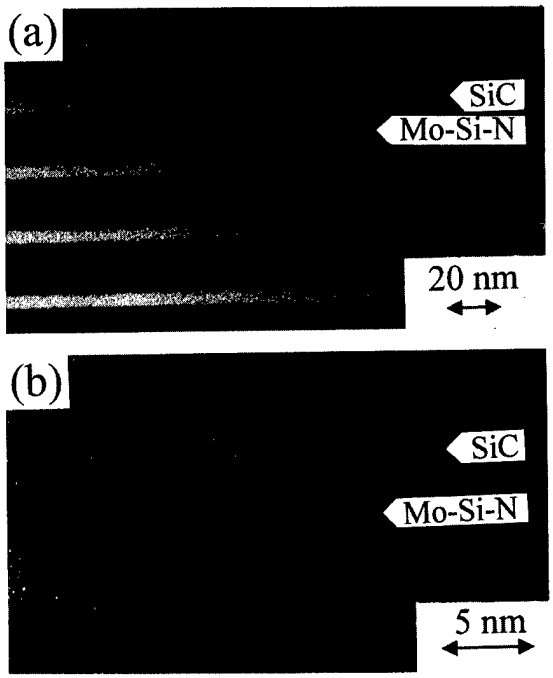


FIG. 3. TEM image from 50×(18+6) nm Mo–Si–N/SiC films after 500 °C/3 h annealing (a) and from a 100×(6+6) nm structure after 900 °C/3 h annealing (b). At 900 °C, Mo–Si–N (darker contrast) partly crystallized, but SiC remained amorphous. At 500 °C, no crystallization was detected.

The observed partially nanocrystalline microstructure is in good agreement with our earlier work where the microstructure and mechanical properties of $MoSi_2N_x$ were studied as a function of nitrogen content.[16] In that study, after a 3 h vacuum anneal at 900 °C, $MoSi_2N_x$ ($x \sim 1.5$) films had developed nanocrystalline C40–$MoSi_2$ microstructure, and films with high nitrogen content ($x \sim 3-4.2$) possessed an amorphous structure which was phase separated into two different phases, one phase with a slightly darker bright field image contrast than the other one. The increase in nitrogen content was associated with an increase in crystallization temperature and a refined microstructure. Nitrogen content of the Mo–Si–N film in the present work ($MoSi_{2.2}N_{2.5}$) lies between values of these two structures. A possible explanation for the observed microstructure refinement is that when the nitrogen content is increased, diffusion becomes more sluggish and finally can produce only local phase separation.

In Fig. 5, XTEM images from the four-layer samples ($MoSi_2$/Mo–Si–N/SiC/Mo–Si–N), in the as-deposited (a), 3 h at 500 °C (b), and 3 h at 900 °C (c) vacuum-annealed states are presented. The $MoSi_2$ sublayer undergoes the crystallization sequence from amorphous to the metastable hexagonal C40 structure (3 h at 500 °C) and to the stable tetragonal $C11_b$ (3 h at 900 °C). A detail of this four-layer system is the thin light contrast phase between $MoSi_2$ and Mo–Si–N which is visible in all TEM micrographs, even in as-deposited films. So far we have no explanation of what this thin amorphous looking layer is. SiC and Mo–Si–N have the

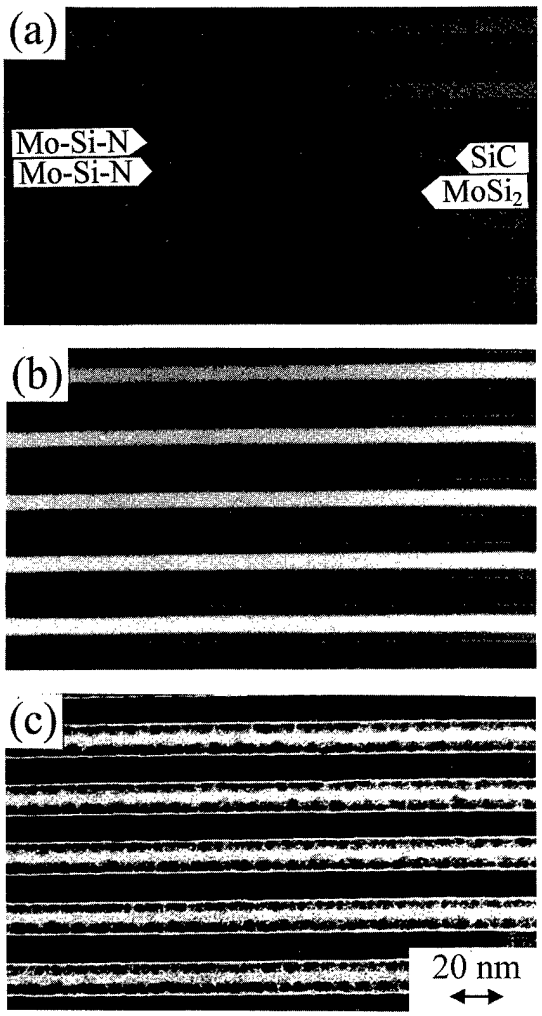


FIG. 5. TEM images of a 50×(10+4+6+4) nm four-layer $MoSi_2$/Mo–Si–N/SiC/Mo–Si–N structure in the as-deposited state (a), after 500 °C/3 h vacuum annealing (b), and after 900 °C/3 h vacuum annealing (c).

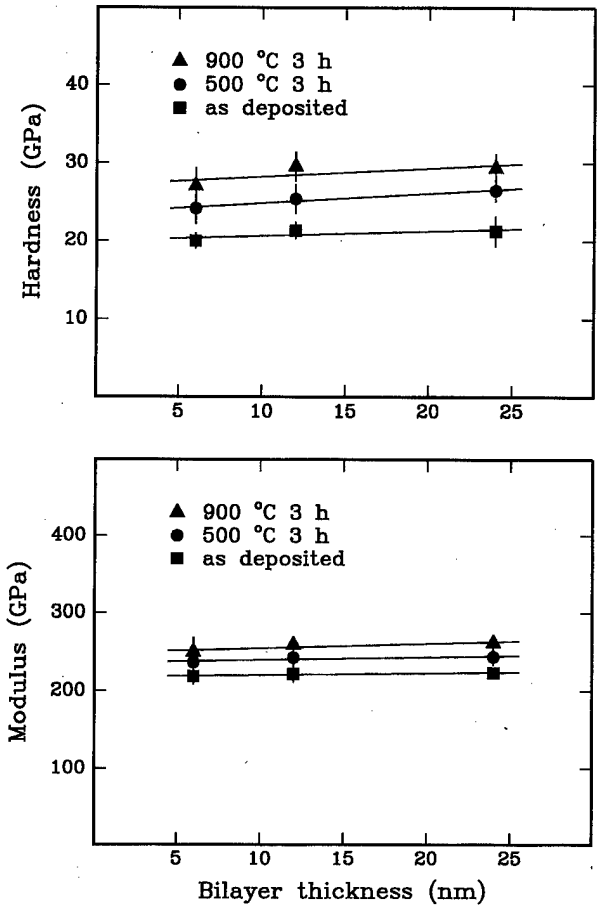


FIG. 6. Nanoindentation hardness and elastic modulus as a function of bilayer thickness for Mo–Si–N/SiC multilayers with 25 vol. % of SiC, i.e., the 200×(4.5+1.5) nm, 100×(9+3) nm and 50×(18+6) nm structures.

same microstructure evolution in the four-layer system as observed in single layer films and multilayered Mo–Si–N/SiC samples.

In our previous work, crystallization of SiC in $MoSi_2$/SiC nanolayered structures was observed to occur after crystallization of $MoSi_2$ during 1 h annealing at 700 °C.[7] This behavior was explained to be due to adjacent crystalline material ($MoSi_2$) which assists the nucleation of SiC. Accordingly, the absence of such material in contact with SiC in the present system prevents the nucleation of SiC at temperatures up to at least 900 °C. Nanocrystalline grains in the amorphous matrix of Mo–Si–N seem not to be sufficient to assist nucleation of SiC. On the other hand, the formation of crystalline $MoSi_2$ in the four-layer sample has no visible effects on the crystallization behavior of amorphous Mo–Si–N. Also, the nanolayered structure remains intact with no signs of layer mixing at least up to 900 °C annealing for 16 h (longest time and highest temperature tested and analyzed with TEM).

C. Nanoindentation hardness and elastic modulus

Nanoindentation hardness and elastic modulus of as-deposited and vacuum-annealed films are presented in Fig. 6 as a function of bilayer thickness (25 vol. % of SiC) and in Fig. 7 as a function of SiC volume fraction (12 nm bilayer thickness, 100 bilayers). Values are 100 nm depth values from continuous stiffness measurements with 500 nm total indent depth. Error bars represent standard deviation of the nine indents. When keeping bilayer thickness constant (12 nm) and changing the Mo–Si–N to SiC ratio (Fig. 7), a simple volume fraction rule of mixture on both hardness and modulus can be seen. Bilayer thickness (or number of interfaces) seems to have no significant effect on the elastic modulus (Fig. 6). The multilayer with the shortest bilayer thickness (6 nm) is slightly softer than the others (12 and 24 nm), but within statistical error no real trends can be deduced. In general, the hardness and modulus both increase with increasing annealing temperature and increasing SiC content.

The hardness and modulus increase in Mo–Si–N films as a result of 900 °C annealing (Fig. 7) can be attributed to crystallization of some part of the material. However, large increases in these properties already take place at 500 °C annealing where the material is still amorphous. No crystal-

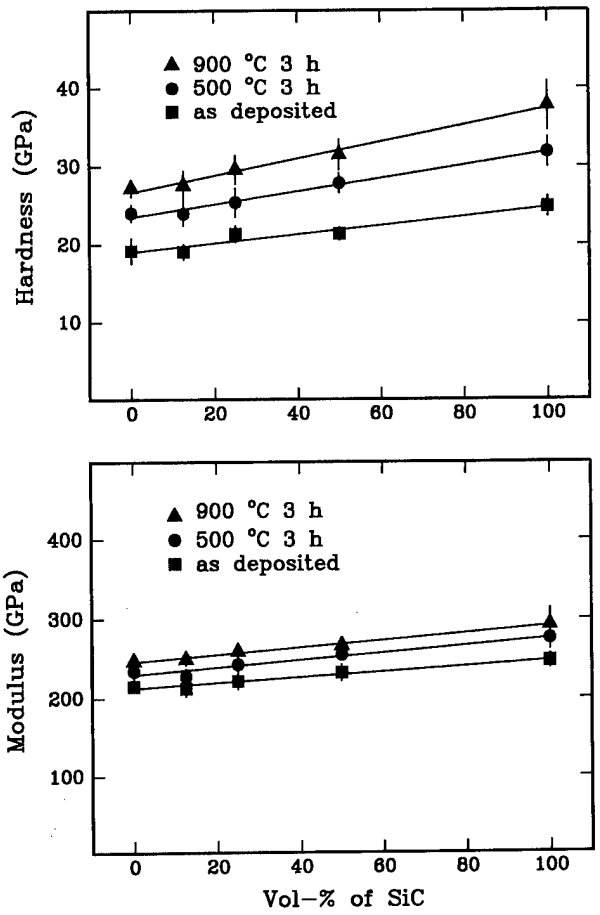


FIG. 7. Nanoindentation hardness and modulus as a function of volume fraction of SiC with constant 12 nm bilayer thickness.

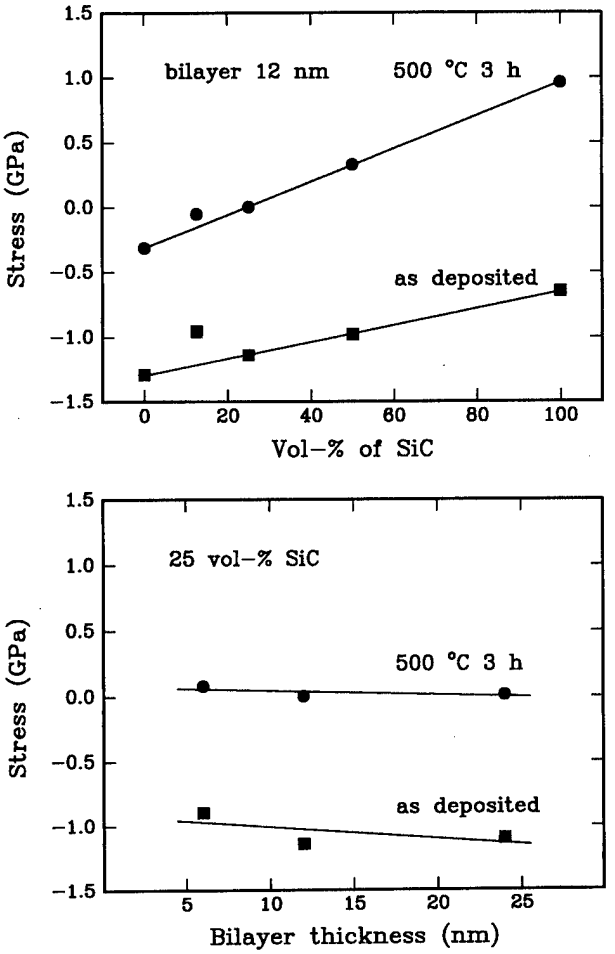


FIG. 8. In-plane residual stress as a function of SiC volume fraction (12 nm bilayer thickness) and as a function of bilayer thickness (25 vol. % of SiC).

lization of SiC was detected in this experiment, either as a single layer or as a sublayer in a nanolayered composite. For this reason, the high increase, especially in the hardness of amorphous films, is most likely due to the densification of the films which retards plastic flow in these amorphous materials.

Hardness and modulus of the four-layer system are listed in Table I. For comparison, we have added values from nanolayered $MoSi_2$/SiC composite (~15 vol. % of SiC) films with 90×(15+2.7) nm geometry and almost similar heat treatments from our earlier work.[7] The relative hardness and modulus during prolonged annealing at 900 °C compared to 1 h annealing (100%) is also presented in brackets. Even though the $MoSi_2$/Mo–Si–N/SiC/Mo–Si–N and $MoSi_2$/SiC samples have almost the same hardness (see 900 °C, 1 h), the elastic modulus of the $MoSi_2$/SiC multilayer is significantly larger. When comparing both transmission electron microscopy (Fig. 5) and nanoindentation results, it is evident that Mo–Si–N can act as a diffusion barrier between $MoSi_2$ and SiC and prevent or at least remarkably slow down deterioration of the layered structure and mechanical properties at 900 °C. The increased modulus of $MoSi_2$/SiC is apparently due to the crystallization of both $MoSi_2$ and SiC; in the $MoSi_2$/Mo–Si–N/SiC/Mo–Si–N samples crystallization of SiC is suppressed.

TABLE I. Nanoindentation hardness and elastic modulus of two different nanolayered composite films. Values for $MoSi_2$/SiC are from our earlier work (Ref. 7).

Condition	Hardness (GPa)	Modulus (GPa)
$MoSi_2$/Mo–Si–N/SiC/Mo–Si–N 50×(10+4+6+4) nm		
as deposited	14.7	200
500 °C/3 h	22.3	250
900 °C/1 h	24.8(100%)	269 (100%)
900 °C/3 h	23.8 (96%)	260 (97%)
900 °C/16 h	22.5 (91%)	254 (94%)
$MoSi_2$/SiC 90×(15+2.7) nm		
as deposited	11.5	217
500 °C/1 h	20.8	290
900 °C/1 h	25.5(100%)	382 (100%)
900 °C/2 h	21.6 (85%)	346 (91%)
900 °C/16 h	18.6 (73%)	342 (90%)

D. Stress evolution

The average in-plane residual stress in as-deposited films and films after vacuum annealing for 3 h at 500 °C is presented in Fig. 8. In addition, residual stress after vacuum annealing at 350 and 500 °C is presented in Fig. 9 as a func-

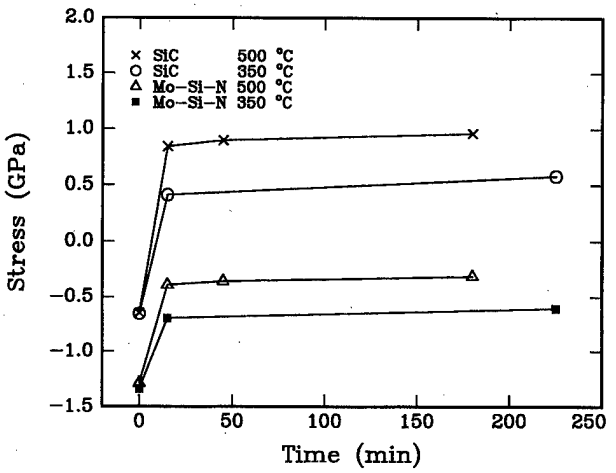


FIG. 9. In-plane residual stress in single layer SiC and Mo–Si–N films as a function of annealing time at 350 and 500 °C.

tion of annealing time for single layer SiC and Mo–Si–N films. Both single layer films and all multilayers follow similar stress evolution as a function of annealing time; after an initial rapid transition period, an almost constant stress level is achieved. All measurements were made at room temperature, and stress values are averaged from nine individual radii of curvature measurements from the same coated silicon beam.

The data of Fig. 9 show that for both SiC and Mo–Si–N, the magnitude of stress change during the first 15 min annealing is almost directly proportional to the temperature difference between room temperature (20 °C) and annealing temperature. From these data, it would be tempting to assume that at elevated temperature stress is totally relaxed and the residual stress is only due to the difference between the thermal expansion coefficient of the film and the silicon substrate. However, even if we use an overestimated value of 440 GPa (hot pressed bulk material value) for the elastic modulus of SiC and linear thermal expansion coefficients of 4.8×10^{-6} for SiC and 3.5×10^{-6} for the silicon substrate, we can estimate, when neglecting in-plane Poisson effects, that only 0.27 GPa stress should result from the differences in thermal expansion coefficient at 500 °C. For SiC single layer films, the stress changes from almost 0.7 GPa compressive to almost 1 GPa tensile as a result of annealing at 500 °C for 3 h (Fig. 8). It is evident from this analysis that other factors such as densification or microstructural changes in the nature of the amorphous material must exist to produce the final stress state in SiC films. Clearly the SiC films are not fully dense as also shown by density measurements on RBS. On the other hand, less than 1% lateral contraction of the film would be enough to produce detected residual stress. Therefore direct measurement of the densification is not possible using combined RBS and thickness measurements.

The residual stress in the multilayered films seems to obey a simple volume fraction rule of mixture of both components either in the as-deposited state or after annealing at 500 °C (Fig. 8). The only exception is the $50\times(11.5+1.5)$ nm sample (12.5 vol. % of SiC). The compressive stress of this sample after deposition is about 20% less than predicted by the rule of mixture, and the magnitude of stress change as a result of annealing is also smaller. At the present time, the reason for this deviation is uncertain. As a function of bilayer thickness, all samples follow similar stress evolution, which indicates that interfaces do not have any significant role in the stress evolution process.

IV. CONCLUSIONS

Nanoindentation mechanical properties, high-temperature thermal stability and residual stress evolution of nanolayered Mo–Si–N/SiC thin films have been investigated. Modulus, hardness and residual stress follow a volume fraction rule of mixture of both components, but are independent of the bilayer thickness. Complete relaxation of residual stress can be observed after annealing of the Mo–Si–N/SiC structures with 25 vol. % of SiC at 500 °C. Annealing of a single layer SiC film, however, results in a change from compressive to tensile residual stress. Crystallization of SiC did not occur at least up to 900 °C annealing in $MoSi_2$/Mo–Si–N/SiC/Mo–Si–N layered structure in contrast to what occurs in nanolayered $MoSi_2$/SiC. This was explained by the lack of SiC nucleation sites. A Mo–Si–N layer appears to provide an excellent diffusion barrier between $MoSi_2$ and SiC, so that the nanolayered structure remains intact at least up to 900 °C annealing for 16 h.

[1]A. K. Vasudevan and J. J. Petrovic, Mater. Sci. Eng., A **155**, 1 (1992).
[2]T. A. Kircher and E. L. Courtright, Mater. Sci. Eng., A **155**, 67 (1992).
[3]J. J. Petrovic and A. K. Vasudevan, Mater. Res. Soc. Symp. Proc. **322**, 3 (1994).
[4]J. Cook, A. Khan, E. Lee, and R. Mahapatra, Mater. Sci. Eng., A **155**, 183 (1992).
[5]D. A. Berztiss, R. R. Cerchiara, E. A. Gulbransen, F. S. Pettit, and G. H. Meier, Mater. Sci. Eng., A **155**, 165 (1992).
[6]J. J. Petrovic, Mater. Sci. Eng., A **192/193**, 31 (1995).
[7]H. Kung, T. R. Jervis, J.-P. Hirvonen, J. D. Embury, T. E. Mitchell, and M. Nastasi, Philos. Mag. A **71**, 759 (1995).
[8]J.-P. Hirvonen, P. Torri, R. Lappalainen, J. Likonen, H. Kung, T. R. Jervis, and M. Nastasi, J. Mater. Res. **13**, 965 (1998).
[9]H. Kung, T. R. Jervis, J.-P. Hirvonen, T. E. Mitchell, and M. Nastasi, J. Vac. Sci. Technol. B **13**, 1126 (1995).
[10]J.-P. Hirvonen, I. Suni, H. Kattelus, R. Lappalainen, P. Torri, H. Kung, T. R. Jervis, M. Nastasi, and J. R. Tesmer, Surf. Coat. Technol. **74–75**, 981 (1995).
[11]J.-P. Hirvonen, in *Handbook of Modern Ion Beam Materials Analysis*, edited by J. R. Tesmer and M. Nastasi (Materials Research Society, Pittsburgh, PA, 1995), p. 167.
[12]T. Osipowicz, K. P. Lieb, and S. Brussermann, Nucl. Instrum. Methods Phys. Res. B **18**, 232 (1987).
[13]W. D. Nix, Mater. Sci. Eng., A **234–236**, 37 (1997).
[14]F. R. Brotzen, Int. Mater. Rev. **39**, 24 (1994).
[15]*Phase Diagrams of Ternary Boron Nitride and Silicon Nitride Systems*, edited by P. Rogl and J. C. Schuster (ASM International, Materials Park, OH, 1992), p. 169.
[16]H. Kung, T. R. Jervis, J.-P. Hirvonen, T. E. Mitchell, and M. Nastasi, Nanostruct. Mater. **7**, 81 (1996).

Characterization of bending in single crystal Si beams and resonators

J. W. Weigold,[a] W. H. Juan,[b] and S. W. Pang[c]
Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan 48109-2122

J. T. Borenstein[d]
Charles Stark Draper Laboratory, Cambridge, Massachusetts 02139-3563

(Received 26 March 1998; accepted 5 April 1999)

Optical interferometry has been applied to determine the displacement of p^{++} Si beams. Clamped-clamped Si beams and cantilevered beams were fabricated with short and long B diffusion processes and characterized. Measurements of beam bending for released Si structures with length varying from 50 to 1000 μm, width varying from 5 to 15 μm, and thickness varying from 6 to 37 μm were obtained. By taking advantage of an etch-diffusion process, thicker beams can be fabricated which have less bending due to stress gradients. A 6.0-μm-thick cantilevered beam had a deflection of 11.2 μm due to stress gradients, while a 36.7-μm-thick beam had a deflection of only 0.3 μm. Beams fabricated using a dissolved wafer process with a 12 h B diffusion were found to bend the same amount as those fabricated with a 4 h diffusion. This indicates that bending in doped Si beams not only depends on the gradients in the B concentrations, it could also be related to the distribution of dislocations. Using the deep-etch shallow-diffusion process, resonating elements that are 20 μm long, 4 μm wide, and 28 μm thick were found to be perfectly flat without any bending.
© *1999 American Vacuum Society.* [S0734-211X(99)00104-3]

I. INTRODUCTION

Due to the excellent mechanical properties of single crystal Si, it has continued to be used as a structural material in the fabrication of a wide variety of devices.[1] Often, the structures are fabricated from B doped Si because this material serves as an etch stop in several wet Si etches.[2] These etches can then be used to release the B doped Si structures. Materials often can have large intrinsic stress gradients which can cause undesirable effects in devices. There are many things done during the processing of materials which can have large effects on their eventual intrinsic stress. Stress arises in B doped Si due to the fact that the B atom is a substitutional impurity with an atomic radius which is smaller than that of Si.[1,3–7] This stress depends on the temperature and the ambient environment in which the diffusion is done, as well as any subsequent annealing steps.[8,9] Stress can even arise due to plastic deformation of the structure when materials with different stresses are in contact.[10]

Stress gradients in microstructures can cause changes in the shape of the structures which can influence device characteristics, stability, and performance. Therefore, it is important to characterize and reduce intrinsic stresses in released microstructures. There are many methods currently used to determine stress in released structures. Often, one uses the deformation of the structure due to stress gradients to evaluate the intrinsic stress in that structure. Deformation due to stress gradients can be measured using profilometry, scanning electron microscopy, or determined by deflection of laser light.[11,12] Surface profilometry can scratch delicate structures and often is time consuming. Measurements using scanning electron microscopy are usually time consuming and often are imprecise. In this study, a WYKO interferometer is used to measure the surface profile.[13–15] The WYKO system utilizes white light passed through a beam splitter to form interference fringes as a function of height on the sample; objective height is controlled by a piezoelectric transducer. Two- and three-dimensional maps of the surface profile can be obtained from the measurements to provide information on displacement and stress in the structure. Using this method, bending due to stress gradients can be measured quickly in a nondestructive manner.

In this work, we have investigated the intrinsic stresses in structures with different geometries, as well as structures fabricated with different processes. A dissolved wafer process was studied, in which a long B diffusion is performed from the surface of the wafer.[16] Structures fabricated using this process are limited to a thickness of about 13 μm. In this process, bending arises due to a variation in doping concentration throughout the structures.[17,18] The B dopant concentration decreases towards the bottom of structures and this difference in B concentration from the surface to the bottom of the structures causes the deformation of the structure.[12] The second process we have used is a deep-etch shallow-diffusion process or a frontside-release etch-diffusion process.[19,20] In these processes, B diffuses from the sides of the beams so the B concentration is nearly constant throughout the beam. In addition, the B diffusion time required is reduced substantially. The thickness of structures fabricated in this process are not limited by the diffusion time but by the etch selectivity. Structures thicker than 50 μm have been fabricated using this process.

[a]Electronic mail: weigold@eecs.umich.edu
[b]Electronic mail: whjuan@wafertech.com
[c]Electronic mail: pang@eecs.umich.edu
[d]Electronic mail: jborenstein@draper.com

II. EXPERIMENT

Samples were fabricated using the dissolved wafer process with arrays of cantilevered Si beams bonded to glass and released. In addition, cantilevered beams, clamped-clamped beams, and resonators were fabricated using a frontside-release etch-diffusion process.[20] These samples were then measured with a WYKO interferometer to determine the amount of bending of the beam due to intrinsic stress gradients. The WYKO system has a vertical resolution better than 10 nm and a horizontal resolution of about 5 μm.[13–15]

The effects of geometry and fabrication conditions on the intrinsic stress in the beams were evaluated. The beam lengths were varied from 20 to 1000 μm. The beam widths ranged from 2 to 20 μm and the beams were fabricated with thicknesses varying from 6.0 to 36.7 μm. This allowed the effects of geometry and dopant profile on beam displacement to be investigated. The application of most interest for these single crystal Si beams is capacitively sensed and driven structures. In this work, all structures are designed to move laterally in the plane of the substrate and out of plane movement is undesirable. Therefore, the primary effect of bending is an unpredictable capacitive plate overlap resulting in a decreased or unpredictable drive force or sense current. Thus, the most useful quantitative result in this case is a measurement of beam tip deflection or beam displacement and this is what has been measured throughout this work.

III. RESULTS AND DISCUSSION

A. Fabrication

Released cantilevered beams were fabricated using a dissolved wafer process. First, a recess etch is performed in KOH to remove 4 μm from the entire Si wafer except the masked anchor areas. This defines the regions which will serve as anchors when bonded to glass. Then, a long diffusion is performed in a furnace at 1175 °C using a solid boron nitride source in an O_2/N_2 ambient. The flow rate for N_2 was 3 s*l*m and 250 sccm for O_2. The doping time used was dependent on the process and desired thickness of the doping layer. All diffusions were then followed by a 5/20/5 min dry/wet/dry thermal oxidation at 1000 °C to drive the dopants from the surface oxide formed during the diffusion and facilitate removal of the borosilicate glass. The thickness of the structures is determined by the diffusion time. Next, a deep dry etch in an electron cyclotron resonance (ECR) source is used to define the structure and etch through the heavily B doped Si layer. The frontside of the wafer is then bonded to a glass substrate and the wafer is subsequently dissolved in a long wet etch in ethylenediamine pyrocatechol (EDP). The structures are not dissolved because the EDP does not attack structures which are doped with B concentrations greater than about 7×10^{19} cm^{-3}.

In order to investigate the effects of stress caused by the long diffusion from the surface, clamped-clamped beams, cantilevered beams, and resonators were fabricated using an etch-diffusion process. In this process, a deep Si etch is performed in an ECR source. The diffusion to convert the structure to p^{++} Si is then performed. This allows a short diffusion to be used regardless of how thick the structures are. The reason for this is that the B diffuses into the structures from both sides throughout the whole depth of the dry etched trench. This should give a much more uniform doping distribution throughout the structure. Following the diffusion, a maskless Si etch is performed in the ECR source to etch through the continuous B layer at the bottom of the trenches allowing the mechanical structures to separate from one another after they are released. High aspect ratio structures can be fabricated which are very stiff in the direction out of the plane of the substrate. This causes a negligible amount of bending due to stress which will be further detailed later.

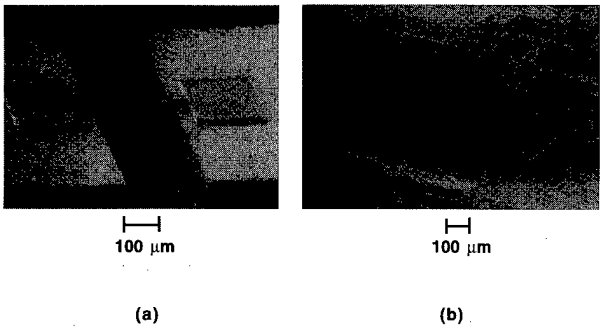


FIG. 1. (a) Micrograph of 20-μm-thick, 300-μm-long, and 5-μm-wide clamped-clamped beam fabricated with the etch-diffusion process, and (b) micrograph of a 37.5-μm-thick array of 5-μm-wide cantilevered beams with lengths varying from 50 to 1000 μm fabricated with the etch-diffusion process.

B. Clamped-clamped beams

Figure 1(a) shows a micrograph of a beam clamped on both sides which is 300 μm long, 5 μm wide, and 20 μm thick bonded to a glass substrate. The laser beam from the interferometer was scanned across the length of the doubly clamped beam. Figure 2 shows a scanned profile of a 300-μm-long, 15-μm-wide, and 20-μm-thick clamped-clamped beam which has a variation in height of 250 nm. A positive beam tip deflection indicates a displacement, away from the original Si substrate, and a negative beam tip deflection in-

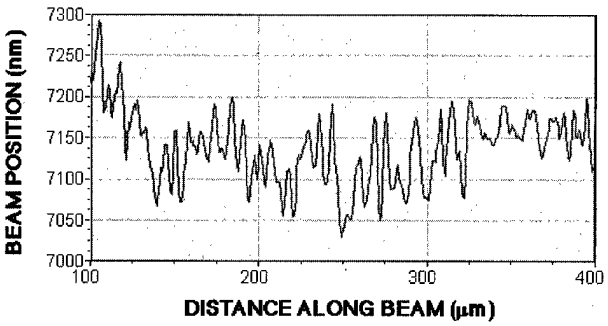


FIG. 2. Height profile as a function of distance along the 300-μm-long, 15-μm-wide beam fabricated using the etch-diffusion process, obtained from WYKO interferometric scan.

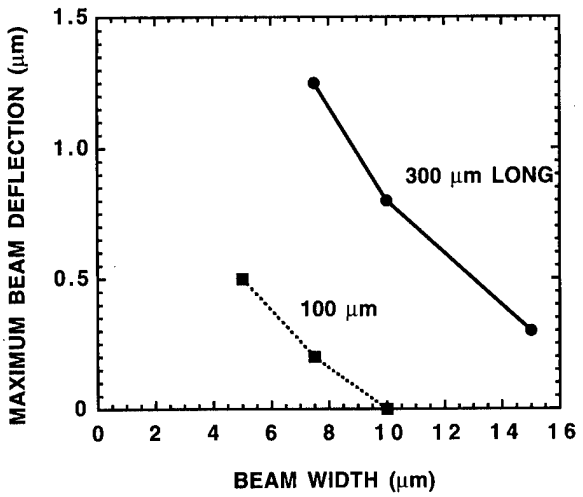


FIG. 3. Variations of beam deflection with beam width and length for 20-μm-thick beams fabricated using the etch-diffusion process.

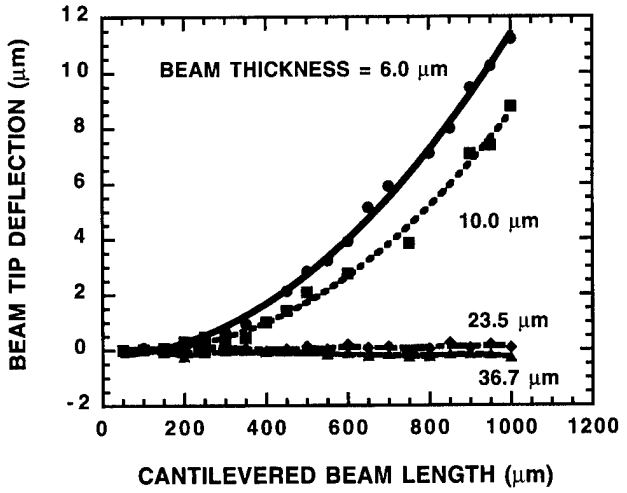


FIG. 4. Plot of cantilevered beam tip deflection variations with thickness and length. The beams were 10 μm wide and fabricated using the etch-diffusion process.

dicates a displacement toward the original Si substrate. All clamped-clamped beams were fabricated using the etch-diffusion process which provides Si structures bonded upside down to a glass substrate. Variation in maximum beam deflection for 20-μm-thick beams with different length and width fabricated by the etch-diffusion process was also measured and the results are shown in Fig. 3. Diffusion times for beams of different widths were calculated based on previous measurements so that the beams were fully diffused with additional time added to guarantee complete diffusion of the beam. In Fig. 3, maximum beam deflection refers to the maximum height difference between the fixed ends of the beam and any other point on the beam. The 7.5-μm-wide, 300-μm-long beam has a height variation of 1.25 μm, but when the length of the beam is decreased to 100 μm, the height variation is decreased to 0.2 μm. For 100-μm-long, 20-μm-thick beams, but with the width increased from 5.0 to 7.5 μm, the membrane curvature is reduced from 0.5 to 0.2 μm. The amount that the beam bends due to stress gradients is dependent on the stiffness of the beam. The spring constant is a measure of this stiffness. The larger the spring constant in a given direction, the less it will bend when a given force is applied. The spring constant of a beam is related to its moment of inertia divided by its length. By increasing the beam's length, its spring constant decreases and the beam becomes more compliant. Therefore, longer beams tend to bend more due to stress gradients. On the other hand, the moment of inertia in the direction of interest (out of the plane of the substrate) is related to the cube of the thickness of the beam, multiplied by the width of the beam. As the beam width is increased, its moment of inertia and the spring constant increase, which makes the beam stiffer in this direction. If the stress gradients caused by doping gradients were uniform across the width of the beam, then a wider beam would bend more due to stress gradients and this would balance the wider beam's increased stiffness, causing the bending to be independent of width. However, experimental results show that the bending is dependent on width which suggests that the stress gradients across the width of the beam are not uniform and therefore do not track with the increase in beam stiffness. This means that the stiffness of the wider beam dominates and thus, it bends less even with stress gradients across the beam. Also, since the bonding is performed at an elevated temperature, small differences in the thermal expansion coefficient between the Si and the glass could give rise to stress gradients in the beams. However, this effect is minimized by using Corning No. 7740 glass, which has a thermal expansion coefficient closely matched to that of the Si substrate.

C. Cantilevered beams

Arrays of cantilevered beams were fabricated and a 37.5-μm-thick array is shown in Fig. 1(b). These beams were fabricated using the frontside-release etch-diffusion process so that B doped Si beams are formed which are anchored to a Si substrate. The displacement of cantilevered beams of varying lengths and thicknesses was measured and is shown in Fig. 4. The lengths of the beams varied from 50 to 1000 μm in 50 μm steps, and the beam width was 5 μm. The tip of the cantilevered beam was scanned and the tip height could then be determined, referenced to the plane of the Si surface. A positive beam tip deflection indicates an upward displacement, away from the substrate, and a negative beam tip deflection indicates a downward displacement toward the substrate. For the 6.0- and 10.0-μm-thick beams, as the length of the beam increased, the beam tip deflection also increased. For the thicker 23.5- and 36.7-μm-thick beams the tip deflection was much smaller and stayed relatively constant as the length of the beam increased.

Figure 4 also shows that beam tip deflection decreased as the thickness of the beam increased. For 1000-μm-long beams, as the thickness of the beam increased from 6.0 to 36.7 μm, the deflection at the tip decreased from 11.2 to -0.3 μm. That is, the thicker beams are much less affected by the intrinsic stress gradients due to their increased stiff-

ness and tend to stay flatter than thin beams. This agrees with conventional beam theory which states that the deflection at the tip of a cantilevered beam is given by

$$y \propto \frac{L^2}{T^3 W}, \tag{1}$$

where y is the deflection of the beam tip and L, T, and W are the length, thickness, and width of the beam, respectively.[21] This decreased bending is desirable since the positions of the fabricated beams can be more accurately predicted if they do not bend due to stress gradients. In addition, large displacement can cause large variations in device performance and reproducibility. Therefore, devices with higher aspect ratios are desirable to minimize the displacement due to intrinsic stress gradients.

D. Comparison of stress among different fabrication methods

Two methods of fabrication were used to fabricate cantilevered beams as described previously. The intrinsic stress in the single crystal silicon arises from the B dopant which is used to provide etch selectivity over the lightly doped Si substrate in the EDP etch. This dopant substitutes into the Si lattice and gives rise to stresses because the B atom has a smaller atomic radius than the Si atom. Typically, the B is diffused into the Si from the upper surface, giving rise to a concentration gradient of the B in the Si from the surface to the bottom of the cantilever. It has been reported that the greater the difference in concentrations at the top and bottom of the cantilever, the greater the deflection of the beam.[12] This explanation is in good agreement with experimental data and assumes a linear B distribution from the top of the cantilever to the bottom, and no B concentration gradient across the width of the beam. Devices fabricated using a long B diffusion could have this type of variation in B concentration from the top surface to the bottom of the cantilever.

For the short B diffusion, with diffusion from the top and sides of the cantilever, the cantilever should be more uniformly doped and have less beam deflection due to stress gradients. Figure 5 shows that the beam tip deflection does not depend on the diffusion time. Two sets of 10-μm-thick, 10-μm-wide cantilevers were measured with varying lengths. The beam deflection was similar for beams made with 4 h of B diffusion from the top and sides in the etch-diffusion process and 12 h of B diffusion from the top only using the dissolved wafer process. The 4 h diffusion results in a junction depth of 6.6 μm while the 12 h diffusion results in a junction depth of 10.0 μm. Junction depth as used here describes the depth at which the concentration falls below that serving as an etch stop in EDP which is typically 7×10^{19} cm^{-3}. In the deep etch-short B diffusion process, the B concentration in the structure should be highest at the top and side surfaces of the beam. The B concentration should decrease as the center of the beam is approached. Since there is now a lateral B concentration gradient across the width of the beam, the previous model may not hold. However, this lateral concentration gradient should be symmetrical. At high

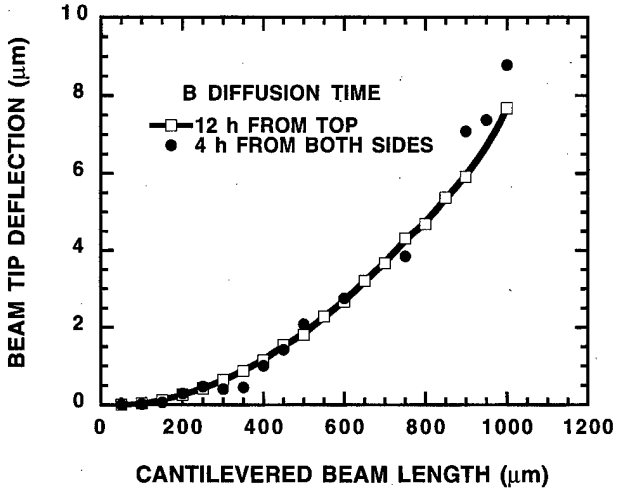


FIG. 5. Comparison of beam tip deflection between 10-μm-wide cantilevered beams fabricated using the dissolved wafer process with a 12 h B diffusion and those fabricated using the etch-diffusion process with a 4 h B diffusion.

B concentrations, which are typical of an etch stop, simple functions to describe the dopant profile may not hold true.[22] In addition, there will still be variations in the B concentration from the top of the beam to the bottom, due to the larger B concentration at the top surface than below the surface. Also, at the bottom of the structure where the etch stop occurs, there will be a small doping concentration gradient where B concentration decreases from top to bottom. This may also give rise to stress that could cause deflection of the cantilevered beam. Typically, however, the diffusion time is lengthened to ensure that the whole beam is converted to p^{++}Si. This should allow the B to diffuse further and serves to make the doping profile quite uniform throughout the beam.

This suggests that something other than a gradient in B concentration through the beam is giving rise to the deflection of the beam. One possible explanation is that the deflection arises from a gradient in dislocation density through the beam. Dislocations follow the diffusion of B into the Si, in order to relieve the stress created by the B.[23] However, the distribution of dislocations through the beam is not necessar-

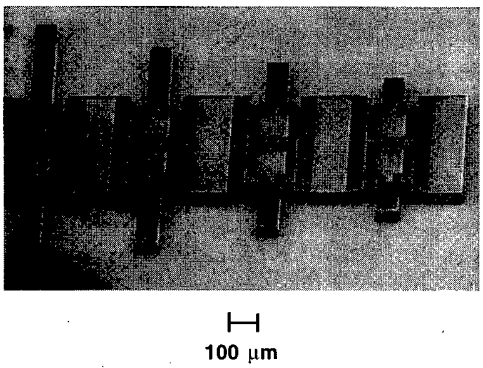


FIG. 6. Micrograph of 50-μm-thick comb drive resonator with 2-μm-wide and 30-μm-long fingers fabricated with the etch-diffusion process.

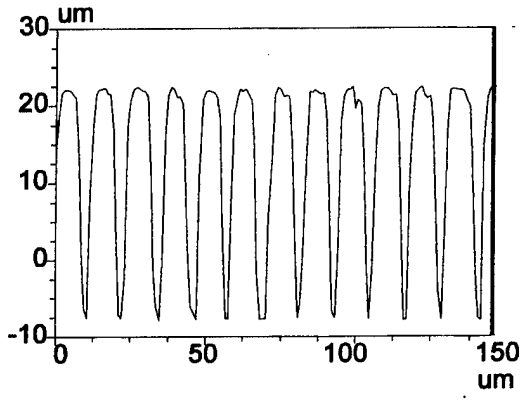


FIG. 7. Interferometric scan across the resonating elements of a 28-μm-thick comb drive fabricated with the etch-diffusion process.

ily the same as the distribution of B through the beam, since dislocations are formed close to the surface, where the B concentration is above the threshold level for dislocation formation. The gradient in dislocation concentration through the beam may be what is giving rise to the bending of the beam, and this gradient may not change with the two diffusion methods used here. This would explain the similar amount of bending observed using two different methods of diffusion. Although the same amount of bending is seen with the long B diffusion, compared to the short diffusion deep-etch process, the latter process allows thicker structures to be fabricated which would then have less bending due to stress gradients.

E. Resonators

The displacement of resonators fabricated using the etch-diffusion process was also evaluated and Fig. 6 shows typical comb drives with resonating elements. For the fingers of a 20-μm-long, 4-μm-wide, and 28-μm-thick comb drive, the interferometric measurements indicate that they are perfectly flat without any bending as shown in Fig. 7. For clamped-clamped beam resonators with the same width and thickness, but 325 μm long, beam displacement up to 1.7 μm was observed. Figure 7 shows a scan across the elements of comb drive fingers. This scan was performed across released comb drive fingers with the fingers oriented perpendicular to the resonant beam, and the scan performed parallel to the resonant beam. This shows the flexibility of the interferometric system to be able to scan many types of structures such as a comb drive with thick fingers and very narrow gaps.

IV. CONCLUSION

Optical interferometric techniques were used to characterize the displacement of structures due to intrinsic stress gradients. This method allows precise measurements to be made in a simple, fast, nondestructive manner. This precise measurement allows the displacement in beams to be monitored and reduced which can increase device performance and reproducibility. Clamped-clamped beams, cantilevered beams, and resonators were fabricated and displacement due to stress gradients was measured. Quantitative measurements were obtained showing that bending of beams due to stress gradients decreased for wider, shorter, and thicker structures as would be predicted by conventional beam theory. Thus the latitude in the etch-diffusion process in fabricating thick structures has the advantage of decreasing the amount of bending of beams of similar geometries. No difference in bending due to stress gradients was observed for beams fabricated using a 12 h B diffusion process, compared to those fabricated with a short B diffusion of 4 h, in a deep-etch shallow-diffusion process. However, the deep-etch, short B diffusion process allows much thicker devices to be fabricated with larger stiffness and less bending due to stress gradients.

ACKNOWLEDGMENTS

This work is supported by the Defense Advanced Research Project Agency under Contract No. DABT63-C-0111, the Charles Stark Draper Laboratories, and the Ford Motor Company.

[1]K. E. Petersen, Proc. IEEE **70**, 420 (1982).
[2]J. C. Greenwood, J. Electrochem. Soc. **116**, 1325 (1969).
[3]V. N. Erofeev, V. I. Nikitenko, and V. B. Osvenskii, Phys. Status Solidi **35**, 79 (1969).
[4]V. N. Erofeev and V. I. Nikitenko, Sov. Phys. Solid State **13**, 116 (1971).
[5]I. N. Smirnov, Sov. Phys. Dokl. **20**, 224 (1975).
[6]I. N. Smirnov, Sov. Phys. Solid State **19**, 859 (1977).
[7]L. S. Milevskii and Y. N. Chuvilin, Sov. Phys. Solid State **22**, 1536 (1980).
[8]B. S. Berry and W. C. Pritchet, J. Vac. Sci. Technol. A **9**, 2231 (1991).
[9]W.-H. Chu and M. Mehregany, IEEE Trans. Electron Devices **40**, 1245 (1993).
[10]F. Maseeh and S. D. Senturia, Sens. Actuators A **A21–23**, 861 (1990).
[11]E. I. Bromley, J. N. Randall, D. C. Flanders, and R. W. Mountain, J. Vac. Sci. Technol. B **1**, 1364 (1983).
[12]X. Ding, W. H. Ko, and J. M. Mansour, Sens. Actuators A **21–23**, 866 (1990).
[13]P. J. Caber, Appl. Opt. **32**, 3438 (1993).
[14]P. J. Caber, S. J. Martinek, and R. J. Niemann, Proc. SPIE **2088**, 195 (1993).
[15]J. T. Borenstein, P. Greiff, J. B. Sohn, and M. S. Weinberg, Proc. SPIE **2879**, 116 (1996).
[16]Y. B. Gianchandani and K. Najafi, IEEE J. Microelectromech. Syst. **1**, 77 (1992).
[17]C. Cabuz, K. Fukatsu, T. Kurabayashi, K. Minami, and M. Esashi, Transducers'95 News, Stockholm, Sweden, 1995, p. 61.
[18]C. Cabuz, K. Fukatsu, T. Kurabayashi, K. Minami, and M. Esashi, IEEE J. Microelectromech. Syst. **4**, 109 (1995).
[19]W. H. Juan and S. W. Pang, IEEE J. Microelectromech. Syst. **5**, 18 (1996).
[20]J. W. Weigold and S. W. Pang, Transducers'97, Chicago, IL, 1997, pp. 1435–1438.
[21]W. C. Young, *Roark's Formulas for Stress and Strain*, 6th ed. (McGraw-Hill, New York, 1989), pp. 100–112.
[22]R. B. Fair, J. Electrochem. Soc. **122**, 800 (1975).
[23]K. C. Wu, P. A. Shay, J. T. Borenstein, and E. A. Fitzgerald, Mater. Res. Soc. Symp. Proc. **444**, 197 (1997).

In situ ellipsometric study of the formation process of metalorganic vapor-phase epitaxy-grown quantum dots

Jeong-Sik Lee,[a] Shigeo Sugou, Hong-Wen Ren, and Yasuaki Masumoto
Single Quantum Dot Project, ERATO, JST, Tsukuba Research Consortium, 5-9-9 Tokodai, Tsukuba 300-2601, Japan

(Received 30 December 1998; accepted 23 April 1999)

Our *in situ* ellipsometry study of metalorganic vapor-phase epitaxy-grown quantum dot (QD) structures showed that the $\Delta-\Psi$ trajectory of the ellipsometric signal in Stranski–Krastanow QD growth sharply differs from that in layer-by-layer growth. When QD formation starts, Δ rapidly decreases, inflecting the $\Delta-\Psi$ trajectory. This indicates increased scattering loss, deduced from increased surface roughness produced in the transition from two- to three-dimensional surface morphology. *Ex situ* atomic force microscopy and photoluminescence results correspond well to the ellipsometric signal at the start of QD formation. Based on these results, we discuss growth-dependent QD formation processes such as QD formation onset and growth mode transition.
© *1999 American Vacuum Society.* [S0734-211X(99)02304-5]

I. INTRODUCTION

Nanostructures showing spatial confinement of carriers in three dimensions have attracted increasing attention due to their physical properties[1] and potential device applications. Self-assembled islands formed in the Stranski–Krastanow (SK) growth mode have attracted particular attention as quantum dot (QD) structures. In the SK mode, island structures are self-formed on a two-dimensional (2D) wetting layer, as a result of the transition of the growth mode, namely, from 2D to three-dimensional (3D) one at a certain layer thickness. The growth and optical characterization of self-assembled QDs have been widely studied, mostly in the In(Ga)As/GaAs system by molecular beam epitaxy (MBE)[2] and metalorganic vapor-phase epitaxy (MOVPE).[3,4] Problems awaiting solutions include nonuniform size distribution and uncertainties on QD height, which have led to difficulty in comparing the observed quantization effect with theoretical one. The precise control of growth condition and determination of the optimum growth parameters are crucial to achieving well-defined QD structures. *In situ* monitoring is an important tool for ensuring reproducibility and feedback on the structure formation. In high vacuum, growth information is obtainable by means of the well-developed electron or ion beam methods, such as reflection high-energy electron diffraction used with MBE.[5] However, these methods cannot be used for reactive ambience or pressures outside the ultrahigh vacuum range. Problems remain in reproducible quantum dot fabrication by means of MOVPE due to the lack of *in situ* characterization methods. Considerable attention has recently been directed to developing optical *in situ* monitoring technique applicable to MOVPE,[6–8] but the development is still in the elementary stage, because optical probes interact only weakly with the growth surface and because the optical signal is difficult to be interpreted.

In this article, we present an ellipsometric study of QD formation. The trajectory of the ellipsometric signal tends to be inflected in the early stage of the $In_{0.5}Ga_{0.5}As$ growth. The inflection point corresponds to the rapid increase in the imaginary part of pseudodielectric function ϵ_2, deduced from increased scattering loss due to surface roughening by the QD formation. The good agreement among the data taken by *in situ* ellipsometry, *ex situ* atomic force microscopy (AFM), and photoluminescence (PL) implies that this inflection point corresponds to the start of QD formation. Results indicate that this technique precisely determines the growth mode transition from the 2D layer growth to the 3D island growth. Based on these results, we study QD formation onset and growth mode transition in QD formation process.

II. EXPERIMENT

Our growth system is based on a conventional Emcore Discovery 75 MOVPE system. The working pressure was 60 Torr. Using 1400 rpm substrate rotation throughout the growth ensured the uniform source gas flow. GaAs (001) nominally oriented substrates and 0.5° off substrates misoriented toward the [010] direction were used in our experiment. After chemical etching, oxide layers on substrates were removed by heating at 680 °C for 5 min in an arsine (AsH_3) ambient, then the substrate temperature was elevated to growth temperature T_g. Triethylgallium (TEG), trimethylaluminum (TMA), trimethylindium (TMI), and AsH_3 were used as source materials. TEG, TMA, TMI, and AsH_3 flow rates were 1.33×10^{-6}, 2.6×10^{-6}, 1.19×10^{-6}, and 1.83×10^{-3} mol/s, respectively. The $In_{0.5}Ga_{0.5}As$ growth rate was 0.1 monolayer (ML)/s and that of $In_{0.5}Al_{0.5}As$ 0.2 ML/s.

Conditions of the experiment were as follows: A Faraday-modulated self-nulling ellipsometer (Waterloo Digital Electronics: EXACTA 2000) was used to monitor MOVPE growth *in situ*.[9] A 632.8 nm He–Ne laser was used for the probe light. The laser has a 4 mW power output (spot size: 1 mm in diameter) and is linearly polarized with a polarization

[a]Author to whom correspondence should be addressed; electronic mail: jslee@obl.cl.nec.co.jp

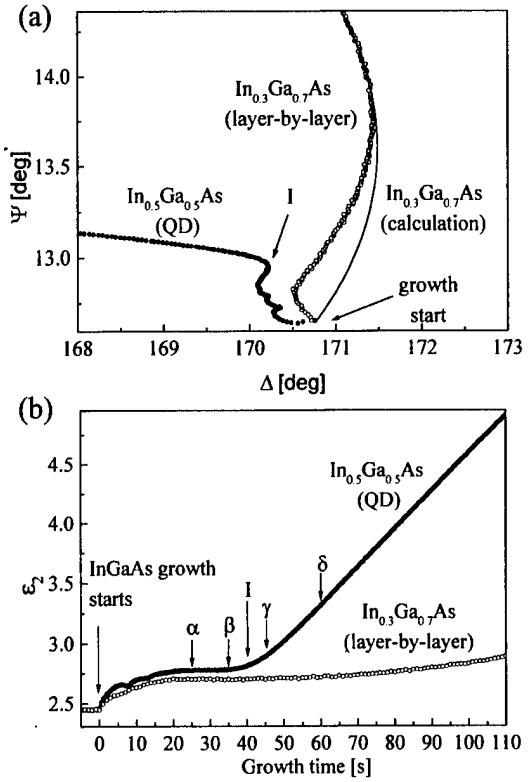


FIG. 1. (a) $\Delta-\Psi$ trajectories with increasing film thickness, measured for InGaAs film deposition on GaAs substrates at 495 °C and analogous calculation for $In_{0.3}Ga_{0.7}As$ layer-by-layer growth. (b) The ϵ_2 derived from the ellipsometric signal in (a) is shown as a function of growth time. The arrows α, β, γ, and δ correspond to $In_{0.5}Ga_{0.5}As$ growth stoppage in nominal 2.5, 3.5, 4.5, and 6 ML thickness.

ratio of 500:1. The incident angle was set to 70°. Data were acquired by means of lockin technique (time constant: 100 ms) and the interval between data acquisitions was 500 ms. Noise levels during growth were typically 0.02° for Δ and 0.01° for Ψ. *Ex situ* surface morphologies were studied by AFM in air. Room temperature (RT) PL measurements were done under the excitation by an Ar^+ laser tuned to 488 nm.

III. RESULTS AND DISCUSSION

Typical examples of the ellipsometric signal during MOVPE growth for InGaAs/GaAs systems are shown in Fig. 1(a). The figure shows the trajectory on the $\Delta-\Psi$ plane with increasing layer thickness of InGaAs in the dot formation growth and in the layer-by-layer growth on GaAs (001) substrates at 495 °C.

The effective complex dielectric functions of the InGaAs films $\epsilon=\epsilon_1+i\epsilon_2$, which are obtained by effective medium theory, are connected with the complex dielectric functions of GaAs. Investigations showed that the Bruggeman effective medium theory expression follows from an aggregate model, where particles of phases are mixed on a random basis.[6] Hence, the Bruggeman effective medium approximation is suitable for the InGaAs films. Accordingly, the effective complex dielectric function, ϵ, of InGaAs films is obtained from the following equations:

$$f_{\mathrm{InAs}}\frac{\epsilon_{\mathrm{InAs}}-\epsilon}{\epsilon_{\mathrm{InAs}}+2\epsilon}+f_{\mathrm{GaAs}}\frac{\epsilon_{\mathrm{GaAs}}-\epsilon}{\epsilon_{\mathrm{GaAs}}+2\epsilon}=0, \tag{1}$$

$$f_{\mathrm{InAs}}+f_{\mathrm{GaAs}}=1, \tag{2}$$

where f_i and ϵ_i are the volume fraction (composition) and dielectric function of each component. Since the dielectric functions of GaAs and InAs depend on temperature, the values of ϵ were not cited from the literature but obtained from the experimental data by the homoepitaxial growth at 495 °C on semi-insulated GaAs and InAs substrates, respectively. The ϵ_1 and ϵ_2 values for GaAs are 17.093 and 2.462 and those for InAs are 17.798 and 6.277. The GaAs and InAs samples having smooth surfaces in atomic order and good crystalline quality were confirmed by AFM and x-ray diffraction measurements. Consequently, the ϵ_1 and ϵ_2 values for the samples do not include any effects of surface roughness, microvoids, etc. The composition of InGaAs layer was determined based on the results of x-ray diffraction measurements. In compositions of QD and layer-by-layer samples obtained from x-ray measurements were 0.5 and 0.3, respectively. The ϵ_1 and ϵ_2 values for $In_{0.3}Ga_{0.7}As$ derived from Eqs. (1) and (2) are 17.305 and 3.348. The simulation program used for *in situ* measurements is commonly called the "paint" approximation, i.e., each part of a layer as it is deposited has a fixed value of index of refraction $n-ik$ which does not change as subsequent parts of the layer are deposited.[9] The model assumes that one can change the index of refraction of sublayers of a layer, but once that sublayer has been deposited its index cannot change subsequently. Consequently, the latest part of the $\Delta-\Psi$ trajectory is associated with the properties of the sublayer that is being deposited at that moment. The result of analogous calculation for $In_{0.3}Ga_{0.7}As/GaAs$ assuming layer-by-layer growth was shown in Fig. 1(a).

In Fig. 1(a), the $In_{0.3}Ga_{0.7}As$ layer-by-layer growth trajectory shows continuous change, mainly with increase in Ψ as the layer thickness increases. In wider $\Delta-\Psi$ plane, the trajectory spirals, indicating a continuous increase in layer thickness.[10] Observed and simulation curves are in good agreement in $In_{0.3}Ga_{0.7}As$ case. The deviation of the observed growth curve from the simulation curve at initial stages of InGaAs growth is due to the island structure[11] or strain effect caused by large lattice mismatch. On the other hand, the $In_{0.5}Ga_{0.5}As$ trajectory growth does not spiral but inflects at an early growth stage [I in Fig. 1(a)]. Correlation between the inflection point and the QD formation is suggested, because the inflection point appears at a critical layer thickness and does not appear in $In_{0.3}Ga_{0.7}As$ layer-by-layer growth. The rapid growth transition (within a second) from 2D to 3D at a certain layer thickness in SK growth mode may cause the inflection, whereas only the layer thickness gradually increases with growth time in layer-by-layer growth case. To obtain intuitive comprehension, we plotted pseudodielectric functions ϵ_2, derived from the ellipsometric signal in Fig. 1(a), as a function of growth time. ϵ_2 increases rapidly when the $\Delta-\Psi$ trajectory passes through the inflection point whereas ϵ_1 changes little at around the inflection

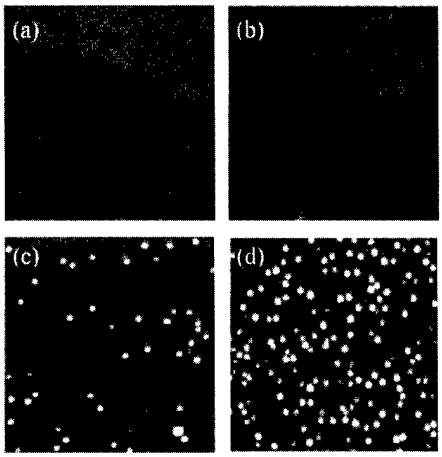


FIG. 2. Surface AFM images of $In_{0.5}Ga_{0.5}As$ layers corresponding to the surface of the samples for which growth was stopped (a) at α (b) at β, (c) at γ, and (d) at δ in Fig. 1(b). The area of all images in 1 μm×1 μm.

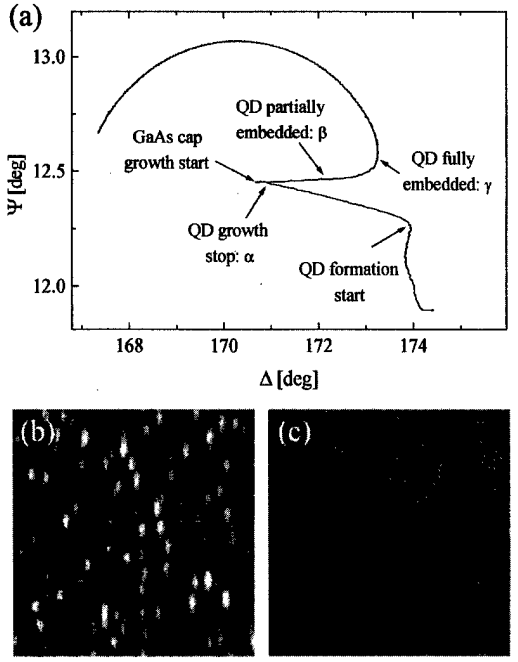


FIG. 3. (a) Δ–Ψ trajectory of GaAs capping layer growth following the $In_{0.5}Ga_{0.5}As$ QD formation. 1 μm surface AFM images of $In_{0.5}Ga_{0.5}As$ QD samples (b) at the point β and (c) at γ in (a).

point (data not shown). The rapid increase in ϵ_2 in spite of continuous and fixed source gas supply is considered to be due to rapidly increased surface roughness because of the QD formation. To study this more precisely, we conducted *ex situ* AFM measurement of the $In_{0.5}Ga_{0.5}As$ sample surfaces whose growth was stopped before and after the inflection point. Figures 2(a)–2(d) correspond to $In_{0.5}Ga_{0.5}As$ growth stoppage α (2.5 ML in nominal thickness), β (3.5 ML), γ (4.5 ML), and δ (6 ML) in Fig. 1(b). AFM clearly proved that the deviation from the spiral trajectory comes from growth mode transition. At the start of growth [α in Fig. 1(b)] ϵ_2 increases slowly and the smooth surface consists of 1 ML high islands and steps [Fig. 2(a)]. Both the ellipsometric signal and surface morphology show the same smooth 2D growth mode. As the $In_{0.5}Ga_{0.5}As$ layer thickness increases, the number of islands increases and islands are piled upon each other. At the point β in Fig. 1(b), small island formation occurs on larger islands as is shown in Fig. 2(b). Though all island sidewalls are 1 ML in height, the piled up $In_{0.5}Ga_{0.5}As$ layer thickness locally exceeds the critical layer thickness for the QD formation, becoming the center of nucleation. As $In_{0.5}Ga_{0.5}As$ becomes thicker, the island height increases and QDs are formed. This drastic change in surface morphology reflects the occurrence of inflection points in the trajectory. QDs are formed on the surface after the inflection point seen in Fig. 2(c). Thus, the rapid increase in ϵ_2 originates from the increase in surface roughness caused by morphological surface change. For the origin of increase in ϵ_2, Mie scattering that dominates when the island size becomes larger than λ (wavelength of probe light)/10 is suggested.[8] Taking the size of the islands [350 Å in diameter and 100 Å in height, in Fig. 2(c)], probe light wavelength of (6328 Å), and refractive index of InGaAs (n: approximately 4) into account, an increase in ϵ_2 by scattering is plausible. The monotonic increase in ϵ_2 observed after the inflection indicates a continuous increase in surface roughness. The increase in QD density and/or the increase in size with thicker $In_{0.5}Ga_{0.5}As$ layer enhance surface roughening. The linear increase in ϵ_2 with growth time increase indicates an increase in the number of QDs rather than the increase in size because scattering efficiency is linearly proportional to the number of dots and the fourth power of size $[(R/\lambda)^4$, R: radius of particle].[12] AFM results also show that the increase in the number of QDs is dominant compared with QD size increase at the initial stages of QD formation [Fig. 2(d)].

PL was also measured to confirm the correlation between QD formation and the ellipsometric signal. Since cooling a sample in MOVPE growth takes a few minutes, surface structure may change. So, to maintain the $In_{0.5}Ga_{0.5}As$ structure as grown, the GaAs capping layer was grown as soon as possible. Practically, a 2 s growth interruption was inserted between the $In_{0.5}Ga_{0.5}As$ and GaAs capping layer to avoid source gas mixing. Surface morphology during the GaAs capping layer growth was also monitored with ellipsometry [Fig. 3(a)]. When capping layer growth starts, Δ increased but Ψ scarcely changed, indicating surface smoothing starts in capping layer growth. The completion of QD embedding is estimated from the ellipsometric signal at the point the trajectory begins to spiral. The surface AFM images of the samples whose GaAs capping layer growths were stopped at points β and γ are shown in Figs. 3(b) and 3(c). The surface morphology of as-grown QD sample corresponds to Fig. 1(d). As seen in the figures, AFM results show good agreement with the ellipsometric signal. The image shows the surface morphology QDs partially embedded at the point β, and flat and smooth surface was observed at the point γ the trajectory begins to spiral. In the case of quantum well structure, the signal begins to spiral at the beginning of GaAs capping layer growth and surface remains smooth throughout the process.

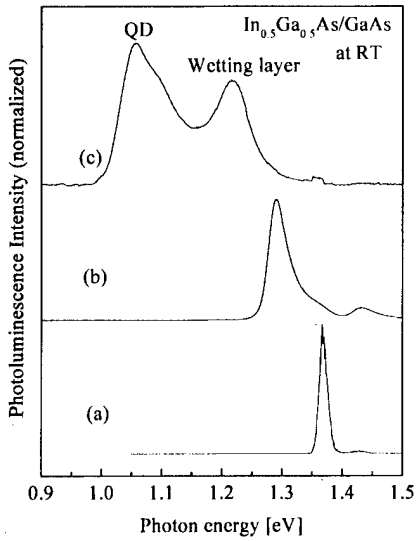


FIG. 4. Room temperature PL of $GaAs/In_{0.5}Ga_{0.5}As/GaAs$ structures. Spectra in (a) correspond to the sample in which $In_{0.5}Ga_{0.5}As$ layer growth was stopped at α, (b) at β, and (c) at γ in Fig. 1(b).

Figure 4 shows the RT PL spectra of the $GaAs/In_{0.5}Ga_{0.5}As/GaAs$ structures. The $In_{0.5}Ga_{0.5}As$ layer growth was stopped at α, β, and γ in Fig. 1(b). A sharp PL peak at 1.37 eV corresponding to radiation from the quantum well [Fig. 4(a)] shows a redshift and PL broadening with increasing $In_{0.5}Ga_{0.5}As$ layer thickness [Fig. 4(b)]. Peak broadening shown in Fig. 4(b) originates in layer thickness fluctuation, confirmed by AFM [Fig. 2(b)]. PL spectra for the sample in which $In_{0.5}Ga_{0.5}As$ growth was stopped after the inflection point show a double peak feature corresponding to radiation from the wetting layer and the QD region. The results of AFM and PL measurement thus show good agreement with the ellipsometric signal, confirming that ellipsometric signals describe surface structures well.

The ellipsometric signal thus provides the precise information on surface morphology during MOVPE growth. Consequently, *in situ* ellipsometry monitoring is concluded to be useful in determining QD formation onset and evaluating growth mode transition. Furthermore, we studied the dependence of QD formation onset on substrate orientation. It has been reported that QD formation on misoriented substrates occurs at thinner InGaAs layer than for an exactly oriented substrate under the same growth condition.[13] This phenomenon was confirmed in our present experiment, as is shown in Fig. 5(a). Looking at the ϵ_2 for samples on (001) exactly oriented and 0.5° misoriented GaAs substrates, the ϵ_2 trajectory on the misoriented substrate shows an inflection point earlier than that for an exactly oriented substrate, implying earlier QD formation. To confirm the QD formation onset, AFM measurement was performed. Figure 5(b) shows the surface AFM images of the sample in which $In_{0.5}Ga_{0.5}As$ growth was stopped at α and β in Fig. 5(a). On the sample surface for which $In_{0.5}Ga_{0.5}As$ growth was stopped at α, QD is not formed on the exactly oriented substrate, but is formed on the misoriented one, in good agreement with the results of ellipsometry. QDs are formed selectively on step edges, suggesting some relation between steps and QD formation. However, on the exactly oriented substrate, which also has steps, there is no QD formation. QD formation thus depends not only on the steps themselves but also on the distance between steps. The shorter distance between the steps leads to localization of QDs at step edges.[13] Low surface diffusion and increase in area dot density are plausible for a high surface step density. The larger ϵ_2 after the inflection point indicates a higher density of QDs as seen in AFM images [Fig. 5(b)].

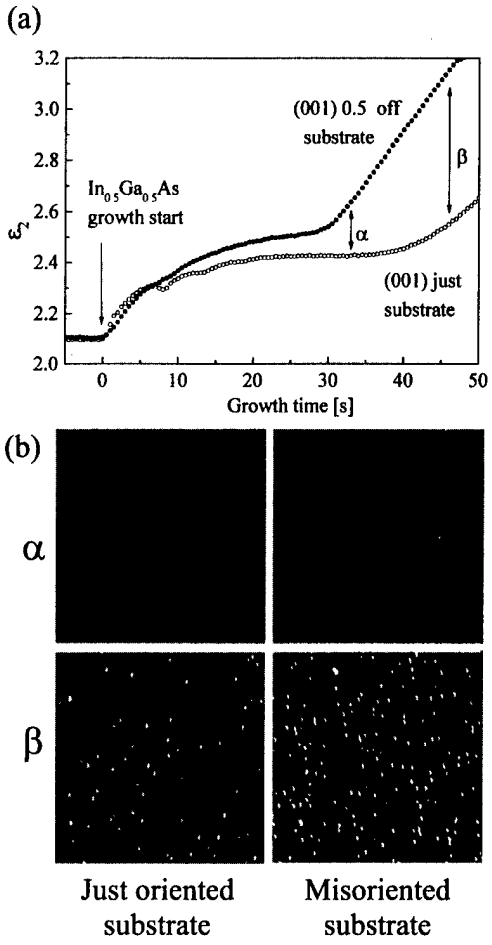


FIG. 5. (a) Pseudodielectric functions ϵ_2 derived from the ellipsometric signal as a function of $In_{0.5}Ga_{0.5}As$ growth time on a GaAs (001) exactly oriented substrate and a 0.5° off substrate misoriented toward the [010] direction. (b) 1 μm×1 μm surface AFM images of samples for which $In_{0.5}Ga_{0.5}As$ growth was stopped at α (3.3 ML in nominal) and β (4.6 ML) in (a).

The ellipsometric signal also shows a difference in QD formation. Figure 6(a) shows ϵ_2 for the $In_{0.5}Al_{0.5}As$ layer growth on the GaAs (001) substrate at 480 and 550 °C. *Ex situ* AFM measurements were also performed for the sample whose $In_{0.5}Al_{0.5}As$ layer growth was stopped at α and β in Fig. 6(a). The different ϵ_2 is due to a difference in T_g. At 550 °C, the trajectory has an inflection point indicating the growth mode transition just as in the InGaAs QD growth. The transition from 2D to 3D growth mode was confirmed by AFM as seen in Fig. 6(b). The surface showing flat fea-

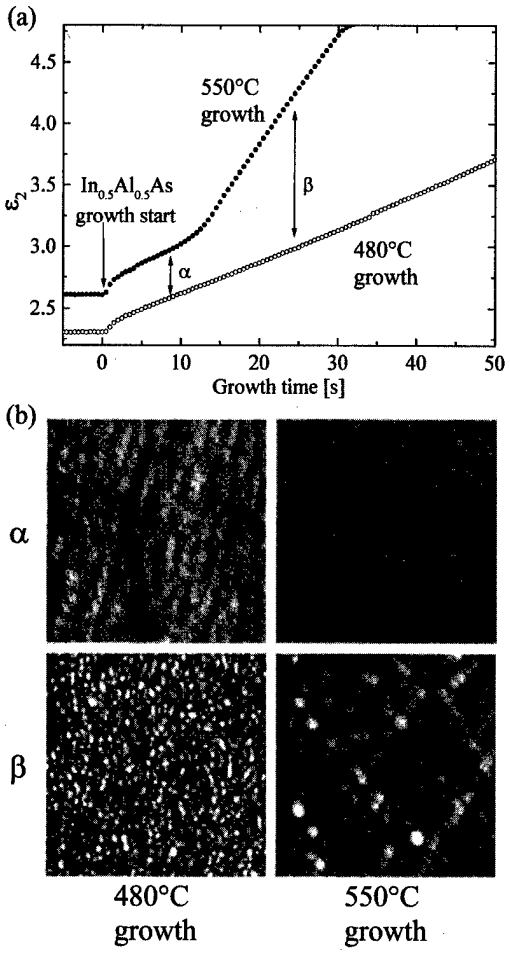


FIG. 6. (a) Pseudodielectric functions ϵ_2 derived from the ellipsometric signal during $In_{0.5}Al_{0.5}As$ QD growth at 480 and 550 °C. (b) 1 μm×1 μm surface AFM images of samples for which $In_{0.5}Ga_{0.5}As$ growth was stopped at α and β in (a).

ture at the point α changes at point β and QDs averaging 80 Å in height and 400 Å in diameter are formed.

On the other hand, ϵ_2 increases monotonously from the start of $In_{0.5}Al_{0.5}As$ growth at 480 °C indicating no obvious growth mode transition at lower T_g. In wider Δ–Ψ plane, the trajectory spirals, indicating a continuous increase in layer thickness. A comparison of surface AFM images for each sample clearly shows the difference in QD formation process. The AFM results for the samples grown at lower T_g indicate that the QD size increases continuously in contrast to the drastic structural change in the SK growth mode. Small islands less than 100 Å in diameter and a few monolayers in height are formed at the first stage of the $In_{0.5}Al_{0.5}As$ growth. The volume of each QD increases gradually as the growth proceeds, and the increase rate of QD height is almost the same as the nominal $In_{0.5}Al_{0.5}As$ layer growth rate. However, the QD height is limited to 30 Å because of small migration length of adatoms due to low T_g. In this case, the strain may be relieved by introduction of dislocation. Since the size of QDs grown at lower T_g is below the scattering detection range, the trajectory shows no difference between that of layer-by-layer growth. Thus, the ellipsometry signal shows a different feature reflecting a difference in the growth mode.

As mentioned above, *in situ* ellipsometry monitoring provides information on the QD formation onset, the difference in island formation, and can further help in understanding the QD formation mechanism.

IV. CONCLUSION

In situ ellipsometric monitoring of $In_{0.5}Ga_{0.5}As$ QD formation shows that the ellipsometric signal markedly changes its trajectory before and after QD formation. A comparison of results obtained from *ex situ* AFM and PL measurement confirmed that the inflection point of the ellipsometric signal corresponds to QD formation onset. This technique enables us to precisely determine the mode transition from 2D layer to 3D island growth. In addition to the information on the QD formation onset, the ellipsometric signal informs us of the different QD formation process depending on growth conditions. It thus turns out to be very useful in understanding QD formation and is applicable to optimizing growth parameters and ensuring the reproducible QD structure fabrication.

[1]H. Sakaki, Jpn. J. Appl. Phys. **19**, L735 (1980).
[2]L. Goldstein, F. Glas, J. Y. Marzin, M. N. Charasse, and G. L. Roux, Appl. Phys. Lett. **47**, 1099 (1985).
[3]R. Notzel, J. Temmyo, and T. Tamamura, Jpn. J. Appl. Phys., Part 2 **33**, L275 (1994).
[4]J. Oshinowo, M. Nishioka, S. Ishida, and Y. Arakawa, Appl. Phys. Lett. **65**, 1421 (1994).
[5]Y. Nabetani, T. Ishikawa, S. Noda, and A. Sasaki, J. Appl. Phys. **76**, 347 (1994).
[6]D. E. Aspnes, Am. J. Phys. **50**, 704 (1982).
[7]J. M. Olson and A. Kibbler, J. Cryst. Growth **77**, 182 (1986).
[8]E. Steimetz, F. Schienle, J.-T. Zettler, and W. Richter, J. Cryst. Growth **170**, 208 (1997).
[9]H. X. Tran, Ph.D. thesis, University of Waterloo, Waterloo, Ontario, Canada, 1996.
[10]D. E. Aspnes, Surf. Sci. **307-309**, 1017 (1994).
[11]M. Yamamoto and T. Namioka, Appl. Opt. **31**, 1612 (1992).
[12]C. F. Bohren and D. B. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983), p. 136.
[13]M. Kimura, M. Nishioka, J. Oshinowo, and Y. Arakawa, Appl. Phys. Lett. **66**, 3663 (1995).

High temperature reaction of nitric oxide with Si surfaces: Formation of Si nanopillars through nitride masking and oxygen etching

K. Prabhakaran[a] and T. Ogino
NTT Basic Research Laboratories, 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa-ken 243-0198, Japan

(Received 13 October 1998; accepted 5 April 1999)

In this article, we report the fabrication of Si nanopillars by surface reaction with nitric oxide (NO) at high temperature. NO was leaked onto clean Si surfaces, at a temperature of ~850 °C, in ultrahigh vacuum chamber and examined *in situ* by x-ray and ultraviolet photoelectron spectroscopy, Auger electron spectroscopy, low energy electron diffraction, and *ex situ* by atomic force microscopy. NO molecules dissociate on the surface and nitrogen atoms thus produced form nitride islands. These islands act as protective masks for the etching of Si by the oxygen atoms, through the desorption of SiO species. Occurrences of these two simultaneous processes results in the formation of nanometer sized Si pillars (typical basewidth 100–150 nm and height 5–15 nm) capped by silicon nitride. These pillar structures are separated by clean Si areas as shown by the clear presence of two domains. We also show that the height of these Si pillars can be increased by *ex situ* chemical etching. © *1999 American Vacuum Society.* [S0734-211X(99)00304-2]

I. INTRODUCTION

Semiconductor nanostructure formation in general and that of Si in particular have been gaining alot of attention both by the theoreticians and the experimentalists. The main intention behind this interest is to explore new avenues of application such as in optical device and quantum effect devices.[1–3] If these structures can be fabricated in a three-dimensionally periodic fashion, then it may also find application as photonic crystals.[4] In our laboratory, we have been working on various ways to control the formation of nanostructures by suitably modifying the substrates.[5] Another approach to achieve formation of nanostructures is through controlling surface chemical reactions. We call this approach "chemical bond manipulation." The main advantage of this approach is the possibility to achieve formation of the desired structures on the whole surface of the wafer, while the controllability of the size and distribution has been a challenge. The important step in this approach is to identify the reaction pathways, selectively with regard to surface bond formation, bond cleavage, and desorption, which are governed by the thermodynamics and kinetics of the system. Following the above guidelines, we recently succeeded in employing the bonding partner change reaction of oxygen from Ge to Si to fabricate multiperiod, nanometer sized layered structure consisting of $Si/SiO_2/Ge$ units on a Si(111) wafer.[6] Tabe and Yamamoto[7] reported fabrication of ultrafine Si structures on a Si(111) substrate through a two step process involving nitridation with N_2, which was followed by local oxidation of Si. Their study, based on scanning tunneling microscopy and x-ray photoelectron spectroscopy (XPS), demonstrated the fabrication of Si nanopillars with an average diameter of 5 nm and a height of 0.3–0.6 nm. Here, we demonstrate that surface chemical reaction and nitric oxide (NO) with Si surface can be effectively utilized for the formation of Si pillars, in a single step process. There are several reports on the reaction of NO with Si surfaces, addressed to oxynitridation.[8–10] This is a report on Si nanostructure formation by employing the surface reaction. The main advantage of using NO instead of N_2 is that, under the present experimental conditions, the dissociation probability and therefore the sticking coefficient of nitrogen is higher in the case of NO compared to that of N_2, due to the difference in their bond strengths.[11] Since the nitride islands act as a sort of template for the formation of Si pillars, the island size is crucially related to the size of the Si pillars.

II. EXPERIMENT

Si wafers (*n*-type 1–5 Ω cm) were cleaned by thermally decomposing an ultrathin *ex situ* prepared chemical oxide,[12] inside the ultrahigh vacuum (UHV) chamber. Samples prepared in this manner were very clean (carbon level less than 1%) and exhibited sharp low energy electron diffraction (LEED) patterns and strong surfaces states in ultraviolet photoelectron spectroscopy (UPS). The UPS and XPS measurements are performed using an electrostatic analyzer (VG CLAM II, equipped with three channeltrons) with He I and Al/Mg $K\alpha$ sources, respectively, in the normal emission geometry. All the spectroscopic measurements were carried out at room temperature. NO was leaked onto clean Si surfaces from a gas bottle (50 cc) attached to the UHV chamber through a needle valve. In the case of etching with oxygen, molecular oxygen was dosed onto the nitrided surface at a temperature of 850 °C. Sample morphology was examined, outside the UHV chamber, using a DI atomic force microscope (AFM). We performed the study on (100) and (111) orientations of the Si wafer. All the sample handling and the measurements were carried out in a clean room. High pressure reactions (typically $\sim 1 \times 10^{-5}$ Torr) with NO were carried out in the reaction chamber attached to the main chamber.

[a]Electronic mail: prab@will.brl.ntt.co.jp

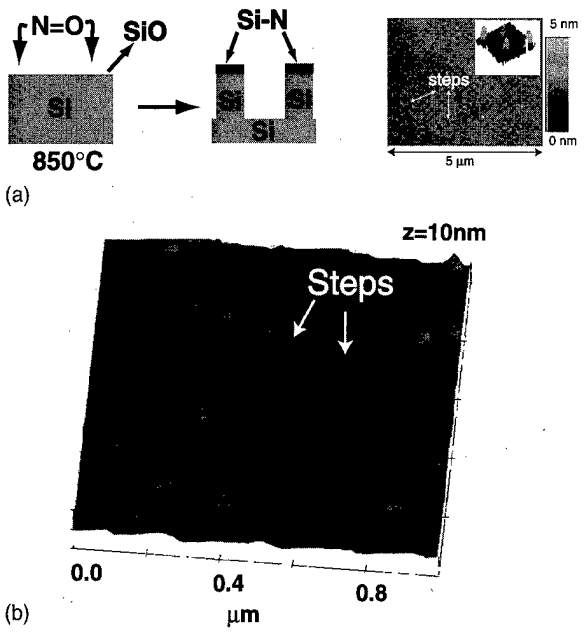


FIG. 1. Schematic of the high temperature reaction of NO with Si surface, which results in the formation of Si nanopillars capped by nitride islands. Also shown is an AFM image showing the morphology in the early stage of the reaction (10 L NO at 850 °C). Surface steps become irregular in the etching process. (b) shows the Si(111) surface reacted with NO at 1×10^{-6} Torr for 1 h. Etching occurs at step edges and is pinned by the nucleation of nitride islands.

III. RESULTS AND DISCUSSION

In Fig. 1(a) we show schematically, the principle behind the process of Si nanopillar formation. When NO molecule reaches the surface, it dissociates and nitrogen atoms stick to the surface forming nitride islands. Since the substrate temperature is high, oxygen atoms etch off the surface Si atoms through the desorption of SiO species and nitride islands act as protective masks for this etching process. It should be noted that both the island formation and the etching processes take place simultaneously. Continuation of both the processes by long exposure to NO gas, results in the formation of nanometer sized Si pillars capped by nitride islands. The advantage with this single step process is that the nitride capping also grows thicker and therefore its protective action is significantly enhanced and even becomes resistant to chemical etching performed outside the UHV. The region between the pillars seems to be preferentially etched by the oxygen rather than nitridized during the later stages of the reaction. This selectivity may be linked to the thermodynamics and kinetics of the two competitive processes. In Fig. 1(a) we also show the AFM in the initial stages of the etching mode reaction on Si(111). We can see the corroded step edges where the step edges become irregular (due to partial etching) and this phenomenon seems to be similar to the step flow growth occurring in the case of homoepitaxy of Si.[13] This means that atom attachment/detachment process takes place selectively at step edges. The AFM image clearly indicates that the etching process starts at the step edges and is sporadically pinned by the formation of nitride islands at step edges. This is clearly shown in Fig. 1.

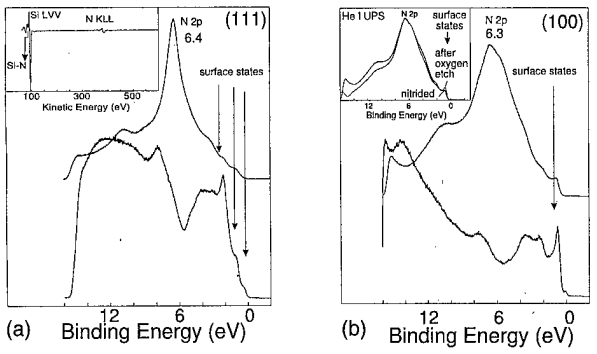


FIG. 2. He I UPS recorded before and after the reaction with NO at ~850 °C in the case of (a) Si(111) and (b) Si(100) surfaces. The spectra clearly indicate the formation of domains consisting of clean Si areas as well as nitride islands. Inset to (a) represents typical AES recorded after the reaction with NO, showing the peaks due to surface nitride species. Inset to (b) shows a comparison of the UPS of the nitrided sample (containing Si pillars) and after etching the same sample with oxygen [corresponding AFM image is shown in Fig. 5(a)]. The signal due to dangling bond states increases due to the emergence of clean Si areas.

Figure 2 plots the He I UPS of clean surface as well as the surface reacted with NO at 850 °C. Both the (111) and (100) orientations are separately shown. As can be seen in the spectra, the clean surfaces are characterized by the presence of sharp emission from the dangling bond states of the atoms in the termination layer. After the reaction, surface states continue to be observed, with diminished intensity and clear peaks (~6.4 eV due to N 2*p* derived orbitals) due to surface nitride species are also observed. This indicates the formation of two different domains consisting of bare Si areas and the Si nanopillars capped by the nitride species. In the insert to Fig. 2(a), a typical Auger electron spectrum recorded after the reaction is shown and strong signals due to surface nitride species are observed at kinetic energies of 84 and 387 eV. Figure 3 shows the corresponding XPS spectra. Si 2*p* shows a chemically shifted component, by ~2.8 eV typical of the formation of Si_3N_4 type of nitride and the N 1*s* is observed at 398 eV.[14,15] One important difference between (111) and (100) surfaces is the nature of the N 2*p* derived peak in the He I UPS. The peak is narrower in the case of (111) surface, compared to that of (100) surface. We believe

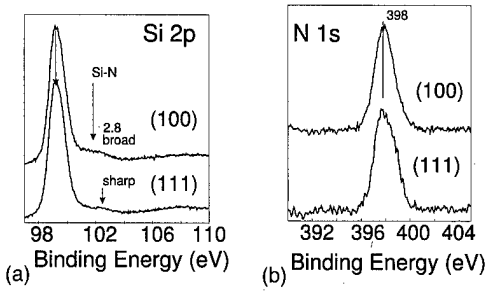


FIG. 3. XPS spectra in the (a) Si 2*p* and (b) N 1*s* regions recorded after the high temperature reaction with NO, for both (111) and (100) surfaces.

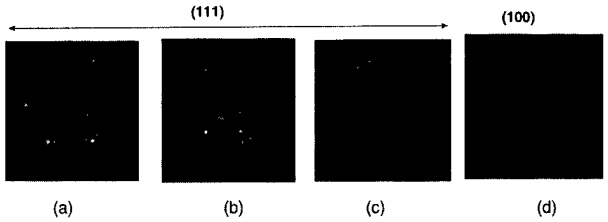


FIG. 4. LEED patterns during various stages of the reaction with NO. (a) Diffuse 7×7 with submonolayer coverage of N, $E_p = 38$ eV. (b) Mixture of 7×7 and emergence of quadruplet patterns at ~1 ML coverage, $E_p = 58$ eV. (c) Mixture of 1×1 and 7×7 patterns at above monolayer coverages of N, $E_p = 59$ eV. (d) Weak (1×1) at above monolayer coverages on Si(100) surface $E_p = 52$ eV.

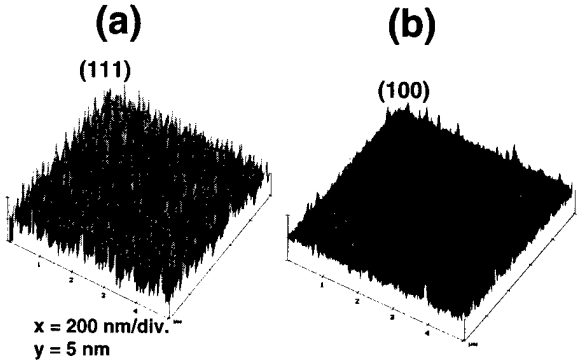


FIG. 5. AFM image after dosing the Si (111) [image (a)] and Si(100) [image (b)] surfaces with NO (1 h at 1×10^{-8} Torr) at a substrate temperature of 850 °C.

it is due to the better ordering in the nitride layer. For the same reason, the signal due to nitride in the Si 2*p* spectrum is broad in the case of (100) compared to that in (111) (see Fig. 3). The LEED observation also substantiates this view (Fig. 4). In the case of (111) surface, during the early stages of the reaction with NO, 7×7 pattern becomes diffuse [Fig. 4(a)]. Similarly, in the case of (100), a diffuse (2×1) pattern is observed during the early stages of NO exposure [pattern not shown in Fig. 4(a)]. However, as nitrogen coverage increases (approximately 1 ML) a mixture of 7×7 and quadruplet structures are seen in the case of (111) surface [Fig. 4(b)]. There is some evidence for an 8×8 pattern also at this stage. When the high temperature reaction is carried out for a longer period of time (1 h at 1×10^{-8} Torr) thicker nitride islands are formed (2–3 ML), and we observe a sharp pattern consisting of 1×1 and quadruplet structures [Fig. 4(c)]. The quadruplet pattern has been explained as due to triangular domains of nitride, rotated by an angle, with respect to the underlying substrate.[19] In the case of (100) surface, at higher nitrogen coverages, a diffuse 1×1 pattern is observed [Fig. 4(d)]. There has been a debate in the literature that the quadruplet pattern is due to possible carbon contamination.[16–18] The carbon level was less than 1% in our experiment [see inset to Fig. 2(a)] and therefore we believe that the quadruplet pattern formation may not be related to the presence of surface carbon. Moreover, the fact that the UPS exhibits a single sharp peak due to N 2*p* derived orbital, after the high temperature reaction [see Fig. 1(a)], with no additional signal due to carbon, indicates that the quadruplet pattern is originating from the surface nitride layer. These results are in agreement with the report by Schrott and Fain.[19]

Figure 5 shows the typical AFM image after the high temperature reaction with NO in the case of (111) and (100) surfaces. The images clearly show high density pillar formation in both cases. The nature of the pillars and their interaction between one another appear to be quite sensitive to the surface temperature. In Fig. 6 we illustrate that the height of these pillars can be increased by performing further etching of Si through additional means. For example, the pillars formed by reaction of NO at 850 °C were subjected to an etching reaction with oxygen (1×10^{-8} Torr for 1 h) at 850 °C. After this process we find that the average pillar height is increased by approximately 1 nm. It is worth noticing that the nature of the nitride species and its amount is unaffected by this etching process, as the N 1*s* as well as Si 2*p* signals recorded before and after the etching process were identical in intensity and peak position. Therefore, the nitride islands formed through the dissociative adsorption of NO are very stable and function effectively as a protective capping layer. The signal due to the dangling bond states in the He I UPS is found to increase after the oxygen etching. We show a comparison of the UPS before and after the oxygen etching in the inset to Fig. 2(b) and there is ~10%–20% increase in

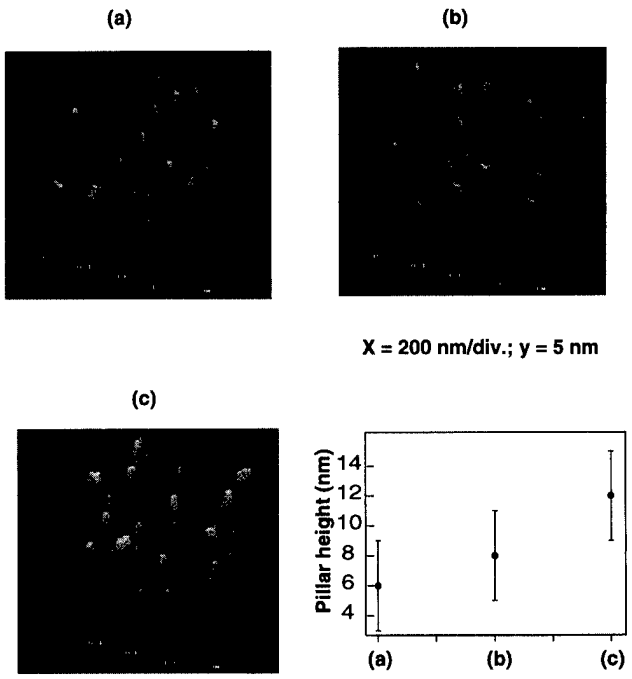


FIG. 6. Experimental demonstration of the fact that the pillar height can be increased through *in situ* and *ex situ* etching processes. Image (a) is from a Si(100) exposed to NO at 850 °C (1×10^{-8} Torr for 1 h) and subsequently etched with O_2 at 850 °C (1×10^{-8} Torr for 1 h). The comparison of UPS shown as an inset to Fig. 2(b) corresponds to this stage. (b) The same sample was taken out of the UHV chamber and etched with very dilute (1%–2%) HF for 25 s. (c) After three cycles of oxidation in very dilute (~5%) HNO_3 and HF etch. (d) is a plot of the typical pillar height after treatments (a), (b), and (c).

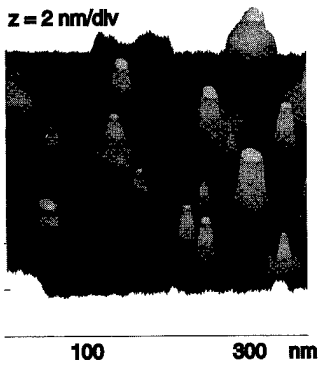


FIG. 7. AFM image of the pillar containing sample after it was left open in a clean room for a couple of days. The base and wall side are selectively oxidized and the capping nitride appears to be resistant to oxidation. This demonstrates the efficient protective action by the nitride cap on the Si pillars.

the intensity of the signal due to the surface states originating from the dangling bonds. This indicates that subsequent etching with oxygen exposes more of the clean Si areas. We speculate that these clean Si regions are possibly located around the wall side of these pillars. Figure 6(b) is the image of the same sample [same as Fig. 6(a)], which was given a mild etch *ex situ* (1%–2% HF, ~25 s). The silicon oxide formed due to the exposure to air is expected to be removed in this process. The average pillar height is increased by around 1–2 nm due to the removal of Si from the region between the pillars. Figure 6(c) is after the same sample was subjected to three cycles of oxidation and etching in very dilute HNO_3 and HF, respectively. The average pillar height increased quite significantly, by 3–4 nm after this process. Moreover, the small pillars are found to be etched off completely. We believe that the wall side of the pillar is more reactive from the fact that the AFM image of the pillar formed sample, which was left in air for long periods of time (typically a couple of days), showed the growth of features, possibly due to oxide and/or other contaminants, on the wall side of the pillars (Fig. 7).

IV. CONCLUSIONS

To summarize, we show that the reaction of No with Si surfaces at high temperatures results in the formation of nanometer sized Si pillars capped by silicon nitride, in a single step process. We also show that the height of these pillars can be increased by *ex situ* chemical etching. The main challenge ahead is to control the distribution of these structures which is mainly decided by the initial nucleation centers of the nitride islands.

[1]T.-C. Shen, C. Wang, G. C. Abeln, J. R. Tucker, J. W. Lyding, Ph. Avouris, and R. E. Walkup, Science **268**, 1590 (1995).
[2]R. C. Ashoori, Nature (London) **379**, 413 (1996).
[3]P. B. Fischer and S. Y. Chou, J. Vac. Sci. Technol. B **11**, 2524 (1993).
[4]E. Yablonovitch, T. J. Gmitter, R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, Phys. Rev. Lett. **67**, 3380 (1991).
[5]T. Ogino, H. Hibino, and K. Prabhakaran, J. Vac. Sci. Technol. B **14**, 4234 (1996).
[6]K. Prabhakaran, T. Matsumoto, T. Ogino, and Y. Masumoto, Appl. Phys. Lett. **72**, 3169 (1998).
[7]M. Tabe and T. Yamamoto, Appl. Phys. Lett. **69**, 2222 (1996).
[8]A. Kamath, D. L. Kwong, Y. M. Sun, P. M. Blass, S. Whaley, and J. M. White, Appl. Phys. Lett. **70**, 63 (1997).
[9]H. C. Lu, E. P. Gusev, T. Gustafsson, E. Garfunkel, M. L. Green, D. Brasen, and L. C. Feldman, Appl. Phys. Lett. **69**, 2713 (1996).
[10]R. I. Hegde, P. J. Tobin, K. G. Reid, B. Maiti, and S. A. Ajuria, Appl. Phys. Lett. **66**, 2882 (1995).
[11]*CRC Handbook of Chemistry and Physics* (Chemical Rubber, Boca Raton, FL, 1989), p. F-177.
[12]A. Ishizaka and Y. Shiraki, J. Electrochem. Soc. **133**, 666 (1986).
[13]H. Hibino and T. Ogino, Appl. Phys. Lett. **67**, 915 (1995).
[14]L. Kubler, J. L. Bischoff, and D. Dolmont, Phys. Rev. B **38**, 13113 (1988).
[15]M. Tabe and T. Yamamoto, Surf. Sci. **376**, 99 (1997).
[16]M. Yoshimura, E. Takahashi, and T. Yao, J. Vac. Sci. Technol. B **14**, 1048 (1996).
[17]E. A. Khramtsova, A. A. Saranin, and V. G. Lifshits, Surf. Sci. Lett. **280**, L259 (1993).
[18]M. D. Wiggins, R. J. Baird, and P. Wynblatt, J. Vac. Sci. Technol. **18**, 965 (1981).
[19]A. G. Schrott and S. C. Fain, Jr., Surf. Sci. **111**, 39 (1981).

Low energy electron beam decomposition of metalorganic precursors with a scanning tunneling microscope at ambient atmosphere

H. Brückl[a)]
University of Bielefeld, 33615 Bielefeld, Germany and Institute of Solid State and Materials Research, IFF, 01177 Dresden, Germany

J. Kretz
Institute of Solid State and Materials Research, IFF, 01177 Dresden, Germany

H. W. Koops
Deutsche Telekom AG, 64295 Darmstadt, Germany

G. Reiss
University of Bielefeld, 33615 Bielefeld, Germany

(Received 11 February 1999; accepted 21 May 1999)

Electron beam induced decomposition of metalorganic precursor substances with a scanning tunneling microscope is a convenient way to create nanometer-sized structures. Up to now, the application was limited due to the necessity of vacuum conditions. In this contribution, we report experiments to form metallic nanostructures by the decomposition of dimethylgold(III)-trifluoro-acetylacetonate and cyclopentadienylplatin(IV)-trimethyl on graphite and indium–tin–oxide coated glass substrates at ambient atmosphere. Small hillocks with diameters down to 30 nm and heights of 10 nm and wire-shaped structures were produced. The amount of the deposited material depends on the energy and the dose of the electrons. A threshold voltage was found for both materials. © *1999 American Vacuum Society.* [S0734-211X(99)07804-X]

I. INTRODUCTION

The potential of scanning probe microscopy as a suitable tool for laboratory-scale nanofabrication was demonstrated with a variety of different techniques. With scanning tunneling microscopes (STMs) single atoms were moved at low temperatures,[1] or small clusters were deposited by field evaporation from the tip.[2] The mechanical modification at the nanometer scale with a scanning force microscope (SFM) was also pursued.[3] Mask lithography was done by resist patterning[4] or selective oxidation.[5,6] Another promising method for the production of nanostructures is the direct writing with a STM or a SFM with a conducting tip by electron beam induced chemical vapor deposition (CVD).[7–9] In STM CVD low energy electrons (3–30 eV) from the tip decompose precursor molecules either from a gaseous environment or from adsorbates at the surface, in order to form directly small structures on the substrate surface. Mainly organometallic molecules such as nickel tetracarbonyl,[10] trimethylaluminum[11] or iron pentacarbonyl[12] served as precursor materials up to now. This class of material possesses organic ligands and metal–carbon bonds at the metal atom. The latter are not present in metalorganic precursors,[13] which are employed for additive lithography in scanning electron microscopy (SEM),[14–16] a similar technique to STM CVD. An inorganic precursor (tungsten hexafluoride) was tested once.[17] According to our knowledge, all experiments were performed up to now in an ultrahigh vacuum (UHV) environment, mainly on silicon substrates. In this article, we demonstrate the possibility of STM CVD of metalorganic precursors in ambient environment. The dependence of the structure sizes on the parameters of the deposition is presented.

II. EXPERIMENTAL SETUP

Patterning was done with two different instruments: a low-cost commercial STM (STMage GbR, Germany) and a home-built one. With these instruments, the access to the source code of the imaging software allows the application of our own versatile lithography software, simplifying the adjustment of patterning parameters, e.g., delay times. The latter instrument is placed in an environment chamber, which can be pumped down to a pressure of approximately 1×10^{-5} Torr and can subsequently be filled with a protective or reactive gas.

For the STM CVD process sharp tips were either cut from Pt/Ir or etched from W wires. The quality of the tips was controlled by SEM and by convolution analysis of images of rough surfaces.[18] In order to recognize deposited nanostructures and distinguish them from other surface bumps, we choose rather flat substrates, i.e., highly oriented pyrolytic graphite (HOPG) and indium–tin–oxide (ITO) thin film on glass.

In order to confine the vapor of the metalorganic molecules, a trough was formed by gluing a viton ring onto a glass plate. The substrates were fixed on this glass plate in the middle of the viton ring (Fig. 1). A small amount of solid precursor material was placed inside the ring. We used as precursors dimethylgold(III)-trifluoro-acetylacetonate (Me_2Au-tfac) and cyclopentadienylplatinum(IV)-trimethyl ($CpPtMe_3$) with low vapor pressures of 40 and 56 mTorr,

[a)]Electronic mail: brueckl@physik.uni-bielefeld.de

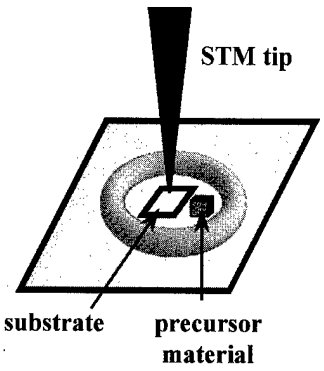


FIG. 1. Schematic of the setup for the STM CVD deposition in air. The substrate together with the solid precursor material is placed in a trough (viton ring) confining the vapor of the precursor molecules.

respectively, at room temperature. These materials slowly sublime at room temperature and spread over the area limited by the ring which prevents escaping. Due to the large molecular weight the metalorganic molecules separate from the "light" air and deposit as sediments.[19] The materials and conditions are summarized in Table I.

III. RESULTS AND DISCUSSION

A subtle point of the described procedure is the waiting time between placing the solid precursor material in the ring and the first attempts for CVD. This time is required to obtain proper deposition conditions. It turned out that the dot formation at atmospheric pressure was only possible after 2–3 h. This time depends on the vapor pressure of the precursor material. We suggest therefore that the precursor material covers the surface to a reasonable amount at least after some hours. This additionally implies that mainly adsorbed molecules on the surface contribute to pattern formation, and not gaseous molecules. After the waiting time, STM CVD deposition works reproducibly at least for a few hours in dry air, until the adsorbate layer becomes too thick for a reasonable STM operation. The growth rate of the adsorbate layer depends on the kind of substrate and the air humidity due to the hygroscopic properties of most of the precursor materials. On the hydrophobic graphite STM CVD works reproducibly for a few days. It should be mentioned that the waiting time can be shortened by pumping the STM and flooding again with gas after a few minutes.

A first scan with a low tunneling current of about 5 nA and a voltage of 100 mV determines the local topography before patterning. For example, in Fig. 2(a), a flat HOPG surface with one terrace step and a small defect in the upper

TABLE I. Parameters for STM CVD.

Precursor	Vapor pressure[a]	Substrate	Deposition technique	Environment	Threshold voltage
Me_2Au-tfac	40 mTorr	HOPG	Pulse mode	Air	−5 V
$CpPtMe_3$	54 mTorr	ITO glass	Pulse mode	N_2	−4 V

[a]At room temperature.

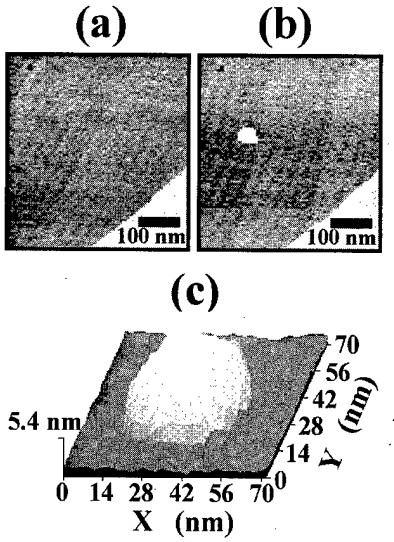


FIG. 2. Deposit from Me_2Au-tfac on HOPG substrate at a tip voltage of −6 V, current of −300 nA and a pulse time of 0.5 ms (ten times); (a) before, (b) after deposition, (c) landscape plot of a close-up of (b).

left corner is imaged prior to the patterning process. Then, the tip is moved to the desired location on the sample, and short current and voltage pulses are applied to the tip. The feedback is not active during the pulses, and, therefore, the tip is not drawn back to the field emission regime. The tip stays about 1 nm from the surface during the whole patterning procedure. Pulses shorter than 1 ms minimize external disturbances by mechanical or electromagnetic noise. The patterned nanostructures were subsequently characterized with the same tip. In our example [Figs. 2(b) and 2(c)] a small dot with a diameter of 30 nm and a height of 10 nm was deposited on the flat area between the terrace step and the defect with Me_2Au-tfac as precursor. The lateral extension of the dot in the image, however, is overestimated due to the convolution with the tip geometry and is therefore actually smaller than 30 nm. The size and volume of the deposits depend on the electron energy and dose in a systematic way as will be shown below.

For a reliable patterning process a single deposition step contains ten subsequent pulses at one location. This enhancement of reliability was also described by Rubel, Wang, and Delozanne for the in-vacuum experiments.[20] The number of pulses usually increases the amount of the deposited material, too, which we have not yet investigated systematically. The volume of the deposit also depends on the pulse time, applied voltage and current. The dependence of the deposit diameter and height on the tip voltage is shown in Fig. 3 for a fixed tunneling resistance and pulse time. Material deposition occurs above a negative threshold tip voltage of about −5 V. From the deposited volume, a deposition rate of about 100 atoms per μs can be calculated at −6 V. Furthermore, if a gold content of 50 at. % is assumed in the deposits (see text below), a yield of one gold atom per 2×10^4 impinging electrons[21] can be estimated from the current and the pulse duration at −6 V. The diameter (and height) of the deposits increases rapidly with larger voltages. Electrons of higher

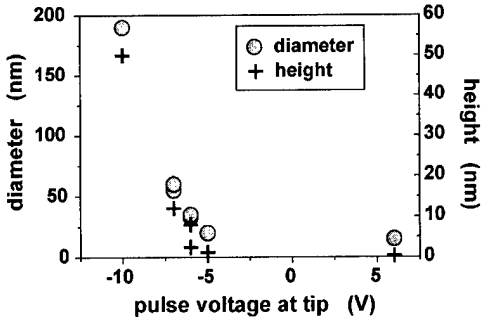


FIG. 3. Plot of the diameter and height of deposits from Me_2Au-tfac on HOPG substrate vs tip voltage (ten pulses of 0.5 ms) at a constant tunneling resistance of 20 MΩ.

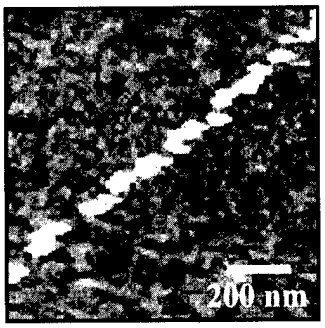


FIG. 5. Row of single dots deposited from $CpPtMe_3$ on ITO at a tip voltage of −6 V, current of −50 nA and a pulse time of 15 μs (single pulse).

energies are more effective: we calculated about 100 e^- which are necessary for the deposition of one gold atom at −10 V. Since in simple activated processes the yield γ_{Au} depends exponentially on the energy, $\gamma_{Au}(E) \sim \exp(E/E_A)$, we estimated an activation energy E_A of about 1 eV, which is reasonable in comparison to usual binding energies of organic compounds of 0.5–1.5 eV.

For positively biased tip voltages we found small and very flat (a few Å) deposits, which are nearly independent on the voltage strength. In this case, we suggest the deposition of carbon-based compounds like in Ref. 6, in which the authors also found a mound production for positively biased atomic force microscopy (AFM) cantilever tips.

The determination of the purity of the deposited substances is a difficult task. We succeeded in an Auger analysis of a large dot deposited at −10 V from Me_2Au-tfac. The spectra showed an approximate composition of 50 at. % gold and 50 at. % carbon. Oxygen and fluorine signals were insignificantly low. The gold content of the precursor material itself is only 5%. The same 50/50 composition was found by Folch and co-workers[22] by electron beam decomposition in an environmental SEM at a high partial pressure of 2 mTorr O_2. They showed a reduction of C content with increasing O_2 partial pressure. We suggest therefore that the presence of O_2 and H_2O in our ambient setup also favors a high Au content, even at room temperature. The 50/50 composition is conductive and can be scanned without crashing the tip at any voltage. The topography of larger deposits was investigated in detail and was found to consist of small grains with diameters of about 20 nm which look rather similar to crystallites (Fig. 4). It should be stressed that there are no studies of the growth mechanism at STM CVD processes up to now. The occurrence of grain growth is typical for a diffusion-controlled, low energy deposition process with a relatively high mobility of the atoms of the growing materials.[23]

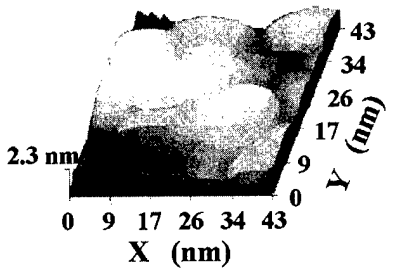


FIG. 4. STM topography of the grainy structure of a large deposit from Me_2Au-tfac on HOPG (at −10 V, −600 nA, 10×0.5 ms).

The deposition procedure described above is not restricted to Me_2Au-tfac and simple dot pattern. A more complicated pattern is created by exposing single sites in a linear row in order to form a platinum line from $CpPtMe_3$ on ITO substrate (Fig. 5). In this case, a N_2 atmosphere provides a more stable deposition condition than air. Similar to the Au deposition, the threshold voltage was about −4 V.

IV. CONCLUSION

In conclusion, we have succeeded in the deposition of small nanometer-sized structures by decomposing metalorganic precursor molecules with low energy electrons from a STM tip at ambient atmosphere. This STM CVD was formerly described only for in-vacuum applications. Our experiments differ from the previously reported attempts in the use of short current and voltage pulses instead of continuous electrical fields. The amount of the deposited material depends on the energy and the dose of the electrons. A threshold and an activation energy of about 5 and 1 eV, respectively, were determined for the Me_2Au-tfac decomposition, which are quite reasonable in comparison to usual binding energies. Accordingly, the yield increases with tip voltage. Thus, about 2×10^4 impinging electrons deposit one Au atom at a tip voltage of −6 V, for example. The smallest structures had 30 nm in diameter and 10 nm in height. The approximate atomic composition of the deposits was 50 at. % gold and 50 at. % carbon.

ACKNOWLEDGMENTS

The authors thank T. Dietrich (Merck Balzers AG, Liechtenstein) for courteously supplying ITO-coated glass substrates. This work was supported by the DFG.

[1]E. J. Heller, M. F. Crommie, C. P. Lutz, and D. M. Eigler, Nature (London) **369**, 464 (1994).
[2]H. J. Mamin, P. H. Guethner, and D. Rugar, Phys. Rev. Lett. **65**, 2418 (1990).
[3]H. Brückl, R. Rank, H. Vinzelberg, I. Mönch, J. Kretz, and G. Reiss, Surf. Interface Anal. **25**, 611 (1997).

[4]Y. Bruynseraede, C. Strunk, G. Neuttiens, L. Stockman, C. van Haesendonck, and V. V. Moschalkov, Nanostruct. Mater. **6**, 169 (1995).
[5]E. S. Snow, P. M. Campbell, and P. J. McMarr, Appl. Phys. Lett. **63**, 749 (1993).
[6]T. Mühl, H. Brückl, G. Weise, and G. Reiss, J. Appl. Phys. **82**, 5255 (1997).
[7]R. M. Silver, E. E. Ehrichs, and A. L. Delozanne, Appl. Phys. Lett. **51**, 247 (1987).
[8]A. L. de Lozanne, Jpn. J. Appl. Phys., Part 1 **33**, 7090 (1994).
[9]M. McCord, D. P. Kern, and T. H. P. Chang, J. Vac. Sci. Technol. B **6**, 1877 (1988).
[10]S. Rubel, M. Trochet, E. E. Ehrichs, W. F. Smith, and A. L. de Lozanne, J. Vac. Sci. Technol. B **12**, 1894 (1994).
[11]A. Laracuente, M. J. Bronikowski, and A. Gallagher, Appl. Surf. Sci. **107**, 11 (1996).
[12]S. Gider, J. Shi, D. D. Awschalom, P. F. Hopkins, K. L. Campman, A. C. Gossard, A. D. Kent, and S. von Molnar, Appl. Phys. Lett. **69**, 3269 (1996).
[13]M. J. Hampden-Smith and T. T. Kodas, Chem. Vap. Deposition **1**, 9 (1995).
[14]S. Matsui and K. Mori, J. Vac. Sci. Technol. B **4**, 299 (1986).
[15]K. L. Lee and M. Hatzakis, J. Vac. Sci. Technol. B **7**, 1941 (1989).
[16]H. W. P. Koops, J. Kretz, M. Rudolph, M. Weber, G. Dahm, and K. L. Lee, Jpn. J. Appl. Phys., Part 1 **33**, 7099 (1994).
[17]E. E. Ehrichs and A. L. de Lozanne, J. Vac. Sci. Technol. A **8**, 571 (1990).
[18]G. Reiss, F. Schneider, J. Vancea, and H. Hoffmann, Appl. Phys. Lett. **57**, 867 (1990).
[19]H. W. P. Koops, J. Kretz, and H. Brückl, Patent pending, No. 19,733,795.3.
[20]S. Rubel, X. D. Wang, and A. L. Delozanne, J. Vac. Sci. Technol. B **13**, 1332 (1995).
[21]Another possible cracking mechanism is the molecule decomposition by hydrogen ions, which can be generated by the tunneling electrons in the presence of a thin water film on the surface. H. W. P. Koops, L. Dobisz, and J. Urban, J. Vac. Sci. Technol. B **15**, 1369 (1997).
[22]A. Folch, J. Tejada, Ch.H. Peters, and M. S. Wrighton, Appl. Phys. Lett. **66**, 2080 (1995).
[23]J. Vancea, G. Reiss, F. Schneider, K. Bauer, and H. Hoffmann, Surf. Sci. **218**, 208 (1989).

Two-dimensional dopant profiling of patterned Si wafers using phase imaging tapping mode atomic force microscopy with applied biases

M. W. Nelson, P. G. Schroeder, R. Schlaf,[a] and B. A. Parkinson[b]
Department of Chemistry, Colorado State University, Ft. Collins, Colorado 80523

(Received 10 November 1998; accepted 28 May 1999)

Tapping mode atomic force microscopy with applied bias was used to spatially resolve areas of different doping type on Si wafers patterned with photolithography and subsequent ion implantation. The application of a direct current bias between cantilever and sample during the measurement produces Coulomb (electrostatic) forces, whose magnitude depends on the spatial variation of the doping density. This effect was utilized to detect areas of different doping type by monitoring the phase angle between the driving frequency and the cantilever response while scanning areas of different doping density. In this article we present a series of measurements at various bias voltages demonstrating that the observed phase contrast between differently doped areas is directly connected to the bias induced surface potential (band bending) present on these areas. To investigate the contrast mechanism quantitatively, we also measured deflection (force), amplitude and phase versus distance curves for a typical cantilever with an applied bias on a gold thin film. This allowed correlation between phase contrast observed and the actual Coulomb force measured. © *1999 American Vacuum Society.* [S0734-211X(99)09104-0]

I. INTRODUCTION

High spatial resolution methods for two-dimensional characterization of doping profiles in submicron patterned semiconductor circuits are becoming increasingly valuable as device structures continue to shrink. Selective etching, scanning secondary ion mass spectroscopy and several scanning probe microscopy (SPM) related techniques are currently being developed as high spatial resolution methods for dopant characterization.[1–3] SPMs provide the lateral resolution required for characterization of structures down to atomic dimensions and have the additional advantage of being nondestructive. Scanning capacitance microscopy (SCM),[4–6] scanning Kelvin force microscopy (SKFM),[7,8] scanning spreading resistance probe microscopy (nano-SRP or μ-SRP),[1] and current imaging tunneling spectroscopy (CITS)[9] were identified as promising SPM related techniques for dopant profiling. These methods, however, require considerable experimental effort; such as additional electronic circuitry and sample preparation in the case of SCM and SKFM or the requirement of ultrahigh vacuum for reliable measurements in the case of CITS. The atomic force microscopy (AFM) imaging process, which does not require conductive samples, makes it desirable for measuring doping level profiles in real semiconductor process wafers that are often covered by an insulating layer, rendering scanning tunneling microscopy measurements difficult or impossible.

It has been recently shown that Coulomb forces can have a significant influence on the imaging mechanisms in SPM[10] and that single dopant atoms in layered semiconductor materials can be imaged with tapping mode AFM (TMAFM) with a direct current (dc) bias applied.[11] Herein we resort to the use of AFM to resolve doping profiles in patterned semiconductor surfaces. Recently we introduced a simple new method for two-dimensional spatial imaging of doped silicon using TMAFM.[12] This method uses standard TMAFM with an applied dc bias. It was demonstrated that the phase signal in TMAFM is sensitive enough to detect bias induced Coulomb force. This allows us to resolve minute surface charge variations between areas of different doping type and to map out the surface doping density. In this article we will investigate the measurement process more closely and present quantitative results regarding the magnitude of the Coulomb forces involved.

II. EXPERIMENT

The experiments were carried out using a commercially available Nanoscope III® multimode SPM (Digital Instruments, Santa Barbara, CA) in which a bias was fed into the sample stage by connecting its ground connector pin to a dc power supply. Bias voltages are reported with respect to the grounded cantilever. To avoid high currents between sample and cantilever in case of point contact, a 10 MΩ resistor was put in series between power supply and sample. Conductive cantilevers of the same type (Digital Instruments, Santa Barbara, CA, SESP Ni_2Si_3 coated Si, spring constant range: 2–7 N/m, tip radius 20–40 nm, quality factor $Q \approx 300$) were used in all experiments.

The sample investigated was a p-doped silicon wafer (boron; $1\times10^{17}\,cm^{-3}$) containing counter n-doped (arsenic, $5.9\times10^{15}\,cm^{-3}$) 1.3 μm diam circular regions prepared by photolithography and subsequent ion implantation. After implantation the sample was annealed to activate the dopants. Prior to imaging, the sample was dipped into a dilute (pH 3.8) solution of HF to remove the surface oxide. This was followed by rising in distilled water and drying under a

[a]Author to whom correspondence should be addressed; current address: Department of Physics, State University of New York, Binghamton, NY 13902-6016; electronic mail: rschlaf@binghamton.edu
[b]Author to whom correspondence should be addressed; electronic mail: parkinson@mail.chm.colostate.edu

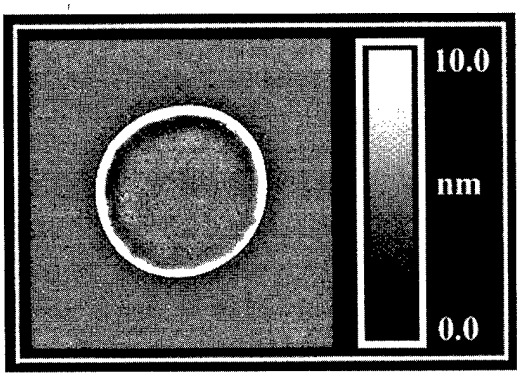


FIG. 1. TMAFM height image (3×3 μm^2 area) of one of the *n*-doped ion implanted circular areas. For details see the text.

stream of nitrogen. Phase contrast was measured using a Digital Instruments "extender" electronics module. Measurements under applied dc bias were carried out using the Nanoscope's "interleave" feature in lift mode. In this mode every scan line is measured twice: the first scan measures the topography using the standard TMAFM height mode while the second scan retraces the same line following the trajectory established during the first scan, but at a user defined height above the sample. This allows us to measure the phase contrast during the second scan at an approximately constant cantilever sample distance. All phase contrast data shown in this article were measured at an interleave height of either 30 or 100 nm. The sample used for the force calibration measurements was an epitaxial gold film grown on mica. The Au film showed flat terraces and islands of about 300 nm diameter. The curves were measured after the cantilever tip was positioned in the center of one of these islands.

III. RESULTS AND DISCUSSION

Figure 1 shows the topographical image of a 3×3 μm^2 area of a *p*-doped silicon wafer which has *n*-doped 1.3 μm diam circular regions prepared by photolithography and subsequent ion implantation. The image was measured with TMAFM in height mode and shows one of the circular *n*-doped regions surrounded by a 5–6 nm high ring (ion-damaged photoresist). It is important to note, for the following evaluation of the phase contrast data, that the *n*-doped areas inside and the *p*-doped areas outside the rings are of similar topographic height. In Fig. 2, phase contrast images of such an area are shown as a function of the applied dc bias. Since our measurements depend on the precise measurement of the Coulomb force induced phase change of the cantilever oscillation, tip sample distance changes must be avoided during the measurement. This is accomplished by using the interleave mode which automatically compensates for the topography. It also allows us to avoid contact with the surface during the phase contrast measurement which would result in erroneous data, since physical contact causes dampening which can also result in a phase change.[13]

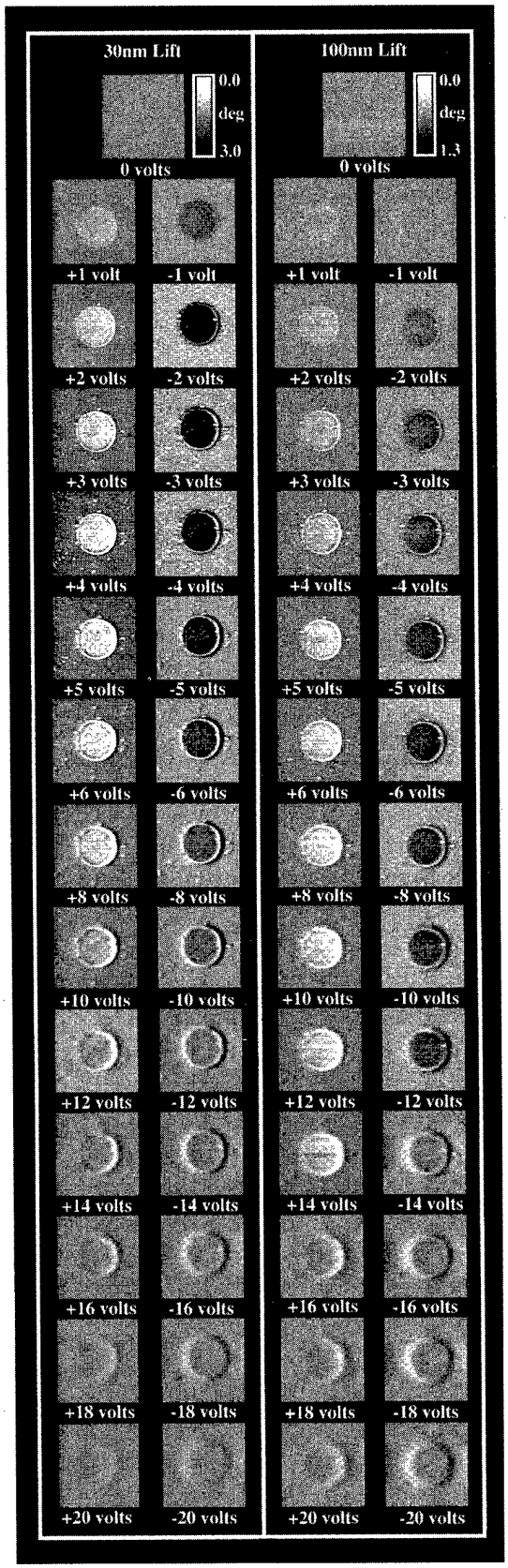


FIG. 2. TMAFM phase images at different sample biases, on the same area of a silicon sample prepared by arsenic ion implantation of 1.3 μm circular regions (5.9×10^{15} cm^{-3}) into a boron doped (1×10^{17} cm^{-3}) silicon substrate. Images on the left were taken at 30 nm interleave lift height, while images on the right were taken at a lift height of 100 nm.

The phase contrast images in Fig. 2 are shown in the units given by the Nanoscope software and were only plane fit. In order to demonstrate the influence of the interleave height, two sets of data obtained at 30 and 100 nm height are shown on the left and the right in Fig. 2. Both sets of images were measured at biases ranging from −20 to +20 V. It is evident that the unbiased images (sample grounded) at both scan heights appear completely featureless and flat, demonstrating the absence of phase changes from influences other than the applied bias. It should be mentioned that the cantilever center-of-oscillation moves towards the surface upon application of a bias, potentially causing problems due to contact oscillation dampening with the surface. However, we will demonstrate below, using calibration curves measured on a gold film, that this effect can also be ruled out.

Phase contrast between the *n*-type circle and the *p*-type surrounding appears when the bias voltage is applied and a contrast reversal occurs upon switching the sign of the applied bias. The appearance of the residue ring in all the images, as an area of smaller phase lag, could be due to charging effects in the insulating photoresist material (residue). It is evident that at both distances the contrast between *n*- and *p*-type regions first increases with the absolute value of the applied bias and then, after peaking at a certain voltage, successively decreases until it has almost completely vanished. Comparison between the 30 and 100 nm lift height images shows that at 30 nm the maximum phase contrast is reached at lower biases and that the contrast diminishes faster than in the 100 nm case.

The magnitude of the phase contrast for both sets of images and at positive and negative biases is shown in Fig. 3(a). In this graph the phase contrast of the *n*-type region relative to the *p*-type region (which is defined as zero) is plotted with respect to the applied bias. As will be demonstrated below a larger phase angle between cantilever and driving frequency corresponds to a larger Coulomb force between cantilever and sample. Therefore, in Fig. 3(a) a positive phase contrast value corresponds to a stronger force on the *n*-type area, while a negative value indicates a weaker force relative to the *p*-type area. In this context, both of the curves demonstrate that upon bias reversal the force difference between *n*- and *p*-type areas also reverses. This is a first indication that the contrast mechanism is connected to band bending in the Si surface induced by the applied bias. Since the band bending will be affected oppositely on *p*- and *n*-type regions at the same bias voltage, contrast reversal is to be expected when the bias is inverted. Figure 3(a) also shows that the phase contrast between *p* and *n* regions reaches a maximum at about +5 and −3 V at 30 nm interleave height and at +7 and −5 V at 100 nm cantilever sample separation.

The asymmetry between positive and negative bias maximum phase shifts in each series is probably related to the initial contact potential difference between cantilever tip and the sample surface due to the difference in work functions of up to 1 eV (work functions of Ni_2Si (Ref. 14) and Si(intrinsic),[15] Si (*p*-type $10^{16}\,cm^{-3}$),[15] and Si (*n*-type $10^{16}\,cm^{-3}$)[15] are 4.96, 4.85, 5.3, and 4.4 eV, respectively).

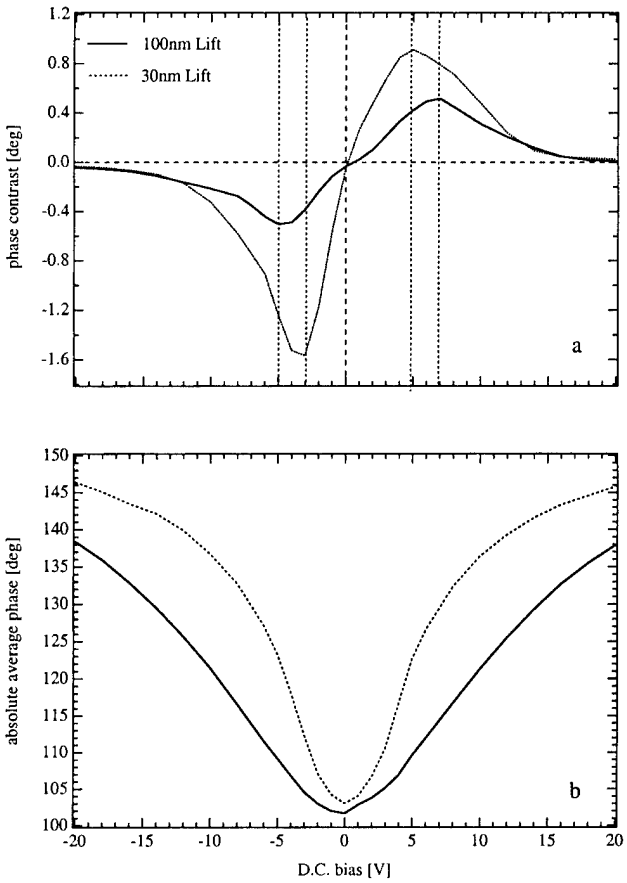


FIG. 3. (Top) TMAFM phase lag on the *n*-type region relative to the phase lag on the *p*-type region as a function of bias. (Bottom) Absolute phase angle, averaged over each image, of the cantilever relative to the cantilever drive frequency plotted against the applied dc bias. Phase angle data in both images were taken from images presented here.

The contact potential acts as a constant voltage offset superimposed upon the applied bias voltage. The wider peak widths of the 100 nm curve, when compared to the peaks of the 30 nm curve, are caused by the fact that changing the surface potential at larger tip–sample separation requires higher voltages to produce the same change as observed at smaller bias and smaller separation. The difference between the absolute contrast values between both series can be explained by the distance dependence of the Coulomb forces. We want to emphasize here that the contrast observed is the phase angle difference between *p*- and *n*-type regions. The dependence of the *absolute* phase angle change on the bias is shown in Fig. 3(b). Comparison of the absolute phase angle values in this graph with the phase contrast values between *p*- and *n*-type areas demonstrates that the phase contrast between these areas is only a small fraction of the total phase change caused by the bias voltage. The absolute phase angle values were determined from the zero order plane fit coefficient as explained in Appendix A. After the plane fit, the image shows the phase contrast between the different regions in the imaged area. This small contribution to the absolute phase angle change contains the information about the local doping density.

The overall shape of the contrast curves in Fig. 3(a) is

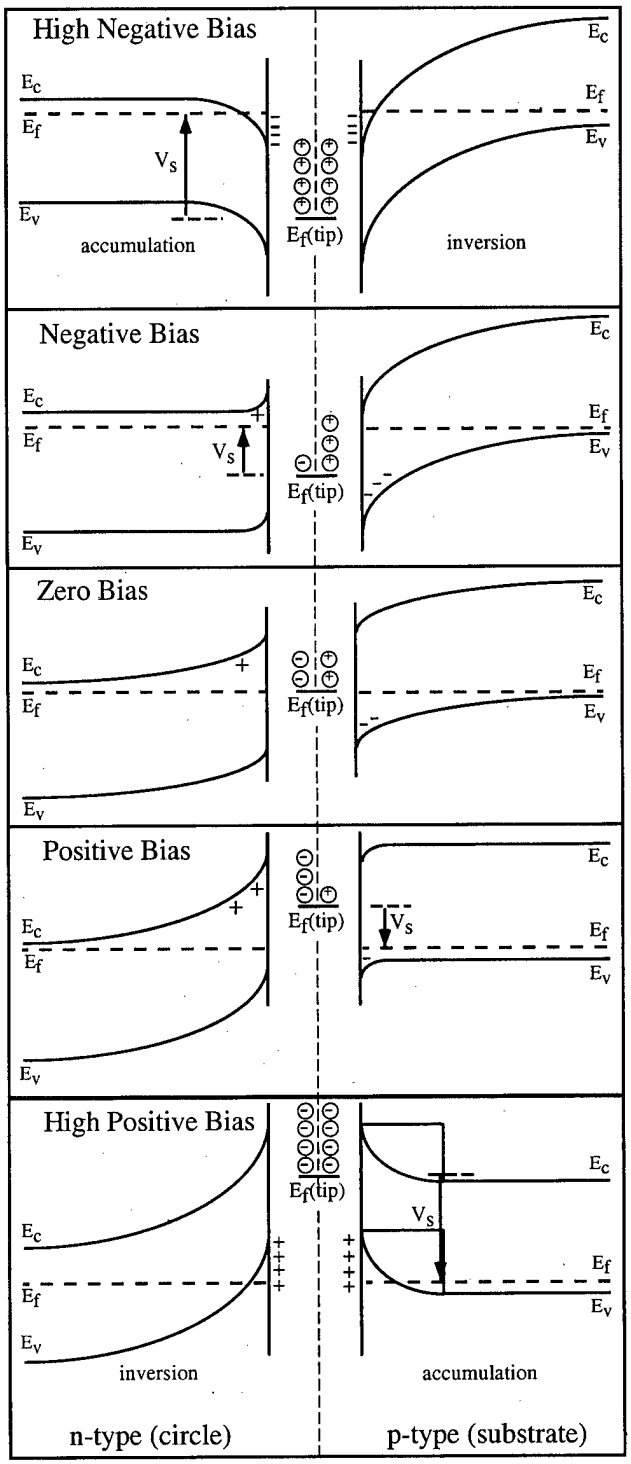


FIG. 4. Energy level diagrams of the TMAFM cantilever/air/Si junction for both the *n*-type area (circle) and *p*-type substrate at positive, high positive, negative, high negative, and zero biases. E_c is the conduction band edge, E_v is the valence band maximum, E_f^s is the Fermi energy of the Si, E_f (tip) is the cantilever Fermi level, and V_s is the sample bias.

explained by a model where the surface potential induced in the sample surface by the applied bias is the mechanism responsible for the observed phase contrast between *n*- and *p*-doped areas. Figure 4 summarizes the changes in surface potential resulting from the applied bias depending on the dopant type. The five scenarios shown in Fig. 4 range from highly positive to highly negative sample biases. At 0 V only Coulomb forces caused by the contact potential (due to the work function difference between cantilever and sample areas) are acting between cantilever and the differently doped sample areas. The contact potential is relatively small when compared to the applied bias voltages and could be nearly the same for both areas if the cantilever work function is somewhere in between the work functions of the *p*- and *n*-type areas. Little or no contact potential difference results in no significant phase contrast. Positive or negative biases change the surface potentials on both of the doped areas opposite each other. Moderate negative biases increase the band bending on the *p*-type areas, while the *n*-type area is driven towards the flatband condition. This results in more depletion on the *p*-type region which needs to be compensated by image charges on the cantilever, while there is less depletion on the *n*-type region. This means that the Coulomb force between cantilever and *p*-type region will be larger than that between cantilever and the *n*-type region, resulting in a larger phase lag for the *p*-type than for the *n*-type region. At *moderate* positive biases the reverse happens, causing stronger band bending on the *n*-type area driving the bands more towards flatband conditions on the *p*-type area. This results in a larger phase lag on the *n*-type region than on the *p*-type region. At *high* positive or negative biases (top and bottom diagrams) the *n*-type area is driven onto accumulation while the *p*-type region goes into inversion or vice versa. Both accumulation and inversion conditions are dramatically different from the depletion case since the Fermi level is driven into the bands. This means electronic states at the valence and conduction band edges become emptied or filled, which in both cases results in either a high positive or negative surface sheet charge density. Differences in doping no longer play a role in the phase contrast at high biases since the surface charge density under accumulation or inversion is primarily determined by the density of states at the band edges. Since both the *n* and *p* regions are either in accumulation or inversion at higher biases, the surface charge density in both regions becomes approximately equal, resulting in the disappearance of the phase contrast between these regions.

The relationship between applied bias, Coulomb force, and induced phase angle changes were investigated by measuring phase, amplitude and force versus distance curves on a gold coated mica substrate. The data evaluation procedures used are discussed in Appendix B. The gold film was used since it provides a good reference due to the absence of any surface potential from surface charging or band bending found in semiconductors. Figure 5 shows cantilever phase (a) and amplitude versus tip–sample separation (b) curves measured simultaneously in tapping mode (oscillating cantilever). The bottom graph (c) shows standard force calibration curves measured with a nonoscillating cantilever. All curves shown were measured as the tip approaches the sample surface. The force calibration curves (c) directly demonstrate the influence of the bias induced Coulomb force on the can-

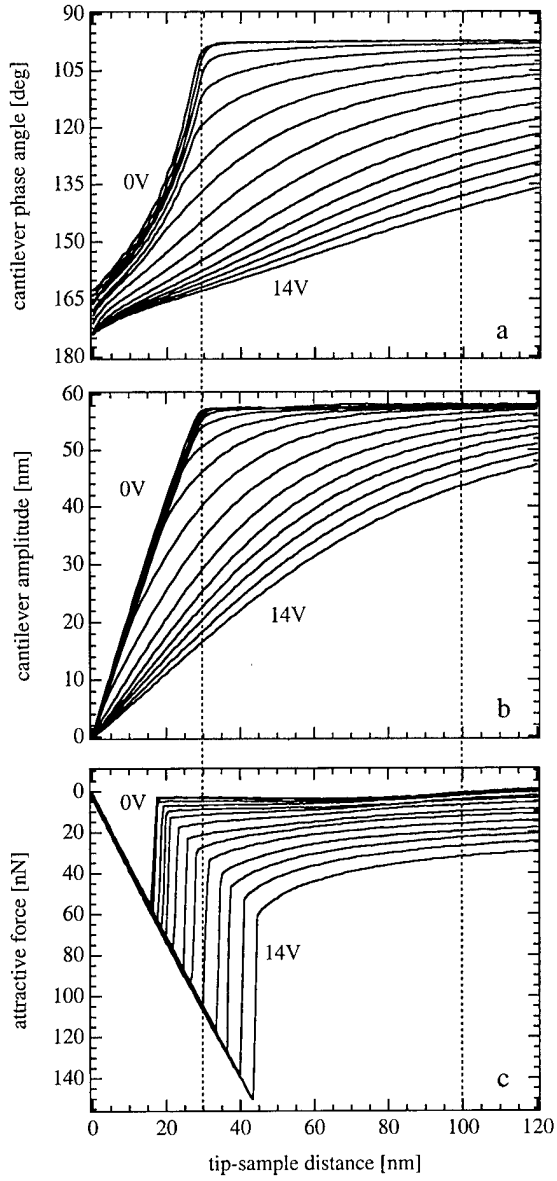


FIG. 5. Cantilever calibration curves taken at various dc biases (1 V increments between 1 and 14 V) on a gold evaporated on mica sample: (top) TMAFM phase calibration, (middle) TMAFM cantilever amplitude curves, (bottom) contact mode force calibration curves (force units were obtained by multiplying the cantilever deflection data with the force constant of the cantilever).

tilever deflection. The topmost curve corresponds to zero applied bias. The curve remains essentially unchanged as the cantilever approaches the surface until a tip–sample distance of about 17 nm. The cantilever then snaps towards the surface as the capillary forces acting on the cantilever exceed the spring constant of the cantilever. The curves beneath the zero bias curve were taken with applied bias. The bias was applied in 1 V steps up to 14 V. The graph shows that the cantilever start positions at the 120 nm distance already change dramatically upon application of the various biases. As the distance is reduced the cantilever bends even more toward the surface due to the increasing Coulomb force. The biased curves shape resemble a quadratic function expected from the influence of Coulomb forces. As the bias increases the "snap to contact" points shift to distances further away from the surface since the electrostatic force adds to the capillary forces when the bias is applied.

The phase and amplitude calibration curves in Figs. 5(a) and 5(b) demonstrate how the bias induced Coulomb force affects the cantilever oscillation in TMAFM measurements. The curves presented in Fig. 5(b) show the variation of the amplitude depending on bias and tip–sample separation. The top curve in this graph corresponds to the zero bias case. As the cantilever moves towards the surface (starting at 120 nm) the amplitude of about 58 nm remains unchanged at first, then as the initial contact occurs the amplitude of the oscillation becomes dampened until it is completely suppressed at zero tip–sample distance. As the bias increases the amplitude curves also bend down, similar to the force calibration curves in Fig. 5(c). Comparison between amplitude and force calibration curves reveals rather amazing behavior: the vertical line drawn at 30 nm allows comparison of all curves at the point where the 0 V amplitude curve just starts to tap on the surface. When increasing the bias it should be expected that the cantilever would bend down toward the surface as is shown in the force calibration curves below. Therefore, the amplitude should be attenuated earlier (at larger tip–sample separation) than in the unbiased case. However, the data show the opposite. As the bias is increased, the amplitude becomes attenuated smoothly as the tip approaches the surface and contact is only made after the curve intersects the 0 V curve. It is evident that at biases exceeding 8 V the cantilever hits the surface after the amplitude becomes zero. These curves do not intersect the contact region of the 0 V curve. This behavior can be explained by the fact that the application of the bias changes the resonance frequency of the cantilever so strongly that the amplitude is attenuated faster than the tip–sample distance shrinks. This is obvious from Fig. 5(a) where the phase angle, depending on bias and tip–sample separation, is plotted. As the bias is increased the resonance frequency of the cantilever decreases, which means the driving frequency "runs ahead", resulting in a phase angle increase. The data shown in Fig. 5(a) clearly prove that the cantilever phase angle is extremely sensitive to the bias induced Coulomb force. The 0 V curve shows that the phase angle also increases upon contact between cantilever and surface. However, the bend of the phase curves is opposite in the contact case when compared to the Coulomb force case. This is especially evident in the 5, 4 and 3 V curves where inflection points can be recognized at about 18, 23 and 28 nm tip–sample distances, respectively. We believe that these inflection points indicate the moment the tip starts touching the surface. These curves also show that at biases exceeding 5 V no contact occurs before the cantilever is at zero distance due to the amplitude attenuation mentioned above.

IV. CONCLUSIONS

We performed TMAFM measurements on patterned Si wafer samples with an additional dc bias applied between

cantilever and sample. Our results show that areas of different doping density cause a bias dependent variation in the phase angle between the cantilever oscillation and its driving frequency. This phase angle contrast was used to characterize the lateral doping profile of Si wafers patterned by photolithography with subsequent ion implantation. The contrast mechanism was investigated by calibrating the force measurements and amplitude/phase relationships with applied biases on a gold coated mica substrate. The impact of the Coulomb force induced by the applied bias on the cantilever oscillation can then be determined. The results of these measurements clearly demonstrate that the phase of the cantilever oscillation depends on the Coulomb force acting between the cantilever and sample. A model was developed in which the bias-induced surface potential (band bending) variation between differently doped areas of the surface is identified to be the cause of the image contrast observed in the phase images. Our results clearly demonstrate that this measurement method has the potential to be a nondestructive method for nanoscale two-dimensional dopant profiling of semiconductor devices. Further research is in progress to investigate the limits of the spatial resolution and quantitative doping level sensitivity in order to develop this method into a practical tool for precise lateral resolution of doping densities.

ACKNOWLEDGMENTS

The authors are grateful to C. W. Almgren for providing the ion implanted Si sample and to J. Shoemaker-Parry for the preparation of the gold on mica samples. They thank Digital Instruments for upgrading their scanning probe instrumentation. R. S. acknowledges support by a Deutsche Forschungsgemeinschaft (DFG) Forschungsstipendium fellowship. Partial support for this work from DOE-BES under contract No. DE-F603-96ER14625 is gratefully acknowledged.

APPENDIX A

The main results presented in this article are so called "phase images". These images show the phase angle between the cantilever driving signal (alternating current) that drives the cantilever via oscillation of a piezo to which the cantilever is attached and the actual cantilever oscillation as displayed relative to the scan position. The above described interleave mode topography and phase angle can be measured simultaneously allowing a correlation between topographic and phase images. Once a phase image is measured it is plane fitted to remove scan artifacts. After completing this procedure the image contains only relative phase contrast variation between different regions of the image (i.e., *p*- and *n*-type regions in our case). To obtain the average total phase angle (bottom graph of Fig. 3) of the image data relative to the drive signal, the plane fit parameters, which are stored in the file header, were used. Depending on the order of the polynomial used for the plane fit, up to three values are stored in the header after the plane fitting. The first of the three values corresponds to the zero order fit parameter that represents the average phase angle of the entire image. This value was used to determine the absolute phase angle of the cantilever oscillation upon which the phase contrast values are superimposed. The average phase angle values are assigned by the computer software as "twos complement 16 bit numbers" which range from +32768 to −32767. This range corresponds to phase angles between 0° and 180°. This assignment adheres to the usual convention to describe the forced oscillator. In this convention, the driving frequency is fixed and the phase angle defines the phase of the oscillator relative to the driver. It should be pointed out that in a forced elastic harmonic oscillator the oscillator frequency is always the same as the driving frequency, while amplitude and phase can change in response to changes in the driving frequency. Due to the relatively small amplitude at which the cantilevers are typically resonated, it is assumed that the oscillation is close enough to the harmonic case to use the terminology of the harmonic oscillator. In this convention a 0° phase angle corresponds to the extreme case of a driving frequency that is much lower than the resonance frequency, resulting in oscillator movement that directly follows the driver. The 90° phase angle corresponds to the condition when the driving frequency is the same as the resonance frequency, hence the oscillator lags 90° behind the motion of the driver. A 180° phase angle corresponds to the high frequency extreme case where the driving frequency is much higher than the resonance frequency, causing the oscillator to move opposite the driver. As was demonstrated in Sec. III, an increase of the phase angle corresponds to an increase of the Coulomb force between cantilever and sample. This results from the attractive Coulomb force reducing the resonance frequency of the cantilever while the driving frequency remains constant. Thus, an attractive Coulomb interaction corresponds to the situation where the driver is faster than the resonance frequency, resulting in a phase angle between 90° and 180°.

APPENDIX B

Any force calibration curves (except the phase lag versus distance curves) taken with Nanoscope equipment is given in relative *Z*-piezo displacement units and in "twos complement 16 bit" data that correspond to the deflection of the cantilever. Relative *Z*-piezo displacement means that the absolute position of the cantilever is not given by the software but has to be determined from the curves themselves. Therefore, in order to convert these data into real deflection, force or amplitude values, the curves were calibrated as follows:

Standard force calibration curves were obtained from cantilever deflection versus *Z*-piezo displacement curves by transforming the original 16 bit deflection data into true deflection values in nanometers. This was accomplished by multiplying the spectra with a constant that yielded a slope of 1 for the contact region of the curves (the left part of the curve where the cantilever is in contact with the sample surface, where its deflection directly equals the cantilever displacement). The obtained deflection values were then multiplied by the force constant of the cantilever resulting in true force values. The spring constant of 4.5 N/m, used in the

force calculations, represents the average of the minimum and maximum values stated by the manufacturer (2–7 N/m). The x axis of these curves was calibrated to obtain true tip–sample distances by determining the intersection of the non-contact region of the zero bias curve (right part of the curve which is nearly horizontal; the weak wavy superimposed shape is an artifact which is always present in these curves measured on both of our Nanoscope III setups) with the contact region of the curve. This intersection point was also defined as the origin of the force scale since it represents the point where the cantilever would have touched the sample surface had it not been attracted by capillary/Coulomb forces before it touched. The origin of the amplitude graphs was determined similarly by using the point where the contact region (left part of the graphs) bends off into a horizontal line corresponding to complete suppression of the vibration due to direct contact with the sample surface. The amplitude scale was calibrated by multiplying the curves with a constant so that the contact region would exhibit a slope of 2. This value was chosen in order to account for the fact that amplitudes correspond to the distance between both elongation maxima. This only approximates the real measurement conditions, since the driving force of the cantilever will accelerate the cantilever after being dampened due to surface contact during its upwards motion away from the surface. Therefore, the true slope can be expected to be slightly different from 2. Phase calibration curves were measured simultaneously with the amplitude curves. Therefore, the x-axis calibration of the amplitude curves was taken from the amplitude graphs. The phase scale shows the values as given by the Nanoscope software.

[1]A. C. Diebold, M. R. Kimp, J. J. Kopanski, and D. G. Seiler, J. Vac. Sci. Technol. B **14**, 196 (1996).
[2]C. Spinella, V. Raineri, F. La Via, and S. U. Campisano, J. Vac. Sci. Technol. B **14**, 414 (1996).
[3]W. Vandervorst, V. Privitera, V. Raineri, T. Clarysse, and M. Pawlik, J. Vac. Sci. Technol. B **12**, 276 (1994).
[4]Y. Huang, C. C. Williams, and M. A. Wendman, J. Vac. Sci. Technol. A **14**, 1168 (1996).
[5]A. Erickson, L. Sadwick, G. Neubauer, J. Kopanski, D. Adderton, and M. Rogers, J. Electron. Mater. **25**, 301 (1996).
[6]C. C. Williams, J. Slinkman, W. P. Hough, and H. K. Wickramasinghe, Appl. Phys. Lett. **55**, 1662 (1989).
[7]H. D. Liess, R. Mäckel, H. Baumgärtner, and J. Ren, Sens. Actuators B **B14**, 739 (1993).
[8]T. Hochwitz *et al.*, J. Vac. Sci. Technol. B **14**, 440 (1995).
[9]H.-A. Lin, R. Jaccodine, and M. S. Freund, Appl. Phys. Lett. **72**, 1993 (1998).
[10]R. Schlaf, P. G. Schroeder, W. M. Nelson, R. Stübner, S. Tiefenbacher, H. Jungblut, and B. A. Parkinson, Thin Solid Films **331**, 203 (1998).
[11]R. Schlaf, D. Louder, M. W. Nelson, and B. A. Parkinson, J. Vac. Sci. Technol. A **15**, 1466 (1997).
[12]M. W. Nelson, P. G. Schroeder, C. W. Almgren, A. N. Erickson, R. Schlaf, and B. A. Parkinson, Appl. Phys. Lett. **74**, 1421 (1999).
[13]S. N. Magonov, V. Elings, and M.-H. Whangbo, Surf. Sci. **375**, L385 (1997).
[14]E. Bucher, S. Schulz, M. C. Lux-Steiner, P. Munz, U. Gubler, and F. Greuter, Appl. Phys. A: Solids Surf. **40**, 71 (1986).
[15]D. R. Lide, *CRC Handbook of Chemistry and Physics* (Chemical Rubber, New York, 1998).

Fabrication of a microcavity structure with a polyimide thin film prepared by vacuum deposition polymerization

Y. Sakakibara[a)]
Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba, Ibaraki 305-8568, Japan

T. Tani
Department of Applied Physics, Tokyo University of Agriculture and Technology, 2-24-16 Naka-machi, Koganei, Tokyo 184-8588, Japan

(Received 23 October 1998; accepted 23 April 1999)

A microcavity structure has been fabricated in which a polyimide layer with thickness on the order of the wavelength of light was sandwiched in between dielectric mirrors by using vacuum deposition polymerization (VDP). We deposited a 330-nm-thick polyimide layer through a polymerization reaction between two mutually reactive monomers evaporated onto a dielectric mirror consisting of TiO_2/SiO_2 multistacks. In the middle of the VDP process we deposited tris(8-hydroxyquinoline) aluminum (ALQ) so that a sublayer 30%-doped with ALQ was inserted within the polyimide layer. Upon the polyimide layer we coated a dielectric mirror consisting of Ta_2O_5/SiO_2 multistacks. The mirrors were designed to have a broad high reflection band (stop band) with a reflectivity maximum of ~90% at 530 nm. Fabrication of a microcavity was considered to be successful based on its optical properties. Within the stop band a sharp transmission band with a width of 8 nm appeared due to Fabry–Perot resonance. Correspondingly, ALQ photoluminescence became narrowed by the same width. We characterized the cavity resonance modes using a model considering penetration of the optical field into the dielectric layers. © *1999 American Vacuum Society.* [S0734-211X(99)01804-1]

I. INTRODUCTION

Microcavities have recently attracted a great deal of attention since they can be useful tools to control emission of light and to enhance optical nonlinearity effectively by the confinement of light. In a planar microcavity structure, a thin spacing layer with thickness on the order of the wavelength of light separates two mirrors, forming a Fabry–Perot resonator. The layer is usually composed of inorganic semiconductors or organic materials. In the latter case, it has been made of such materials as dye solution,[1,2] Langmuir–Blodgett films,[3,4] vacuum-deposited films of molecular materials,[5–11] and spin-coated polymer films.[12–19] These materials must satisfy the following conditions: (1) thickness controllability on the order of the wavelength of light, and (2) ability to fabricate a layer that will be sandwiched between the mirrors.

As spacing layers, polymer films are useful since they have mechanical durability, processability, and optical transparency with small loss. So far, however, they have been prepared by a wet process such as spin coating using solvent, thus residual solvent is not easily removed. If a vacuum process can be used to deposit the polymer spacer, such problems do not occur. We propose that vacuum deposition polymerization (VDP) may provide a new candidate for the spacing layer. The VDP method is a way by which amorphous polymer films are prepared in high vacuum through a polymerization reaction between two mutually reactive monomers evaporated and deposited onto the substrate surface.[20,21] This method is good at controlling film thickness on the order of the wavelength of light. Also, it can deposit polyimide films, the thermal durability of which enables stable overcoating of inorganic dielectric mirrors or metal mirrors. Furthermore, other materials with such photofunctions as luminescence and optical nonlinearity can be doped in the polymer layer by codeposition.[22,23] These properties seem to provide attractive possibilities for various device applications.

Thus, in this article, we demonstrate our first attempt at fabricating a microcavity structure using a polyimide layer prepared by VDP. Fabrication of a microcavity was successful judging from its optical properties such as sharp transmission and emission spectra due to Fabry–Perot resonance, as will be described in the following sections.

II. FABRICATION OF MICROCAVITY STRUCTURES

Figure 1(a) illustrates a microcavity structure fabricated in our laboratory. A thin polyimide layer of thickness d was sandwiched in between dielectric mirrors. The polyimide layer contained a sublayer 30%-doped with a photoluminescent material tris(8-hydroxyquinoline) aluminum (ALQ: molecular structure is shown in the inset of Fig. 6). The dielectric mirrors and the total thickness of the polymer layer were designed so that cavity resonance would match with the emission region of ALQ. Thickness of the sublayer was 1/10 times the full thickness of the polyimide layer. The sublayer was set in the middle of the polymer layer.

This structure was fabricated in the following three steps. (1) Upon a glass substrate (10 mm×10 mm×1 mm), a dielectric mirror consisting of $\lambda/4$ stacks of TiO_2 and SiO_2 layers was commercially coated; the SiO_2 layers were depos-

a)Electronic mail: e8809@etl.go.jp

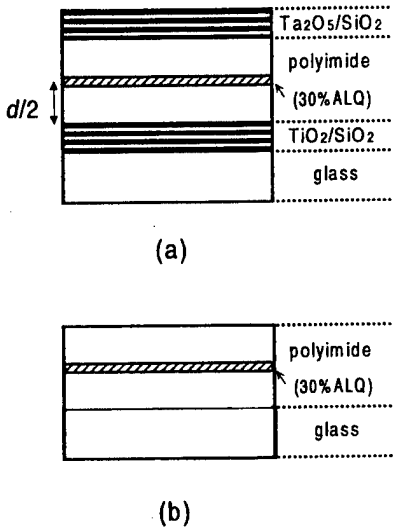


FIG. 1. Illustrations of (a) a microcavity structure and (b) a reference sample. The black layers indicate TiO_2 or Ta_2O_5 layers.

ited by heating SiO_2 with an electron gun in a vacuum of 10^{-4} Pa and the TiO_2 layers were deposited by heating Ti_2O_3 with an electron gun in an O_2 atmosphere of 10^{-2} Pa. The mirror was designed to have a reflectivity maximum of ~90% at 530 nm. Figure 2(a) shows the transmission spectrum. Because of the ~90% reflectivity, transmittance at 530 nm is lowered to ~10%. (2) The mirror was set on the substrate holder of a VDP apparatus[23] as illustrated in Fig. 3. The source cells were in precise heat control. First, we coevaporated two precursor monomers 6FDA, hexafluoroisopropylidene-2,2-bis(phthalic acid anhydride) (Hoechst), and BIS–AF–A, 2,2-bis[4-(4-aminophenoxy) phenyl] hexafluoropropane (Central Glass Co.)—whose molecular formulas are shown in Fig. 4, onto substrates kept at room temperature, keeping 1:1 stoichiometric balance of the evaporation rates. Then, a polyamic acid layer was grown on the mirror at a deposition rate of ~3.5 nm/min through a polymerization reaction (see scheme 1 in Fig. 4). During this deposition the shutter of the ALQ source was closed but the source was preheated in order for ALQ material to have an evaporation rate ~3/7 times the polymer deposition rate. At

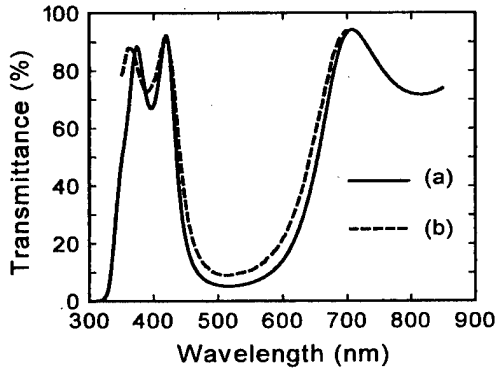


FIG. 2. Transmission spectra of (a) the TiO_2/SiO_2 mirror and (b) the Ta_2O_5/SiO_2 mirror coated on glass substrates.

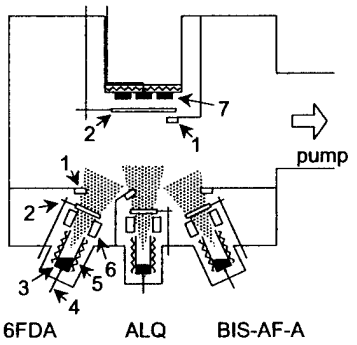


FIG. 3. A VDP apparatus to prepare the polyimide layer doped with ALQ: (1) quartz thickness monitor, (2) shutter, (3) source cell, (4) thermocouple, (5) heater, (6) cold trap, and (7) substrate. Base pressure is 3×10^{-5} Pa.

thicknesses corresponding to the boundaries of the ALQ-doped zone we opened and closed the shutter of the ALQ source, so that a sublayer 30%-doped with ALQ was deposited. Successively, we deposited a polyamic acid layer until the full thickness reached the desired value. After the deposition, we raised the substrate temperature gradually and kept it constant at 300 °C for 2 h to convert the polyamic acid to polyimide according to reaction scheme 2 in Fig. 4. The whole process was carried out in a vacuum at a pressure of 10^{-4} Pa. The layer thickness d that we intended to prepare was 330 nm. This corresponds to an optical length of 530 nm, assuming a polymer refractive index of 1.6. (3) We moved the sample into a different chamber for upper mirror coating. We coated a dielectric mirror consisting of $\lambda/4$ stacks of Ta_2O_5 and SiO_2 layers upon the polyimide layer kept at ~300 °C by ion beam assisted deposition in an O_2 atmosphere of 10^{-2} Pa. The mirror was designed to have a reflectivity maximum of ~90% at 530 nm. To check optical property of the Ta_2O_5/SiO_2 mirror, we deposited a mirror directly on a glass substrate in the same coating process and measured the transmission spectrum [Fig. 2(b)].

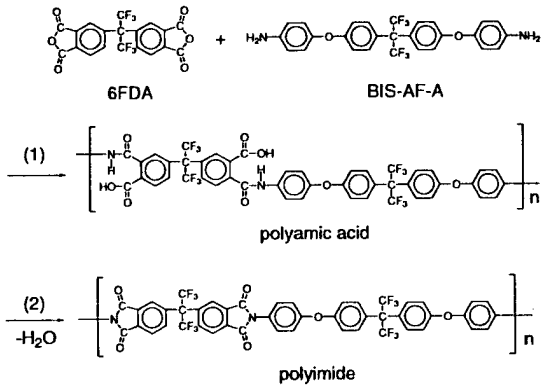


FIG. 4. Molecular structures of precursor monomers and polymerization reaction schemes. Reaction (1) proceeds on the substrate through the deposition process. Reaction (2) proceeds by heating the deposited polyamic acid film at 300 °C.

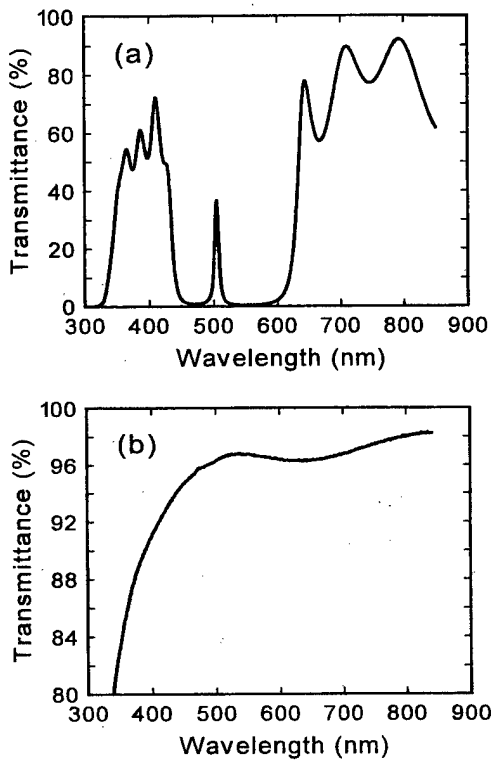


FIG. 5. Transmission spectra of (a) the microcavity sample and (b) the reference sample.

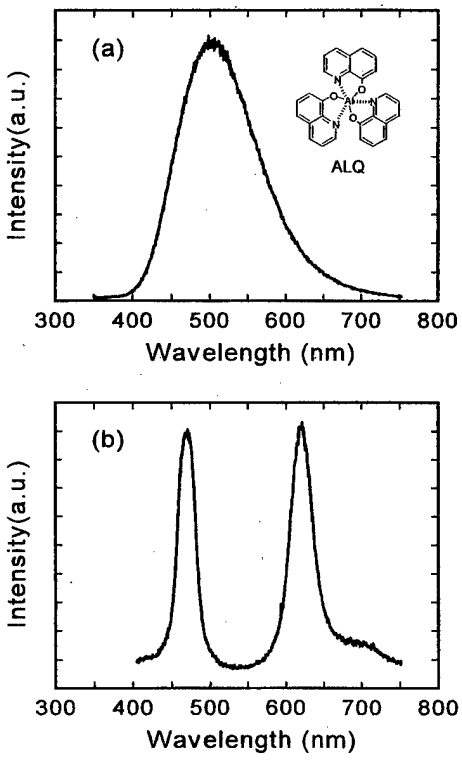


FIG. 6. Coarse spectra of ALQ photoluminescence emitted from (a) the reference sample and (b) the microcavity sample. Molecular structure of ALQ is shown in the inset of (a).

For comparison purpose a reference no-cavity sample was prepared consisting only of an ALQ-doped polyimide layer deposited directly on a glass substrate in the same VDP process [see Fig. 1(b)].

III. OPTICAL PROPERTIES

All the deposited layers exhibited good adhesion over the whole area of the substrate and a flat optical device was achieved. We investigated its microcavity performance with transmission and emission spectroscopies, and characterized its resonance modes.

A. Transmission spectra

Figure 5(b) shows that within the high reflection region (or so-called stop band) of the dielectric mirror (430–650 nm), the reference sample had only a little absorption giving a loss to the cavity. (An interference pattern is superimposed on the absorption spectrum. The increasing absorption for λ less than 400 nm is due to the absorption of the glass, the polymer, and ALQ.) So, if the cavity resonance comes into the stop band, a sharp transmission band should occur due to Fabry–Perot resonance. Figure 5(a) clearly shows such a sharp band at 500 nm. (The spectrum was measured along the normal to the surface with an optical beam diameter of 0.9 mm.) Its full width at half maximum (FWHM) was 8.0 nm. Compared to the reference sample and the dielectric mirrors, the microcavity device showed a complicated spectrum, which will be analyzed in Sec. III C.

B. ALQ photoluminescence spectra

ALQ—a well-known organic electroluminescent material—has an absorption band with a peak around 400 nm. Since the device had a transparent window around the absorption peak [see Fig. 5(a)], we photo-excited the doped ALQ molecules with a 386 nm N_2-dye laser (Laser Photonics LN120C) with an incident angle of 30° and with a beam diameter of 1 mm. Using a polychromater (Chromex 250IS Imaging Spectrograph) with two sets of gratings (50 and 600 grooves/mm), we measured coarse and fine photoluminescence spectra along the normal to the surface. For comparison we also measured an emission spectrum of the reference sample.

While the reference sample showed a broad emission [Fig. 6(a)], the microcavity device showed two sharp emission bands [Fig. 6(b)]. This spectral narrowing is a characteristic performance of the microcavity in which a broadband luminescent material is contained. The two bands correspond well to the transmission peaks shown in Fig. 5(a), although they are located at different wavelengths due to the slight thickness difference of the polymer layers. Figure 7 shows a spectrum of the sharp band in the middle of the stop band. The FWHM bandwidth was 8.1 nm being in good agreement with that of the transmission band (8.0 nm).

C. Mode characterization of cavity resonance

To understand optical modes of microcavities, cavity length L_{cav} is an important parameter because resonance

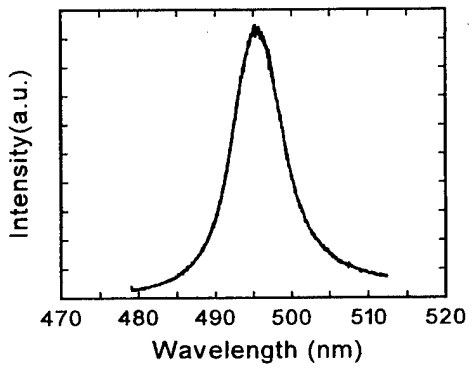


FIG. 7. Fine spectrum of ALQ photoluminescence emitted from the microcavity sample.

modes are given by the relation $m\lambda = 2L_{cav}$, where m is the mode index and λ is the resonance wavelength.[6] According to the ordinary Fabry–Perot theory, L_{cav} is equal to nd where n and d are the refractive index and the thickness of the spacing layer respectively. In the present case, $n \approx 1.6$ and $d \approx 330$ nm and thus the second mode ($m=2$) should appear around 530 nm. Actually, the reference sample has a corresponding interference peak at around 530 nm [Fig. 5(b)] in good agreement with the theory. However, the theory has difficulties in explaining the experimental results of the microcavity device. As shown in Figs. 5(a) and 6(b), the sharp resonance band of the microcavity device was located at a wavelength region around 500 nm, which deviates from the region estimated by theory. Also, the theory cannot explain the existence of the neighboring band observed at 640 or 620 nm in the same figures. [There is not a corresponding interference peak in the reference sample spectrum shown in Fig. 5(b)]. This neighboring band cannot be ascribed to the first ($m=1$) resonance mode because it should be located in a near-infrared region around 1000 nm.

These difficulties of the ordinary Fabry–Perot theory might be solved if the penetration of the optical field into the dielectric layers is taken into account because effective cavity length L_{cav} becomes different from nd.[6,24] Hence, we examine the resonance modes of the microcavity structure taking account of the penetration with an equation for L_{cav} that is modified from the equation presented by Dodabalapur *et al.* for a microcavity consisting of a dielectric mirror and a metal mirror.[6] Replacing the penetration depth term of the metal mirror in Eq. (2) of Ref. 6 by that of the dielectric mirror, L_{cav} is given by

$$L_{cav} \approx \frac{\lambda_{c,\mathrm{upper}}}{2}\left(\frac{n_{\mathrm{av,upper}}}{\Delta n_{\mathrm{upper}}}\right) + \frac{\lambda_{c,\mathrm{lower}}}{2}\left(\frac{n_{\mathrm{av,lower}}}{\Delta n_{\mathrm{lower}}}\right) + nd, \qquad (1)$$

where $\lambda_{c,\mathrm{upper}}$ (or $\lambda_{c,\mathrm{lower}}$) is the center wavelength of the stop band of the upper (or lower) mirror, $\Delta n_{\mathrm{upper}}$ (or $\Delta n_{\mathrm{lower}}$) is the index difference between the stack layers of the upper (or lower) mirror, $n_{\mathrm{av,upper}}$ (or $n_{\mathrm{av,lower}}$) is the average index of the upper (or lower) mirror. Using $\lambda_{c,\mathrm{upper}} = \lambda_{c,\mathrm{lower}} = 530$ nm, $n(TiO_2) = 2.3$, $n(SiO_2) = 1.47$, $n(Ta_2O_5) = 2.26$, $n = 1.6$, and $d = 330$ nm, L_{cav} is estimated to be 1750 nm. Thus, the resonance wavelength of the cavity, λ, is given by $\lambda = 3500/m$ (nm), and the fifth, sixth, seventh, and eighth resonance modes are located at 700, 583, 500, and 438 nm. So we may attribute the sharp bands at around 500 nm seen in Figs. 5(a) and 6(b) to the seventh mode. The additional bands on both sides [e.g., the bands at 430 and 640 nm in Fig. 5(a)] may be attributed to the eighth and the sixth modes. Consequently, the mode analysis using Eq. (1) seems to better explain the resonance wavelength of the sharp band (500 nm) and the existence of the additional bands than ordinary Fabry–Perot theory.

IV. DISCUSSION

Judging from the optical properties described earlier, fabrication of a microcavity was successful. So far, in fabrication of organic-layer-based microcavity structures, in many cases the upper mirror has been a vacuum-coated metal mirror[4–19] and in a few cases a dielectric mirror deposited on another glass substrate that is in mechanical contact with the organic layer.[1–3] Hence, to our knowledge, the present work is the characteristic demonstration of fabricating a microcavity structure in which an *organic* spacing layer is sandwiched by *dielectric* mirrors from *both* sides *without* mechanical support. This was possible because the upper mirror could be deposited on the polyimide layer adhesively by vacuum deposition, tracing any surface roughness and keeping a constant spacing thickness. So the fabrication method can produce compact, self-standing and large area devices. The structure with dielectric mirrors on both sides will be advantageous particularly in transmissive uses because absorption loss of dielectric mirrors is much smaller than that of metal mirrors. (Peak transmittances of the organic microcavities using metal mirror(s) reported so far[4,12,13] are much smaller than that of our microcavity.) The possibility of doping photofunctional materials at the intended position will be useful in fabricating devices that utilize the optical field intensity distribution of resonant standing waves formed in the cavity. This doping-position controllability is not easily attained by other fabrication methods of polymer layers reported so far. Of course, because all the processes are carried out in vacuum, the device is free from solvent and the opportunity to contain impurity in the polymer layer is much reduced. From these features, we believe the presented method will be useful in various applications.

ACKNOWLEDGMENTS

The authors are grateful to N. Takada, T. Nakayama, T. Tokizaki, and H. Kawashima for useful discussions. They are also grateful to T. Shimizu for his x-ray photoelectron spectroscopy analysis in determining the kind of material of the top layer of the TiO_2/SiO_2 mirror.

[1]H. Yokoyama, M. Suzuki, and Y. Nambu, Appl. Phys. Lett. **58**, 2598 (1991).
[2]M. Osuge and K. Ujihara, J. Appl. Phys. **76**, 2588 (1994).
[3]M. Suzuki, H. Yokoyama, S. D. Brorson, and E. P. Ippen, Appl. Phys. Lett. **58**, 998 (1991).
[4]S. E. Burns, N. Pfeffer, J. Grüner, M. Remmers, T. Javoreck, D. Neher, and R. H. Friend, Adv. Mater. **9**, 395 (1997).

[5]A. Dodabalapur, L. J. Rothberg, T. M. Miller, and E. W. Kwock, Appl. Phys. Lett. **64**, 2486 (1994).
[6]A. Dodabalapur, L. J. Rothberg, R. H. Jordan, T. M. Miller, R. E. Slusher, and Julia M. Phillips, J. Appl. Phys. **80**, 6954 (1996).
[7]T. Nakayama, Y. Itoh, and A. Kakuta, Appl. Phys. Lett. **63**, 594 (1993).
[8]N. Takada, T. Tsutsui, and S. Saito, Appl. Phys. Lett. **63**, 2032 (1993).
[9]N. Takada, T. Tsutsui, and S. Saito, Optoelectronics **8**, 403 (1993).
[10]T. Tsutsui, N. Takada, S. Saito, and E. Ogino, Appl. Phys. Lett. **65**, 1868 (1994).
[11]A. Zhang, L. Zhuang, Y. Lin, Z. Xia, Y. Ma, X. Ding, S. Wang, D. Zhou, and C. Huang, Solid State Commun. **97**, 445 (1996).
[12]W. Guss, H. Vestweber, M. Hopmeier, Y. H. Tak, M. Deussen, and R. F. Mahrt, Synth. Met. **83**, 257 (1996).
[13]J. Grüner, F. Cacialli, and R. H. Friend, J. Appl. Phys. **80**, 207 (1996).
[14]U. Lemmer *et al.*, Appl. Phys. Lett. **66**, 1301 (1995).
[15]D. G. Lidey, M. A. Pate, D. M. Whittaker, D. D. C. Bradley, M. S. Weaver, T. A. Fisher, and M. S. Skolnick, Chem. Phys. Lett. **263**, 655 (1996).
[16]D. G. Lidey, D. D. C. Bradley, M. A. Pate, J. P. R. David, D. M. Whittaker, T. A. Fisher, and M. S. Skolnick, Appl. Phys. Lett. **71**, 744 (1997).
[17]G. R. Hayes, F. Cacialli, and R. T. Phillips, Phys. Rev. B **56**, 4798 (1997).
[18]H. F. Wittmann, J. Grüner, R. H. Friend, G. W. C. Spencer, S. C. Moratti, and A. B. Holmes, Adv. Mater. **7**, 541 (1995).
[19]N. Tessler, G. J. Denton, and R. H. Friend, Nature (London) **382**, 695 (1996).
[20]Y. Takahashi, M. Iijima, K. Inagawa, and A. Itoh, J. Vac. Sci. Technol. A **5**, 2253 (1987).
[21]J. R. Salem, F. O. Sequeda, J. Duran, W. Y. Lee, and R. M. Yang, J. Vac. Sci. Technol. A **4**, 369 (1986).
[22]Y. Sakakibara, M. Iijima, K. Tsukagoshi, and Y. Takahashi, Jpn. J. Appl. Phys., Part 2 **32**, L332 (1993).
[23]Y. Sakakibara, H. Matsuhata, and T. Tani, Jpn. J. Appl. Phys., Part 2 **32**, L1688 (1993).
[24]R. P. Stanley, R. Houdré, U. Oesterle, M. Gaihanou, and M. Ilegems, Appl. Phys. Lett. **65**, 1883 (1994).

Low voltage electron beam lithography in PMMA

A. Olkhovets[a] and H. G. Craighead
School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853

(Received 30 December 1998; accepted 14 May 1999)

We have studied low voltage (1–2 kV) electron beam lithography processes in PMMA and compared them to conventional high voltage processing. We looked at the deposited metal after liftoff as well as directly imaging resist profiles by atomic force microscopy. As expected, the proximity effects were greatly reduced. The forward scattering was found to increase at low voltage. The study of developed resist profiles showed that linewidth versus dose has a single Gaussian functional form, proving that forward scattering plays the major role in line broadening. The effective Gaussian linewidth is 60 nm at 1 kV in a 50 nm resist layer. Modeling of the lithographic process showed a significant increase in resolution and process latitude for thinner resists. © *1999 American Vacuum Society.* [S0734-211X(99)06304-0]

I. INTRODUCTION

Future types of microelectronics will require sub-100 nm structure dimensions. Conventional optical lithography is not expected to achieve such resolution. Currently several alternative next generation lithography schemes are being considered. One of them is an array of microcolumn electron beam sources. This array has the required resolution provided by the electron beam, and should achieve high throughput by employing many sources in parallel. The development of this lithography system is promising, with a prototype electron column demonstrating a 10 nm beam spot size at 1 keV.[1]

The trend in electron beam lithography has been to increase the electron accelerating voltage, up to 100 kV, to achieve high resolution. However, microcolumns need to be operated at relatively low beam energy, typically 1–2 keV. Another promising low voltage application is the use of a scanning probe microscope for local exposure of thin electron sensitive resist layers.[2] The development of fast scanners and miniaturized scanning tunneling microscopes (STMs) suggests that viable lithographic writing speeds may be achieved.[3] Operating at low voltage offers both advantages and disadvantages. The focused beam size is usually increased and the luminosity is decreased at low voltage. However, electron energy deposition is more efficient at low voltages, which offsets decreased luminosity. The electron penetration depth is decreased, and requires thin ($\sim\leqslant$50 nm) resist processing. The lower energy and shorter scattering lengths also result in greatly reduced proximity effects[4] and reduce substrate damaging compared to higher energy exposure. Possible resist charging is also a concern.

Here we report the results of experiments and modeling to explore the issues associated with 1–2 keV lithography. For the experiments we used poly(methylmethacrylate) (PMMA) resist and looked at two types of patterns: deposited metal after liftoff and developed resist profiles.

II. EXPERIMENT

A layer of PMMA 50–120 nm thick was spun onto a Si substrate, and prebaked at 170 °C for 1 h. The thickness was determined by an interferometric microscope and a scanning atomic force microscope. The exposure tool was a thermally assisted field emission digital scanning electron microscope (SEM) (Leo DSM982) controlled by an external pattern generator. The spot size of the focused electron beam can be less than 10 nm at 1 kV, and less than 3 nm at 30 kV.[5] Alternating-current interference was reduced by using a short working distance (<5 mm). The beam current was measured with a Faraday cup mounted on a sample stage. The exposed resist was then developed in 1:3 solution of methylisobutylketone: isopropyl alcohol (MIBK:IPA) for 1 min. Then 20 nm of Cr was thermally evaporated, and lifted off in a roughly 1:1 solution of methylene chloride: acetone for 10 min with ultrasonic agitation.

A. Metal after liftoff

1. Resolution test

We looked at 20 nm of Cr after liftoff. An array of dots with 150 nm pitch was patterned on an 80 nm thick layer of PMMA, with electron beam energies ranging from 1 to 20 keV. The dose for each voltage was adjusted to minimize the dot size. The SEM images of the patterns are shown in Fig. 1. The resolution increases with increasing voltage; at 1 and 1.5 keV the metal did not lift off, and we have continuous films instead of separate dots. This is consistent with earlier work showing that the decrease in resolution is due to increased scattering and decreased penetration depth of low voltage electrons in the resist. Monte Carlo simulations show that 1 keV electrons are scattered laterally up to 25 nm, much more than the higher energy electrons.[6] In addition, the electron range is only about 55 nm at 1 keV, increasing to about 90 nm at 1.5 keV,[7] so the resist did not develop completely. The penetration depth at 2 keV is increased to ~120 nm,[8] thus clearing the resist. This test shows that it is possible to achieve sub-200 nm resolution for 2 kV in 80 nm of PMMA, but thinner resists are needed for higher resolution or lower voltage. This will be addressed in more detail later.

[a]Electronic mail: ago2@cornell.edu

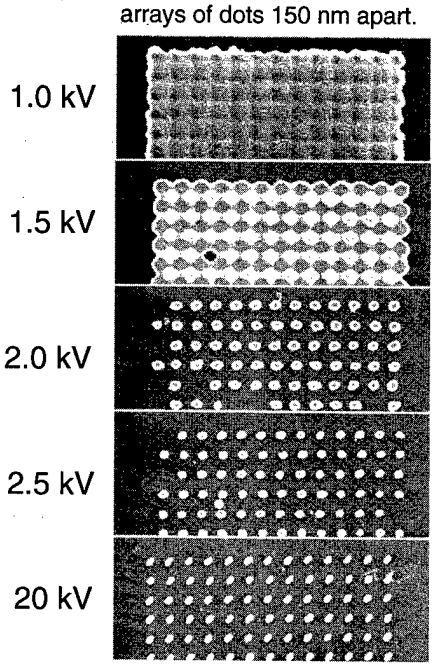


FIG. 1. SEM images. An array of dots with 150 nm pitch was patterned on an 80 nm layer of PMMA. Then 20 nm of Cr was thermally evaporated, followed by liftoff. The dose was adjusted for each voltage to minimize the dot size.

2. Proximity effects

A pattern consisting of a square with overlapping single pass lines was used to test the proximity effects. The SEM images of this are shown in Fig. 2. At 2 keV, after a threefold increase in the dose the overall shape of the pattern looks essentially the same. However, there is a major difference between the low and high dose pattern for 20 keV exposure. The square becomes pillow shaped, and the lines become a factor of 2–3 wider near the square. Even at low dose the linewidth increases by 40%–60% near the square compared to that several microns away. This illustrates the advantages of low voltage—it does not require a proximity correction or use of complicated multilayer schemes to suppress secondary electrons. Another method to reduce proximity effects is to go up to 75 or 100 keV, where scattered electrons provide a uniform background due to their long range.

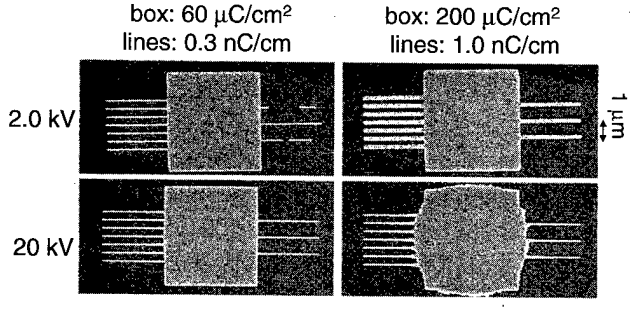


FIG. 2. SEM images. A box with overlapping single pass lines was patterned on an 80 nm thick layer of PMMA. Then 20 nm of Cr was thermally evaporated, followed by liftoff.

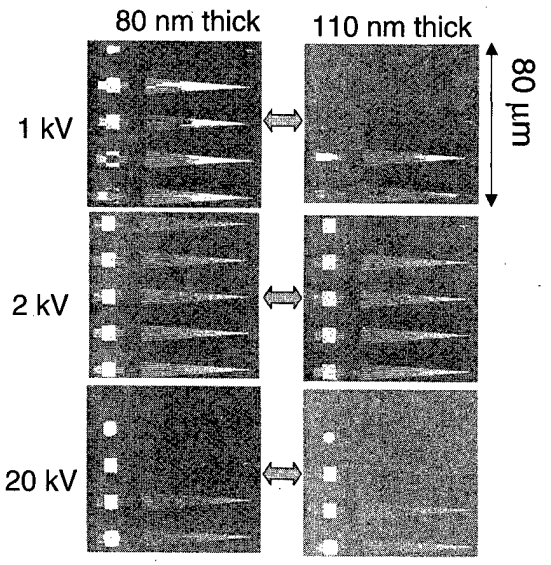


FIG. 3. SEM images of liftoff of five rows of the same pattern of converging lines and a square, with every row dose approximately twice the dose of the row above. The same pattern was exposed in 80 and 110 nm thicknesses of PMMA, with the same dose used for the same voltage.

3. Charging effects

Charging is expected to be more severe at low voltages due to more efficient electron stopping in the resist. Nevertheless, we did not observe any obvious charging effects, which would show up in bending of the single pass lines near a square—a large charged area (the lines in Figs. 2 and 4 are straight). We believe that the increased sensitivity to charging is offset by reduced resist thickness, which does not support a large surface voltage. Electron diffusion in thin films should further eliminate charging.[7]

4. Resist thickness effects

We used a pattern consisting of five rows of the same subpattern of lines and a square, with every row dose approximately twice the dose of the above row. The same pattern was exposed in 80 and 110 nm thicknesses of PMMA, with the same dose used for the same voltage. There is almost no difference among the patterns at high voltage (see Fig. 3). This is not surprising, since the 20 keV electron penetration depth is several microns,[9] and 80 and 110 nm resist exposures should look essentially the same. For 2 keV, however, there is some difference; the second row developed for the 80 but not for the 110 nm thick resist. And the greatest distinction is at 1 keV: the second and third rows developed for the thinner resist and did not develop for the thicker. This is consistent with the electron penetration depths discussed in Sec. II A 1. This test shows that the resist thickness is critical at low voltages, and that thinner resist should be used for reliable patterning.

B. Developed resist profiles

A SEM is limited as a tool for looking at developed resist profiles due to the high sensitivity of PMMA to electron beam damage and contamination. Lowering the current less-

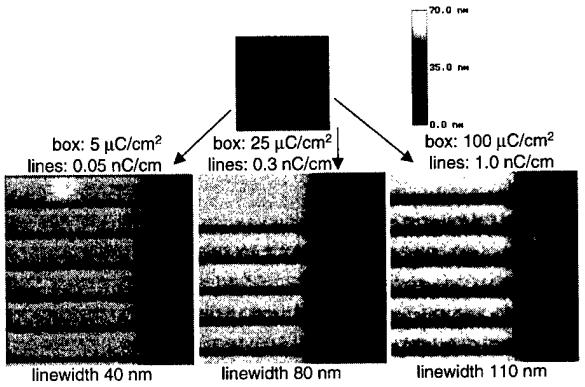


FIG. 4. AFM tapping mode image. A square with overlapping single pass lines was patterned on a 50 nm thick layer of PMMA with 1 keV electrons, and developed. The square and line dose increase from left to right.

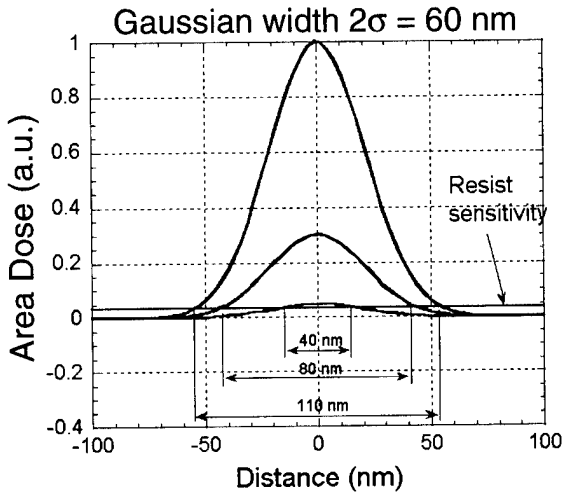


FIG. 5. Plot of the deposited dose for three different exposure times for the single pass lines shown in Fig. 4. The resist is exposed completely in the region where the curve is above the resist sensitivity level. Higher exposure time results in higher dose and thus wider lines.

ens the degradation, but reduces the image contrast. Therefore, we used atomic force microscopy (AFM) to image the resist patterns. We used a Dimension 3000 scanning probe microscope operated in tapping mode to reduce contact force between the tip and the sample.

We exposed a pattern consisting of a square and overlapping lines to study proximity effects and linewidth versus dose information. The developed resist profile of the exposed 50 nm layer of PMMA is shown in Fig. 4. Note that the proximity effects are negligible. Lines as narrow as 40 nm were reliably patterned. The edge of a line was determined as a place with a high height gradient. Measuring such narrow lines with an AFM tip is not very accurate due to tip size effects and the fact that the tip probes the top, not the base, of the resist. However, the relative magnitude of this error is less significant for wider lines, which provide most information for the fits, as will be discussed in Sec. II C.

C. Thin resist modeling

Using resists thinner than 50 nm presents difficult fabrication problems due to increased numbers of defects. A model can be useful to estimate the properties of thin resist layers, thus providing guidance for process optimization. Here we develop a simple model and estimate the process latitude of single pass line fabrication.

We can analyze the linewidth versus dose dependence to obtain information about scattering in the resist. The functional form of the energy density $E(r)$ at a distance r from a point irradiated with a charge Q has often been described by a two-Gaussian model (see, for example, Ref. 3):

$$E(r)=QK[(1/\sigma_f)^2\exp(-r^2/\sigma_f^2) + (\eta/\sigma_b^2)\exp(-r^2/\sigma_b^2)], \quad (1)$$

where σ_f and σ_b are the characteristic forward- and backscatter radii, η is the ratio of the total energy deposited in the resist by backscattered electrons to the total energy deposited by forward-scattered electrons in the resist, and K is a constant of proportionality. For high energies, where we have many fast secondary electrons generated by the interaction between the primary electron beam and the substrate, σ_f is usually small (tens of nm) and σ_b is large (μm). The low voltage regime is markedly different from that at high voltage, there is almost no backscattering and therefore η is small, so the formula can be reduced to a single Gaussian:

$$E(r)=QK[(1/\sigma_f)^2\exp(-r^2/\sigma_f^2)]. \quad (2)$$

The area where $E(r)$ is higher than the critical dose required to clear the resist determines the linewidth. We take width versus dose information from Fig. 4, fit it to formula (2), and obtain σ_f and resist sensitivity for 1 keV exposure in a 50 nm thick resist layer. When we linearize the Eq. (2), the most significant data points for the fit are the ones furthest from the origin, that is, the ones with large **r**. Therefore, a relatively large linewidth measurement error in narrow lines due to tip size effects does not influence the fit results significantly.

A plot of the deposited dose is shown in Fig. 5. The functional form of the linewidth agrees with Eq. (2) even for a 20-fold increase in line dose (from 0.05 to 1 nC/cm), which proves that we indeed have very little backscattering at low energies. The Gaussian width $2\sigma_f$ was found to be 60 nm, which is close to the 50–55 nm obtained from Monte Carlo simulations.[6,7] We did similar fits for 2 kV exposure, also with 20-fold variation in dose, and a Gaussian width $2\sigma_f$ was found to be 125 nm in 80 nm of PMMA. Similar σ_f for 2 kV was also observed by Lee *et al.*[9] (their data produce fits with $2\sigma_f=105$ nm for 66 nm of PMMA).

We use these numbers in our model, developed to give simplified analytical estimates of process parameters. We split two-dimensional (2D) electron travel into vertical and lateral motion. We consider the resist layer to be made of many thin layers. The total lateral electron scattering is the sum of independent scattering in all layers. If the total scat-

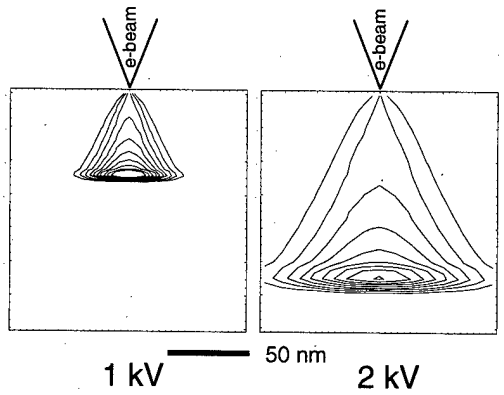


FIG. 6. Contour plot of modeled deposited energy profiles for 1 and 2 keV exposures in PMMA. Electrons are coming from the top and they scatter in the resist. The model parameters are taken from the linewidth vs dose fits and from known electron range values.

tering angle is not very large, this can be modeled as a random walk in lateral direction: an electron can scatter with equal probability either to the right or to the left in each layer. The energy loss can be approximated by the Bethe formula, $dE/dz \propto -1/E$.[10] The Bethe formula and small angle approximation are valid unless the resist thickness is close to the electron range, which is usually not the case for lithography patterning. As is known, the total amount of average horizontal travel increases as a square root of the number of layers, or of depth. Therefore, to calculate σ_f for an arbitrary resist thickness, it is enough to know σ_f at some particular thickness, which we take from the fits described above. To calculate the deposited energy profile in the vertical direction we take an electron range to be 55 nm for 1 keV (Ref. 7) and 120 nm for 2 keV.[8] The overall deposited energy profile in PMMA is plotted in Fig. 6. It has a fruit-like shape, like that obtained in simulations elsewhere.[9,11]

Calculated process latitudes of patterning single pass lines are plotted in Fig. 7 for 1 and 2 keV exposures. As expected, one can pattern wider lines more reliably. Choosing a reasonable requirement, such as *change in linewidth <20%*, one can pattern lines as narrow as 20 nm with 1 keV and 30 nm with 2 keV exposure for a 20 nm resist layer. The resolution is about a factor of 2 better than that for 50–70 nm resists with the same process latitude requirement. Although the resolution usually increases with increasing accelerating voltage, here we observe the opposite effect. From Fig. 6 we see that the "fruit" for 2 keV is both longer and slightly wider than the 1 keV one. Electrons at 2 keV have larger range, which results in a longer profile. Both 1 and 2 keV electrons interact strongly with resist, which results in large scattering and in wider spreading at 2 keV, possibly due to a larger range of 2 keV electrons.

In our simplified model we relied on single-Gaussian behavior for the energy distribution profile. However, this does not limit wider applicability of the model; one can "map" the energy distribution profile for any process, recording linewidth versus dose information, and use it for fitting to the theoretical formula. Forward- and backscatter radii, incoming electron beam diameter, resist sensitivity, and other important process parameters can also be inferred from the fitting. This method enables us to measure these parameters directly from experiment as opposed to taking them from simulations.

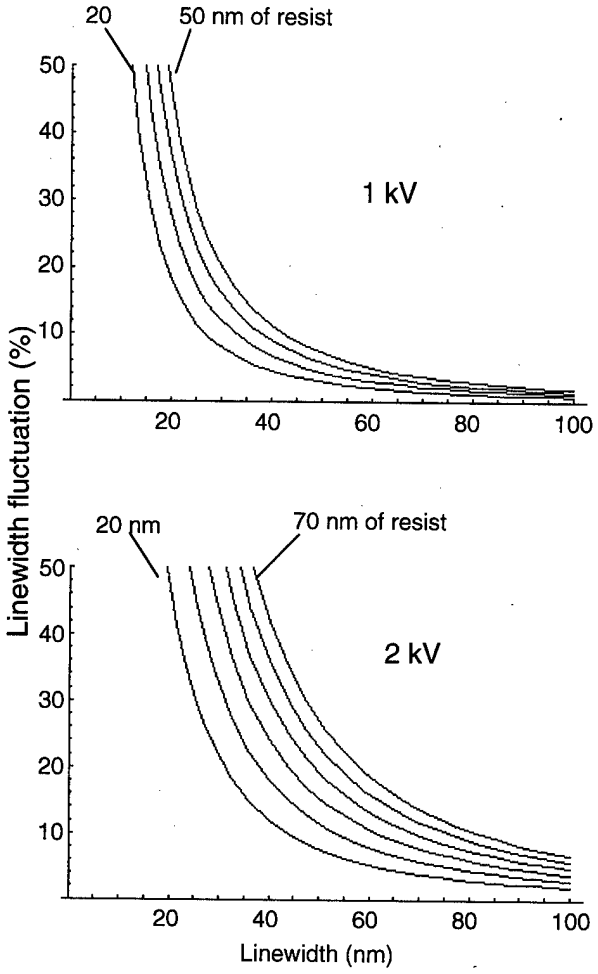


FIG. 7. Plots of modeled process latitudes of patterning single pass lines for 1 and 2 keV exposures for thin resists. 10% fluctuation in resist sensitivity (or deposited dose) is assumed. For other dose fluctuations the vertical axis should be scaled accordingly. Different curves correspond to different resist thicknesses, with 10 nm steps.

III. SUMMARY

We studied low voltage (1–2 kV) lithography in PMMA, and compared it to 20 kV lithography. The proximity effects and charging were negligible at low voltage. Resist thickness of ~50 nm or smaller is required for reliable patterning. The linewidth versus dose functional behavior agrees with a single-Gaussian model for 1 and 2 keV electrons, proving that the major cause of line broadening at these energies is forward scattering. We modeled thin resist behavior and showed that the resolution increases significantly with thinner resists. Overall, we showed that 1 and 2 kV lithography in PMMA is a viable alternative for sub-100 nm lithography.

ACKNOWLEDGMENTS

The authors thank Dustin Carr, other Cornell Nanofabrication Facility staff, and members of the Craighead research group for help and suggestions. This work was supported by DARPA/ITO.

[1]E. Kratschmer, H. S. Kim, M. G. R. Thomson, K. Y. Lee, S. A. Rishton, M. L. Yu, and T. H. P. Chang, J. Vac. Sci. Technol. B **13**, 2498 (1995).
[2]K. Kragler, E. Günther, R. Leuschner, G. Falk, and H. von Seggern, J. Vac. Sci. Technol. B **14**, 1327 (1996).
[3]E. A. Dobisz, C. R. K. Marrian, R. E. Salvino, M. A. Ancona, F. K. Perkins, and N. H. Turner, J. Vac. Sci. Technol. B **11**, 2733 (1993).
[4]K. J. Polasko, Y. W. Yan, and R. F. W. Pease, Proc. SPIE **333**, 76 (1982).
[5]M. J. Lercel, H. G. Craighead, A. N. Parikh, K. Seshadi, and D. L. Allara, Appl. Phys. Lett. **68**, 1504 (1996).
[6]D. M. Tanenbaum, C. W. Lo, M. Isaacson, H. G. Craighead, M. J. Rooks, K. Y. Lee, W. S. Huang, and T. H. P. Chang, J. Vac. Sci. Technol. B **14**, 3829 (1996).
[7]M. A. McCord and T. H. Newman, J. Vac. Sci. Technol. B **10**, 3083 (1992).
[8]C. W. Lo, M. J. Rooks, W. K. Lo, M. Isaacson, and H. G. Craighead, J. Vac. Sci. Technol. B **13**, 812 (1995).
[9]Y.-H. Lee, R. Browning, N. Maluf, G. Owen, and R. F. W. Pease, J. Vac. Sci. Technol. B **10**, 3094 (1992).
[10]H. A. Bethe, Ann. Phys. (Leipzig) **5**, 325 (1930).
[11]C. W. Lo, W. K. Lo, M. J. Rooks, M. Isaacson, H. G. Craighead, and A. E. Novembre, J. Vac. Sci. Technol. B **13**, 2980 (1995).

Surface roughness development during photoresist dissolution

Lewis W. Flanagin,[a] Vivek K. Singh,[b] and C. Grant Willson[c]
Department of Chemical Engineering, The University of Texas at Austin, Austin, Texas 78712
and Technology CAD Department, Intel Corporation, Hillsboro, Oregon 97124-6497

(Received 16 February 1999; accepted 14 May 1999)

The minimization of nanoscale roughness in patterned images has become a priority for the process of photolithography in the production of microprocessors. In order to probe the molecular basis for surface roughness, the development of photoresist has been simulated through application of the critical-ionization model to a three-dimensional molecular lattice representation of the polymer matrix. The model was adapted to describe chemically amplified photoresists of the sort now commonly used in microlithography. Simulations of the dependence of the dissolution rate and surface roughness on the degree of polymerization, polydispersity, and fractional deprotection agree with experimental results. Changes in surface roughness are shown to correlate with the length of the experimentally observed induction period. Model predictions for the effect of void fraction and developer concentration on roughness are also presented. Observations of differences in the effect of developer concentration on top-surface and sidewall roughness are explained by a critical development time predicted by the simulation. © *1999 American Vacuum Society.*
[S0734-211X(99)06704-9]

I. INTRODUCTION

The tremendous gain in computational speed and storage capacity afforded by miniaturization of the integrated circuit impels the semiconductor industry forward in its quest for smaller device features. The demands placed on microlithography grow more arduous with each new generation of microprocessors. The minimization of roughness associated with the surface and edges of photoresist images now stands as one of the challenges to continued advances in lithographic technology.

The problems of top-surface and line-edge roughness have drawn a considerable amount of attention within the last year, and several recent atomic force microscopy (AFM) studies have yielded notable insights into the process dependency of photoresist roughness. He and Cerrina[1] have studied the relationship between surface roughness and exposure dose over a range of postexposure bake times for a positive-tone chemically amplified photoresist. Their results indicate that systems having the same overall average degree of deprotection, but different process histories, exhibit similar surface morphologies, but different degrees of roughness. Reynolds and Taylor[2–4] have explored the responses of top-surface and sidewall roughness for chemically amplified photoresists to a variety of processing parameters, including exposure dose and developer concentration. They find that higher exposure doses and lower developer concentrations lead to decreased surface roughness, but their measurements of sidewall roughness show no correlation with their surface roughness results and reveal a negligible dependence of sidewall roughness on either exposure dose or developer concentration. Yoshimura *et al.*[5] have characterized the effects of the polymer structure and molecular weight distribution on edge roughness. They show that a photoresist based on cresol novolac exhibits a rougher surface than one based on polyvinylphenol and attribute this difference in roughness to the greater structural rigidity of cresol novolac that results from the presence of aromatic rings in the backbone of the polymer. They observe that polymers having lower molecular weight and lower polydispersity produce less roughness.

The need for robust lithographic simulators that can model and predict the generation of roughness during photoresist development grows imperative. Most of the earlier simulations of photoresist development do not adequately describe surface roughness because they considered the photoresist a uniform structure. Guerrieri and Neureuther[6] have studied the time evolution of the development etch front using a simplified material crack model in which development proceeds faster along highly exposed filaments (cracks) than through the background matrix, and they have found that the surface roughness increases with crack length. Trefonas[7] alludes to the production of top-surface roughness during his molecular cell-based simulations of percolational development. Scheckler *et al.*[8] have used an even more advanced molecular-scale photoresist development simulation in which a realistic polymer chain length distribution is represented to demonstrate excellent agreement between their model and AFM measurements for the dose dependence of surface roughness. These studies have made significant advancements in the simulation of roughness in photoresist development, yet they still rely heavily on empirical data for the dependence of the dissolution rate on such fundamental quantities as molecular weight and degree of deprotection.

Our simulations represent the polymer matrix as a three-dimensional rectangular lattice in which each lattice cell corresponds exactly to a polymer repeat unit. The lattice cells are strung together to form polymer chains via random walks as described previously.[9] The dissolution of these chains is governed by the critical-ionization model,[10] which proposes

[a]Present address: The University of Texas at Austin.
[b]Present address: Intel Corporation.
[c]Corresponding author; electronic mail: willson@che.utexas.edu

that a critical fraction of segments on a chain must be in the ionized state in order for the chain to dissolve. The lattice model has been adapted to represent a chemically amplified photoresist system in which each cell begins as either a chemically protected (blocked) group or a chemically deprotected (unblocked) group. In a previous paper, we presented simulations of photoresist dissolution on a two-dimensional (2D) grid.[9] Because percolation models have demonstrated a dependence on the dimensionality of the system,[6,7] we have extended the moving front from a line to a surface, which begins parallel to the substrate. Adding the third dimension alters the quantitative results, but not the qualitative behavior, of our system. The third dimension (3D) allows a greater number of possible chain conformations and a greater number of cell faces that may be exposed to developer (six instead of four).

The simulations described in this article allow polymer blends so that we can study the effects of polydispersity. An initial void fraction, which can represent either the inherent polymer free volume or residual casting solvent, is also specified. The molecular-level model correctly predicts trends in the responses of the dissolution rate and roughness to changes in the degree of polymerization, polydispersity, degree of deprotection, polymer free volume, residual casting solvent, and development time.

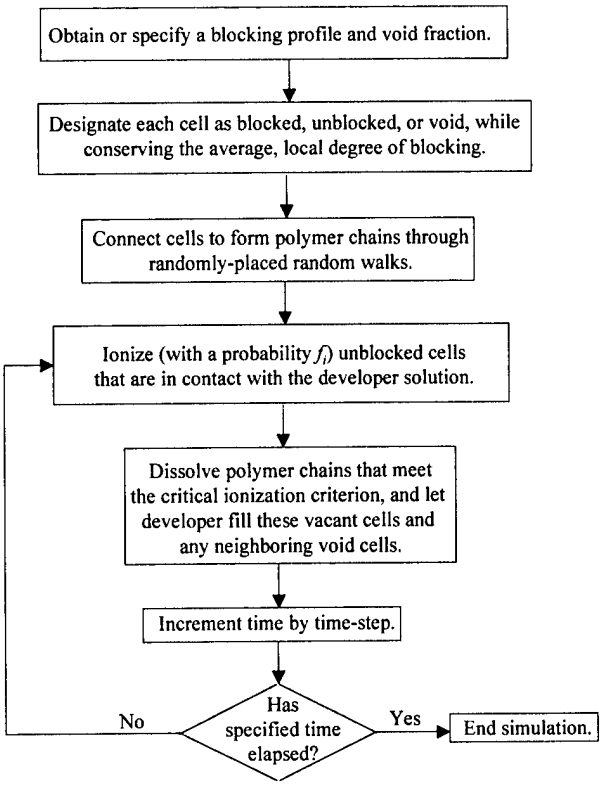


FIG. 1. Flowchart for the three-dimensional molecular simulation.

II. MODEL DESCRIPTION

A three-dimensional lattice of cubic cells is used to represent the polymer matrix. The number of cells in each orthogonal direction is 73. Each lattice cell, having sides of 0.7 nm in length, corresponds exactly to a single repeating unit of a polymer chain and may have one of the following states: blocked, unblocked, ionized, developed, or void. The initial degree of blocking, specified by the user, is designated f_{b0}. For a blanket (uniform) exposure, the value of f_{b0} represents the average degree of blocking present in the entire lattice. For a patterned photoresist image, f_{b0} represents the average degree of blocking prior to exposure; the spatial variation in blocking, $f_b(x,y,z)$, prior to dissolution is the product of f_{b0} and $p(x,y,z)$, a function with values between 0 and 1. The function $p(x,y,z)$ describes the relative amount of protection remaining after exposure and the postexposure bake and is supplied by FINLE Technology's PROLITH, one of several commercially available lithographic simulators. The initial void fraction, also specified by the user, is designated f_v. The void fraction can represent either the polymer's inherent free volume or residual casting solvent. A flowchart describing the steps in the molecular model simulation is shown in Fig. 1.

Prior to development, a fraction f_v of the cells are selected at random to be void. Each of the remaining cells is designated randomly as either blocked or unblocked, where the probability of being blocked is given by $f_b/(1-f_v)$. This expression for the probability of being blocked ensures that the average local concentration of blocked sites is preserved through the random selection process. The accuracy of the discrete representation of the continuous function $f_b(x,y,z)$ is checked by examining the average amount of blocking over an area and comparing this against the scalar-field value supplied. (The number of cells used in the average is chosen so that their combined volume is the same as the cellular volume in the PROLITH blocking profile.) As in our 2D simulations,[9] statistical variation is introduced through the transformation of the continuous polymer matrix into a molecular grid. However, such roughness is expected from the statistical variation in the energy that is deposited into the photoresist during exposure.[11]

The simulation requires having polymer chains on the grid. The user specifies values (DP_i) for the degree of polymerization and the fraction of chains (f_{DPi}) that are supposed to have each degree of polymerization. To form these chains, the cells are strung together[12] via random walks, which are begun in randomly chosen blocked or unblocked cells that have not already been used in the formation of other chains. The number of steps taken in each walk equals one of the nominal degrees of polymerization specified by the user in the input file. The fraction of walks with DP_i steps is given by f_{DPi}. The random walks avoid the cells designated as void (or filled with solvent), but more than one chain may share any given blocked or unblocked cell (Fig. 2). A cell may be counted as belonging to any single chain only once; thus, there are no repeated cells within any given chain. This precaution leads to a distribution of actual chain lengths. The formation of chains continues until all blocked or unblocked cells have become a part of at least one chain. If only one degree of polymerization is specified, the random walk process produces a Gaussian distribution of actual chain lengths

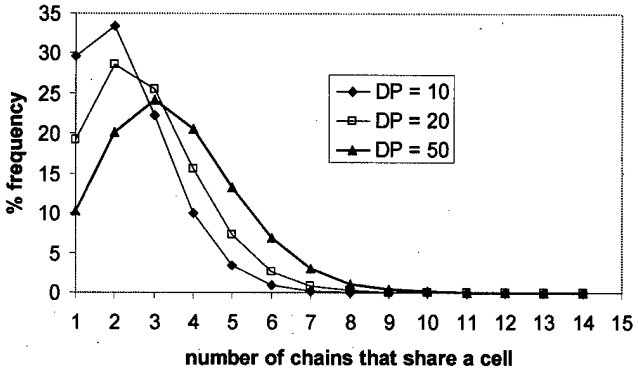


FIG. 2. Distribution in the number of chains that share a cell as a function of the specified degree of polymerization.

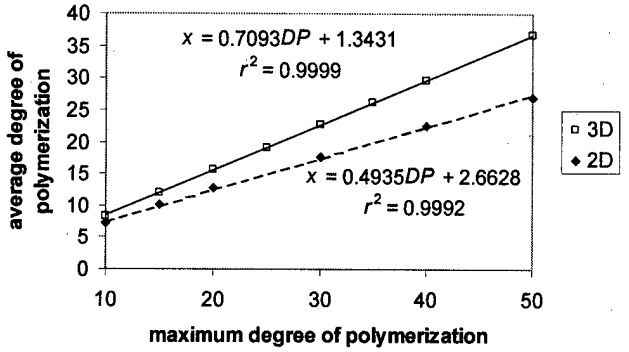


FIG. 4. Relationship between the maximum degree of polymerization (DP) and the average degree of polymerization (x) from the random walk process. Note that the r^2 values are the squares of the linear correlation coefficients for the least-squares linear regression of the data.

[Fig. 3(a)]. The chain length distribution for 2D is shown also for comparison. The additional third dimension narrows the chain length distribution and shifts it towards greater chain lengths (which will presently be explained). If two degrees of polymerization are specified, a bimodal distribution results [Fig. 3(b)], where the distribution about each mode is Gaussian.

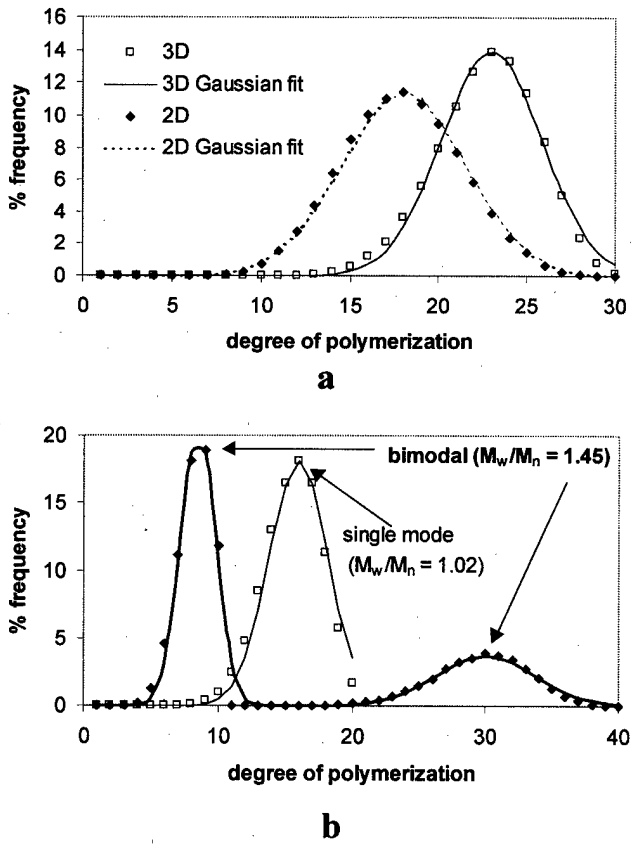


FIG. 3. (a) Polymer chain length distributions produced through the random walk process when a nominal (maximum) chain length of 30 is specified for 3D and 2D simulations. (b) Two polymer chain length distributions having the same average degree of polymerization (15.5) but different polydispersity (DP=20 for the single-mode distribution; $DP_1=10$, $DP_2=40$, and $f_{DP1}=0.6589$ for the bimodal distribution).

The average (or mean) degree of polymerization is easily obtained from the chain length distribution. In Fig. 4, the average degree of polymerization is plotted versus the maximum degree of polymerization, which is identically equal to the specified number of steps to be taken in the formation of each chain. The corresponding plot for two dimensions is also shown. The range of chain lengths that is shown here, 10–50 repeat units, is representative of the oligomeric, phenolic polymers used to formulate photoresists. Note that the process produces a linear relationship between the average and maximum degrees of polymerization. The third dimension increases the average degree of polymerization because the extra degree of freedom in the third dimension reduces the average number of redundant steps taken during the random walks and increases the probability for a growing chain to encounter an available cell. The random walk process produces a polymer size distribution that is uniform throughout most of the film, but that has slightly smaller values towards the ends (Fig. 5).

The mean-squared radius of gyration, $\langle R_g^2 \rangle$, and mean-squared end-to-end distance, $\langle R^2 \rangle$, are common measures for the polymer chain length. Calculated values for $\langle R_g^2 \rangle$ and $\langle R^2 \rangle$ from our simulations, shown in Fig. 6, exhibit the linear dependence on the degree of polymerization expected for dense polymer networks.[13] The ratio $\langle R^2 \rangle / \langle R_g^2 \rangle$ (Table I) is close to the theoretical value of 6.[13]

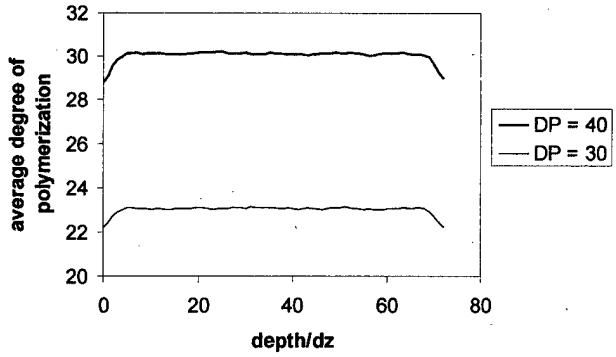


FIG. 5. Variation in the average degree of polymerization with depth into the film.

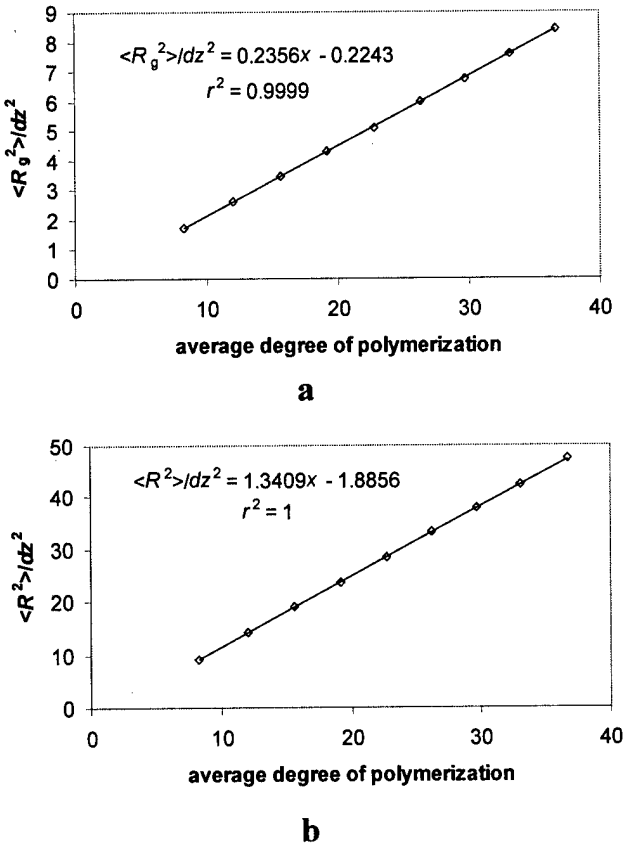


FIG. 6. Measures of polymer chain length as a function of the average degree of polymerization (x): (a) mean-squared radius of gyration, $\langle R_g^2\rangle$, and (b) mean-squared end-to-end distance, $\langle R^2\rangle$. Note that the r^2 values are the squares of the linear correlation coefficients for the least-squares linear regression of the data.

Once polymer chains have completely filled the grid, the simulator begins the process of ionization. All cells on the top layer of the lattice are always exposed to developer. Ionization is possible for all top-layer cells that are in the unblocked state, and any top-layer cells that are in the void state automatically fill with developer. Unblocked cells that are adjacent to developed cells may also undergo ionization. Whether ionization actually occurs depends on the probability of ionization (f_i), a factor specified by the user that depends on the concentration of the developer and the pK_a of the resin.[14,15]

TABLE I. Ratio of mean-squared end-to-end distance to mean-squared radius of gyration as a function of maximum degree of polymerization.

DP	$\langle R^2\rangle/\langle R_g^2\rangle$
10	5.36
15	5.46
20	5.50
25	5.53
30	5.60
35	5.58
40	5.61
45	5.58
50	5.64

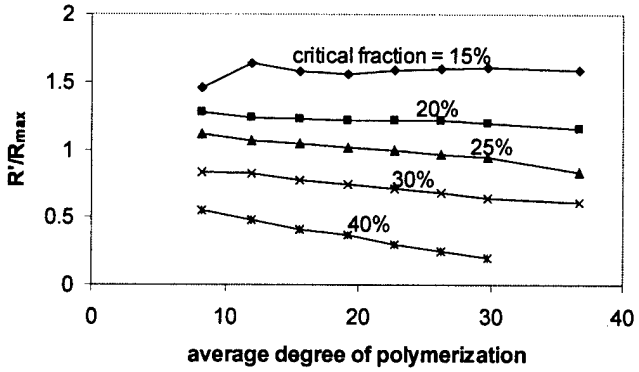


FIG. 7. Correction factor by which the time step must be multiplied in order to obtain the specified $R_{\max}$ when $f_{b0}=0$, $f_v=0$, and $f_i=0.5$.

After the ionization step, each chain is examined to see if it has the requisite fraction of ionized sites for dissolution (f_c). At the end of this accounting, all chains meeting the solubility criterion are dissolved (erased). If a cell is shared by at least one other chain, it remains in its current state (blocked or unblocked); otherwise, the cell is counted as developed. The cycle of ionization followed by dissolution continues until the time specified by the user has elapsed. The time (t_c) that each cycle represents is given by

$$t_c=\frac{dz}{R_{\max}}(R'/R_{\max}). \tag{1}$$

Here dz is the cell height, $R_{\max}$ is the experimental dissolution rate observed under complete exposure, and $R'/R_{\max}$ is a correction factor added to ensure that the rate calculated by the simulator when $f_{b0}=0$ is $R_{\max}$. (For an ideal monomer system with $f_{b0}=0$ and $f_i=1$, $R'/R_{\max}=1$ because one entire layer dissolves during each ionization/dissolution cycle. $R'/R_{\max}$ is a measure of the deviation from this ideality.) Figure 7 shows how $R'/R_{\max}$ varies with the critical degree of ionization and average degree of polymerization. A detailed description of $R'/R_{\max}$ is given elsewhere.[9]

The spatial average and the standard deviation of the thickness of remaining photoresist are computed during the simulation. The surface roughness is defined throughout this article as the standard deviation in the spatial variation of the photoresist thickness. Results for surface roughness, r, and remaining thickness, θ, versus time, t, are presented as averages from multiple simulations using different seeds for the random number generator that fills the lattice. The purpose of running multiple simulations is to sample a larger subspace of the total ensemble of possible spatial configurations, and standard deviations of the results are provided to show reproducibility. Dimensionless variables are used wherever possible. Lengths are scaled by the cell height (dz), and time is scaled by the time per ionization/dissolution cycle [Eq. (1)].

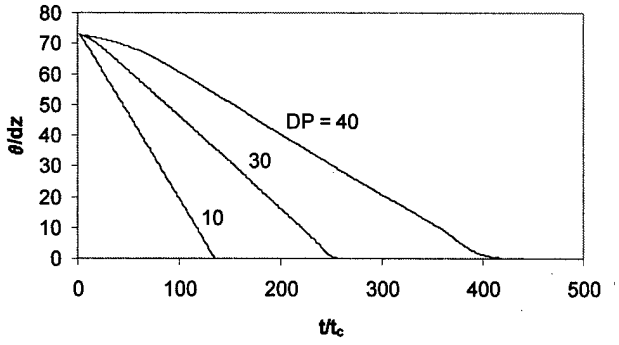


FIG. 8. Effect of the nominal degree of polymerization, DP, on a dimensionless plot of the average thickness of polymer film remaining, θ, vs time, t ($f_c=0.4$, $f_{b0}=0$, $f_v=0$, $f_i=0.5$). Each curve is the result of averaging seven simulations.

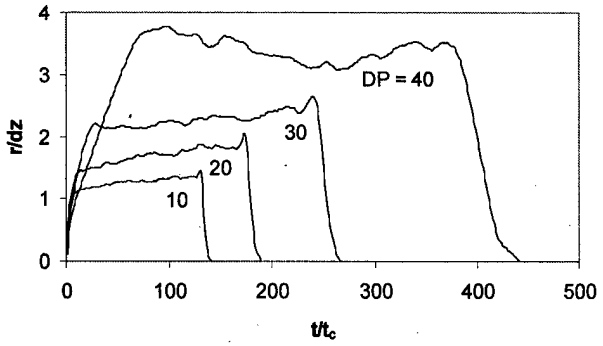


FIG. 9. Effect of the nominal degree of polymerization, DP, on a dimensionless plot of roughness, r, vs time, t ($f_c=0.4$, $f_{b0}=0$, $f_v=0$, $f_i=0.5$). Each curve is the result of averaging seven simulations.

III. SIMULATION RESULTS

A. Effect of the degree of polymerization

In Fig. 8, the remaining thickness of a fully deprotected polymer is plotted versus development time for several nominal degrees of polymerization. The model predicts that the dissolution rate decreases with increasing degree of polymerization. Many experimental studies confirm this result.[16–20] The instantaneous dissolution rate is given by the negative of the slope from a plot of thickness versus time ($-d\theta/dt$, where θ = thickness and t = time). The instantaneous rate is constant over most of the course of dissolution, but interesting deviations in $d\theta/dt$ occur at the start and end of the process. The time that the system takes to achieve a constant dissolution rate is called the induction period. This phenomenon, which is observed experimentally, appears "naturally" in the simulation. Because the instantaneous rate is smaller at both the top and bottom surfaces of the film, surface rate inhibition is said to occur at both interfaces. Dissolution rates are customarily defined as the instantaneous rate when half of the initial film thickness remains. This arbitrary definition minimizes the effect of surface inhibition, and is adopted for the purposes of this article.

From Fig. 8, it is apparent that surface rate inhibition becomes more pronounced as the degree of polymerization increases. All of these systems display an induction period, but the length of the induction period increases faster than a linear dependence on the degree of polymerization would predict. The surface rate inhibition at the bottom film surface also grows with the degree of polymerization, but the dependence is not as steep at the bottom as it is at the top.

The variation in average thickness across simulations (not shown) increases until developer reaches the region where bottom-surface rate inhibition occurs, and the most rapid increase occurs during the top-surface induction period. The results indicate that the variation among runs rises significantly with an increase in the degree of polymerization.

Surface roughness as a function of development time and degree of polymerization is plotted in Fig. 9 for the system described in Fig. 8. In agreement with the experimental results of Yoshimura *et al.*,[5] the simulation results show that surface roughness increases with increasing degrees of polymerization. For all degrees of polymerization, the same general behavior is observed: roughness rises sharply at first, reaches a plateau with a general slight incline, spikes, and then rapidly falls. This trend is compatible with AFM measurements that we have taken for the time evolution of surface roughness (Table II). Variation in surface roughness across multiple simulations (not shown) increases with degree of polymerization, and a spike occurs at the same location as the spike in surface roughness. A comparison between Figs. 8 and 9 reveals that the initial, sharp rise in roughness corresponds exactly to the induction period of the dissolution rate and that the final spike in roughness occurs where bottom-surface rate inhibition begins.

The correspondence between surface rate inhibition and surface roughness in our simulations helps to explain the surface rate inhibition phenomenon. An increase in roughness, by definition, occurs when the surface area exposed to developer increases. When dissolution begins, the film surface is flat, and the area exposed to developer is the smallest it will be during the entire process. As chains dissolve away, a greater surface area is exposed, which leads to increases in the surface roughness and the instantaneous dissolution rate. The amount of surface area that may be exposed at any time reaches a nearly constant value, after which point the instantaneous dissolution rate and surface roughness stay relatively constant. When developer finally reaches the bottom of the film, the surface area in contact with developer decreases rapidly, and both the surface roughness and dissolution rate

TABLE II. Surface roughness as a function of development time[a] (rms=root mean square).

Development time (s)	Thickness remaining (nm)	rms roughness (nm)
0	1089	0.3
25	999	3.8
50	807	3.8

[a]Polymer: *m*-cresol novolac, cast with PGMEA at 3000 rpm for 30 s, baked at 120 °C for 120 s, developed with 0.182N TMAH for the time indicated, rinsed with distilled water, and blown dry with nitrogen. The surface roughness was calculated from AFM images taken on a Park Scientific Instruments Autoprobe.

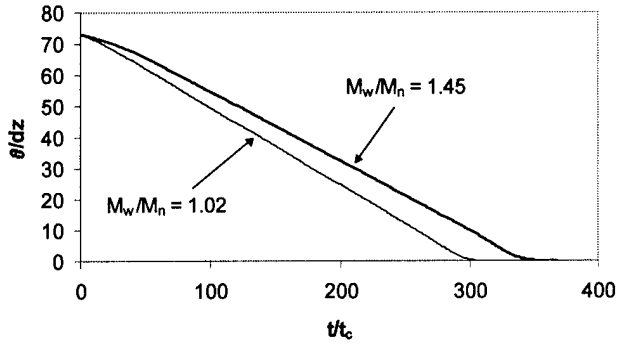


FIG. 10. Effect of polydispersity, M_w/M_n, on a dimensionless plot of the average thickness of polymer film remaining, θ, vs time, t ($f_c=0.4$, $f_{b0}=0.1$, $f_v=0$, $f_i=0.5$). The polymer chain length distribution for each curve is shown in Fig. 3. Each curve is the result of averaging seven simulations.

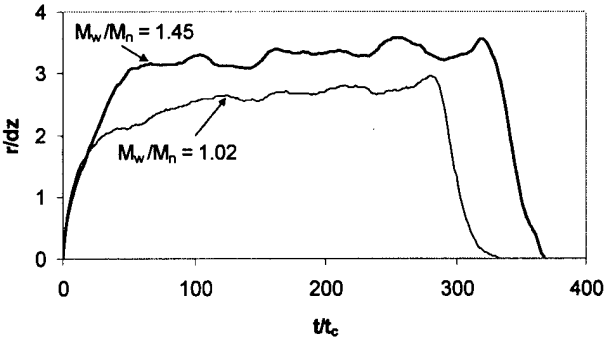


FIG. 11. Effect of polydispersity, M_w/M_n, on a dimensionless plot of roughness, r, vs time, t ($f_c=0.4$, $f_{b0}=0.1$, $f_v=0$, $f_i=0.5$). The polymer chain length distribution for each curve is shown in Fig. 3. Each curve is the result of averaging seven simulations.

decline. Thus, both the induction period and bottom-surface rate inhibition arise from rapid changes in surface roughness. Similar arguments have been used to explain surface rate inhibition in percolation models.[7]

B. Effect of polydispersity

In Fig. 3(b), a single-mode and a bimodal polymer chain length distribution, which have the same average degree of polymerization but different polydispersity, are compared. The single-mode distribution has a narrower chain length distribution and a lower polydispersity ($M_w/M_n=1.02$) than the bimodal distribution ($M_w/M_n=1.45$).

Experimental efforts to determine the effect of polydispersity on the dissolution rate have led to conflicting results. Tsiartas *et al.*[1] measured the dissolution rates of blends of fractionated novolacs and found that the dissolution rate of the novolac blends decreased with increasing polydispersity. A similar study by Barclay *et al.*[2] concluded that increasing the polydispersity of (higher molecular weight) poly(hydroxystyrene) leads to higher dissolution rates. Figure 10 depicts the predictions from our simulations for the dissolution rate of the chain distributions shown in Fig. 3(b). According to our model, the lower-polydispersity polymer dissolves faster than the polymer blend, which has a wider molecular-weight distribution. In agreement with the conclusions of Tsiartas *et al.*, the simulations suggest that higher molecular-weight fractions have a disproportionately large influence on the overall dissolution rate.

Figure 11 shows how surface roughness develops in these two systems. In agreement with the experimental results of Yoshimura *et al.*,[5] the higher-polydispersity sample achieves a greater degree of surface roughness. The time that the system requires for the degree of surface roughness to level increases with polydispersity and corresponds to the induction period in Fig. 10. Thus, increasing the polydispersity enhances the effect of surface rate inhibition. Those who formulate resists have recognized this phenomenon for some time, but there has been no satisfying explanation for the observation.

C. Effect of the degree of deprotection

The time evolution of surface roughness as a function of the degree of blocking, f_{b0}, is depicted in Fig. 12. The simulations predict that roughness should increase with the degree of blocking as long as the film is able to clear. The dissolution rate decreases with increasing degrees of blocking. After a critical degree of blocking is exceeded (e.g., $f_{b0}=0.23$ for DP=10), the film cannot completely clear, and a measure of roughness remains even after the film stops developing. For a completely protected system (i.e., $f_{b0}=1$), none of the film clears, and no surface roughness develops. Thus, surface roughness begins at zero, passes through a maximum, and returns to zero as the fractional deprotection is varied from 0 to 1. This point is illustrated by Fig. 13, where the predictions of our model for roughness as a function of deprotection at a set development time are compared to the experimental data of He and Cerrina.[1] The relationship between exposure dose and fractional deprotection has been quantified previously.[21] Figure 14 demonstrates that very little clearing occurs at low doses (or at high degrees of blocking) in both simulation and experiment.

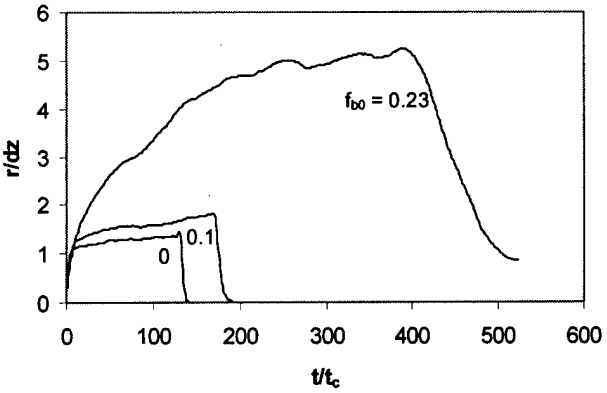


FIG. 12. Effect of the degree of blocking, f_{b0}, on a dimensionless plot of roughness, r, vs time, t ($f_c=0.4$, DP=10, $f_v=0$, $f_i=0.5$). Each curve is the result of averaging seven simulations.

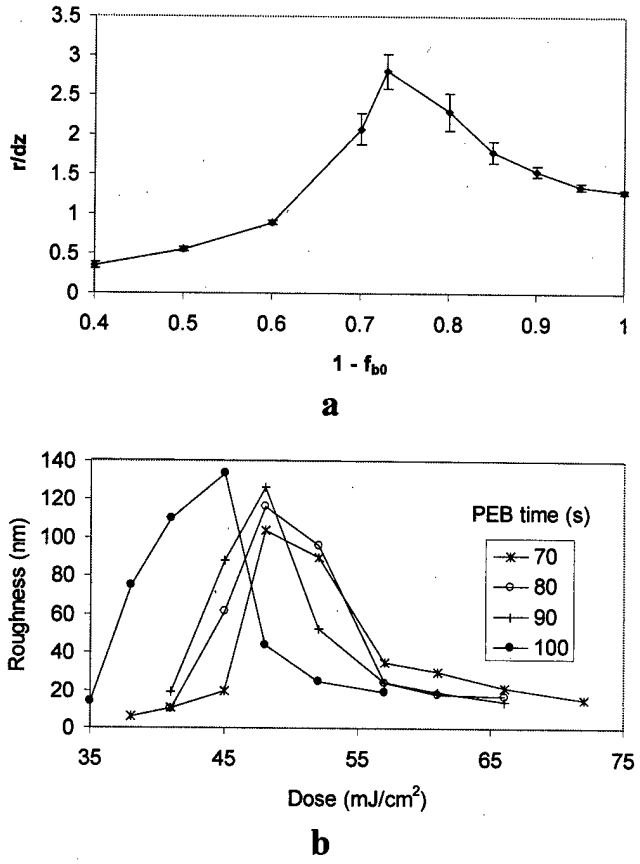


FIG. 13. Comparison of (a) our simulation predictions to (b) experimental results of He and Cerrina (Ref. 1) for the effect of deprotection on surface roughness for a fixed development time. Simulation conditions: $f_c=0.4$, DP$=10$, $f_v=0$, $f_i=0.5$, $t/t_c=60$.

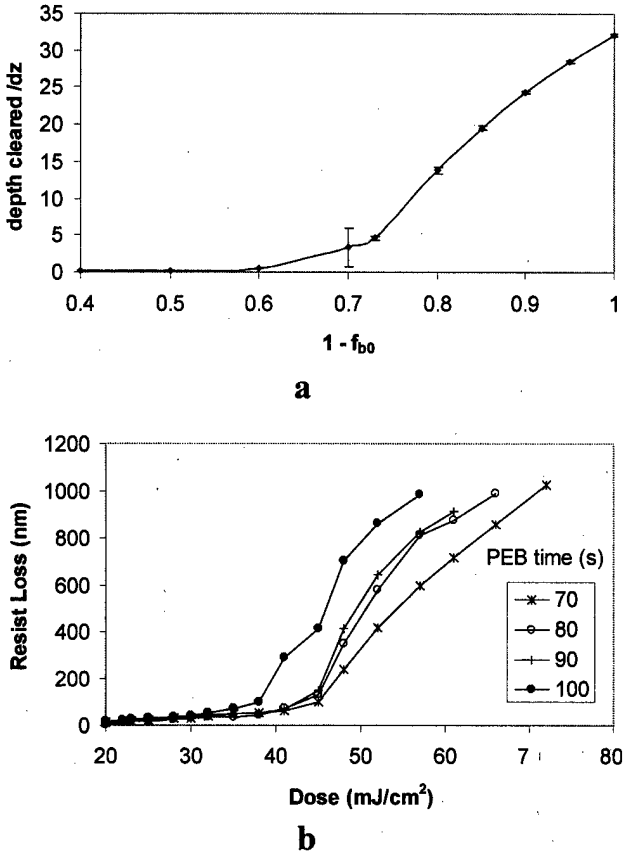


FIG. 14. Comparison of (a) our simulation predictions to (b) experimental results of He and Cerrina (Ref. 1) for the effect of deprotection on depth cleared for a fixed development time. Simulation conditions: $f_c=0.4$, DP $=10$, $f_v=0$, $f_i=0.5$, $t/t_c=60$.

D. Effect of the void fraction

The simulations described in this article allow a portion of the polymer matrix to be set aside as void space. The specified initial void fraction, f_v, may represent either free volume or residual casting solvent. Void cells automatically convert to developed cells whenever one of their neighboring cells develops.

In Fig. 15 the effect of void space in the polymer matrix on surface roughness is considered. [For reference, the free volume for polymers at the glass transition temperature is 0.025 according to the Williams–Landel–Ferry (WLF) equation.[22]] Increasing the initial void fraction causes the development rate to increase and the surface roughness to decrease. The standard deviation in surface roughness across multiple simulations (not shown) decreases as the initial void fraction increases.

E. Effect of the developer concentration

In Reynolds and Taylor's initial AFM examination of the effect of developer concentration on roughness,[2] the surface roughness of the unexposed portion of a chemically amplified resist was measured following a development time long enough for complete etching of the exposed portion of the photoresist. In later studies,[3,4] the same investigators examined the effect of developer concentration on roughness as a function of development time. Under both the earlier and later sets of conditions, surface roughness was observed to

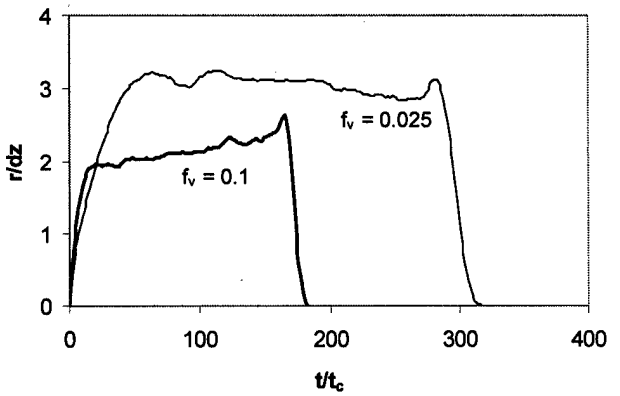


FIG. 15. Effect of void fraction, f_v, on a dimensionless plot of roughness, r, vs time, t ($f_c=0.4$, DP$=40$, $f_{b0}=0$, $f_i=0.5$). Each curve is the result of averaging seven simulations.

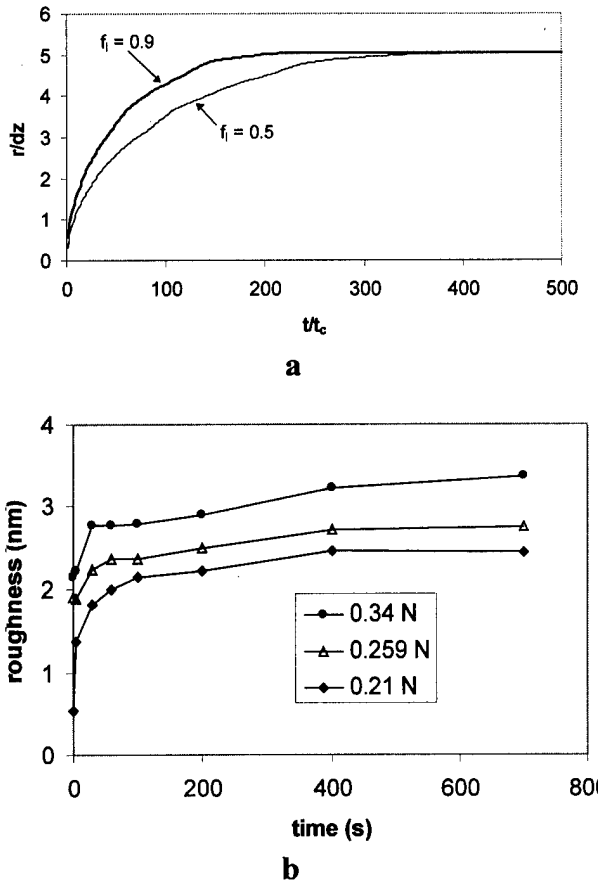


FIG. 16. Comparison of (a) our simulation predictions to (b) experimental results of Reynolds and Taylor (Refs. 2–4) for the effect of developer concentration on roughness as a function of development time. Simulation conditions: $f_c=0.4$, DP=10, $f_{b0}=0.27$, $f_v=0$. Each curve is the result of averaging seven simulations.

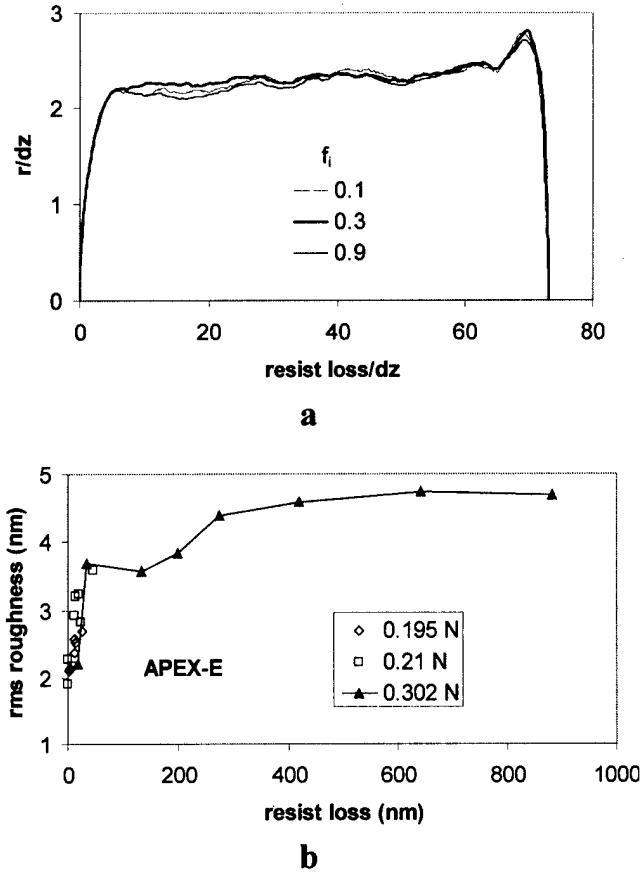


FIG. 17. Comparison of (a) our simulation predictions to (b) experimental results of Reynolds and Taylor (Refs. 2–4) for the effect of developer concentration on roughness as a function of resist loss. Simulation conditions: $f_c=0.4$, DP=30, $f_{b0}=0$, $f_v=0$. Each curve is the result of averaging seven simulations.

increase with developer concentration, whereas the sidewall roughness appears to be independent of developer concentration.

For comparison, our simulation uses a partially blocked matrix to portray the unexposed portion of the photoresist. The simulation parameter analogous to developer concentration is f_i, which is the probability of ionization for a cell in contact with developer. Greater values for f_i correspond to higher developer concentrations. In Fig. 16, the simulated predictions for roughness as a function of development time and f_i are compared to recent AFM measurements[3] of roughness as a function of development time and developer concentration. Increasing f_i causes the roughness in the simulations [Fig. 16(a)] to develop at a faster rate, but roughness eventually reaches an asymptotic value that is independent of f_i. If development is interrupted before a critical time has elapsed, surface roughness in the simulations appears to increase with increasing developer concentration. If the surface roughness measurements are taken after the critical time, surface roughness in the simulations appears to be independent of developer concentration. One would expect the critical time to be less for a lower degree of blocking. Sidewall roughness may demonstrate the same behavior, and, if so, the model would explain that sidewall roughness has been observed to be independent of developer concentration[2] because the development time has exceeded the critical time for the exposed portion of the photoresist.

Recent AFM measurements by Reynolds and Taylor[4] suggest that roughness increases sharply at first and reaches a level roughness, but the final roughness appears greater for higher developer concentrations [Fig. 16(b)]. Whether these experiments show that the developer concentration increases the rate at which roughness develops remains uncertain.[4] Our model predicts that surface roughness, as a function of resist loss, is independent of developer concentration [Fig. 17(a)]. A plot of the surface roughness data of Reynolds and Taylor[4] versus resist loss also produces a single curve that is independent of developer concentration [Fig. 17(b)]. The results of Fig. 17(b) help to explain Fig. 16: the differences in final resist roughness seen in Fig. 16(b) coincide with varying degrees of resist loss. The dependence of surface roughness on developer concentration arises indirectly through the effect of developer concentration on resist loss.

IV. SUMMARY AND CONCLUSIONS

Photoresist development has been simulated as the molecule by molecule removal of polymer chains from a three-

dimensional cubic lattice with a critical-ionization fraction as the criterion for removal. Model predictions for the dissolution rate and surface roughness as a function of development time are presented. The model correctly predicts surface rate inhibition, as evidenced by an induction period, and this phenomenon is shown to correspond to a change in surface roughness (a measure of the area of photoresist exposed to developer). Surface roughness is generally inversely related to the dissolution rate. Polymers with a lower degree of polymerization, narrower polydispersity, and greater void fraction are shown to produce less surface roughness. The model predicts that, as the degree of blocking increases, surface roughness passes through a maximum, and experimental studies confirm this prediction. Recent AFM measurements have confirmed model predictions for the effect of development time on surface roughness. The model predicts that surface roughness develops faster with increasing developer concentration and eventually reaches a maximum that is independent of developer concentration. These model predictions are used to explain differences that have been observed in the responses of top-surface and sidewall roughness to changes in developer concentration.

ACKNOWLEDGMENTS

The authors thank Professor Chih-Kang Shih and Rachel Mahaffy of the University of Texas for assistance with the operation of their AFM. They are grateful to Dongxing He, Professor Francesco Cerrina, Geoffrey Reynolds, and Professor James Taylor of the University of Wisconsin for supplying them with the AFM data. They thank Intel Corporation and the Semiconductor Research Corporation (96-LP-409) for financial support.

[1]D. He and F. Cerrina, J. Vac. Sci. Technol. B **16**, 3748 (1998).
[2]G. Reynolds and J. Taylor, Proc. SPIE **3333**, 916 (1998).
[3]G. W. Reynolds and J. W. Taylor, J. Vac. Sci. Technol. B **17**, 334 (1999).
[4]G. W. Reynolds and J. W. Taylor (personal communication).
[5]T. Yoshimura, H. Shiraishi, J. Yamamoto, and S. Okazaki, Jpn. J. Appl. Phys., Part 1 **32**, 6065 (1993).
[6]R. Guerrieri and A. R. Neureuther, IEEE Trans. Comput.-Aided Des. **7**, 755 (1988).
[7]P. Trefonas III, Proc. SPIE **1086**, 484 (1989).
[8]E. W. Scheckler, S. Shukuri, and E. Takeda, Jpn. J. Appl. Phys., Part 1 **32**, 327 (1993).
[9]L. W. Flanagin, V. K. Singh, and C. G. Willson, J. Polym. Sci., Part B: Polym. Phys. **37**, 2103 (1999).
[10]P. C. Tsiartas, L. W. Flanagin, C. L. Henderson, W. D. Hinsberg, I. C. Sanchez, R. T. Bonnecaze, and C. G. Willson, Macromolecules **30**, 4656 (1997).
[11]A. R. Neureuther and C. G. Willson, J. Vac. Sci. Technol. B **6**, 167 (1988).
[12]For filling the lattice with polymer chains, one may use alternative methods that allow control of the chain length distribution and more accurately account for excluded volume effects. The authors are currently studying the application of this technique to the systems described in this article.
[13]K. Kremer and G. S. Grest, in *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, edited by K. Binder (Oxford University Press, New York, 1995), p. 199.
[14]C. L. McAdams, L. W. Flanagin, C. L. Henderson, A. R. Pawloski, P. Tsiartas, and C. G. Willson, Proc. SPIE **3333**, 1171 (1998).
[15]L. W. Flanagin, C. L. McAdams, W. D. Hinsberg, I. C. Sanchez, and C. G. Willson, Macromolecules (in press).
[16]S. Nonogaki, M. Hashimoto, T. Iwayanagi, and H. Shiraishi, Proc. SPIE **539**, 189 (1985).
[17]R. A. Arcus, Proc. SPIE **631**, 124 (1991).
[18]T. Long and F. Rodriguez, Proc. SPIE **1466**, 188 (1986).
[19]P. C. Tsiartas, L. L. Simpson, A. Qin, C. G. Willson, R. D. Allen, V. J. Krukonis, and P. M. Gallagher-Wetmore, Proc. SPIE **2438**, 261 (1995).
[20]G. G. Barclay, C. J. Hawker, H. Ito, A. Orellana, P. R. L. Malenfant, and R. F. Sinta, Proc. SPIE **2724**, 249 (1996).
[21]G. M. Wallraff, J. Opitz, G. Breyta, H. Ito, and B. Fuller, Proc. SPIE **2724**, 149 (1996).
[22]M. L. Williams, R. F. Landel, and J. D. Ferry, J. Am. Chem. Soc. **77**, 3701 (1955).

Relations between the solubility speed and the electrical conductivity of phenol novolak polymer solutions

T. Takeda[a)]
Electronic Device Materials Research Laboratory, Sumitomo Bakelite Company, Limited, Kiyohara, Utsunomiya-city 321-3231, Japan

M. Saka
Department of Mechanical Engineering, Tohoku University, Sendai 980-8579, Japan

(Received 4 February 1999; accepted 21 May 1999)

Relations between the solubility speed of phenol novolak resins which are typical photoresist base polymers, with the ac-electrical conductivity of their polymer solutions were studied. Strong correlations between these quantities were found. These relations were found to be affected by the concentration of the solution and the measured temperature. It is suggested to use such relations to predict the solubility speed of photosensitive novolak polymers by measuring their ac-electrical conductivity instead of measuring the solubility speed directly. © *1999 American Vacuum Society.* [S0734-211X(99)08004-X]

I. INTRODUCTION

In a previous article, a strong relationship was found between the development speed of polybenzoxazol (PBO) polymer and the ac-electrical conductivity of its polymer solution.[1] A calibration equation was proposed in another previous article to describe the influence of the temperature and the concentration of the solution.[2] If similar relations exist for typical photoresist polymers, then they might be used for in-line monitoring systems for controlling the development speed of the photoresist in alkaline developers during the manufacturing process of photoresists.

In the present article, relations between the solubility speed of phenol novolak resins, as typical photoresist polymers, and the electrical conductivity of their polymer solutions are studied. The possible application of such relations for predicting the development speed of photoresist polymers by measuring their electrical conductivity is discussed.

II. EXPERIMENTAL PROCEDURE

The cresol novolak resins were supplied by Sumitomo Durez Co., Ltd., as commercial photoresist base polymer products. Their chemical structure is shown in Fig. 1, and their polymer properties are listed in Table I. Since the novolak resins are synthesized by condensation polymerization, the control of their molecular weight distribution is considered to be difficult. In fact, the reaction temperature is kept constant within ±2 K, and the mol ratio of reactants is controlled precisely for that purpose. If industrial raw materials are used, the range of variation in molecular weight cannot be narrower than ±2000. For this reason, the solubility speed must be adjusted by composition of various photoresists having different molecular weight distributions. The molecular weight of the novolak resins was measured using gel permeation chromatography (GPC) as a polystyrene standard. The free cresol contents were measured by gas chromatography (GC). The solubility speed in an alkaline developer was directly measured using the same method as in the previous article.[1] A 2.38 wt % aqueous solution of tetramethyl ammonium hydroxide was used as alkaline developer. Since novolak resin itself has no functional groups, it cannot bring about the change in molecular weight even after pre-baking on wafers and thus the molecular weight of the surface part of the photoresist layer on the wafer is assumed to be the same as that of the bottom photoresist layer. The relative dielectric constant, ϵ_r, and the dissipation factor, $\tan\delta$, of the N-methyl-2-pyrrolidone (NMP) solution of novolak resins were measured by a low capacitance resonator (LCR) meter (type 4284A: Hewlett-Packard) with the electrode for liquid specimen (type LE-22: Andoh Electric) in the frequency range of 50 Hz–1 MHz. As for measurement temperature, ϵ_r and $\tan\delta$ were measured at 296, 313 and 333 K. Furthermore, they were measured at the concentration about 20.12 and 24.70 wt %. Conductivity σ was calculated from ϵ_r and $\tan\delta$ with the following equation:

$$\sigma = \omega \epsilon_0 \epsilon_r \tan\delta, \tag{1}$$

where $\omega = 2\pi f$ (f: frequency) and $\epsilon_0 = 8.86 \times 10^{-12}$ F/m which means the dielectric constant in vacuum.

They were measured at various frequencies as shown in Table II. An example of $\sigma/(\omega\epsilon_0)$ data obtained for 50 Hz is plotted in Fig. 2 instead of $\epsilon_r \times \tan\delta$ against development speed.

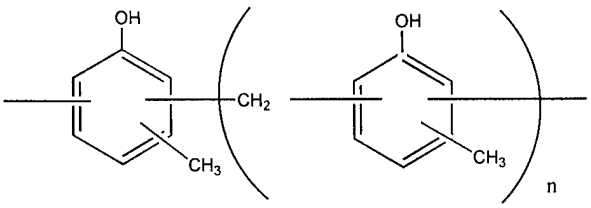


FIG. 1. Chemical structure of cresol novolak resin.

a)Electronic mail: ttakeda@sumibe.co.jp

TABLE I. Properties of novolak resins.

	Solubility speed (nm/s)	Molecular weight (GPC, Pst[a] standard)	Free cresol (%) (GC)
Sample A	20.33	12 563	0.4
Sample B	70.42	5 697	0.6
Sample C	34.72	8 967	0.7
Sample D	18.35	13 078	0.4

[a]Pst: polystyrene.

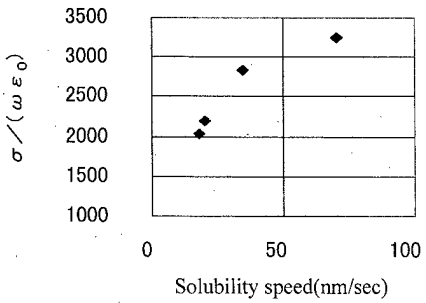


FIG. 2. Relationship between $\sigma/(\omega\epsilon_0)$ (50 Hz) and solubility speed of novolak resin solution.

III. RESULTS AND DISCUSSION

A. Relationship between solubility speed of novolak resins and electrical conductivity of their NMP solution

Figure 2 shows that there is also a certain correlation between the electrical conductivity measured at 50 Hz and the solubility speed of novolak resin solutions at almost the same measurement temperature and almost the same solution concentration. Therefore, there is a possibility that the solubility speed of phenol novolak resin can be predicted from the electrical conductivity of its solution. This is the first time to consider the relationship between the solubility speed of photoresist base polymer, novolak resin, in an alkaline developer and the electrical conductivity of its solution.

Figures 3 and 4 show the frequency dependence of the relative dielectric constant and the dissipation factor of a sample of novolak resin having a solubility speed of 20.33 nm/s. Figure 5 shows the relationship between $\ln(\epsilon_r \times \tan\delta)$ and $\ln(f)$. A linear relationship with negative slope in logarithmic plots is recognized between $\epsilon_r \times \tan\delta[=\sigma/(\omega\epsilon_0)]$ and the frequency. These tendencies are almost the same as the case of polybenzoxazol (PBO) base photoresist polymer described in the previous article.[1] The relationship between the solubility speed of novolak resin in the alkaline developer and electrical conductivity measured at 100 Hz–1 MHz is shown in Figs. 6–10. In these figures, a certain correlation is obviously recognized between the solubility speed and the electrical conductivity of the novolak base photoresist polymer, except in Figs. 9 and 10. These phenomena at high frequency were also observed in the case of PBO polymer. The reason is thought to be the difficult detection of polymer movement at high frequency measurement.

It is found that the higher electrical conductivity of a polymer solution, the greater its solubility in alkaline developer. This means that the higher the electrical conductivity generated by cresol novolak resin movement with electrical charge in organic solvent is, the more the solubility speed of the novolak resin base polymer in alkaline developer. It is suggested that if a base polymer has a bigger electrical charge and that if OH groups in the polymer are easier to dissociate, the acidity of the base polymer becomes strong and then the base polymer can easily be soluble in alkaline solutions. The structure of the novolak resin and the dissociation of OH groups are schematically shown in Fig. 11.[3,4] This mechanism for the relationship between solubility speed and electrical conductivity could be the same as the case of PBO discussed in the previous article.[1]

Electrical conductivity depends not only on the strength of the electrical charge of the polymer, but also on the ease of movement of the polymer molecule chain, even if it has the same electrical charge. In the case of a polymer that has the same electrical charge, molecular movement that is more active in a solution results indeed in higher electrical conductivity. It is well-known that the ease of molecular movement of the polymer chain in solution largely depends on its molecular weight, as already mentioned in the previous article.[1] The relationship between solubility speed in alkaline developer and molecular weight of novolak resin polymer is shown in Fig. 12 using the data of Table I. From this figure, it is obviously clear that the solubility speed in an alkaline developer decreases with increasing molecular weight of the novolak resin polymer. The higher the molecular weight of

TABLE II. Solubility speed and electrical conductivity of novolak resins.

Experimental No.	1	2	3	4	5	6	7	8
Solubility speed (nm/s)	20.33	20.33	20.33	70.42	34.72	18.35	18.35	18.35
Concentration (wt %)	20.12	24.7	29.82	30.27	29.93	30.03	30.03	30.03
Measurement temp. (K)	296	296	296	296	296	296	313	333
ϵ_r (50 Hz)	1 034.93	95.13	37.23	36.74	36.49	41.24	96.18	195.23
$\tan\delta$ (50 Hz)	42.5675	98.4627	59.09	88.4805	77.6477	49.2245	101.386	89.576
$\epsilon_r \times \tan\delta$ (50 Hz)	44 054.38	9366.76	2200.01	3250.77	2833.36	2030.02	9751.31	17 487.92
$\epsilon_r \times \tan\delta$ (100 Hz)	21 924.71	4624.44	1258.3	1598.85	1402.88	1074.93	4798.43	8 690.38
$\epsilon_r \times \tan\delta$ (1 kHz)	2 209.53	460.73	124.2	158.137	140.334	106.394	477.37	883.79
$\epsilon_r \times \tan\delta$ (10 kHz)	222.46	46.42	12.503	16.138	14.636	10.015	47.885	89.528
$\epsilon_r \times \tan\delta$ (100 kHz)	23.44	4.982	1.368	1.848	1.8081	1.992	4.98	9.225
$\epsilon_r \times \tan\delta$ (1 MHz)	2.804	0.8802	0.3954	0.6815	0.7204	0.7697	0.9501	1.412

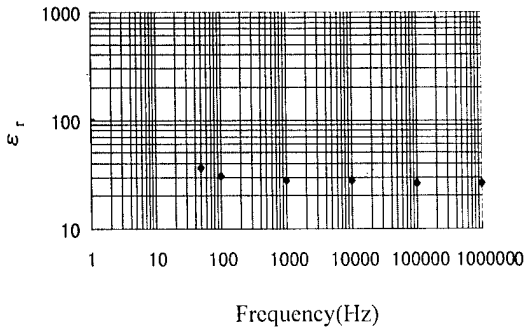


FIG. 3. Frequency dependence of ϵ_r of novolak resin solution.

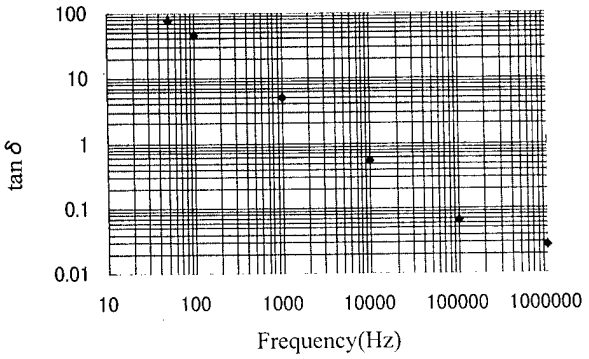


FIG. 4. Frequency dependence of tan δ of novolak resin solution.

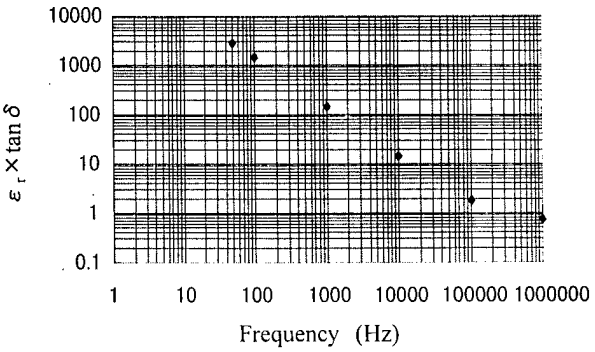


FIG. 5. Frequency dependence of $\epsilon_r \times \tan\delta$ of novolak resin solution.

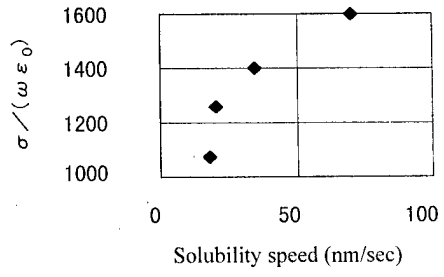


FIG. 6. Relationship between $\sigma/(\omega\epsilon_0)$ (100 Hz) and solubility speed of novolak resin solution.

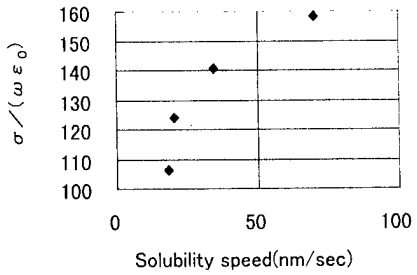


FIG. 7. Relationship between $\sigma/(\omega\epsilon_0)$ (1 kHz) and solubility speed of novolak resin solution.

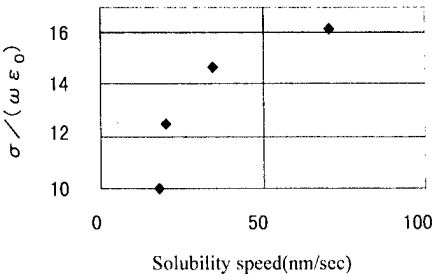


FIG. 8. Relationship between $\sigma/(\omega\epsilon_0)$ (10 kHz) and solubility speed of novolak resin solution.

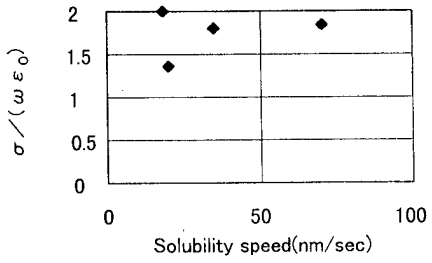


FIG. 9. Relationship between $\sigma/(\omega\epsilon_0)$ (100 kHz) and solubility speed of novolak resin solution.

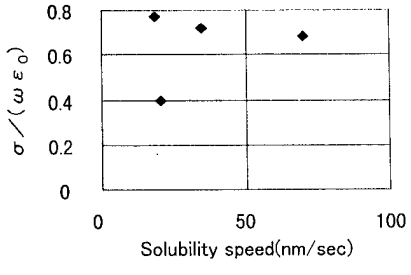


FIG. 10. Relationship between $\sigma/(\omega\epsilon_0)$ (1 MHz) and solubility speed of novolak resin solution.

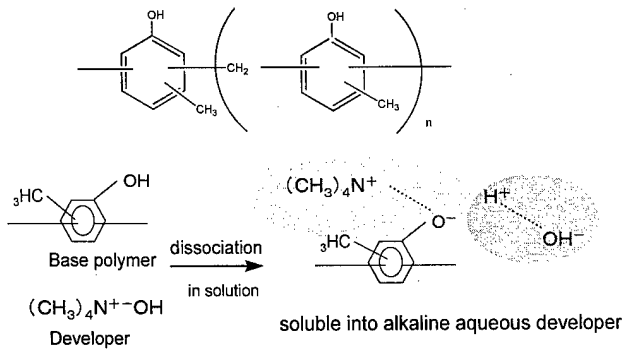


FIG. 11. Cresol novolak resin structure and its solution mechanism into alkaline developer.

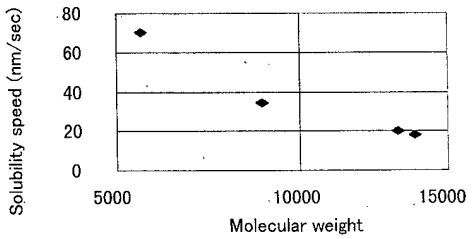


FIG. 12. Relationship between solubility speed and molecular weight of novolak resin.

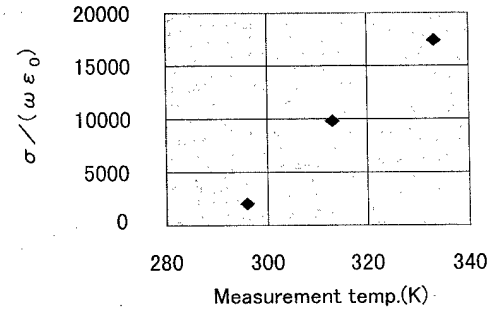


FIG. 13. Relationship between electrical conductivity of novolak resin solution and measurement temperature.

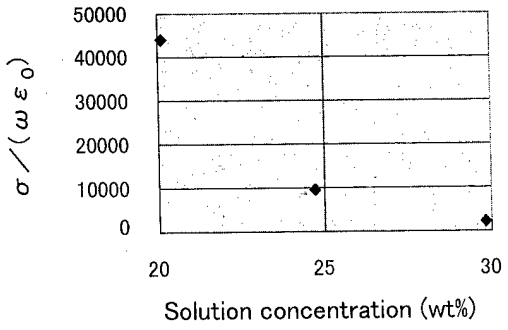


FIG. 14. Relationship between electrical conductivity of novolak resin solution and solution concentration.

the polymer, the more difficult it is for it to move fast in the solution.

B. Dependence of electrical conductivity of a novolak resin solution on temperature

In Fig. 13, the electrical conductivity of a novolak resin polymer is plotted over the sample temperature. It shows that the electrical conductivity of the novolak solution decreases with increasing temperature. This result is almost the same as that obtained for PBO base polymer solutions. It may be interpreted by the general hypothesis that the electrical conductivity increases because the molecular movement of the polymer in solution is activated by increasing temperature of the sample solution. Knowing that the electrical conductivity of a photosensitive base polymer depends on the temperature of the sample, it is evident that careful attention should be paid to make all measurements at the same temperature in order to avoid errors in the interpretation of the results of measurements of the electrical conductivity of polymer solutions.

C. Dependence of electrical conductivity of a novolak resin polymer solution on the concentration of the solution

In an earlier article[1] the electrical conductivity of PBO polymers was found to depend on the concentration of the solution. The electrical conductivity of three novolak polymer solutions with different concentrations (20–30 wt %) was measured. The results plotted in Fig. 14 show clearly that the conductivity depends on the concentration of the solution, i.e., it decreases with increasing concentration. This dependence is qualitatively the same as that found with PBO polymers. It is nonlinear, and stronger than in the case of PBO polymers. This difference is thought to be caused by the essentially different properties of OH groups in PBO and in novolak polymers. It is important to keep the temperature and the concentration of the solution constant in order to obtain reliable results of measurements of the relation between electrical conductivity and the solubility speed of photosensitive polymers in alkaline developers.

IV. CONCLUSIONS

It was found out that there is some correlation between the solubility speed of a photosensitive base polymer, novolak resin, in a developer and the electrical conductivity of an NMP solution, provided the concentration of the solution and the temperature are kept constant. Although the correlation appears to be nonlinear but rather parabolic, it appears possible to predict the solubility speed of a photosensitive polymer by measuring its electrical conductivity in the solution state. The correlation between solubility speed and conductivity is explained by the assumption that the solubility of a novolak resin in the developer depends more strongly on

the ease of its electrical charge movement in the solution state than in the case of PBO base polymers.

[1]T. Takeda and M. Saka, J. Vac. Sci. Technol. B **16**, 3055 (1998).
[2]T. Takeda and M. Saka, J. Vac. Sci. Technol. B **16**, 2974 (1998).
[3]T. Ushiroguchi, Y. Onishi, and A. Kumagae, J. Photopolym. Sci. Technol. **2**, 75 (1989).
[4]A. Furuta and M. Hanabata, J. Photopolym. Sci. Technol. **2**, 383 (1989).

Low-stress W/Cr films for SCALPEL® mask scattering layers

David L. Windt[a)]
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

(Received 28 January 1999; accepted 24 May 1999)

I describe the development of low-stress W/Cr bilayer films, for use as SCALPEL® mask scattering layers. These films are produced by dc magnetron sputtering in argon, and consist of 25–50-nm-thick W layers deposited onto 5–10-nm-thick Cr layers. X-ray reflectance analysis is used to measure the thicknesses of the individual W and Cr layers with subangstrom precision; surface and interface roughnesses, film densities, and also the thickness of the tungsten–oxide overlayer which forms after exposure to air are determined by this technique as well. Film stress, which is measured using the wafer curvature technique, is controlled by adjusting the deposition conditions such that the Cr layers are in tension while the W layers are in compression (and thus have high density and low surface roughness), so that the net stress in the bilayer is balanced near zero. I present data that illustrates how the net stress in these films varies with argon pressure, background pressure (i.e., partial pressure of residual gases present in the vacuum chamber), and Cr layer thickness. I also show how the stress depends on the composition of the substrate: i.e., stresses measured in films deposited onto Si wafers are systematically higher (by several hundred MPa) than the stresses measured for the same films deposited onto silicon–nitride-coated Si wafers. I discuss the implication of these results with regard to the production of high-quality SCALPEL® mask blanks for sub 0.12 μm lithography. © *1999 American Vacuum Society.* [S0734-211X(99)07604-0]

I. INTRODUCTION

The SCALPEL®[1] electron beam projection lithography system requires masks consisting of a patterned "scattering" layer, i.e., composed of material having a relatively large scattering cross section for 100 keV electrons (e.g., W) formed on a supporting membrane composed of material having a relatively small scattering cross section for electrons (e.g., silicon–nitride.) The supporting silicon–nitride membrane is in slight tension, in order to maintain flatness, and the stress in the W scattering layer must be minimized as much as possible in order to reduce pattern-placement errors. The polycrystalline W layers should ideally consist of small columnar grains, in order to minimize line-edge roughness, and should have a high density, in order to minimize voids, and also to minimize the film thickness (and thus the aspect ratio of patterned features) required for high electron contrast.

Previous efforts to produce low-stress W films, for use in masks for x-ray lithography,[2] for instance, attempted to exploit the widely known dependence of stress with sputter-gas (e.g., argon, typically) pressure observed in metal films prepared by sputtering.[3] However, this approach proved unsatisfactory, for the following reasons. First, the variation in stress with argon pressure is quite large near the zero stress point, and so it is difficult to control the pressure with sufficient accuracy. Second, the microstructure of the film also varies strongly with argon pressure,[4] essentially for the same reasons that the stress does: at low argon pressure, because of the large amount of kinetic energy (~100 eV/Ar atom) delivered to the surface of the growing film by neutral argon atoms reflected from the sputtering target, through the so-called "atomic peening" effect,[5] the film consists of tightly packed columnar grains, with high density and low surface roughness, and consequently has a large compressive stress; at higher pressure, the reflected neutral argon atoms suffer more collisions in the gas phase and thus deliver less kinetic energy to the surface of the growing film, which is thus more porous, consisting of loosely packed grains and a large number of voids, with low density and high surface roughness, and consequently has a large tensile stress. Because of this variation in stress and microstructure with argon pressure, the sputter-gas pressure for which the stress is close to zero results in films having a microstructure that is somewhere between the two extreme cases just described, which is undesirable with regard to the requirement of minimal line-edge roughness in the finely patterned features needed for sub 0.12 μm lithography.

The approach presented here to produce low-stress scattering layers for SCALPEL® masks also makes use of the variation in film stress with argon pressure just described, but with one crucial difference: rather than attempting to find the deposition conditions (i.e., argon pressure) that give rise to near-zero stress in the W layer, instead we attempt to balance the stress in the W layer with the stress in an underlying layer of Cr (which also serves as a dual purpose as an etch stop layer, and acts to mitigate membrane charging during e-beam exposure). This technique takes advantage of the fact that the "transition pressure," i.e., the pressure at which the stress in a sputtered film changes from compressive to tensile, also depends on the atomic mass of the adatom relative to the atomic mass of the working gas.[3] Specifically, in the case of W and Cr films sputtered in argon, this transition pressure is larger for W than it is for Cr, and thus there exists (in general) a range of argon pressures for which the W is in

[a)]Electronic mail: windt@bell-labs.com

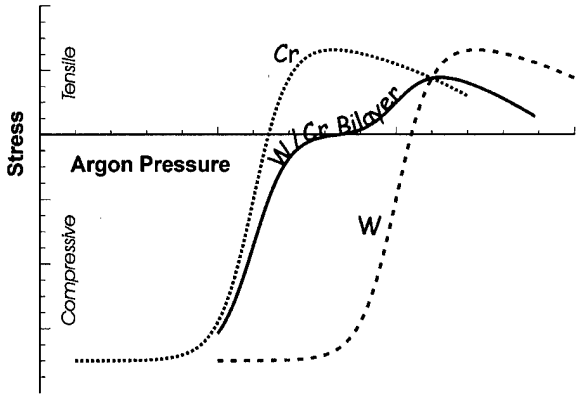


FIG. 1. Diagram showing conceptually how the stress varies with argon pressure in single-layer W and Cr layer films, and in a W/Cr bilayer film.

compression while the Cr is in tension; with a suitable choice of W and Cr layer thicknesses (consistent with the requirements on electron contrast[1]), the net stress in the bilayer can be made to balance. Although the bilayer stress still depends strongly on the argon pressure, the variation in stress with pressure is much less sensitive than in the case of a single-layer film, as illustrated schematically in Fig. 1. Furthermore, the W layer remains in compression and thus consists of the more desirable fine-grained, high-density, low-roughness material.

In Sec. II, I describe the preparation of such low-stress W/Cr bilayer films, and the characterization of these films using x-ray reflectance analysis for the precise determination of layer thicknesses, roughnesses and densities. In Sec. III, I present experimental results illustrating the dependence of film stress on argon pressure, Cr layer thickness, and background pressure (i.e., partial pressure of residual gases present in the vacuum system during deposition). I also present results showing that the stress in these films depends strongly on the composition of the substrate on which they are grown. Finally, in Sec. IV, I discuss the implications of these results with regard to the production of high-quality SCALPEL® mask blanks for sub μm lithography.

II. FILM PREPARATION AND CHARACTERIZATION

The films described here are grown by dc magnetron sputtering in argon of 99.999% purity, using a deposition system having sub-angstron film thickness control for the production of multilayer x-ray optics that has been described previously.[6] The system is cryo-pumped, and the argon pressure is maintained with a closed-loop gas-flow system using a capacitance manometer and a mass-flow controller. In all cases, the power to each of the two 50-cm-long× 9-cm-wide planar magnetrons, i.e., one for W (99.95% purity), one for Cr (99.99% purity), is fixed at 200 W, and the individual film thicknesses are adjusted by varying the rotational velocity of the substrate (which faces downward) as it travels over the sources (which face upward, 10 cm below the plane of the substrate). The deposition rates are typically of order 0.18 nm/s for W and 0.24 nm/s for Cr (with a slight dependence on argon pressure), and are computed from the film thicknesses determined by x-ray reflectance analysis, described below. Except where noted, the samples described here were grown on (unheated) 14 mm×14 mm, 100-μm-thick Si wafer substrates having a thin (~2–3 nm) native oxide layer.

X-ray reflectance measurements are made as a function of grazing incidence angle at a fixed wavelength, using a four-circle diffractometer with a rotating anode x-ray source having a Cu target, and a pyrolytic graphite monochromator tuned to the CuKα line near 8 keV (1.54 Å). Reflectance measurements are typically made for incidence angles in the range $0° < \theta < 4°$, which in most cases discussed here corresponds to a span of roughly seven orders of magnitude in reflected intensity. The angular resolution of the diffractometer is ~0.02°, and measurements are typically made every 0.01°, conditions that are just sufficient to resolve the thickness fringes for the thickest films (~50 nm W/10 nm Cr) discussed here.

Fits to the x-ray reflectance data, performed with the IMD software package,[7] are used to determine film thicknesses, densities and roughnesses, for both the W and Cr layers, as well as for the tungsten–oxide overlayer (assumed to be WO_3) that apparently forms on these films after exposure to air. With this technique, the measured reflectance versus incidence angle data are compared with a theoretical reflectance curve computed using an algorithm based on recursive application of the Fresnel equations; the formalism described by Stearns[8] is used to account for the effects of interfacial roughness (or diffuseness). Nonlinear, least-squares curve fitting can be used to fit the data, but in many cases simply taking advantage of IMD's ability to "manually" vary several adjustable parameters simultaneously is a more efficient approach, particularly for the highly oscillatory reflectance data typical of these films. The effect on the resultant reflectance curve is significantly different for each of the adjustable parameters (i.e., film thicknesses, roughnesses, and densities) used to fit the data, and so each of these parameters can be determined uniquely, in general. The sensitivity of the x-ray reflectance analysis technique to layer thicknesses and roughnesses is particularly high—typically these parameters can be determined with a precision of better than 0.1 nm—but the sensitivity to film density is much worse: the precision with film density can be determined is generally no better than about 10%.

To illustrate, shown in Fig. 2 is the measured x-ray reflectance for a W(~20 nm)/Cr(~4.5 nm) bilayer film. The best-fit curve (labeled "σ=3.5 Å," i.e., corresponding to a WO_3/W interface roughness of 3.5 Å) was obtained by fitting eight adjustable parameters: the densities, layer thicknesses, and interface roughnesses of both the W and Cr layers, as well as the thickness and roughness of the WO_3 overlayer; the best-fit parameters are indicated. Also shown are curves computed using other WO_3/W interface roughness values, illustrating the strong effect of this particular parameter on the computed reflectance. The high-frequency and low-frequency modulations in the reflectance data (having periods of ~0.18° and ~1.0°) correspond to interfer-

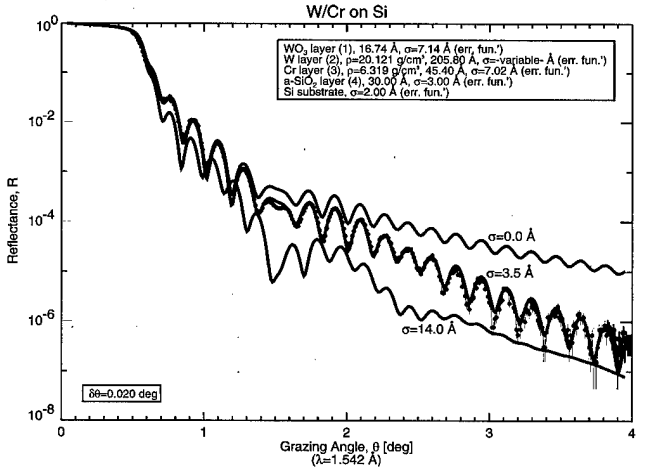


FIG. 2. Typical x-ray reflectance-vs-incidence-angle data for a W/Cr bilayer film. Calculated reflectance curves for three values of the WO_3/W interfacial roughness are shown; the best-fit curve (labeled "σ=3.5 Å") to the measured data (filled circles) was obtained using nonlinear, least-squares curve fitting, with eight adjustable parameters as indicated.

ence due to the total film thickness (i.e., WO_3 + W + Cr layer thicknesses) and the Cr layer thickness, respectively. But note that these thickness modulations shift measurably with WO_3/W interface roughness; in general, although each of the fit parameters can ultimately be determined uniquely, the fact that many parameters are often coupled adds to the complexity of finding the best-fit parameters.

III. FILM STRESS

Film stress is measured using the wafer curvature technique, with an instrument that has been described previously.[9] With this technique, the net film stress is determined using Stoney's equation[10] from the measured change in radius of curvature of the thin substrate; the radius of curvature of the substrate is determined by measuring precisely the deflection of a laser beam as it is scanned along the length of the sample. The experimental uncertainty in the measured film stress is larger on the small wafer sections used here (relative to measurements on full-sized wafers) as a result of the larger relative uncertainty in the substrate thickness, and also because of the larger uncertainty in the determination of the change in radius of curvature of these samples associated with the smaller scan lengths. In all cases presented below, scans were made at several positions on the substrate, resulting in scatter in the data between different measurements for a given sample.

Shown in Fig. 3 is the stress in sputtered W(50 nm)/Cr(25 nm) films as a function of argon pressure, in the range P_{Ar}=2.1–2.8 mTorr. In this case, the background pressure in the vacuum chamber prior to film deposition was $5.0\pm0.1\times10^{-6}$ Torr for all samples. The variation in stress with pressure is approximately linear, ranging from ~200 MPa compressive for $P_{Ar}=2.10\pm0.01$ mTorr, to ~350 MPa tensile for $P_{Ar}=2.80\pm0.01$ mTorr, and the films having the lowest stress were grown at an argon pressure of 2.35 ±0.01 mTorr. The scatter in the stress data for multiple measurements made on multiple samples is roughly ±40 MPa. We note that, although the sputter target voltages can vary with argon pressure, in principle, and the voltages can in turn affect the amount of energy delivered by argon ions to the surface of the growing film (which can thus affect the film stress and microstructure), in this case the target voltages varied by less than ~5 out of ~340 V (i.e., ~1.5%) over this range of argon pressures; given the large amount of energy delivered to the surface of the growing film by reflected neutral argon atoms (i.e., ~100 eV/atom), the variation with argon pressure in energy delivered by argon ions is thus negligible.

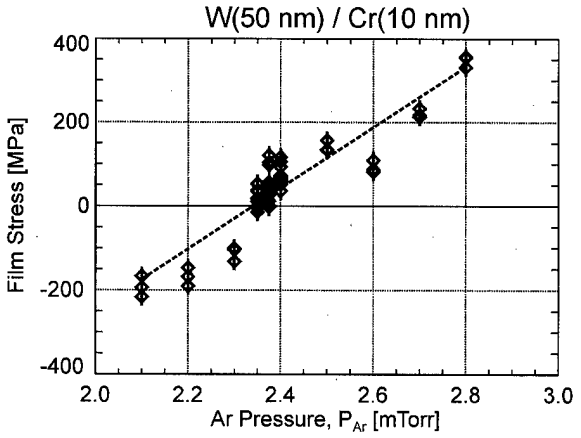


FIG. 3. Measured stress in W(50 nm)/Cr(25 nm) films as a function of argon pressure. The dashed line is a linear fit to the data.

In a previous investigation,[11] the stress in Mo/Si x-ray multilayer films were found to depend strongly on the background pressure of the vacuum system. We find a similar result for W/Cr bilayer films, as indicated in Fig. 4, which shows the stresses measured as a function of background pressure for W(50 nm)/Cr(10 nm) films deposited at an argon pressure of $P_{Ar}=2.40\pm0.01$ mTorr. The background pressure was measured with an ionization gauge just prior to film deposition, and was adjusted by varying the pumpdown time of the vacuum system: the pumpdown time ranged from ~90 min for samples grown at $P_{background}=5.0\times10^{-6}$ Torr,

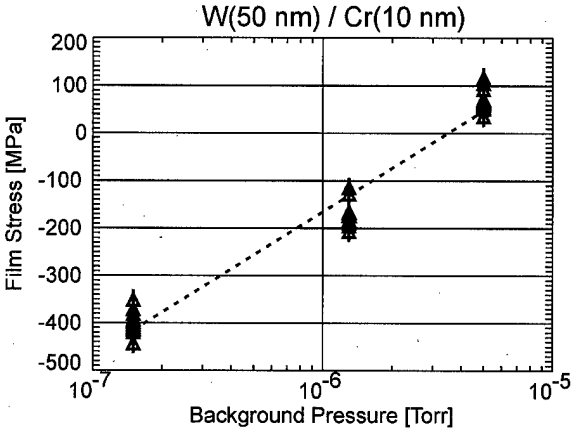


FIG. 4. Measured stress in W(50 nm)/Cr(25 nm) films are a function of background pressure. The argon pressure was fixed at 2.4 mTorr in this case. The dashed line is a linear fit to the data.

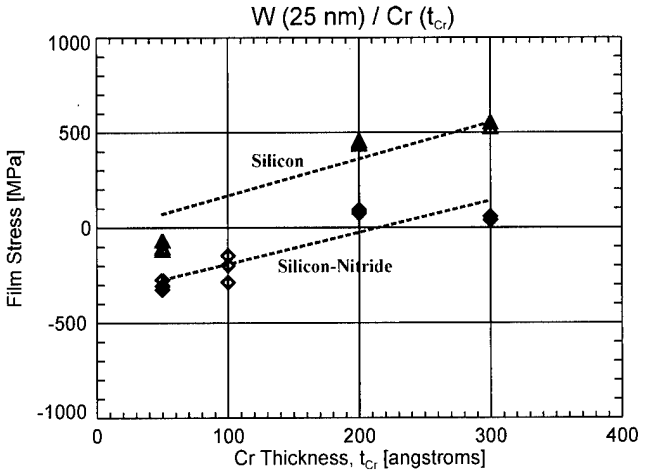


FIG. 5. Measured stress in W/Cr films containing 25-nm-thick W layers, as a function of Cr layer thickness. Stresses are shown for films deposited onto both bare and silicon–nitride-coated silicon substrates, as labeled. The dashed lines are linear fits to the data.

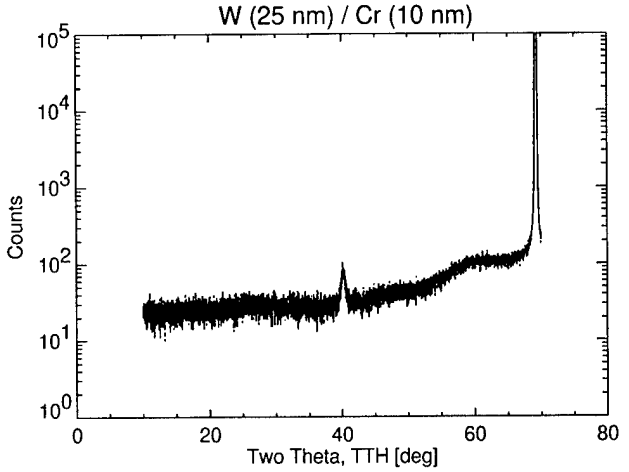


FIG. 6. Large-angle x-ray diffraction data (obtained in the θ-2θ geometry) for W(25 nm)/Cr(10 nm) films grown on both bare (solid lines) and silicon–nitride-coated (dashed lines) silicon substrates. The curves are nearly indistinguishable. The peaks near 2θ=40.5° are due to diffraction from the bcc W (110) lattice planes, while the sharp peaks near 2θ=69° are due to diffraction from the silicon substrate. The broad peaks near 2θ=60° are due to diffraction from the (amorphous) native silicon–oxide layer present on the silicon substrates prior to coating.

to ~24 h for samples grown at $P_{\text{background}}=1.5\times10^{-7}$ Torr. As in the case of the Mo/Si multilayer films reported previously, the stress is highly compressive in W/Cr bilayer films prepared at low background pressure (~−400 MPa), and is tensile (~+60 MPa) for those grown at high background pressure. In the case of Mo/Si multilayers, x-ray diffraction measurements revealed no observable microstructural changes, but forward recoil scattering measurements indicated that the concentration of hydrogen in films varied from ~0.5 at. % for films grown at low background pressure to ~2.0 at. % for those grown at high background pressure; the variation in stress with background pressure was thus attributed to this measured variation in hydrogen concentration, although the precise mechanisms responsible for the stress variations were not identified conclusively. In the case of the W/Cr films shown in Fig. 4, x-ray diffraction also reveals no microstructural variations. Forward recoil scattering measurements were not performed on any of the W/Cr samples discussed here, but it is possible that a similar systematic variation in hydrogen concentration might be found. As in the case of the stress-versus- argon-pressure data shown in Fig. 3, the small (~3%) variation in target voltage with background pressure for the data in Fig. 4 is unlikely to have a significant affect on the bilayer stress.

The variation in stress in W/Cr bilayer films containing 25-nm-thick W layers as a function of Cr layer thickness is shown in Fig. 5, for films deposited onto Si wafer substrates. The stress increases by several hundred MPa as the Cr layer thickness is increased from 5 to 30 nm. Because the net stress in the bilayer depends on the stresses in W and Cr layers, weighted by the respective W and Cr layer thicknesses, this result clearly indicates that the Cr layers are in tension, as thicker Cr layers result in more tensile films. Also shown in Fig. 5 are the stresses measured in identical films (i.e., grown during the same deposition cycle) but deposited onto 100-μm-thick Si wafers that were coated [by chemical vapor deposition (CVD)] with ~120 nm of silicon–nitride. The stresses in the films grown on silicon–nitride are systematically lower—by several hundred MPa—than those grown on Si. The difference in the stresses for films grown on silicon versus silicon–nitride can be due to differences in the interfacial stresses present at the Cr-substrate interface, or possibly due to differences in the microstructure of the two films.[12] However, large-angle x-ray diffraction measurements made on these films reveal no microstructural differences, as shown in Fig. 6, for example, suggesting that the difference in stress is predominantly due to interfacial stress differences.

IV. CONCLUSIONS

I have described an approach to produce low-stress W/Cr bilayer films suitable for use as scattering layers in SCALPEL® masks for sub 0.12 μm lithography. The films are grown by magnetron sputtering in argon: the deposition conditions are adjusted so that the W layers are in compression and the Cr layers in tension, so that for a given set of layer thicknesses (as driven by the requirements on electron contrast), the net stress in the bilayer is minimized by balancing the stresses in the two metal layers, and yet the compressive W layer is composed of the more desirable fine-grained, high-density, low-roughness material. X-ray reflectance analysis can be used to determine with high precision layer thicknesses, roughnesses and densities.

I have shown that the stress in these films depends on the argon pressure, the background pressure, and the Cr layer thickness, and also depends on the composition of the substrate—stresses measured in W/Cr bilayers deposited onto Si wafers are systematically higher (by several hundred MPa) than the stresses measured for the same films deposited onto silicon–nitride-coated wafers; these results have the following implications for the production of high-quality

SCALPEL® mask blanks. First, to produce films with consistently low stress, it is clearly necessary to control precisely both the argon pressure and the background pressure of the vacuum system. Other deposition conditions that are known to affect films stress, such as magnetron power, source-to-substrate distance, etc., will almost certainly need to be controlled with high precision as well. Second, if such films are to be grown on silicon–nitride-coated substrates, then one must take into account the differences in the composition (and undoubtedly the cleanliness as well) of thin-wafer substrates (suitable for wafer-curvature measurements) that might be used to calibrate the dependence of stress with deposition conditions.

ACKNOWLEDGMENTS

I would like to acknowledge G. Weber for providing the silicon–nitride coatings used in this study, W. Brown for a careful review of the manuscript, and the encouragement of the entire Bell Labs SCALPEL® team.

[1]S. D. Berger, and J. M. Gibson, Appl. Phys. Lett. **57**, 153 (1990); see also <http://www.lucent.com/SCALPEL> for a complete bibliography of articles describing SCALPEL technology.
[2]R. R. Kola, G. K. Celler, J. Frackoviak, C. W. Jurgenson, and L. E. Trimble, J. Vac. Sci. Technol. B **9**, 3301 (1991).
[3]H. Windischmann, J. Vac. Sci. Technol. A **9**, 2431 (1991).
[4]R. Messier, A. P. Giri, and R. A. Roy, J. Vac. Sci. Technol. A **2**, 500 (1984).
[5]F. M. D'Heurle, Metall. Trans. **1**, 725 (1970).
[6]D. L. Windt and W. K. Waskiewicz, J. Vac. Sci. Technol. B **12**, 3826 (1994).
[7]D. L. Windt, Comput. Phys. **12**, 360 (1998).
[8]D. G. Stearns, J. Appl. Phys. **65**, 491 (1989).
[9]C. A. Volkert, J. Appl. Phys. **70**, 3251 (1991).
[10]G. G. Stoney, Proc. R. Soc. London, Ser. A **82**, 172 (1909).
[11]D. L. Windt, W. L. Brown, C. A. Volkert, and W. K. Waskiewicz, J. Appl. Phys. **78**, 2423 (1995).
[12]The stress difference could also be due, in principle, to the small difference in the coefficients of thermal expansion between silicon and silicon–nitride; but to explain the measured stress difference of several hundred MPa measured here would require a substrate termperature of roughly 5000 °C during deposition, which is completely unreasonable.

Microelectron gun integrating a point-source cathode

Zoubida Hammadi, Michel Gauch, and Roger Morin[a)]
Centre de Recherche sur les Mécanismes de la Croissance Cristalline, CRMC2–CNRS Campus de Luminy, F-13288 Marseille Cedex 09, France

(Received 20 November 1998; accepted 9 April 1999)

A structure integrating a sharp field emission tip inside a coaxial structure with an overall diameter as small as 60 μm is described. It can emit nA electron current with a minimum kinetic energy of 50 eV. It is demonstrated that this structure behaves like an electron gun and is able to produce a low-energy, divergent and highly coherent electron beam at distances as small as 100 μm from the tip. © *1999 American Vacuum Society.* [S0734-211X(99)00604-6]

I. INTRODUCTION

In many systems using electron beams, electron source characteristics govern the ultimate performance of the whole system. However, these characteristics are often described by too many parameters. In imaging instruments as electron microscopes or electron diffractors, the most important parameters are the energy distribution of the electrons and the brightness of the sources. Brightness hides three independent parameters which are the intensity and the opening angle of the beam and the source size. The importance of these parameters comes from the fact that energy governs the electron interaction strength with materials and that other parameters mainly dictate the ultimate performance of the electron optical system.

In instruments operating at high energy (>1 keV), it is possible to first proceed to electron extraction and thereafter to produce a bright enough electron source for most applications. This separation of the emission part from the other parts of a system finds its origin in the high performance of high energy electron optics. This also has the important consequence of being able to independently monitor intensity and energy. The situation is quite different at low energy (<100 eV) where electron optics of high quality are not available. In this energy range, up until recently only simple surface diffractors existed. These systems enable independent current–energy monitoring but do not enable imaging because of the macroscopic dimension of the source which cannot be demagnified by any appropriate electron optics. Recently point projection microscopy enabled lensless imaging by the use of an electron point source made of an ultrasharp field emission tip.[1] But, with its use the intensity–energy independence is lost. This comes from the diode structure of the system where the sample plays the role of an anode used to extract electrons from the tip.

The aim of this article is to describe a more involved structure where the very localized character of electron emission is preserved and where the intensity and energy are independent. We first describe the structure and its fabrication. Afterward, we show electrical measurements that allow one to estimate the electrostatic influence coefficients for different geometries. Comparisons with computer simulations are also given. Finally, we show by electron optics experiments direct proof of the localized character of the electron emission.

II. EXPERIMENTAL AND METHODS

A. Fabrication of the microelectron gun

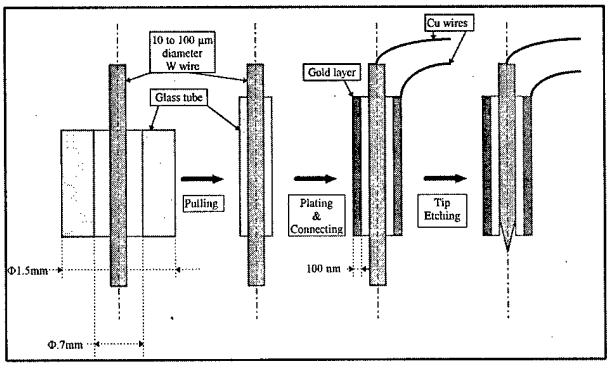
The structure we built (Fig. 1) is made of a sharp metallic tip (the cathode) just emerging from the end of a glass capillary. The outside wall of the capillary is plated with a conductive layer (the sheath) and electric connections are made on the tip and on the sheath. The aim is to get electron field emission from the very end of the tip if it is negatively biased relative to the sheath. Ideally, we are looking for a device where this emission current is independent of the potential and of the position of a third electrode placed in the vicinity of this structure.

Fabrication of the structure starts by pulling a tube of glass. The internal and external diameters of the tube are, respectively, 0.7 and 1.5 mm. This pulling procedure is performed using a small torch (with a millimeter size flame). The pulling speed is controlled by small charges fixed on the lowest part of the tube. Prior to pulling the tube a tungsten wire is introduced inside the tube. The wire diameters used were 100, 25, and 10 μm. With 25 μm diam wires, this leads to structures with an outside diameter of 60 μm.

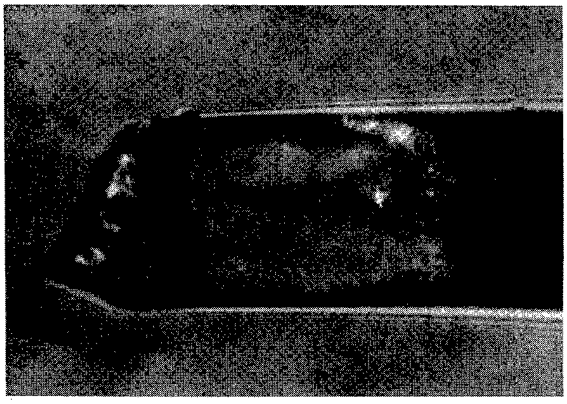
Once this is done, a number of problems has to be solved.

(a) The first one is to smooth the end of the glass tube after cutting it. This is achieved by approaching a hot wire, heated by the Joule effect under binocular optical control and allows one to cure the end of the tube from the cracks, a consequence of capillary break.
(b) The second one is to fabricate the metallic sheath. We make it by vacuum depositing a 100 nm thick layer of gold.
(c) Then the tip and the sheath are electrically connected using dots of silver glue.
(d) Finally, the tip is fabricated by electrolytic etching in an appropriate solution.[2] Presently 25 μm diam wires are the smallest wires in which we have been able to etch sharp enough tips.

[a)]Author to whom correspondence should be addressed.



(a)



(b)

FIG. 1. (a) Different main steps in fabricating a microelectron gun. (b) Optical micrograph of a microelectron gun. The tip is visible through the glass. The diameter of the capillary is about 60 μm.

Note that at each step we take advantage of the fact that the glass capillary does not wet the tungsten wire and therefore it is possible to position the end of the wire relative to the exit plane of the capillary using a dedicated holder and binocular optical control. In this way, the wire is placed inside the capillary in steps (a) and (b) whereas it is placed outside of the capillary in step (d). After step (d), the tip position is adjusted relative to the exit plane of the capillary. Although we succeeded fabricating smaller structures (20–25 μm outside diam capillary on 10 μm wires), the whole process is only successful for devices down to 60 μm for the overall structure diameter using 25 μm diam wires.

B. Characterization of the microgun

After its fabrication, the microelectron gun fixed on its holder is transferred into an ultrahigh vacuum chamber (Fig. 2). This chamber contains a piezoelectric actuated micromanipulator which enables one to position a thin carbon membrane in front of the gun at any distance from about 10 mm down to submicrometric distances. This carbon membrane is supported by a molybdenum grid and exhibits a lace structure ranging from a 100 μm scale down to a nanometric scale. Placed 15 cm from the gun, on the membrane side, is a channel plate-screen assembly which amplifies and makes visible the distribution of the electron intensity after its transmission through the membrane.

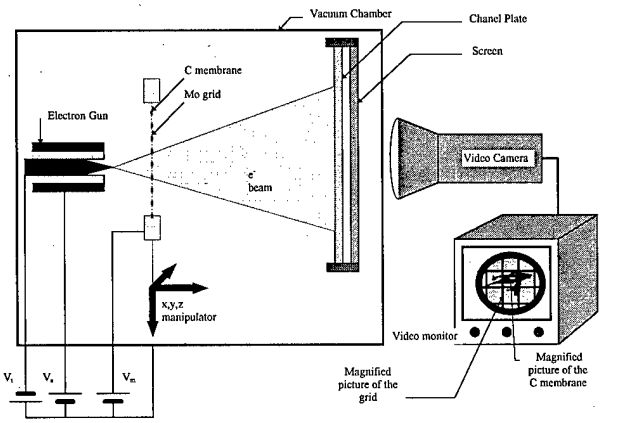


FIG. 2. Experimental arrangement used to characterize the microelectron gun.

This arrangement (which is the one used for projection microscopy[1]) enables two kinds of measurements that are quite difficult to perform by other means.

The first is the measurement of the tip-to-membrane distance, done with great precision. This is made possible using the following procedure. We first look at the size of the projection of an object with a known size, a mesh, for instance [Fig. 3(a)]. This gives us the magnification ratio of the projector. The tip-to-membrane distance is, therefore, the tip-to-channel plate distance divided by the magnification. On the same image a smaller object is selected afterward. Hence its real dimension is known. This object is then approached toward the tip in order to increase the magnification. When its projection is comparable to the screen size [Fig. 3(b)] the new magnification is determined and hence the new tip-to-object distance. This procedure is applied successively to smaller and smaller objects and enables one to measure shorter and shorter distances down to a submicrometer scale.

As long as the size of the holes in the membrane stays small in comparison with the tip-to-membrane distance, an equipotential plane corresponding to the membrane plane can be assumed. The variations of the electron current as a function of the voltages of the different electrodes are recorded and so the electrical parameter of the gun can be determined.

The second parameter of this device is its dimensions. This parameter is important because the size of the sheath limits the sample-to-tip distance without affecting the emission current and the electron trajectories outside the sheath (in the present case the carbon membrane acts as a sample). The smaller this distance is, the higher the resolution of a scanning instrument or the magnification of a projector which takes direct advantage of the small size of the source. The dimensions of the gun are measured by optical microscopy or scanning electron microscopy.

The last important parameter is the virtual size of the source which controls the ultimate coherence of the electron beam. This parameter measures the ability of the electron gun to be used in a coherent optical mode. An estimate of this parameter can be obtained with the projector setup by observing the sharpness of the geometrical projection of the

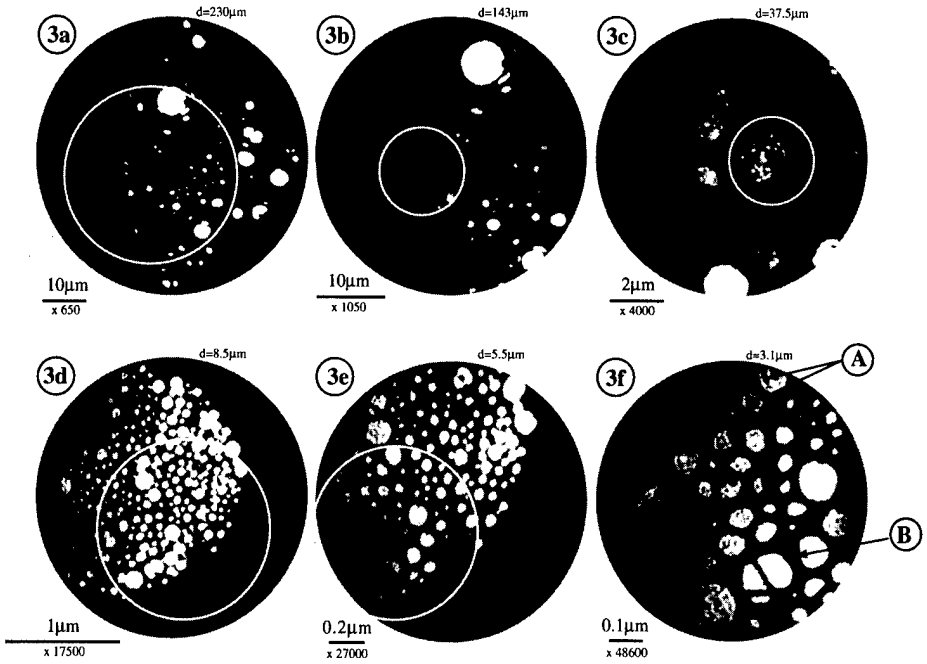


FIG. 3. (a)–(f) Projection images of a carbon membrane for increasing magnifications. The actual diameter of the screen is 4 cm. In (a), the image corresponds to a single square mesh with a side length of 45 μm which enables one to determine the initial magnification. The area seen in (b) is shown by a white circle in (a), the area seen in (c) is shown by a white circle in (b), and so on. The scale and magnification are indicated on each image as is the distance d between the tip and the membrane. (f) Edge fringes (A) and the interference pattern produced by a filament (B) illustrating the coherence of the beam.

membrane and for a smaller source size the interference pattern produced by a small object or an edge.

III. RESULTS AND DISCUSSION

A. Electrical measurements

Our main goal is to determine the domain where the structure produces electrons with low kinetic energy on an electrode which is placed in front of the tip. According to electrostatics, the electric field F at the apex of the tip is a linear combination of the voltages V_t, V_s, and V_m applied to the different electrodes, the tip, sheath, and membrane, respectively,

$$F = C_{tt}V_t + C_{ts}V_s + C_{tm}V_m . \quad (1)$$

Because the field emission current I is essentially determined by the field F (Fowler–Nordheim law),[3] plots of V_t vs V_s for constant I and V_m are linear (note that the electrical insulation provided by the glass capillary is leakage free for the voltages used in these experiments). This is illustrated in Fig. 4 where the different plots correspond to different distances d between the tip and the membrane. Obviously the absolute values of the voltages depend on many microscopic geometrical parameters describing the tip, like the cone angle and the tip apex radius. A number of experiments have been carried out on structures of different sizes. We report measurements on the same tip but when it was placed at two different positions relative to the sheath. We chose to define this position by the parameter Q (Fig. 5) which is the ratio of the distance l the tip gets from the sheath to the radius of the sheath. This experiment is made possible thanks to the tip sliding inside the capillary and it enables one to estimate the importance of this macroscopic positioning.

In order to simplify the analysis, the slope S of these plots versus distance d is shown in Fig. 6.

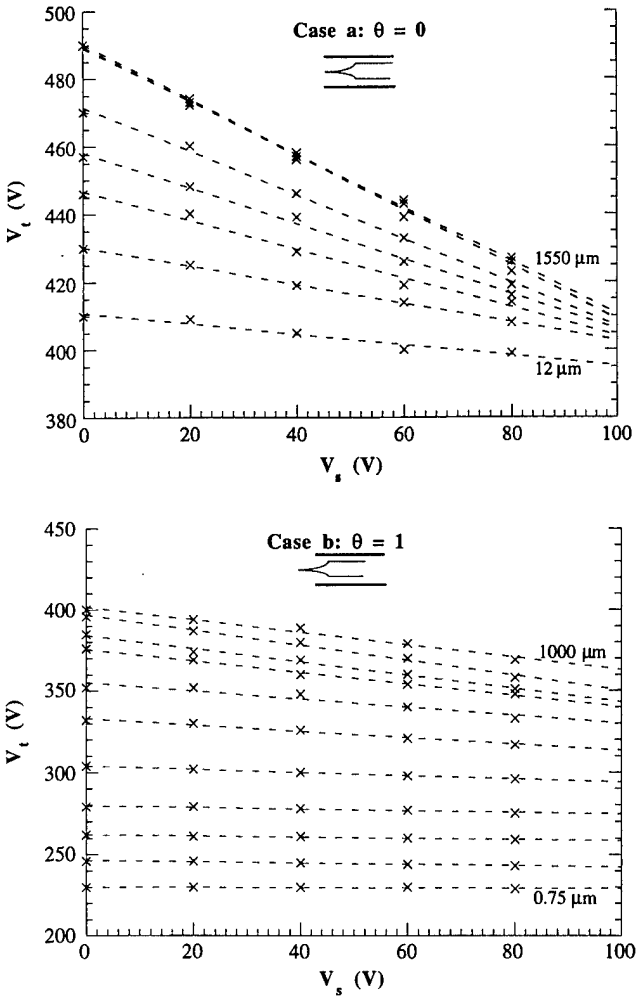


FIG. 4. Variation of the tip voltage V_t vs the sheath voltage V_s for two different positions of the tip relative to the sheath. (a) $Q=0$ and (b) $Q=1$.

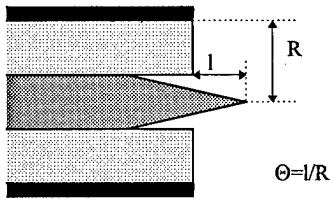


FIG. 5. Schematic showing the macroscopic positioning of the tip relative to the sheath.

From Eq. (1),

$$S = -V_t/V_s = C_{ts}/C_{tt}.$$

This means that S measures the relative influence of the sheath. A value of S close to l characterizes a perfect gun where the emission current is only determined by the voltage difference between the tip and the sheath.

As expected, Fig. 6 shows clearly that varying d from small to large values relative to the radius of the sheath increases the influence of the sheath. Note, however, that the value of d corresponding to the inflexion point does not depend strongly on $Q=0$ or 1 and is roughly equal to the sheath radius, i.e., 40 μm. What seems, however, more dependent on the tip-to-sheath position is the absolute value of the slope S for large tip-to-membrane distances, increasing from about 0.5 to 0.8 when Q changes from 1 to 0.

B. Computer simulations

In order to check these results, we performed a computer simulation[4] of the electric field in such a structure. In this simulation the tip is described (Fig. 7) by two half spheres connected by a cone. One sphere has the diameter of the wire (25 μm) and the length of the cone is determined by the cone angle which is seen by optical microscopy (5° half angle in the present case). The sheath is described by a cylinder with a diameter equal to the outside diameter of the capillary (80 μm). The membrane is described by a 3 mm diam metallic disk, which is the diameter of the grid supporting the membrane. A field of 10^8 V/cm is assumed and the simulation leads, after some trials, to a tip apex diameter of about 40 nm. This is the sole parameter of the simulation. In Fig. 8 are the results of the computer simulation and the experimental results concerning the variation of the influence coefficient C_{tt} vs d with the sheath and the membrane to ground potential. We observe quite good agreement between experiment and simulation. More specifically, note that C_{tt} stays independent of distance d as long as d is larger than the sheath radius (40 μm).

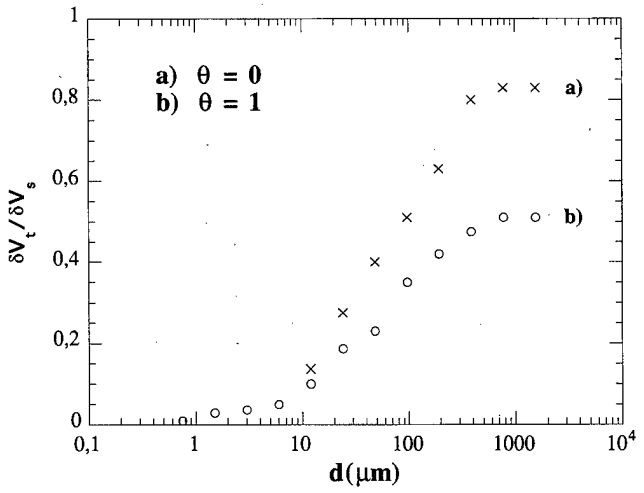


FIG. 6. Relative influence of DV_p/DV_s for two different geometries. (a) $Q=0$ and (b) $Q=1$.

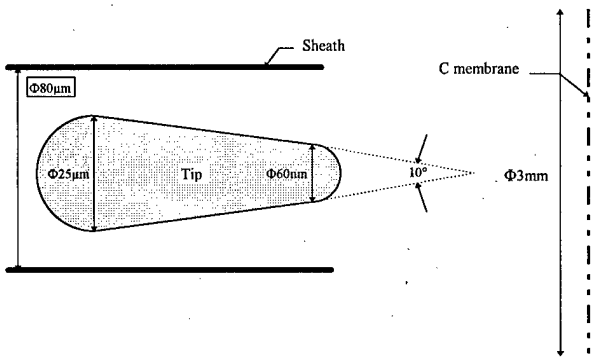


FIG. 7. Description of the shape of the electrodes used in the simulation. The system has cylindrical symmetry around the tip axis.

Simulations of S vs d are very time consuming because the software does not allow one to make a direct computer experiment which consists of fixing the sheath voltage V_s and the field F at the tip apex and computing the corresponding tip voltage V_t. So, we use Eq. (1) and the linear dependence of F with V_t and V_s to study the influence of the tip position relative to the sheath (i.e., Q) with the membrane placed at a large distance (1 mm). The results are shown in Fig. 9. Note that the experimental results are in excellent agreement with this simulation. Note too that placing the tip one sheath radius inside the capillary leads to an almost 100% dependence of the emission current on the tip-to-sheath voltage.

C. Source size

As previously explained in Sec. II B, the projection setup, by the obtainment of sharp magnified pictures of the carbon membrane, shows the small size of the source. This is illustrated in Figs. 3(a)–3(e) for moderate magnification. Increasing the magnification shows the first interference fringes on some parts of the membrane [Fig. 3(f)]. It seems that what prevents larger magnification of these fringes is not the size of the source itself but only mechanical vibrations, which are mainly related to the length of the device (8 mm). We think this can be significantly improved using a better arrangement. Taking into account the magnification (48 600×) of the projector, a fringe spacing of 0.4 mm on the screen can only be seen if the source size is smaller than 0.4/48 600, i.e., 8 nm. This clearly proves that this device does not only deliver a low-energy electron beam it delivers also a highly coherent beam.

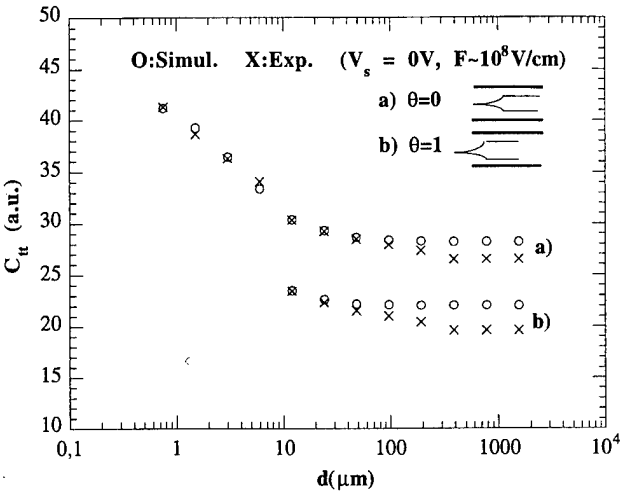


FIG. 8. Comparison between experiment and theory (a) $Q=0$ and (b) $Q=1$. The parameters used for the simulations are those in Fig. 7.

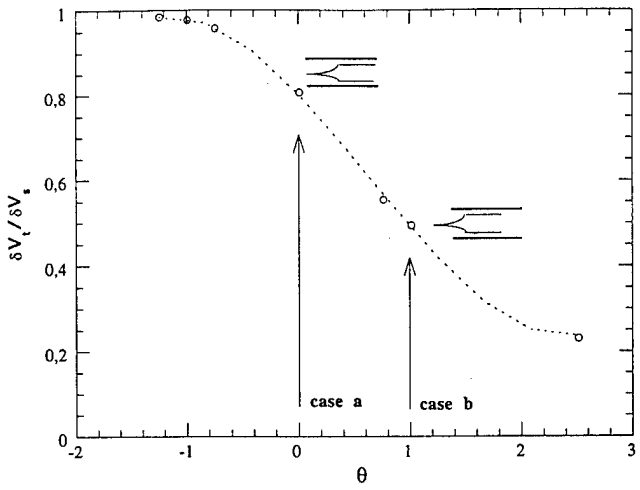


FIG. 9. Relative electrostatic influence of the sheath vs its position relative to the tip. The membrane is placed very far (1 mm) from the gun, at the sheath potential. Note the agreement with the experimental results (shown by arrows).

IV. CONCLUSION AND OUTLOOK

We fabricated a microelectron gun with a coaxial structure made of a metallic tip placed inside a glass capillary with a metal plated outside wall (the sheath). The total diameter of the device can be made as small as 60 μm. By applying a negative voltage to the tip, electrons can be field emitted from the tip. In this way nA currents can be drawn for voltages between 60 and 200 V. We studied the influence of a third conductor, a carbon membrane, placed in front of the tip at a controlled distance between mm and μm. This was essentially possible through the use of an electron projector setup. In this way we showed that this structure behaves like an electron gun as long as the distance between the tip and the carbon membrane is much larger than the radius of the sheath, i.e., a few tens of micrometers in the present case. We point out that, at a distance smaller than 100 μm, the field required to get field emission is almost completely screened by the sheath. Moreover, at such a distance from the tip, the observation of interference fringes proves the high coherence of the electron beam. Altogether, these properties open new perspectives in the field of low-energy coherent electron optics.

[1]W. Stocker, H.-W. Fink, and R. Morin, Ultramicroscopy **31**, 379 (1989). Combinations of low-energy point sources with a microscale lens [G. M. Shed, H. Schmid, P. Unger, and H.-W. Dugner, Rev. Sci. Instrum. **64**, 2579 (1993)] were done but it looks like it is difficult to replace the tip by this kind of setup, especially if it has to be approached close to a sample.
[2]The best results are obtained using the D71 electrolyte from PRECl, Grenoble, France. A NaOH 2 N solution is also possible.
[3]See, for example, R. Gomer, *Field Emission and Field Ionization* (Harvard University Press, Cambridge, MA, 1961).
[4]Computer simulations were performed using EFIELD software from Electric Field and Ion Optics Simulation, Copyright © 1990, 1995 by A. Tolmachev (Chernogolovka, Russia).

Low brightness and high emittance electron gun for a reducing image projection system

Mamoru Nakasuji and Hiroyasu Shimizu
1st Designing Department, IC Equipment Division, IC & LCD Equipment Business Headquarters, Nikon Corporation, Ohi Plant: 6-3, Tokyo 140, Japan

(Received 14 December 1998; accepted 23 April 1999)

An electron gun for an accelerating voltage of 100 kV to be operated under space charge limited condition is designed, and its emittance and brightness are calculated. To obtain low brightness, a Pierce-type electron gun with a Wehnelt electrode and a control anode is adopted. Choosing a low value of the ratio between anode and cathode curvatures reduces the electric field strength at the cathode. A maximum emittance of 1.8 mm mrad and a minimum brightness of 350 A/(cm^2 sr) are obtained under the above conditions with a cathode diameter of 1.66 mm. The radius of the Wehnelt aperture and its axial position are optimized. The effects of deviations of the shapes of the Wehnelt and control electrodes from concentric sphere are studied. It is found that in the space between the anode and the control anode the electrodes must be spherical up to an angle of at least 55°, whereas in the space between the Wehnelt electrode and the control anode the electrodes must be spherical up to an angle of at least 21°. © *1999 American Vacuum Society.* [S0734-211X(99)01904-6]

I. INTRODUCTION

Photolithography, which has historically been the mainstream technology for volume production of integrated circuits, is operated near the limit of linewidth resolution set by diffraction effects of the radiation used. Consequently, radiation with shorter wavelength such as electron and ion beams is investigated extensively for future lithography application. With an ion beam projection system using a dispersive lens designed by computer simulation, a 12.5×12.5 mm^2 field is obtained.[1] Projection electron-beam lithography may benefit from resist process technologies established by scanning electron beam lithography.[2] In projection lithography it is important to obtain a reasonable main field size for, say, a 4G-dynamic random access memory (DRAM) memory device for a high-throughput system, and to obtain a high emittance electron gun. For the lens system the symmetric magnetic doublet (SMD),[3] the PREVAIL,[4] and SMD with deflectors[5] have been studied. Unfortunately, in the SMD the useful beam semiangle is limited by the aberration called field curvature. For the SMD with a deflector system, the main field of 2 mm, subfield of 0.25×0.25 mm^2, and the beam semiangle of 4 mrad are reported.

De Vore and Berger[6] presented a high emittance and low brightness gun which operated under temperature limited conditions. This operating condition has two problems:

(1) The angular distribution of the beam intensity exhibits irregular nonuniformities caused by local variations of the work function or the temperature on the cathode surface.
(2) In the near future shot noise will become a serious problem for the critical dimension (CD) control.

The shot noise i_f^2 for the electron gun with a temperature limited condition is known as in Eq. (1)

$$i_f^2=2ei_0\Delta f. \tag{1}$$

For the electron gun with a space charge limited condition, it generally must be taken account of them together by writing i_f^2 in the form

$$i_f^2=2ei_0\Gamma^2\Delta f \tag{2}$$

with the factor Γ usually varying between about 1 and 0.1, depending on operating conditions.[7]

An electron gun for a variably shaped beam operated under the condition of space charge limited condition was studied experimentally, and a maximum emittance of 0.042 mm mrad at a beam energy of 20 keV[8] was obtained. We have made calculations to study a high-emittance electron gun to be operated under the condition of a space charge limited condition. An emittance exceeding 1 mm mrad and brightness from 4×10^4 to 9×10^4 A/(cm^2 sr) was obtained by computer simulation.[9] Such brightness values are too large for the subfield size of 0.25×0.25 mm^2 and the beam semiangle of 4 mrad. In this article we are trying to design a Pierce-type electron gun in which the electric field strength at the cathode surface is weak, and low brightness may be expected.

II. DESIGN TARGET AND DESIGN

The characteristics of an electron gun for a reducing projection system are defined as follows, where the reticle pattern is projected on the wafer. If the subfield diameter is D_{sf} and the beam semiangle in the image plane is α, then the emittance E_{mi} is defined as

$$E_{mi}=\alpha_g D_g>\alpha D_{sf}, \tag{3}$$

where α_g is defined as the value of the beam half angle at which the angular intensity deviates by more than ±0.5% of its value in axial direction, and D_g is the full width at half maximum (FWHM) gun crossover diameter. If α=4 mrad and $D_{sf}=2^{0.5}\times0.25$ mm, then we have

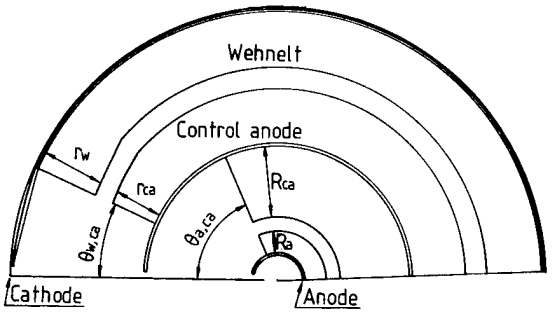


FIG. 1. Pierce-type electron gun with Wehnelt electrode, control anode, and anode showing the parameters $\theta_{w,ca}$, $\theta_{a,ca}$, R_a, R_{ca}, r_{ca}, and r_w used in the optimization.

$$E_{mi} > 1.41 \text{ mm mrad}. \tag{4}$$

The beam current I at the wafer, the brightness B, the beam semiangle α, the subfield area S_{sf}, and a reticle filling factor f (transmission area/scattering area) satisfy the following equation:

$$I = \pi \alpha^2 B S_{sf} f. \tag{5}$$

If a complementary mask is used, the filling factor may be smaller than 0.25. With $S_{sf}=6.25\times10^{-4}\,\text{cm}^2$ and a beam current of $I=20\,\mu\text{A}$ the brightness amounts to

$$B = 2.6\times10^3 \text{ A/(cm}^2 \text{ sr)}. \tag{6}$$

The design targets are chosen as follows:

$$E_{mi} > 1.41 \text{ mm mrad}; \quad B = 2.6\times10^3 \text{ A/(cm}^2 \text{ sr)}.$$

To obtain low brightness in the electron gun operating with a space charge limited condition, the electric field strength at the cathode must be low. To obtain high emittance, the electric field strength at the cathode must be uniform. To satisfy these two conditions, a Pierce-type[10] electron gun is adopted. The electron gun model is shown in Fig. 1. The cathode, control anode, and anode are approximately a concentric sphere with 100, 50, and 10 mm radii, respectively. The Wehnelt electrode has approximately conical shape. The control anode and anode are both 0.5 mm in thickness.

The electrodes are supported by insulators, which must be positioned far from the optical axis. The irregular electric field generated at the insulator surfaces must be shielded from the paraxial region. To achieve this shielding, the following electrode shapes are used. In the space between the Wehnelt electrode and the control anode, the electrodes are shaped as a concentric sphere only for an angle smaller than $\theta_{w,ca}$. For angles exceeding $\theta_{w,ca}$, both electrodes are shaped as shown in Fig. 1. In the space between the control anode and the anode, the electrodes are shaped as a concentric sphere only for angles smaller than $\theta_{a,ca}$, for angles exceeding $\theta_{a,ca}$, both electrodes are shaped as shown in Fig. 1. The parameters $\theta_{w,ca}$, $\theta_{a,ca}$, R_{ca}/R_a, and r_w/r_{ca} will be used for optimization later.

The cathode material is a LaB_6 single crystal. For its work function a value of 2.4 eV is assumed, and for the cathode temperature a value of 1600 K. Figure 2 shows an enlarged view of the region around the cathode and the Wehnelt electrode. The fine solid lines show the mesh pattern for the finite element method. The Wehnelt bore radius r_{w0} and its axial position Z_w are used as parameters for the first optimization step described in the following section.

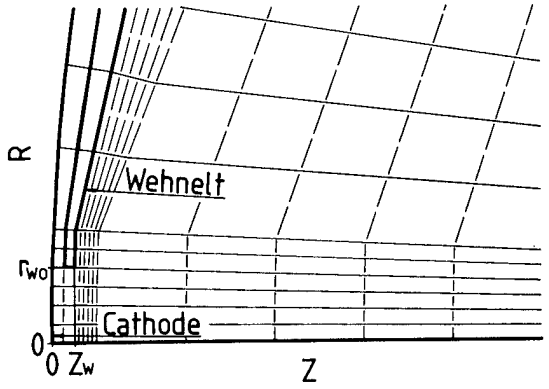


FIG. 2. Enlarged view of the region close to the cathode showing the parameters r_{w0} and Z_w.

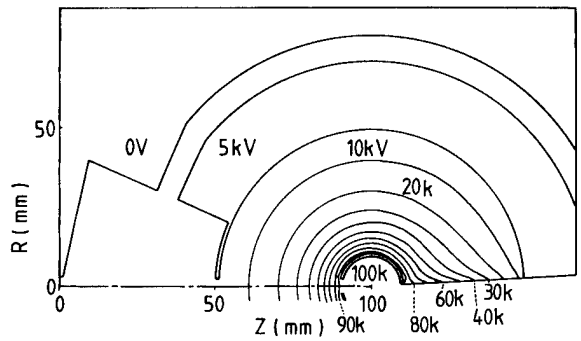


FIG. 3. Equipotentials for $V_w=0$, $V_{ca}=5$ kV, $V_a=100$ kV.

III. CALCULATION AND RESULTS

The cathode and anode potentials are 0 and 100 kV, respectively. The potentials of the Wehnelt electrode and the control anode are varied from 0 to -10 V and from 1 to 10 kV, respectively. Calculations are carried out using commercially available software.[11] For zero Wehnelt potential and a control anode potential of 5 kV, equipotentials are shown in Fig. 3. The equipotentials around the optical axis are sparse near the cathode and dense near the anode. Therefore the electric field at the cathode is weak and then low brightness can be expected.

For the Wehnelt and the control anode potentials of -5 and 5 kV, respectively, calculated examples are shown in Fig. 4. In Fig. 4, equipotentials are shown as solid lines, whereas the broken lines represent trajectories leaving the cathode in axial direction. Trajectories are calculated from the cathode to the end plane at $Z=95$ mm. At the end plane, the radial coordinate and its first derivative with respect to Z are calculated for each trajectory. From these values and the initial energies at the cathode, the position and diameter of the virtual crossover and the angular distribution at the crossover are calculated. Typical trajectories are shown in Figs. 5, 6, and 7.

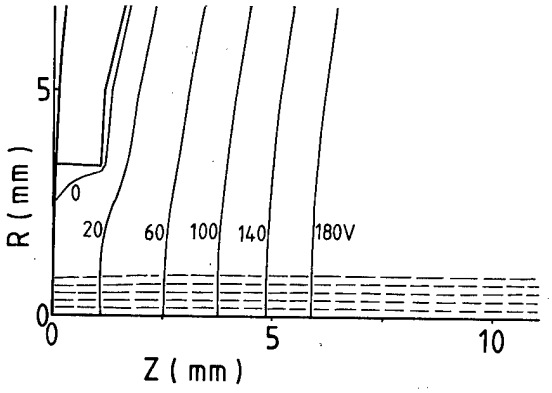


FIG. 4. Equipotentials (solid lines) and trajectories (broken lines).

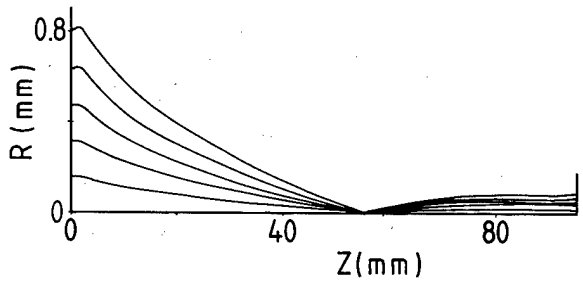


FIG. 5. Trajectories for $V_w=-7$ V, $V_{ca}=1.5$ kV. The crossover is formed close to the control anode bore.

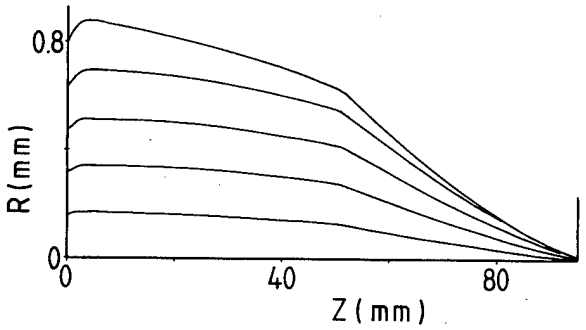


FIG. 6. Trajectories for $V_w=-5$ V, $V_{ca}=4$ kV. The crossover is formed close to the common center of the concentric electrodes.

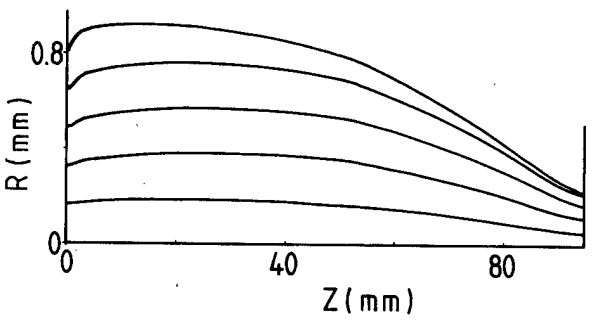


FIG. 7. Trajectories for $V_w=-8$ V, $V_{ca}=8$ kV. The crossover is formed beyond the common center of the concentric electrodes.

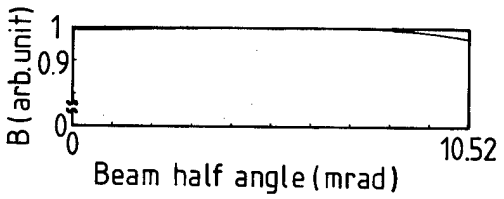


FIG. 8. Angular intensity distribution.

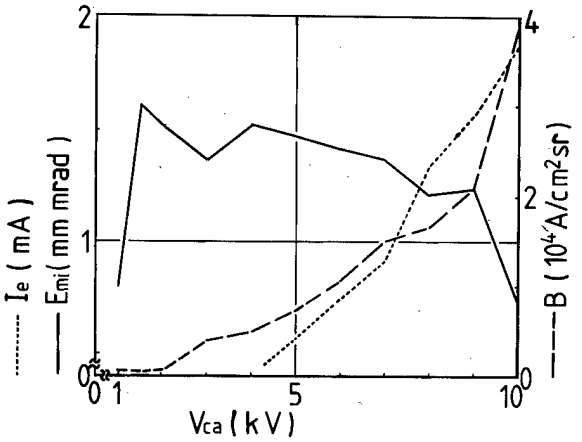


FIG. 9. Emittance, brightness, and emission current vs control anode potential.

Figure 5 is the case where the control anode and Wehnelt potentials are 1.5 kV and −7 V, respectively. This figure shows the beams emitted from the cathode are focused by the convergent lens formed by the Wehnelt electrode, forming a crossover at Z=56 mm. They pass through the convergent lens formed by the aperture in the control anode at small distances from the axis. Thus the aberrations of this lens are small, and a high emittance of 1.6 mm mrad is obtained.

Figure 6 shows the case where the control anode and Wehnelt potentials are 4 kV and −5 V, respectively. The trajectories form a crossover close to the center of the concentric electrodes. The trajectories are more strongly bent near the cathode and the control anode, and an emittance of 1.81 mm mrad is obtained.

Figure 7 shows the case where the control anode and Wehnelt potentials are 8 kV and −8 V, respectively. The trajectories pass the aperture in the control anode at a large distance from the axis but the lens formed by this aperture is weak for this value of the control anode potential, and the trajectories are only slightly bent. No crossover is formed in the region Z<95 mm. Independent of the crossover position. Figs. 5, 6, and 7 indicate that high emittance is obtained.

A typical angular intensity distribution is shown in Fig. 8. In the axial direction, the angular intensity is 99.5% of its maximum value. The emission angle is defined as the value of the beam half angle where the angular intensity deviates by more than 1.0% of its value in axial direction. The crossover diameter D_g is calculated by the program. For different values of the control the anode potential is chosen to yield maximum emittance. The results are shown in Fig. 9. The maximum emittance of 1.6 mm mrad and the minimum brightness of 4×10^2 A/(cm^2 sr) are obtained for the same value of V_{ca} of 1.5 kV. The emission current is also plotted. Too large emission currents would make the power supply too expensive. The lack of smoothness of the curves is at-

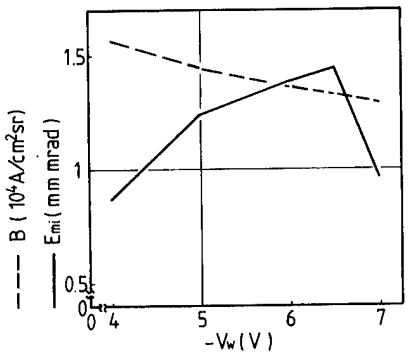


FIG. 10. Emittance and brightness vs Wehnelt potential.

tributed to the fact that different values of the Wehnelt potential are used for each control anode potential. The optimum Wehnelt bias $|V_w|$ increases with increasing control anode potential.

IV. EMITTANCE AS A FUNCTION OF SOME PARAMETERS

For a control anode potential of 6 kV and varying Wehnelt potentials, emittance and brightness are calculated. The results are shown in Fig. 10. Maximum emittance is obtained for $V_w = -6.5$ V, and the brightness decreases with increasing Wehnelt bias. Figure 11 shows for $V_{ca} = 5$ kV and $V_w = -5$ V the dependence of the emittance on the Wehnelt bore radius r_{w0} for different values of the axial position Z_w of the Wehnelt electrode. Maximum emittance of 1.8 mm mrad is obtained for $r_{w0} = 2.5$ mm and $Z_w = 1$ mm. For all other combinations of r_{w0} and Z_w values, the emittance is below 1 mm mrad.

It is important to study how close to the axis the conducting plates shielding the insulator may be. The results of such studies are shown in Fig. 12 where emittance E_{mi} and brightness B are plotted versus $\theta_{a,ca}$ and in Fig. 13 where they are plotted versus R_{ca}/R_a. For $R_{ca}/R_a = 2.5$. Fig. 12 shows that a maximum emittance of 1.8 mm mrad is obtained if $\theta_{a,ca} > 55°$. It is found that in the space between the anode and the control anode the electrodes must be spherical up to an angle of at least 55°. Figure 13 shows that a maximum emittance of 1.8 mm mrad is obtained for $R_{ca}/R_a = 2.5$. For $R_{ca}/R_a > 2.5$, the emittance decreases slowly with increasing values of R_{ca}/R_a. In Figs. 14 and 15, the emittance (as solid

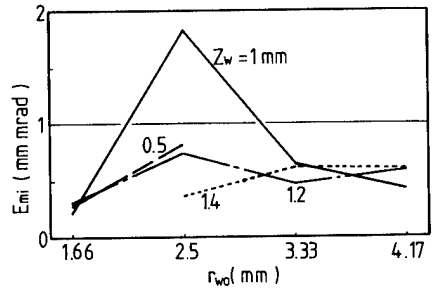


FIG. 11. Emittance vs r_{w0} for different values of the axial position Z_w of the Wehnelt electrode.

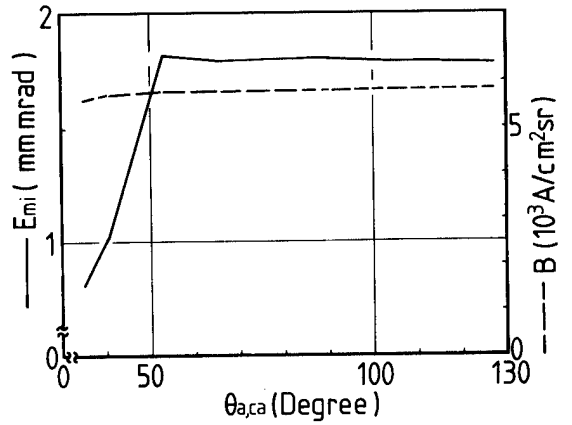


FIG. 12. Emittance and brightness vs $\theta_{a,ca}$.

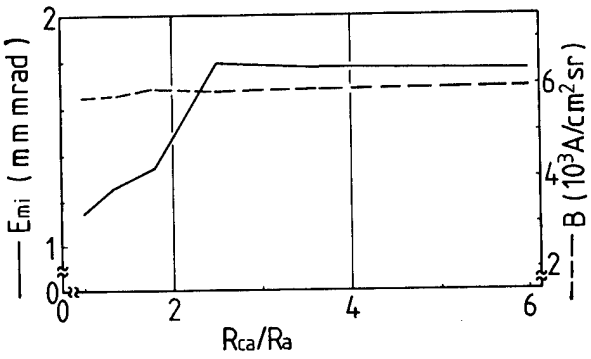


FIG. 13. Emittance and brightness vs R_{ca}/R_a.

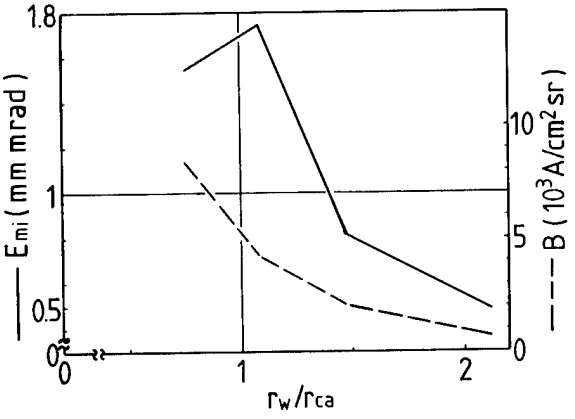


FIG. 14. Emittance and brightness vs r_w/r_{ca}.

line) and the brightness (as broken line) are plotted versus $\theta_{w,ca}$ and r_w/r_{ca}, respectively. In Fig. 14 if r_w/r_{ca} is varied for the constant value $\theta_{w,ca} = 25°$, then a maximum emittance of 1.75 mm mrad is obtained for $r_w/r_{ca} = 1.05$. For values of $r_w/r_{ca} < 1.05$, the emittance decreases rapidly with decreasing r_w/r_{ca}. The brightness decreases with increasing values of r_w/r_{ca}. In Fig. 15, the emittance and the brightness are plotted versus $\theta_{w,ca}$. For $\theta_{w,ca} > 21°$, an emittance of 1.8 mm mrad is obtained. The brightness has a maximum for $\theta_{w,ca} = 30°$. For values $\theta_{w,ca} < 30°$, it decreases with decreasing $\theta_{w,ca}$. Its is found that in the space between the Wehnelt

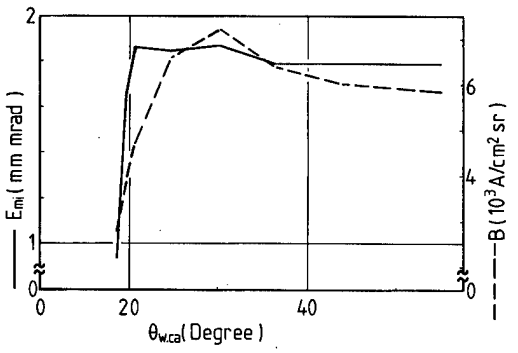


FIG. 15. Emittance and brightness vs $\theta_{w,ca}$.

electrode and the control anode the electrodes must be spherical up to an angle of at least 21°.

V. CONCLUSIONS

An electron gun for an accelerating voltage of 100 kV to be operated under space charge limited emission is designed, and its emittance and brightness are calculated. To obtain low brightness, a Pierce-type electron gun with a Wehnelt electrode and a control anode is adopted. The electric field strength at the cathode is reduced by choosing a low value of the ratio between anode and cathode curvatures. A maximum emittance of 1.8 mm mrad and a minimum brightness of 350 A/(cm^{2}sr) are obtained under the above conditions with a cathode diameter of 1.66 mm. The radius of the Wehnelt aperture and its axial position are optimized. The effects of deviations of the shapes of the Wehnelt and control electrodes from the concentric sphere are studied. It is found that in the space between the anode and the control anode the electrodes must be spherical up to an angle of at least 55°, whereas in the space between Wehnelt electrode and control anode the electrodes must be spherical up to an angle of at least 21°.

ACKNOWLEDGMENTS

The authors thank T. Shimamura, M. Kurosawa, K. Suwa, A. Tanimoto, Dr. K. Okamoto, and S. Suzuki for their useful discussion through this study.

[1]R. Kaesmaier and H. Loschner, The 3rd International Workshop on High Throughput Charged Particle Lithography, Waikoloa Village, Hawaii, 1998.
[2]T. Tamura, K. Nakajima, and H. Nozue, Jpn. J. Appl. Phys., Part 1 **33**, 143 (1995).
[3]M. Nakasuji, J. Vac. Sci. Technol. B **15**, 1833 (1997).
[4]H. C. Pfeiffer, in Ref. 1.
[5]S. Koikari, S. Yoshitake, M. Ogasawara, and S. Tamamushi, "Numerical Calculation on Optical System for EB Projection," Digest of Papers MicroProcess **96**, 236 (1996).
[6]W. DeVore and S. D. Berger, J. Vac. Sci. Technol. B **14**, 3764 (1996).
[7]R. A. Smith, F. E. Jones, and R. P. Chasmaar, *The Detection and Measurement of Infrared Radiation*, 2nd ed. (Oxford University Press, London, 1968).
[8]M. Nakasuji and H. Wada, J. Vac. Sci. Technol. **17**, 1367 (1980).
[9]M. Nakasuji and H. Shimizu, Jpn. J. Appl. Phys., Part 1 **36**, 2404 (1997).
[10]J. R. Pierce, J. Appl. Phys. **10**, 548 (1939).
[11]E. Munro, J. Vac. Sci. Technol. B **8**, 1657 (1990).

Generalization of electrostatic lens characteristics using the Picht ray trajectories

Yoshimi Kawanami
Central Research Laboratory, Hitachi, Ltd., Kokubunji 185-8601, Japan

Tohru Ishitani[a)]
Instruments, Hitachi, Ltd., Hitachinaka 312-8504, Japan

(Received 20 January 1999; accepted 7 May 1999)

Paraxial electron/ion trajectories in rotationally symmetric fields are analyzed in the Picht space to derive analytical forms of thin lens characteristics such as focal length F, magnification M, and chromatic and spherical aberration coefficients C_c and C_s. Dimensionless forms of T/F, $\{(C_c/T)/(T/F)\}/(L_o/T)^2$, and $\{(C_s/T)/(T/F)\}/(L_o/T)^4$, which are determined only by the normalized lens-potential distribution, are derived to evaluate the lens. Here, T is the lens thickness and L_o is the distance from the object to the lens. As to bipotential or unipotential lenses composed of electrodes facing each other symmetrically, it is found that the accelerating mode at $V_2/V_1 = a(>1)$ and the decelerating mode at $V_2/V_1 = 1/a(<1)$ have relationships of $F_{\text{accel}} = F_{\text{decel}}$, $C_{c,\text{accel}} = C_{c,\text{decel}}/a$, and $C_{s,\text{accel}} = C_{s,\text{decel}}/a^{1/2}$, where V_1 and V_2 are the 1st and 2nd electrode potentials, respectively. Under constant lens fields, the known M-dependent C_c and C_s expressions, i.e., $C_c = C_{c,\infty}(1-1/M)^2$ and $C_s = C_{s,\infty}(1-1/M)^4$, are analytically derived also in the present approach. © *1999 American Vacuum Society.* [S0734-211X(99)04304-8]

I. INTRODUCTION

Electrostatic lenses are popularly used in focused ion/electron beam systems. The lens characteristics (i.e., magnification, focusing length, and chromatic and spherical aberrations) are varied both with the distances of source-to-lens and lens-to-image and with beam accelerating voltage. In rough design of the beam optics, analytical expressions of these characteristics are more useful than numerical approaches. In our earlier studies,[1–3] the analytical expressions derived from the Picht transformed ray trajectories have been satisfactorily applied to the coarse design of focused-ion-beam (FIB) systems. In the present study, analytical derivations of the lens characteristics are reviewed in more normalized forms. The lens characteristics in decelerating and accelerating modes are also clarified to show a simple relationship between them. Besides, the lens characteristics are applied for one- and two-lens systems to describe analytically their optical performances.

II. PARAXIAL RAYS IN ROTATIONALLY SYMMETRIC ELECTRIC FIELDS AND LENS FORMULA

Let us consider paraxial electron/ion trajectories in rotationally symmetric fields. If the initial direction of the electrons/ions is in a plane which passes through the z axis, a so-called meridian plane, these trajectories will remain in this plane. The distance of the electron/ion from the z axis is expressed by a linear second-order differential equation,[4,5]

$$r'' + (V'/2V)r' + (V''/4V)r = 0, \tag{1}$$

where V is the beam acceleration voltage. By using a new variable,

$$R \equiv rV^{1/4}, \tag{2}$$

and substituting it into Eq. (1), Picht has introduced a very useful modification,[4,5]

$$R'' + (3/16)(V'/V)^2 R = 0. \tag{3}$$

In the present article, we name the (R, z) space the Picht space in contrast with the (r, z) actual space. Since the R''/R value is always negative or zero at any (R, z) point inside the lens, being independent of a sign of (V'/V), the R trajectories (not r trajectories) have focusing action and both principal planes in the Picht space are within the lens. Therefore, R is approximately constant inside the lens when the lens is thin and weak. The paraxial rays are schematically shown in Fig. 1. Lens formulas in the Picht space are expressed as

$$1/L_o + 1/L_i = 1/F, \tag{4}$$

$$M = -L_i/L_o, \tag{5}$$

$$\alpha_o/\alpha_i = \sigma M \quad (\because a_o L_o V_o^{1/4} = -\alpha_i L_i V_i^{1/4}), \tag{6}$$

where

$$\sigma = (V_i/V_o)^{1/4}. \tag{7}$$

Let us take the following paraxial equation instead of Eq. (3):

$$dR'/dz = -(3/16)(V'/V)^2 R. \tag{8}$$

By integrating this equation between z_0 and z_i, we obtain

$$[R']_{zi} - [R']_{zo} = -(3/16)\int_{z_o}^{z_i} (V'/V)^2 R \, dz. \tag{9}$$

[a)]Mailing address: Instruments, Hitachi, Ltd., Ichige 882, Hitachinaka, Ibaraki-ken 312-8504, Japan; electronic mail: ishitani@cm.naka.hitachi.co.jp

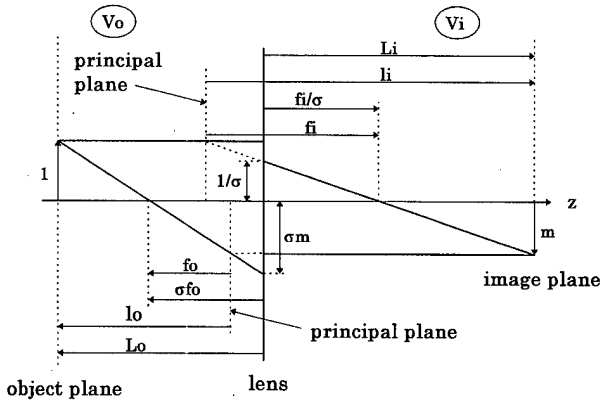


FIG. 1. Paraxial rays when the lens field is concentrated at $z=0$ [$\sigma=(V_i/V_0)^{1/4}$].

The integration range from z_0 to z_i will be omitted below to simplify the expression. For an incident trajectory parallel to the axis (i.e., $R_0'=0$), an image focal length is calculated by

$$1/F=-R_i'/R_o=-R_i'/R_i \quad (\because R_o=R_i), \tag{10a}$$

$$=(3/16)\int (V'/V)^2 dz. \tag{10b}$$

Here, the object focal length is equal to the image focal length.

In the actual space, on the other hand, lens formulas are similarly given by

$$f_o/l_o+f_i/l_i=1, \tag{11}$$

$$m=-f_o/(l_o-f_o)=-(l_i-f_i)/f_i,$$

$$=-(l_i/f_i)/(l_o/f_o), \tag{12}$$

$$\alpha_o/\alpha_i=\sigma^2 m. \tag{13}$$

For an incident trajectory parallel to the axis (i.e., $r_0'=0$), by inserting $R=rV^{1/4}$ and $R'=r'V^{1/4}+(r/4)(V'/V^{3/4})$ into Eq. (10a), we obtain

$$-R_i'/R_o(=1/F)=-r_i'V_i^{1/4}/r_oV_o^{1/4},$$

$$=-(r_i'/r_o)\sigma. \tag{14}$$

Then, the image focal length f_i is given by

$$1/f_i(=-r_i'/r_o)=1/(\sigma F). \tag{15}$$

A similar analysis for the trajectories of an electron/ion entering the lens parallel to the axis from the image side provides the object focal length given by

$$1/f_o(=-r_o'/r_i)=\sigma/F. \tag{16}$$

The actual focal length f_0 and f_i are equal in magnitude only at $V_0=V_i$, while the Picht ones are equal in magnitude even at $V_0\neq V_i$. In addition, the following relations are satisfied:

$$F=\sigma f_o=(1/\sigma)f_i, \tag{17a}$$

$$=(f_of_i)^{1/2}, \tag{17b}$$

$$M=\sigma m. \tag{18}$$

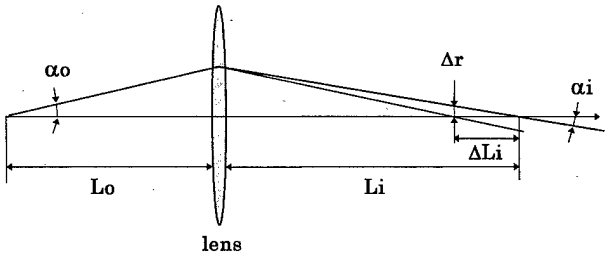


FIG. 2. Schematic aberration disk radius in the image plane.

The present thin-lens approximation is applicable for the condition of $F\gg L_{\text{lens}}$, where L_{lens} is some length characteristic of the lens producing the field $V(z)$, typically a gap or total length.

III. LENS ABERRATIONS

An aberration disk radius in the image plane is schematically shown in Fig. 2 and is expressed as

$$\Delta r=-\alpha_i\Delta L_i. \tag{19}$$

Under the condition of $L_o=$constant, a partial derivative of Eq. (4) with respect to L_i and $1/F$ yields

$$\Delta L_i=-L_i^2\Delta(1/F),$$

$$=-(\sigma m)^2L_o^2\Delta(1/F). \tag{20}$$

By inserting Eq. (20) into Eq. (19) and converting α_i to α_0 through Eq. (13), we obtain

$$\Delta r=mL_o^2\Delta(1/F)\alpha_o. \tag{21}$$

Here, $\Delta(1/F)$ is expressed as a series of both ΔV and $r(=\alpha_0L_o)$ as

$$\Delta(1/F)=\{\delta(1/F)/\delta V\}\Delta V$$

$$+(1/2)\{\delta^2(1/F)/\delta r^2\}(\alpha_oL_o)^2+\cdots. \tag{22}$$

Then, Eq. (21) is rewritten as

$$\Delta r=m[L_o^2\alpha_o\{\delta(1/F)/\delta V\}\Delta V$$

$$+(1/2)L_o^4\{\delta^2(1/F)/\delta r^2\}\alpha_o^3+\cdots],$$

$$=m[C_c\alpha_o(\Delta V/V_o)+C_s\alpha_o^3+\cdots], \tag{23}$$

where

$$C_c=L_o^2V_o\{\delta(1/F)/\delta V\},$$

$$=-2L_o^2(1/F)(V_o/V_{\text{in}}), \tag{24}$$

$$1/V_{\text{av}}=-(1/2)\{\delta(1/F)/\delta V\}/(1/F), \tag{25}$$

$$C_s=(1/2)L_o^4\{\delta^2(1/F)/\delta r^2\},$$

$$=L_o^4(1/F)/r_e^2, \tag{26}$$

and

$$1/r_e^2=(1/2)\{\delta^2(1/F)/\delta r^2\}/(1/F). \tag{27}$$

By differentiating Eq. (10b) with respect to V, we obtain

$$\delta(1/F)/\delta V=-(3/8)\int (1/V)(V'/V)^2dz. \quad (28)$$

Substituting Eqs. (10b) and (28) into Eq. (25), we have

$$1/V_{\rm av}=\left\{\int (1/V)(V'/V)^2dz\right\}\Big/\left\{\int (V'/V)^2dz\right\}. \quad (29)$$

This means that $1/V_{\rm av}$ is an average value of $1/V$ weighted by $(V'/V)^2$, or $d(1/F)/dz$, in the lens. On the other hand, r_e corresponds to a normalized r, which appears in the r-series $1/F$ form of $1/F\approx[1/F]_{r=0}[1+(r/r_e)^2]$ [see Eq. (22)].

In general, using a variable $S(=V'/V)$ the chromatic and spherical aberration coefficients referenced to the object are given in integral expressions:[4]

$$C_c=-(3/8)\int (V_o/V)^{1/2}S^2r^2dz, \quad (30)$$

$$C_s=(1/64)\int (V/V_o)^{1/2}[4S'^2+3S^4-5S^2S'-SS'']r^4dz. \quad (31)$$

These aberration coefficients are referred also to the image and their transformed expressions from the object to the image are expressed as

$$C_{c,\rm image}=m^2\sigma^6C_c, \quad (32a)$$

$$=(M^2\sigma^4)C_c, \quad (32b)$$

$$C_{s,\rm image}=m^4\sigma^6C_s, \quad (33a)$$

$$=(M^4\sigma^2)C_s, \quad (33b)$$

By substituting $r=R/V^{1/4}$ into Eq. (30), taking R ($\approx$constant) out from the integral, and again substituting $R=\alpha_oL_oV_0^{1/4}$ (with $\alpha_0=1$) into it, we obtain

$$C_c=-(3/8)L_o^2V_o\int (1/V)(V'/V)^2dz. \quad (34)$$

Using Eqs. (10b) and (29), we rewrite it as

$$C_c=-2L_o^2(1/F)(V_o/V_{\rm av}), \quad (35)$$

which is identical with Eq. (24). As to C_s, using $S(z_0)=S(z_i)=0$ in the integration of Eq. (31) by parts, we obtain a similar equation,

$$C_s=(1/64)L_o^4G_{\rm int}, \quad (36)$$

where

$$G_{\rm int}=5\int (V_o/V)^{1/2}\{(V''/V)^2-(11/12)(V'/V)^4\}dz. \quad (37)$$

Here, the dimension of $G_{\rm int}$ is m^{-3}. Comparing Eq. (26) with Eq. (36), we obtain

$$1/r_e^2=(1/64)G_{\rm int}/(1/F). \quad (38)$$

Here, L_o and L_i are expressed as functions of f_0, f_i, F, m, and M as

$$L_o=\sigma f_o(1-1/\sigma m), \quad (39a)$$

$$=F(1-1/M), \quad (39b)$$

$$L_i=(f_i/\sigma)(1-\sigma m), \quad (40a)$$

$$=F(1-M). \quad (40b)$$

Substituting Eq. (39b) into Eqs. (35) and (36), we obtain the C_c and C_s forms depending on M and m,

$$C_c=-2F(V_o/V_{\rm av})(1-1/M)^2, \quad (41a)$$

$$=C_{c,\infty}(1-1/M)^2, \quad (41b)$$

$$=C_{c,\infty}(1-1/\sigma m)^2, \quad (41c)$$

$$C_s=(1/64)F^4(1-1/M)^4G_{\rm int}, \quad (42a)$$

$$=C_{s,\infty}(1-1/M)^4, \quad (42b)$$

$$=C_{s,\infty}(1-1/\sigma m)^4, \quad (42c)$$

where

$$C_{c,\infty}=-2F(V_o/V_{\rm av}), \quad (43)$$

$$C_{s,\infty}=(1/64)F^4G_{\rm int}, \quad (44a)$$

$$=F^3/r_e^2. \quad (44b)$$

The $C_{c,\infty}$ and $C_{s,\infty}$ values correspond to the C_c and C_s values under a fixed lens excitation (i.e., $V_o/V_{\rm av}$ and $F=$fixed) with m and $M=\infty$, respectively. These σm-dependent forms [i.e., Eqs. (41c) and (42c)] have been empirically found after rearranging numerical data on their coefficients by Renau and Heddle.[6] Ura has also analytically derived the same forms,[7] but differently from the present approach.

IV. DECELERATING AND ACCELERATING MODES ON BIPOTENTIAL LENS PROPERTIES

First, we will consider the simple case of two electrodes which symmetrically face each other at the distance of $T(=z_i-z_o)$ and have potentials of V_o and V_i. The potential inside the lens is expressed by

$$V(u)=(V_o+V_i)/2+\{(V_i-V_o)/2\}g(u), \quad (45)$$

where

$$u=\{z-(z_o+z_i)/2\}/T \quad (-1/2\leq u\leq 1/2), \quad (46)$$

and $g(u)$ is the dimensionless potential function satisfying $g(u)=-g(-u)$, $g(0)=0$, $g(1/2)=1$, $g(-1/2)=-1$, and $g'(\pm 1/2)=0$. The first and second order differential expressions on V with respect to z are given by

$$dV/dz=(dV/du)(du/dz),$$

$$=\{(V_i-V_o)/(2T)\}(dg/du), \quad (47)$$

$$d^2V/dz^2=\{(V_i-V_o)/(2T^2)\}(d^2g/du^2). \quad (48)$$

Let us consider accelerating and decelerating lens modes, i.e., $V_i/V_o=a(>1)$ and $1/a(<1)$, respectively. By inserting Eq. (47) for the accelerating and decelerating modes into Eq. (13), we obtain the $1/F$ equations;

$$1/F_{\rm accel}=(3/16)(1/T)\int [(a-1)g'/\{(a+1)+(a-1)g\}]^2du, \quad (49a)$$

$$1/F_{\text{decel}}=(3/16)(1/T)\int [(a-1)g'/ \{(a+1)-(a-1)g\}]^2 du. \quad (49b)$$

Here, since $g(u)=-g(-u)$, both of the above integrated values are identical to yield $F_{\text{accel}}=F_{\text{decel}}$. The T/F form is dimensionless and is determined only by a and $g(u)$.

Similarly, the numerator in the right term of Eq. (29) is expressed as

$$(\text{Numerator})_{\text{accel}}=(1/V_o)(2/T)K(a) \quad [\text{for the accelerating mode}], \quad (50a)$$

$$(\text{Numerator})_{\text{decel}}=(1/V_o)(2/T)K(1/a) \quad [\text{for the decelerating mode}], \quad (50b)$$

where

$$K(a)=\int [\{(a-1)g'\}^2/\{(a+1)+(a-1)g\}^3]du. \quad (51)$$

Since $K(a)=K(1/a)/a$, we obtain $(1/V_{\text{av,accel}})=(1/V_{\text{av,decel}})/a$ and $C_{c,\text{accel}}=C_{c,\text{decel}}/a$. Here, Eq. (24) is rewritten as

$$\{(C_c/T)/(T/F)\}/\{L_o/T\}^2=-2V_o/V_{\text{av}}, \quad (52a)$$

$$=-(3/4)K/(T/F), \quad (52b)$$

which is dimensionless and is determined only by a and $g(u)$.

As to G_{int} of Eq. (37), we obtain $G_{\text{int,accel}}=G_{\text{int,decel}}/a^{1/2}$ and $C_{s,\text{accel}}=C_{s,\text{decel}}/a^{1/2}$. Here,

$$G_{\text{int,accel}}=(5/T^3)\int [2/\{(a+1)+(a-1)g\}]^{1/2}[[\{(a-1)g''/ \{(a+1)+(a-1)g\}]^2-(11/12)[(a-1)g'/ \{(a+1)+(a-1)g\}]^4]du. \quad (53)$$

Then, Eq. (36) is rearranged as

$$\{(C_s/T)/(T/F)\}/\{L_o/T\}^4=T^2/r_e^2, \quad (54a)$$

$$=(1/64)(T^3G_{\text{int}})/(T/F). \quad (54b)$$

The relationships of F, C_c, and C_s between the accelerating and decelerating modes for the lens with $g(u)=-g(-u)$ are summarized as

$$F_{\text{accel}}=F_{\text{decel}}, \quad (55a)$$

$$C_{c,\text{accel}}=C_{c,\text{decel}}/a, \quad (55b)$$

and

$$C_{s,\text{accel}}=C_{s,\text{decel}}/a^{1/2}, \quad (55c)$$

where $V_2/V_1=a(>1)$ and $1/a(<1)$ in the accelerating and decelerating modes, respectively.

V. THREE-ELECTRODE LENSES

Let us consider three-electrode lenses, of which both the (first and second) and the (second and third) electrodes face each other symmetrically to satisfy $g(u)=-g(-u)$ in Eq. (45). In the present study we will take three types of potentials, i.e., a linear type, the Butler type, and a bell type. The linear-potential lens is the most simplified lens, in which the potential is constant at both sides of the lens and it increases/decreases linearly inside as shown in Fig. 3(a). Although this lens is far from reality, it provides practical information at the expense of very little work. Since the 1st-2nd-3rd electrode lens is regarded as a combination of the 1st-2nd and 2nd-3rd electrode bipotential lenses, the $1/Vav$ and $1/F$ parameters of the three-electrode lens are simply expressed as

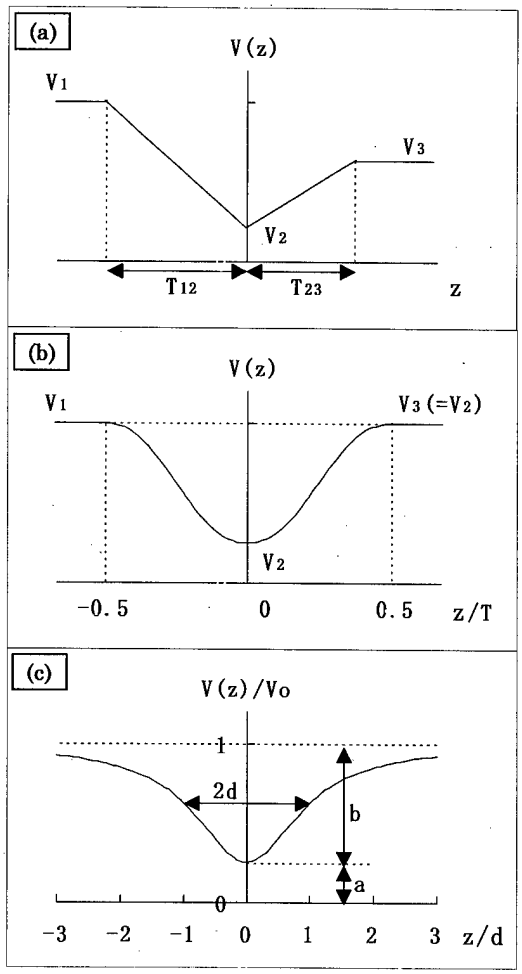


FIG. 3. Potential distributions of the various lenses; (a) a linear type, (b) the Butler type, and (c) a bell type.

$$1/V_{\text{av}}=(1/V_{\text{av},12}+1/V_{\text{av},23})/2, \quad (56)$$

$$1/F=1/F_{12}+1/F_{23}-(T_{12}/2+T_{23}/2)/(F_{12}F_{23}). \quad (57)$$

Here, the $1/V_{\text{av},ij}$ and $1/F_{ij}$ parameters for the i-j electrode lens with linear potentials are given by

$$1/V_{\text{av},ij}=(1/V_i+1/V_j)/2, \quad (58)$$

$$1/F_{ij}=(3/16)(1/T_{ij})(V_i-V_j)^2/(V_iV_j), \quad (59)$$

respectively. When the last term in Eq. (57) is negligible, the V_2 value to focus the beam by a distance L_i under the given V_1, V_3, T_{12}, T_{23}, and L_o parameters is given by

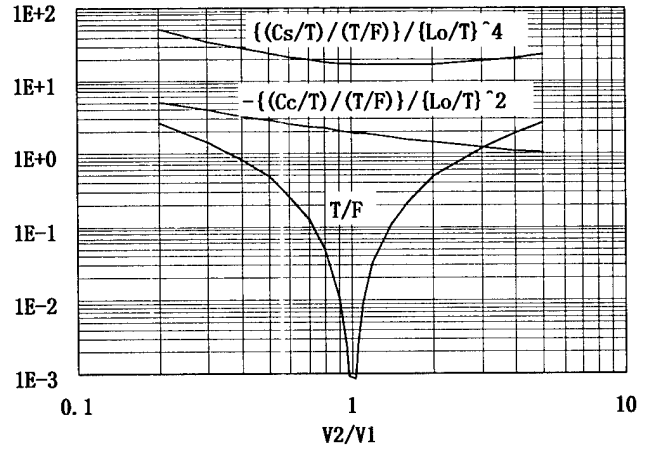


FIG. 4. Dimensionless lens characteristics of T/F, $-\{(C_c/T)/(T/F)\}/\{L_o/T\}^2$, and $\{(C_s/T)/(T/F)\}/\{L_o/T\}^4$ as a function of V_2/V_1 for a unipotential lens with the Butler-shaped electrodes [see Fig. 3(b)].

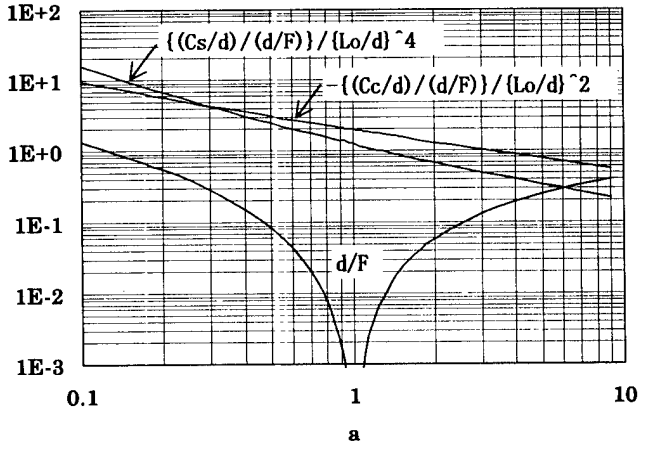


FIG. 5. Dimensionless lens characteristics of d/F, $-\{(C_c/d)/(d/F)\}/\{L_o/d\}^2$, and $\{(C_s/d)/(d/F)\}/\{L_o/d\}^4$ as a function of $a(=1-b)$ for a unipotential lens with the bell-type potential [see Fig. 3(c)].

$$V_2=[(4+A)\pm\{(4+A)^2-4(V_1+V_3)^2/(V_1V_3)\}^{1/2}]/[2(1/V_1+1/V_3)], \tag{60a}$$

where

$$A=(3/16)(1/L_o+1/L_i)(T_{12}/2+T_{23}/2). \tag{60b}$$

The + and − signs in Eq. (60a) correspond to the decelerating and accelerating modes, respectively. It is found that the V_2 values for these modes satisfy the relationships

$$V_{2,\mathrm{decel}}V_{2,\mathrm{accel}}=V_1V_3, \tag{61a}$$

$$V_{2,\mathrm{decel}}+V_{2,\mathrm{accel}}=(4+A)/(1/V_1+1/V_3). \tag{61b}$$

When V_1 and V_3 are of the same potential, the lens becomes a unipotential lens.

The second is the Butler-type unipotential lens,[8] in which the potential distribution is shown in Fig. 3(b). Dimensionless lens characteristics of T/F, $-\{(C_c/T)/(T/F)\}/\{L_o/T\}^2[=2(V_o/V_{\mathrm{av}})]$ and $\{(C_s/T)/(T/F)\}/\{L_o/T\}^4[=(T/r_e)^2]$ are plotted as a function of V_2/V_1 in Fig. 4. The curves satisfy the relationships between accelerating and decelerating lens modes, i.e., Eqs. (55a)–(55c).

The third is the bell-type unipotential lens,[4] in which the potential is expressed by

$$V(z)=V_o\{1-b/(1+(z/d)^2\}, \tag{62}$$

and is plotted in Fig. 3(c). Using Eqs. (10b), (24), (29), and $a[=V(0)/V_o]=1-b$, we successfully obtain analytical forms of d/F and $(C_c/d)/(L_o/d)^2$ as

$$d/F=(3/8)\{\pi/(1-a)\}\{(1+3a)/a^{1/2}-3-a\}, \tag{63}$$

$$(C_c/d)/(L_o/d)^2[=-2(V_o/V_{\mathrm{av}})(d/F)], =-(3/16)\{\pi/(1-a)\}\times\{8-4(1+a)/a^{1/2}+(1-a)^2/a^{3/2}\}. \tag{64}$$

The characteristics curves of d/F, $-\{(C_c/d)/(d/F)\}/(L_o/d)^2$, and $\{(C_s/d)/(d/F)\}/\{L_o/d\}^4$ are plotted as a function of a in Fig. 5. As the potential curves in the regions of $z\geq 0$ and $z\leq 0$ do not strictly satisfy the condition of $g(u)=-g(-u)$, modifications are required in the relationships between accelerating and decelerating lens modes, i.e., Eqs. (55a)–(55c).

Finally, if the V_{av} value in the unipotential lens is approximated to a geometric mean of V_1 and V_2, i.e., $(V_1V_2)^{1/2}$, instead of the harmonic mean, i.e., $(1/V_1+1/V_2)/2$, as obtained from Eq. (56), Eq. (43) yields

$$C_{c,\infty}=-2F(V_2/V_1)^{-1/2}. \tag{65}$$

This is the same expression that Crew has empirically proposed.[9]

VI. SINGLE- AND TWO-LENS OPTICAL SYSTEMS

An ion probe size is given approximately by adding in quadrature the Gaussian image size to the chromatic- and spherical-aberration disk size[1–3]

$$d^2=d_g^2+d_c^2+d_s^2, \tag{66}$$

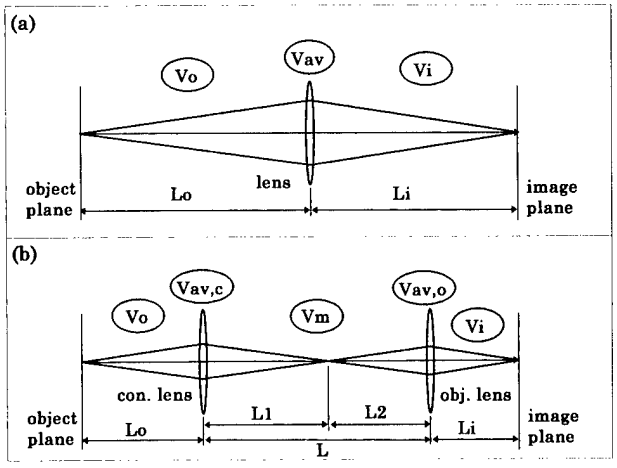


FIG. 6. Schematic diagram of the electrostatic lens systems; (a) a single-lens system and (b) a two-lens system.

where

$$d_g = m\rho, \tag{67}$$

$$d_c = H_c \alpha_0 \Delta V, \tag{68}$$

$$d_s = (1/2) H_s \alpha_0^3, \tag{69}$$

$$H_c = m C_c / V_o, \tag{70}$$

$$H_s = m C_s. \tag{71}$$

Here, ρ is the virtual source size and ΔV is the effective voltage spread of the ions. Equation (66) is applicable also for the electrons when their diffraction effects are negligible in the probe calculation. For the single-lens optical system [see Fig. 6(a)], by inserting Eqs. (35) and (36) into Eqs. (70) and (71), respectively, and by successively eliminating $1/F$ and m using Eqs. (4) and (18), respectively, we obtain

$$H_c = -(2/\sigma)(L_o + L_i)/V_{\mathrm{av}}, \tag{72}$$

$$H_s = (1/64)(1/\sigma)(L_o^3 L_i) G_{\mathrm{int}}, \tag{73a}$$

$$= (1/\sigma)(L_o + L_i)(L_o / r_e)^2. \tag{73b}$$

Combining Eqs. (5) and (7) with (18), we obtain m as

$$m = (V_o / V_i)^{1/4} (L_i / L_o). \tag{74}$$

The axial performance of the FIB optical system is roughly characterized by the maximum current density in the region of $d \approx d_c$,[1-3]

$$J_{\max} = [4(dI/d\Omega)/(\Delta V)^2]/H_c^2, \tag{75}$$

where $(dI/d\Omega)$ is the source angular intensity. Under a constant lens field (i.e., F, σ, V_{av}, and G_{int}=constant), it is found that H_c of Eq. (72) is minimum at $L_o = L_i = 2F$ while H_s of Eq. (73a) is minimum at $L_o = 4F/3$ and $L_i = 4F$.

Next, we consider a two-lens system consisting of condenser and objective lenses at $V_m = V_i$ as shown in Fig. 6(b). A similar analysis to that for the single-lens system has been carried out in the previous articles.[2,3] Only the main results are described here. The total H coefficient is expressed as the sum of each lens's coefficient,

$$H_{c,t} = H_{c,c} + H_{c,o}, \tag{76}$$

where

$$H_{c,c} = 2(V_o / V_i)^{1/4} (L_i / L_2)(L_o + L_1)(1/V_{\mathrm{av},c}), \tag{77a}$$

$$H_{c,o} = 2(V_o / V_i)^{1/4} (L_o / L_1)(L_i + L_2)(1/V_{\mathrm{av},o}), \tag{77b}$$

$$m = (V_o / V_i)^{1/4} (L_i / L_o)(L_1 / L_2). \tag{78}$$

Assuming that $V_{\mathrm{av},c}$ and $V_{\mathrm{av},o}$ depend weakly on L_1 under the conditions of L_o, L_i, L_o and $L_i \ll L$, and $L (= L_1 + L_2)$ = constant, we obtain the optimum L_1 / L_2 to minimize $H_{c,t}$ as

$$L_1 / L_2 = \pm (L_o V_{\mathrm{av},c})^{1/2} / (L_i V_{\mathrm{av},o})^{1/2}, \tag{79}$$

$$H_{c,t,\min} = \pm 4 (V_o / V_i)^{1/4} (L_o L_i)^{1/2} / (V_{\mathrm{av},c} V_{\mathrm{av},o})^{1/2}, \tag{80}$$

and $H_{c,t,\min}/2 = H_{c,c} = H_{c,o}$. The sign $-$ and $+$ denote the collimated (i.e., $L_1 < -L_o$ or $L_1 > L + L_i$) and intermediate crossover (i.e., $0 < L_1 < L$) cases, respectively. Equation (80) indicates that the smaller the $(V_o/V_i)^{1/4}(L_o L_i)^{1/2}/(V_{\mathrm{av},c} V_{\mathrm{av},o})^{1/2}$ value is, the finer the d_c beam is. Particularly, as to the paralleled case (i.e., $L_1/L_2 = -1$), we obtain

$$H_{c,c} = -2(V_o / V_i)^{1/4} (L_i / V_{\mathrm{av},c}), \tag{81a}$$

$$H_{c,o} = -2(V_o / V_i)^{1/4} (L_o / V_{\mathrm{av},o}), \tag{81b}$$

$$m = -(V_o / V_i)^{1/4} (L_i / L_o). \tag{82}$$

VII. CONCLUSION

We have analyzed the paraxial electron/ion trajectories in rotationally symmetric fields in the Picht space to derive the analytical forms of the lens characteristics, i.e., focal length F, magnification M, and chromatic and spherical aberration coefficients C_c and C_s. These characteristics are expressed as functions of L_o, L_i, and V_{av} and the normalized lens potential distribution $g(z)$. The dimensionless forms of T/F, $\{(C_c/T)/(T/F)\}/\{L_o/T\}^2 [= -2(V_o/V_{\mathrm{av}})]$ and $\{(C_s/T)/(T/F)\}/\{L_o/T\}^4 [= (T/r_e)^2]$ are obtained to evaluate the lens. On the bipotential or unipotential lenses composed of electrodes facing each other symmetrically, it is found that the accelerating mode at $V_2/V_1 = a (>1)$ and the decelerating mode at $V_2/V_1 = 1/a (<1)$ have relationships of $F_{\mathrm{accel}} = F_{\mathrm{decel}}$, $C_{c,\mathrm{accel}} = C_{c,\mathrm{decel}}/a$, and $C_{s,\mathrm{accel}} = C_{s,\mathrm{decel}}/a^{1/2}$, where V_1 and V_2 are the first and second electrode potentials, respectively. Besides, on the aberration disk sizes for the one- and two-lens systems, we have derived their analytical forms as functions of L_o, L_i, L, V_o, V_i, and V_{av}. In the single-lens system under the condition of F=constant, d_c and d_s are proportional to $(L_o + L_i)$ and $(L_o^3 L_i)$, respectively, and they are minimum at $(L_o, L_i) = (2F, 2F)$ and $(4F/3, 4F)$, respectively. On the two-lens system, the parameters to minimize d_c have been clarified in Eq. (77) and the optimum L_1/L_2 value is given by Eq. (79). The M-dependent C_c and C_s forms under the constant lens fields are analytically derived as $C_c = C_{c,\infty}(1 - 1/M)^2$ and $C_s = C_{s,\infty}(1 - 1/M)^4$, which have already been reported in other approaches.[6,7]

[1] Y. Kawanami, T. Ohnishi, and T. Ishitani, J. Vac. Sci. Technol. B **8**, 1673 (1990).
[2] T. Ishitani and Y. Kawanami, J. Vac. Sci. Technol. B **13**, 371 (1995).
[3] T. Ishitani and Y. Kawanami, Proc. SPIE **2522**, 430 (1995).
[4] K. Ura, *Electron/Ion Beam Optics* (Kyouritsu, Tokyo, 1994), Chap. 4 (in Japanese).
[5] A. B. El-Kareh and J. C. J. El-Kareh, *Electron Beams, Lens, and Optics* (Academic, New York, 1970), Vol. 1, Chap. II.
[6] A. Renau and D. W. O. Heddle, J. Phys. E: Sci. Instrum. **19**, 288 (1986).
[7] K. Ura, J. Electron. Microsc. **43**, 406 (1994).
[8] K. Kuroda, H. Ebisu, and T. Suzuki, J. Appl. Phys. **45**, 2336 (1974).
[9] A. V. Crew, Optik (Stuttgart) **88**, 118 (1991).

X-ray photoelectron spectroscopy analyses of oxide-masked organic polymers etched in high density plasmas using SO_2/O_2 gas mixtures

C. Monget[a)]
France Telecom, CNET/DTM, BP 98, 38243 Meylan cedex, France

O. Joubert
France Telecom, CNET/DTM, BP 98, 38243 Meylan cedex, France and Laboratoire des Technologies de la Microélectronique, Centre National de la Recherche Scientifique EP 2073, 17 rue des Martyrs, 38054 Grenoble cedex 9, France

(Received 22 December 1998; accepted 5 April 1999)

Top surface imaging processes are considered as attractive schemes for future lithographic processes. A major issue associated with these schemes is the dry development step which implies the obtention of anisotropic resist profiles. This anisotropy is achieve by adding SO_2 to O_2 gas mixtures. In this work, oxide-masked organic resist features are etched in high density plasma sources using SO_2/O_2 gas mixtures. The chemical constituents present on the tops, sidewalls, and bottoms of the etched features are determined by quasi *in situ* x-ray photoelectron spectroscopy (XPS). XPS analyses show that when using SO_2/O_2 gas mixtures, a thin passivation layer is formed on the sidewalls of the polymer features. This layer is mainly formed by sulfur atoms present on the polymer surface in elemental state or bonded to carbon atoms. This film is thick enough to minimize the spontaneous etching reaction of oxygen atoms with the organic resist. Samples are also analyzed after air exposure in order to determine the air-induced modification of the reactive layers formed during the dry development step. XPS analyses show that the sulfur-based passivation layer formed during the etch process interact with water moisture to form sulfur-based acids. © *1999 American Vacuum Society.* [S0734-211X(99)00504-1]

I. INTRODUCTION

Top surface imaging (TSI) and multilayer resist schemes have received much attention as attractive technologies for the future lithography generations (193 nm lithography,[1] extreme ultraviolet lithography,...)[2] as these technologies offer some advantages for the future lithographic processes. In particular, high resolution can be achieved with TSI with a workable process window. Moreover, these processes reduce the influence of local reflection over topography. Top surface imaging using the silylation technique is a technology that has been studied during the last 15 years.[3] In this scheme, the image is created in the uppermost layer of the resist. An alternative technology based on the plasma enhanced chemical vapor deposition of a photosensitive material has been proposed.[4] In this case, a very thin photosensitive layer is formed by the plasma polymerization of methylsilane (PPMS) on top of an organic resist. In both approaches, this initial image is then transferred into the underlying thick organic resist layer using a dry development step based on O_2. During the dry development step, an anisotropic transfer has to be achieved to maintain the initial dimension obtained after the TSI process.

Two limitations of organic film patterning using pure O_2 reactive ion etching (RIE) plasmas have been pointed out over the last few years:[5] significant undercut due to lateral etching under the mask and critical dimension control issues. Two methods have been proposed to minimize these problems. Good profiles are obtained using pure O_2 RIE plasmas at cryogenic temperatures.[6–9] A second method consists in using additives gases to O_2 which may induce the formation of a passivation layer on the sidewalls of the polymer. Hutton, Boyce, and Taylor proposed the addition of CO_2,[10] while other workers used mixtures of O_2 and N_2.[11] Best results were obtained by using mixtures of SO_2 and O_2.[6] Pons, Peileher, and Joubert suggested that sulfur atoms do not spontaneously react with carbon from the polymer to form volatile species when the substrate temperature is below 150 °C and that the anisotropy is obtained by the formation of a passivation layer which minimizes the lateral etching.[12] Several XPS analyses have demonstrated the presence of S atoms on blanket photoresist surfaces,[10,13] but no direct evidence of the presence of sulfur on the sidewall of the features has been reported.

In this work, the dry development of organic films is studied in a high density plasma Helicon source using SO_2/O_2 gas mixtures. We have used quasi *in situ* XPS to determine the chemical composition of SiO_2 masked resist features after etching and validate the sidewall passivation mechanism often invoked to explain the anisotropy enhancement. Influence of air exposure before analysis will also be presented.

II. EXPERIMENTAL PROCEDURES

Etching experiments are conducted in an Helicon source connected to a cluster tool dedicated to etching processes characterization (described in more detail in a previous article).[14] It is composed of a load-lock chamber, a transfer chamber, a reactor chamber, and an XPS analysis chamber. The Helicon source was made by Lucas Labs®[15,16] and is operated at a radio frequency (rf) frequency of 13.56 MHz.

[a)]Electronic mail: cedric.monget@cnet.francetelecom.fr

The sample can be independently biased using a 600 W maximum power supply at 13.56 MHz. The plasma generation region is equipped with two solenoid magnets which are necessary to initiate the plasma, whereas two other solenoid magnets around the diffusion region maintain a high plasma density on the sample. The tuning of the matching network is computer controlled and performed automatically. The etching process is monitored using *in situ* real-time ellipsometry.

XPS analyses of samples etched in the Helicon source are performed quasi *in situ*. After etching, the sample is immediately transferred under vacuum into the XPS analysis chamber. A special mask, already described in detail elsewhere,[17] has been designed for XPS analyses. The areas dedicated to XPS analysis consists of regular arrays of trenches, blanket substrate, and unpatterned mask material. As the area of each array is at least 1.5 mm^2, the x-ray beam (spot size: 1 mm^2) can thus be focused on one individual array, the resulting signal being the average of many identical lines and spaces. XPS analyses are conducted with the lines aligned perpendicularly or in parallel with the electron energy analyzer. In the first case, the analyzed area is the top and, depending on the space between lines, parts of the sidewalls of the features. When the lines are aligned in parallel with the electron energy analyzer, the analyzed areas are the tops and bottoms of the features. An electron gun emitting low energy electrons (less than 80 eV) is used to charge analyzed surfaces, allowing contributions from insulating surfaces (oxide hard mask) and conducting surfaces to be differentiated. Chemical composition are derived from the areas of the different XPS spectra. Spectral deconvolution is performed to extract the Si $2p$, O $1s$, C $1s$, and S $2p$ peak intensities. Individual line shapes are simulated with the convolution of Lorentzian and Gaussian functions. A Shirley function is used to perform background subtraction.[18] The elemental Si $2p$ and S $2p$ peaks are fitted with a doublet constituted of the two spin-orbit components $2p1/2$ and $2p3/2$. Fits are performed forcing an identical full width at half maximum for the two spin-orbit components and an intensity ratio of 0.52 for Si $2p$ and 0.5 for S $2p$. We verified that a doublet separation of 1.2 eV for Si $2p$ is obtained.[19] After XPS analyses, the integrated intensities are divided by the theoretical Scofield cross sections (0.82 for Si $2p$, 1.00 for C $1s$, 1.68 for S $2p$, and 2.93 for O $1s$).[20] As the analysis angle between all the analyzed areas and the electron energy analyzer is 45° whatever is the configuration (parallel or perpendicular modes), direct comparisons can be drawn between the coverage of the different elements present on the tops, sidewalls, and bottoms of the features. The results of the XPS analyses described in Sec. IV represent the tendencies of between 5 and 10 wafers analyzed under the same conditions.

Samples are 200 mm diam wafers covered with three different layers. The trilevel system (see Fig. 1) used consists of a hard baked Novolak resist from Olin as the bottom layer. A 0.7-μm-thick layer is spin coated and baked under air at a temperature of 250 °C during 5 min. Oxide patterns (150 nm thick) are then fabricated using conventional lithographic

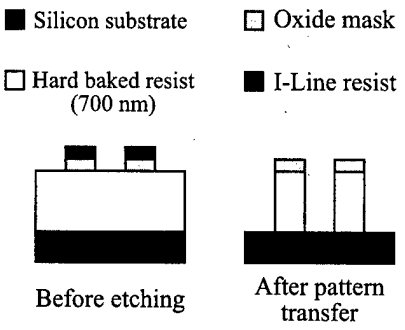


FIG. 1. Schematic view of the trilevel structure.

and etching steps. The thick I-line resist used for the oxide patterns definition is not removed at this point, but is removed during the dry development step.

III. PROCESS DEVELOPMENT

A. Pure oxygen

The Helicon source is operated at a rf power of 1500 W to minimize the O_2 dissociation in the plasma. The pressure is fixed at 2 mTorr, the O_2 flow at 60 sccm. The bias power is fixed to 200 W providing ion energies higher than 150 eV. Endpoint detection is performed using *in situ* real-time ellipsometry. Same conditions are used for the overetch step. The overetch time is fixed to 25%. The duration of the process (taking into account the overetch step) is 85 s (the etch rate of the hard baked resist is around 700 nm/min). Figure 2 shows the etching profile obtained with this process. A severe undercut is observed in the hard baked resist due to spontaneous etching reactions between oxygen atoms and the polymer.

B. SO_2/O_2 gas mixture

SO_2 is added to the gas mixture to promote the formation of a sulfur-based passivation layer on the sidewalls of the

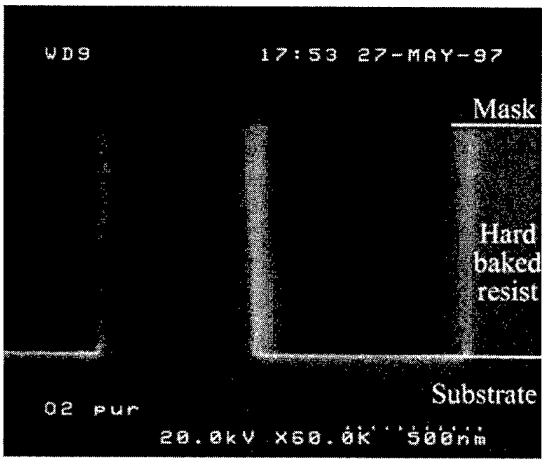


FIG. 2. Oxide-masked hard baked resist features etched in a high density plasma etcher using pure O_2. A severe undercut is observed.

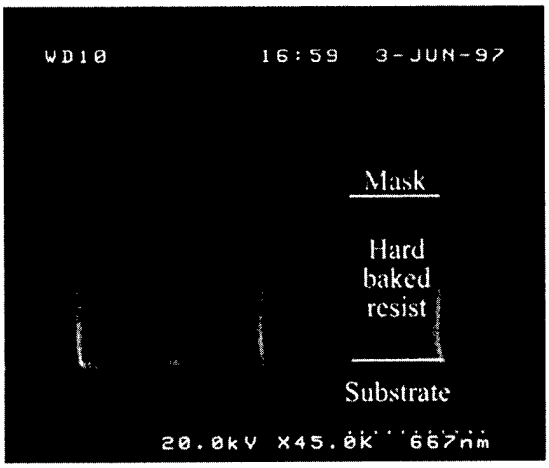


FIG. 3. Oxide-masked hard baked resist features etched in a high density plasma etcher using SO_2/O_2 gas mixture. Anisotropic profile are obtained with only 30% of SO_2 in the gas phase.

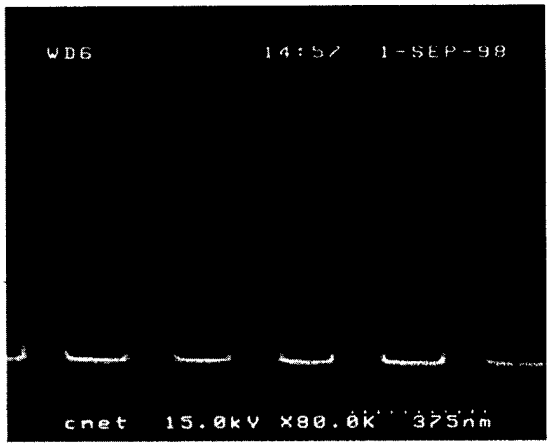


FIG. 4. 0.15 μm l/s features patterned with PPMS resist at 193 nm and developed in a high density plasma etcher using a SO_2/O_2 gas chemistry.

organic materials. Process conditions are identical as before. SO_2 flux (18 sccm) is 30% of the total flux (60 sccm). With SO_2 addition into the gas phase, the etching time increases to 100 s. Even if organic film etch rate decreases to 525 nm/min, anisotropic profiles are obtained as shown in Fig. 3. This experiment shows that adding SO_2 to the gas mixture is a very efficient way to enhance the anisotropy of the etching process. The etching anisotropy obtained is expected to be due to the formation of a passivation layer on the sidewall of the features.

In a top surface imaging scheme, the final resolution obtained is the convolution of the aerial image of the light at the resist surface and the quality of the pattern transfer. To demonstrate that the SO_2/O_2 gas mixture can provide a very high pattern transfer quality, a plasma polymerized methylsilane (PPMS) photosensitive film was deposited on the top of 0.5 μm hard baked resist.[21] Exposures were performed using an experimental tool working at 193 nm.[22] PPMS development was achieved in a high density plasma etcher using pure chlorine and pattern transfer was performed in the same etching tool using the SO_2/O_2 gas mixture. Resolution down to 0.13 μm were obtained at 193 nm at a dose of 15 mJ/cm^2. Figure 4 shows 0.15 μm line/space (l/s) features obtained with this process. The high selectivity between the oxidized PPMS layer and the organic materials allows obtaining high aspect ratio structures with vertical sidewalls. These results show that PPMS is a promising 193 nm negative-tone bilayer resist which combines resolution, sensitivity, and plasma etch resistance.

IV. CHEMICAL DISTRIBUTION ANALYSES OF OXIDE MASKED POLYMER FEATURES DRY DEVELOPED WITH SO_2

In order to analyze the chemical composition of the sidewall passivation layer, samples were etched in the Helicon source using pure SO_2 under the process conditions described earlier (source power: 1500 W, bias power: 200 W, pressure: 2 mTorr). The etching was stopped in the hard baked resist layer 200 nm before reaching the silicon substrate (see Fig. 5). After etching, samples were transferred under vacuum into the XPS analysis chamber. Figure 5 shows the 0.5 μm l/s features etched in the Helicon source and analyzed by XPS.

A. Blanket hard baked resist analyses

XPS analyses are first conducted on a blanket hard baked resist without any plasma exposure. Three elements were detected by XPS (C, O, and S). The percentage of the chemical elements obtained from the XPS spectra after spectral deconvolution are summarized in Table I. The atomic percentage of sulfur detected in the photoresist is only 1% (sulfur is present in the chemical composition of the photoactive compound of the photoresist). The S $2p$ peak is observed around 169 eV, which means that sulfur present in the photoresist is in the highest oxidation state like such as (SO_4^{2-}, $2H^+$) and SO_3 (the chemical shift is around 5 eV relative to its elemental state).[23]

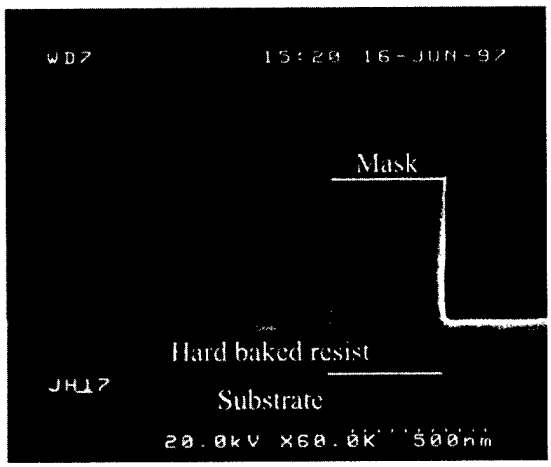


FIG. 5. Oxide-masked hard baked resist features (0.5 μm l/s) partially etched using pure SO_2. The remaining thickness of the resist is about 200 nm.

TABLE I. Element concentrations determined on the surfaces of oxide-masked hard baked resist etched using the pure SO_2 process.

Analyzed area	C 1*s* (in %)	O 1*s* (in %)	S 2*p* (in %)	S 2*p* ox (in %)	Si 2*p* (in %)	SiO_x (in %)
Hard baked resist	62.3	36.7	···	1.0	···	···
Hard baked resist after etching	75.4	12.0	12.1	0.5	···	<0.5
Oxide mask	0.7	63.0	2.3	0.8	···	33.1
Sidewalls of the features	72.1	19.6	8.2	···	···	···

B. Hard baked resist after etching

Blanket hard baked resist surface is also analyzed after the SO_2/O_2 etching process. An XPS survey spectrum reveals that carbon, oxygen, sulfur, and silicon are present on the etched surface after the process. A high resolution spectra of Si 2*p* peak points out that the Si 2*p* contribution is located at 102 eV and is only present as a trace (Si 2*p* signal was at the limit of detection of the XPS system). The presence of silicon on the etched surface can be explained by the erosion of the SiO_2 mask (see Fig. 5). SiO_2 etching products are sputtered by the high energy ion bombardment and redeposited on the blanket resist surface. Figure 6 shows the S 2*p* spectra recorded on the hard baked resist surface with the charge neutralizer turned off. Two contributions are detected in the S 2*p* XPS spectra. The first is the S 2*p* doublet located at a binding energy of 164 eV. This peak indicates that elemental sulfur or C–S bonds are present at the surface of the resist.[19] The other S 2*p* peak located at around 168 eV is attributed to the presence of sulfur oxides. This contribution is already present in the hard baked resist before the etching process (see Sec. IV A), but can also be formed by the plasma oxidation of the sulfur containing polymer formed on the resist surface. O 1*s* and C 1*s* spectra are also recorded (not shown here). The quantitative coverage of the different elements derived from the XPS spectra presented above are summarized in Table I. A decrease in the oxygen concentration (compared to the unetched resist surface) is measured. In fact, O 1*s* signal coming from the hard baked resist is shielded by the passivation layer formed at the surface, which is mainly made of carbon and sulfur.

C. Blanket oxide surface after etching

XPS experiments are also performed on blanket SiO_2 surfaces. O 1*s* and Si 2*p* contributions (not shown here) are detected at 533.3 and 103.3 eV, respectively, which is in good agreement with the theoretical positions. No positive charging is observed on the oxide hard mask. A very weak C 1*s* contribution is also recorded on blanket SiO_2 surfaces. Two contributions are detected in the S 2*p* spectra. The first is located at 164 eV and the other one at 169 eV (see Fig. 7). The 164 eV peak is attributed to elemental sulfur or C–S based compounds present on the oxide hard mask. The peak located at 169 eV corresponds to oxidized sulfur species such as $(SO_4^{2-}, 2H^+)$ species. This contribution is originating from elemental sulfur or C–S compounds oxidized by the oxygen atoms of the plasma.

D. Chemical analyses of the sidewalls of the organic polymers

Si 2*p* XPS spectra (not shown here) are recorded in 0.5 μm l/s features with the lines aligned perpendicularly to the electron energy analyzer (in this case, the signal originates from the top and the sidewall of the oxide hard mask and

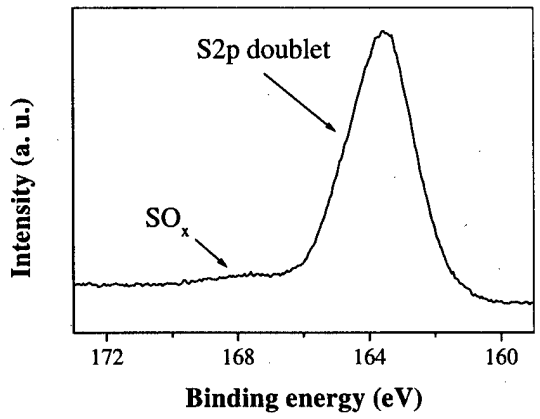


FIG. 6. S 2*p* XPS spectra recorded on a blanket hard baked resist after etching. Sulfur signal is coming from the passivation layer formed on the surface of the resist.

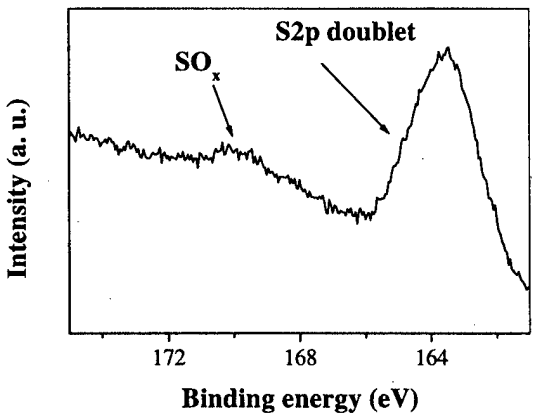


FIG. 7. S 2*p* XPS spectra recorded on blanket oxide-masked area after etching. Two contributions are observed: C–S compound at 164 eV and oxidized sulfur compound at 169 eV.

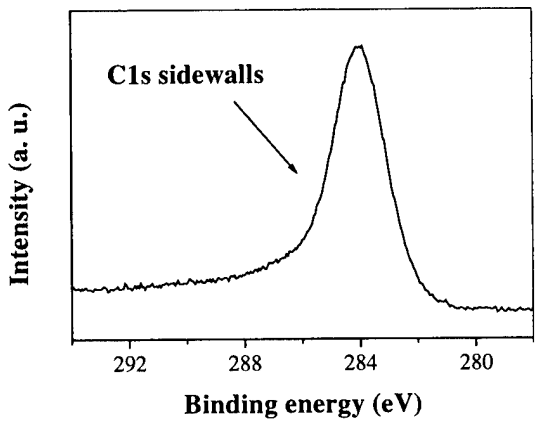


FIG. 8. C 1s XPS spectra recorded on 0.5 μm l/s features with the lines aligned perpendicularly to the electron energy analyzer and the charge neutralizer turned off.

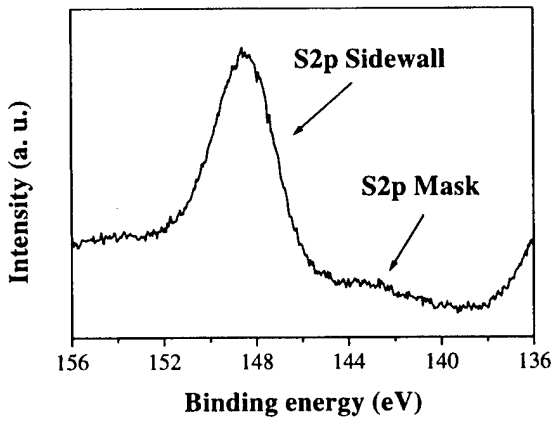


FIG. 9. S 2p spectra recorded in 0.5 μm l/s lines aligned perpendicularly with the electron energy analyzer and with the charge neutralizer turned on. Elemental sulfur or C–S containing species are detected on the sidewalls of the hard baked resist.

from the sidewalls of the hard baked resist). When the charge neutralizer is turned off, a broad peak is detected at a binding energy of 104.8 eV. This peak is shifted to lower binding energy (81.8 eV) and broadened when the charge neutralizer is turned on. No other contributions are detected with the charge neutralizer turned on. This contribution is attributed to the Si 2p signal originating from the top and the sidewall of the oxide hard mask. The shift to the higher binding energy (1.5 eV) when the charge neutralizer is turned off is due to the positive charging of the SiO_2 surface (theoretical position: 103.3 eV). A calibration in energy of all the XPS contribution coming from the mask is performed when the charge neutralizer is off: the positions of XPS peaks are then corrected by taking into account the +1.5 eV shift. When the charge neutralizer is turned on, XPS spectra attributed to the oxide hard mask are calibrated by taking into account the 21.5 eV shift induced by the low energy electrons flux.

Figure 8 shows the C 1s spectra recorded in 0.5 μm l/s features with the lines aligned perpendicularly to the electron energy analyzer. When the charge neutralizer is turned off, the C 1s peak is located at 284 eV. As very low carbon concentrations are detected on the mask (see Sec. IV C), this C 1s signal originates from the sidewalls of the features. When the charge neutralizer is turned on, the C 1s peak is shifted towards lower binding energies (270 eV). The shift in energy due to the charge neutralizer is therefore only 14 eV for the elements present on the sidewalls of the hard baked resist. The higher shift measured on the oxide hard mask (21.5 eV) with respect to the resist (14 eV) is attributed to the higher electron flux collected by the oxide hard mask than by the polymer sidewalls (in our system, the low energy electron flux is flood on the analyzed surfaces with an angle of 30° with respect to the wafer plane: low energy electron mainly reach the oxide hard mask).

The S 2p spectra recorded in 0.5 μm l/s features with the charge neutralizer on is shown on Fig. 9. When the charge neutralizer is turned off, we observe only a broad peak, located at around 164 eV. This peak is a convolution of the signal coming from the mask and from the sidewalls of the features. A perfect differentiation between signals originating from the top of the features (the oxide hard mask) and the sidewalls of the features (hard baked resist) is possible when the charge neutralizer is turned on: assignation of the peaks is made using the shifts in energy determined before (21.5 eV for the signal coming from the mask and 14 eV for the signal coming from the organic polymer sidewalls). The spectra shown in Fig. 9 exhibits two peaks. The first one (located at the shifted binding energy of 149 eV) is attributed to the S 2p doublet originating from the organic resist sidewall whereas the other peak (located at a binding energy of 143 eV) is associated with the S 2p doublet due to sulfur compounds present on the top of the mask. Based on these results, it is now possible to analyze the signal recorded when the charge neutralizer is turned off (see Fig. 10). A key point which needs to pointed out, is that XPS peaks due to sulfur oxides on the top of the mask or on the organic poly-

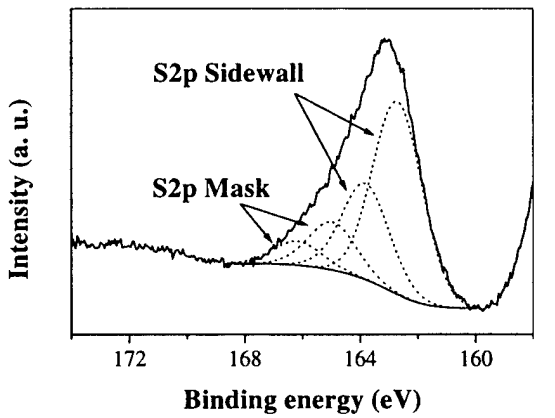


FIG. 10. S 2p spectra recorded in 0.5 μm l/s lines aligned perpendicularly with the electron energy analyzer and with the charge neutralizer turned off. The doublet located around 165 eV is originating from S 2p signal coming from the top of the oxide hard mask whereas the doublet at 164 eV is attributed to the passivation layer formed on the sidewalls of the features.

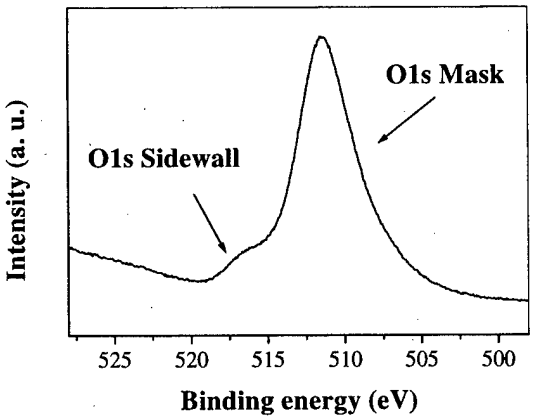


FIG. 11. O 1*s* XPS spectra recorded on 0.5 μm l/s features with the lines aligned perpendicularly to the electron energy analyzer and the charge neutralizer turned on. A clear distinction between the signal coming from the top of the mask and the sidewalls of the features is possible when the charge neutralizer is turned on.

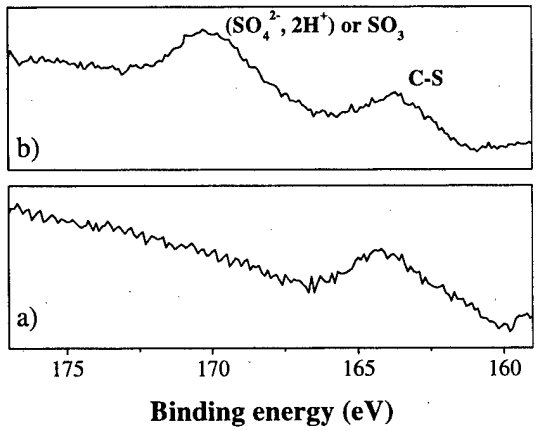


FIG. 12. S 2*p* XPS spectra recorded on 0.5 μm l/s features with the lines aligned perpendicularly to the electron energy analyzer with the charge neutralizer turned off. The sample is transferred under vacuum in the analysis chamber (a) or left 1 h under air before the transfer in the analysis chamber (b).

mer sidewalls are not detected (these species are only present as traces).

As with S 2*p* XPS spectra, a clear separation between O 1*s* signal coming from the mask and the sidewalls of the hard baked resist is obtained when the charge neutralizer is turned on (see Fig. 11).

When the lines are aligned in parallel with respect of the energy analyzer, the top and the bottom of the features are analyzed. At the bottom of the features, a sulfur-based passivation layer is also observed (spectra not shown here) whose composition is close to the layer formed in blanket resist area.

The quantitative composition of the surface derived from the spectra presented earlier are summarized in Table I.

V. CHEMICAL ANALYSES OF THE SIDEWALLS OF THE FEATURES AFTER AIR EXPOSURE

In situ XPS analyzes are interesting because they provide some information on chemical composition of surfaces exposed to the plasma allowing etching mechanisms to be better understood. However, in most cases, etched surfaces are exposed to air after plasma treatment and prior going to the next technological step of integrated circuit (IC) fabrication processes. Under air exposure, chemical reactive layers formed during plasma exposure get oxidized and may also react with water moisture. Air exposure can therefore deeply modify the chemical nature of the reactive layers formed during the plasma exposure that can have some important consequences on the subsequent technological steps. This statement is particularly true for plasma dry development processes which involve resist patterns fabrication. Indeed, resist patterns obtained by TSI techniques or multilayer resist schemes are supposed to be used in IC fabrication for gate, contact holes, or metal patterning.

In order to analyze the air-induced modification of the reactive layers formed on resist patterns after dry development using SO_2/O_2 gas mixtures, the previous sample is left 1 h under air and analyzed again by XPS. XPS spectra are recorded as earlier in blanket resist areas, oxide hard mask areas, 0.5 μm l/s features areas with the lines aligned perpendicularly to the electron energy analyzer. Main changes with respect to the *in situ* analyzes are observed in the S 2*p* spectra. Figure 12 shows the S 2*p* spectra recorded without air exposure [Fig. 12(a)] and after air exposure [Fig. 12(b)]. Only elemental or compounds with C–S bonds (corresponding to the peak located at 164 eV) are detected when the sample is transferred into the analysis chamber without air exposure. After air exposure, two oxidation states of sulfur are observed: the peak located at 164 eV is still detected and a new one (located at 169 eV) appears. This peak is attributed to the highest oxidation state of sulfur (+6, as in SO_4^{2-}) as mentioned earlier. The most likely compound seems to be H_2SO_4 (H is not detected by XPS). This compound is a result of the interaction of sulfur compounds present on the etched surfaces and water moisture.

VI. DISCUSSION

Anisotropic profiles of SiO_2 masked organic materials can be obtained in high density plasmas with the SO_2/O_2 chemistry whereas a severe undercut is observed under the SiO_2 hard mask when pure O_2 is used under the same plasma operating conditions as already reported in several studies.[5,10,12] The mechanism invoked to explain the gain in anisotropy is based on the formation of a passivation layer on the sidewalls of the features. Thermodynamic studies of the carbon/sulfur system predict that spontaneous reaction of sulfur with carbon may form volatile species such as CS_2 only when the temperature is above 150 °C.[24] Using XPS, we have clearly demonstrated that a passivation layer is formed on the sidewalls of the features. The passivation layer is mainly formed of carbon and sulfur atoms and is a result of the dissociation of SO_2 molecules in the plasma and the competitive absorption of sulfur and of oxygen atoms at the surfaces of organic materials. The thickness of this film cannot

be evaluated, but the layer formed on the sidewalls of the features is thick enough to minimize spontaneous etching reactions of oxygen atoms with the hard baked resist. A passivation layer is also formed on the surface of the resist at the bottom of the features. The formation of this layer only induces a decrease of the etching rate with respect to pure O_2 plasmas.

When the samples are transferred in the analysis chamber without air exposure, sulfur is mainly bonded to carbon atoms, but is also observed in its higher oxidation states (+6) in small concentrations. Huang, Melaku, and Nguyen suggested that oxidation of sulfur species is induced by the oxygen atoms present in the gas phase of the plasma.[13]

XPS analyzes after air exposure show that air exposure induces some severe modifications of the etched surfaces. Same elements are detected when the sample is exposed to air (carbon, silicon, oxygen, and sulfur). After air exposure, the XPS peak attributed to oxidized sulfur species now dominate. Taking into account the chemical shift of the S $2p$ peak (+5 eV), these species are attributed to SO_3 or (SO_4^{2-}, $2H^+$) compounds. In fact, as suggested by Huang, Melaku, and Nguyen the most likely compound seems to be sulfuric acid (H_2SO_4) as hydrogen is not detected by XPS.[13] Huang, Melaku, and Nguyen also pointed out that no effects are observed on TSI dry etch developments results if an *in situ* plasma clean is used to minimize the thickness of the film deposited on the reactor walls. However, the presence of H_2SO_4 on every etched surfaces can have an impact on the process steps following the dry development.

VII. CONCLUSION

Anisotropic etching profiles of oxide-masked organic materials can be obtained in a high density etcher using SO_2/O_2 gas mixture. The chemical topography of oxide-masked hard baked resist features etched using this chemistry has been investigated by quasi *in situ* XPS. Sulfur was detected on all the analyzed surfaces, mainly in elemental state or bonded to carbon atoms. A passivation film formed by nonvolatile sulfur containing species is observed on the sidewalls of the features. This film is thick enough to protect the hard baked resist from spontaneous etching reactions with the oxygen atoms of the gas phase of the plasma and explains the enhancement of anisotropy when SO_2 is added to the gas mixture.

When the wafers are exposed to air, the passivation layer interacts with air moisture to form sulfur-based acids. The existence of acids can have some major impact on the following technological steps.

ACKNOWLEDGMENTS

This work has been carried out within the GRESSI consortium between CEA–LETI and France Telecom-CNET.

[1]J. Hutchinson *et al.*, Proc. SPIE **3333**, 165 (1998).
[2]C. Henderson, D. Wheeler, T. Pollagi, G. Cardinale, D. O'Connell, A. Fisher, V. Rao, and J. Goldsmith, J. Photopolym. Sci. Technol. **11**, 459 (1998).
[3]G. N. Taylor, L. E. Stillwagon, and T. Venkatesan, J. Electrochem. Soc. **131**, 1658 (1984).
[4]O. Joubert, T. Weidman, R. Cirelli, S. Stein, J. T. C. Lee, and S. Vaidya, J. Vac. Sci. Technol. B **12**, 3912 (1994).
[5]M. Pons, O. Joubert, C. Martinet, J. Pelletier, J. P. Panabière, and A. Weill, Jpn. J. Appl. Phys., Part 1 **33**, 991 (1994).
[6]O. Joubert *et al.*, Proc. SPIE **1803**, 130 (1992).
[7]S. Tacji, K. Tsujimoto, S. Arai, and T. Kure, J. Vac. Sci. Technol. A **9**, 796 (1991).
[8]W. Varhue, J. Burroughs, and W. Mlynko, J. Appl. Phys. **72**, 3050 (1992).
[9]S. C. Palmateer, R. R. Kunz, M. W. Horn, and A. R. Forte, Proc. SPIE **2438**, 455 (1995).
[10]R. S. Hutton, C. H. Boyce, and G. N. Taylor, J. Vac. Sci. Technol. B **13**, 2366 (1995).
[11]C. W. Jurgensen, R. S. Hutton, and G. N. Taylor, J. Vac. Sci. Technol. B **10**, 2542 (1992).
[12]M. Pons, J. Pelletier, and O. Joubert, J. Appl. Phys. **75**, 4709 (1994).
[13]Z. S. Huang, Y. Melaku, and W. Nguyen, Jpn. J. Appl. Phys., Part 1 **37**, 2373 (1998).
[14]F. Bell, O. Joubert, and L. Vallier, J. Vac. Sci. Technol. B **14**, 96 (1996).
[15]Lucas Labs Inc., Sunnyvale, CA, 94089.
[16]N. Benjamin, B. Chapman, and R. Boswell, SPIE, Santa Clara, CA, 1990 (unpublished).
[17]F. H. Bell, O. Joubert, and L. Vallier, J. Vac. Sci. Technol. B **14**, 1796 (1996).
[18]D. A. Shirley, Phys. Rev. B **5**, 4709 (1972).
[19]C. D. Wagner and D. M. Bickham, NIST XPS Database (NIST, Gaithersburg, MD, 1989), version 1.0.
[20]J. H. Scofield, J. Electron Spectrosc. Relat. Phenom. **8**, 129 (1976).
[21]C. Monget, D. Fuard, O. Joubert, P. Schiavone, O. Toublan, J. S. Cashmore, J. M. Temerson, and R. L. Inglebert, Proceedings of the 4th International Symposium on 193 nm Lithography, Telfs (Austria), 14–17 September 1998.
[22]J. S. Cashmore *et al.*, Proceedings of the 4th International Symposium on 193 nm Lithography, Telfs (Austria), 14–17 September 1998.
[23]A. Fahlman, K. Hamrin, J. Hedman, R. Nordberg, C. Nording, and K. Siegbahn, Nature (London) **210**, 4 (1966).
[24]Thermodata, Thermodynamical Data Bank, BP 66, 38042 St. Martin d'Hères Cedex, France.

Fabrication and characterization of chromium based single-electron transistors with evaporated chromium oxide barrier tunnel junctions

Yu. A. Pashkin[a] and J. P. Pekola
Physics Department, University of Jyväskylä, 40351 Jyväskylä, Finland

L. S. Kuzmin
Physics Department, Chalmers University of Technology, S-41296 Göteborg, Sweden

(Received 22 October 1998; accepted 7 May 1999)

We fabricated chromium based single-electron transistors comprising small-area Cr/CrO_x/Cr tunnel junctions with an evaporated chromium oxide barrier. The transistors are fabricated using e-beam lithography with a bilayer resist and two-angle shadow evaporation. We describe the fabrication process and discuss the device characteristics. © *1999 American Vacuum Society.* [S0734-211X(99)06204-6]

I. INTRODUCTION

In recent years, technological progress in nanofabrication has made it possible to shrink the geometrical size of designed structures to below the submicron level with a characteristic capacitance in the subfemtofarad range. A new class of devices has emerged whose operation is based on Coulomb blockade effects.[1,2] Among them, the single-electron transistor (SET)[3,4] has proved to be one of the most simple and promising for practical applications. It consists of two junctions connected in series and a gate electrode attached to a small island formed between the junctions. In this circuit, a discrete tunnel current through the junctions is controlled by a continuous tiny variation of charge supplied by the gate. As follows from orthodox theory,[1] for this control to be observable, the total capacitance of the device C_Σ must be small compared to e^2/k_BT and the junction tunnel resistances must exceed the resistance quantum $R_Q \equiv h/4e^2 \approx 6.5$ kΩ. Theory predicts extremely low intrinsic charge noise for the SET[5] giving the possibility of using the SET as a highly sensitive electrometer. However, numerous experiments have shown that the performance of SETs (and, in general, most single-electron devices) suffers from background charge fluctuations.[2]

A lot of experimental data has already been accumulated on noise properties of metallic SETs, mostly on Al-based structures.[6–12] It is well known that the charge noise of SETs at low frequencies, $f \lesssim 1000$ Hz, substantially dominates over their intrinsic noise and usually has a *1/f* power spectrum with intensity in the range of 10^{-3}–10^{-4} $e/\mathrm{Hz}^{1/2}$ at 10 Hz. It is generally believed that the origin of excessive background charge noise is in the charge traps randomly distributed around the SET island.

A deep understanding of the nature of the background charge noise seems very important for practical implementation of single-electron devices and it may eventually lead to a substantial reduction in the noise level.

Several substrate materials (Si, SiO_2/Si, Si_3N_4/Si, and Al_2O_3) have been tried for the fabrication of SETs and those made on Si_3N_4/Si were found to have the worst noise properties.[8] This fact clearly indicates that the substrate material does affect the noise level of SETs. In one particular experiment, using a dual spectrum method it was found that the main source of charge noise is located in the substrate in the vicinity of the island.[11] By choosing a SET design where the island was effectively screened from the substrate by the base electrode, the noise level was reduced down to 2.5×10^{-5} $e/\mathrm{Hz}^{1/2}$ at 10 Hz,[12] the best noise figure reported so far for SETs operating at low frequency. The noise level can be further reduced if the SET is operated outside the *1/f* noise range, e.g., by coupling a SET to a high-frequency resonant circuit.[13]

An alternative way to study noise properties of SETs is to use a different (from commonly used aluminum) material for junction electrodes and tunnel barriers. Recently, a Cr-based SET with tunnel junctions having thermally grown native oxide barriers has been reported and its noise properties (noise level and frequency dependence) appeared to be similar to those of Al-based devices.[14]

In this article we present our data on the fabrication and measurement of chromium based SETs made by a conventional angle evaporation technique but with a novel procedure for barrier formation. The use of a normal metal (Cr) instead of a superconductor (Al) for the fabrication of a SET does not require the application of a high magnetic field and may be desirable in many cases.

II. FABRICATION

Small capacitance tunnel junctions forming SETs were fabricated using self-aligned angle evaporation through a suspended resist mask.[15] Fine lines and submicron bridges in the mask were patterned using electron beam lithography. The junctions were made by overlapping the bottom and top chromium electrodes. In contrast to Ref. 14, a tunnel barrier between the electrodes was formed by evaporation of a thin layer of chromium oxide. Below is a detailed description of the fabrication procedure utilized.

[a]On leave from Lebedev Physics Institute, Leninskii Prospekt 53, 117924 Moscow, Russia; present address: NEC Fundamental Research Laboratories, 34 Miyukigaoka, Tsukuba, Ibaraki 305-8501, Japan; electronic mail: pashkin@frl.cl.nec.co.jp

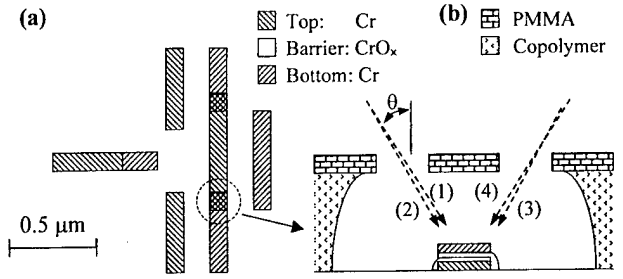


FIG. 1. (a) Layout of the circuit. (b) Schematic view of the evaporation procedure (not to scale). The evaporation sequence is (1) bottom Cr at $\theta=20°$, (2) CrO_x at $\theta=18°$, (3) CrO_x at $\theta=-18°$, (4) top Cr at $\theta=-20°$.

A 3 in. silicon wafer with a thermally grown oxide layer was used as the substrate. The wafer was spin coated by a standard bilayer resist with polymethyl (methacrylate, methacrylic acid) [P(MMA-MAA)] copolymer at the bottom and polymethylmethacrylate (PMMA) on the top. Such a resist is quite suitable for fabrication of nanostructures made of metals having reasonably low evaporation temperatures. The mask made in the resist can successfully withstand deposition of aluminum, gold, copper, silver, chromium, lead, indium, and tin to mention a few.

After the resist system was built up, the wafer was scribed and broken into small chips (~5×10 mm). Each chip was exposed separately in a modified scanning electron microscope (JEOL JSM-840A) which had a PC controlled electron beam. The exposure was performed in two steps. The first step took place immediately after the focus was adjusted, and the fine structure of the pattern with linear dimensions down to about 0.1 μm [see Fig. 1(a)] was exposed with a probe current of 30 pA. After that, the probe current was switched to 30 nA and the rough structure including contact pads (~0.25×0.5 mm) was exposed. By exposing fine and rough structures in one run, we avoid a photolithography step for making contact pads in a separate layer. The total exposure time for one pattern is about 1 min.

After development in two selective developers, the chip is placed into a vacuum system for evaporation of the metallic layers and fabrication of the tunnel junctions. For the formation of a tunnel barrier, we used the property of chromium to getter oxygen during evaporation. This was confirmed by our simple test prior to fabrication of the tunnel junctions. We evaporated thin chromium films (10 nm thick) with a constant rate (≈0.05 nm/s) but at various pressures of oxygen in the chamber and measured their sheet resistance at room temperature. The sheet resistance was about 100 Ω/□ when the residual gas pressure in the chamber was 2×10^{-7} mbar. It increased to 2.5 kΩ/□ at an oxygen pressure of 3×10^{-6} mbar and up to 7 MΩ/□ at 6×10^{-6} mbar. At even higher pressures, the films were completely insulating. We must note that such a method was not very accurate since the settings on the thickness monitor are not absolutely correct when chromium is substituted by chromium oxide during deposition. However, we believe that despite the uncertainty in the film thickness there is a tendency for the resistance to rise with the increase of oxygen pressure. Besides the increase of sheet resistance, we observed a drop of oxygen pressure during evaporation when the initial pressure was higher than 2×10^{-6} mbar. These two effects, the increase of sheet resistance and the decrease of oxygen pressure, clearly indicate that chromium reacts with oxygen during evaporation. Our idea was to use this property of chromium for the fabrication of a tunnel barrier.

The base pressure of the turbomolecular pumped system is typically below 10^{-7} mbar. The system is equipped with a tilting stage and an e-gun evaporator. The first evaporation of chromium is done with the stage tilted to an angle $+\theta=20°$ from normal incidence. The bottom electrode is evaporated from a graphite crucible at a rate of about 0.5 nm/s, to a thickness of 20 nm. After deposition of the bottom electrode, the pumping rate is decreased by 30% and oxygen is let into the chamber. The oxygen pressure is stabilized and kept constant manually in the range of $5\times10^{-5}-10^{-4}$ mbar. After the pressure is stabilized, the crucible is heated gradually. There is an important reason for this. During the heating, oxygen pressure tends to decrease and if there is no feedback it may drop to a very low level rapidly. Simultaneously, the evaporation rate increases and that causes a further pressure drop. This avalanche-like process should be stopped at its initial stage, and one needs some skill to be able to keep both the oxygen pressure and the evaporation rate constant. The typical evaporation rate for chromium oxide was 0.04–0.05 nm/s. An increase of the rate above this level makes the process less controllable and results in a pressure drop below 10^{-6} mbar and eventually in low resistance tunnel junctions. From a comparison of the mean free path corresponding to a pressure of 10^{-4} mbar (~1 m) and the distance between the crucible and the substrate in the vacuum chamber (~20 cm), one can conclude that oxidation of the chromium takes place on its surface in the crucible. Thus, the deposition of chromium oxide occurs as a result of the following sequence: first, the chromium surface oxidizes in the crucible and then the oxide is evaporated onto the substrate.

Chromium is known to have several oxides;[16,17] Cr_2O_3, CrO_3, and CrO_2 are the most common ones. They all have very different chemical and physical properties. Cr_2O_3 has a green color, while CrO_3 is dark red and CrO_2 is black. Cr_2O_3 is a very stable compound with excellent dielectric properties suitable for high quality tunnel junctions, CrO_3 is reported to be an *n*-type semiconductor, and CrO_2 is a ferromagnetic metal. The latter two decompose into Cr_2O_3 and O_2 at temperatures above 250 °C. After evaporation in oxygen, the Cr crucible was always a green color in our experiments and thus we tend to conclude that it was Cr_2O_3 that was evaporated as the tunnel barrier. However, stoichiometric measurements are needed in order to verify the composition.

The idea during the evaporation of chromium oxide was to cover the bottom electrode with Cr_2O_3 in such a way that there were no open parts which could be connected directly to the top electrode and shunt the junction in case of a small misalignment of angles. For this, one half of the oxide layer was evaporated at an angle slightly smaller than φ and another half at an angle slightly smaller than $-\theta$ [Fig. 1(b)] so

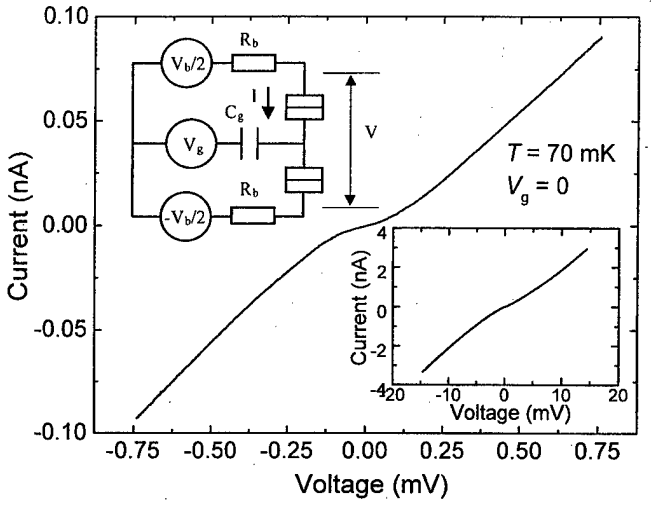


FIG. 2. I–V curve of Cr-based transistor; top inset: measurement circuit; bottom inset: I–V curve at high bias voltage.

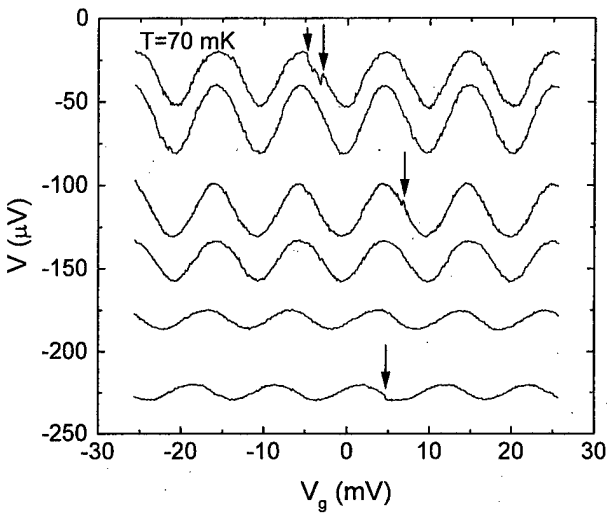


FIG. 3. Modulation curves at various biases. The arrows mark sudden jumps of the V–V_g curves caused by background charge fluctuations. The curves are not shifted in the y axis.

that the edges of the bottom electrode are adequately covered by the oxide layer. The total thickness of the evaporated oxide was 2.5 nm. This gave a tunneling resistance in the range of 10^5–10^6 Ω depending on the bias and sample temperature.

After formation of the barrier, the top chromium electrode was deposited to a thickness of 25 nm at an angle $-\theta$ at a rate of 0.05 nm/s. The use of a thick top resist layer and a higher evaporation rate in comparison to what was reported in Ref. 14 helped to avoid problems with overheating and cracking of the resist mask.

The sheet resistance of the evaporated chromium film was equal to 10 Ω/□, so that the resistance of the leads (~2 kΩ) was much smaller than the tunneling resistance, allowing two-probe measurements in the experiment as well.

III. MEASUREMENTS

Measurements were carried out in a plastic dilution refrigerator capable of reaching a temperature of 20 mK. All the measurement leads were filtered with RC filters located in a helium bath during the measurements. The measurement leads between the sample holder and the mixing chamber were Thermocoax™ cables that had good attenuation at high frequencies.[18] The coax cables entered the sample holder through a small hole and the outer shields of the cables were soldered to the sample holder. The inner conductors hung free in the sample holder. After the chip with a SET transistor was glued onto the copper sample holder, they were soldered with indium to the contact pads using ∅ 0.03 mm silver wires. Several transistors were fabricated and measured, and we report here characteristics of one typical sample.

We used a symmetric bias current as shown in the upper inset of Fig. 2. We could sweep or step the bias voltage V_b and the gate voltage V_g and measure the voltage drop V across and current I through the transistor. An example of the I–V curve at low bias is presented in Fig. 2. Suppression of the tunneling current around zero bias (Coulomb blockade) is clearly seen, however, the suppression does not develop into a perfect blockade. The ratio of zero bias resistance to resistance at, say, 10 mV is only a factor of 6. It is usually much higher in transistors having junctions with thermal oxide barriers, both Cr or Al based. Unlike Al-based devices, the I–V curve of our transistors does not have high bias asymptotes with constant slopes. As seen from the I–V curve taken at ±18 mV (bottom inset of Fig. 2), the differential resistance decreases with bias. Also, asymmetry between positive and negative branches of the I–V curve is visible.

We also measured modulation curves of the transistor. A family of V–V_g curves corresponding to different bias voltages is presented in Fig. 3. Because of the asymmetry of the I–V curve, the modulation amplitude at negative bias was about 10% higher than that at positive bias (not shown). The modulation curves remain quite stable at relatively low bias current but they tend to shift the phase at higher bias current. One can clearly see sudden jumps in the modulation curves, marked by arrows, that are very typical for the transistor. From the modulation curves, we estimated the capacitances of the transistor: $C_1 \approx 190$ aF, $C_2 \approx 160$ aF, and $C_g = 16$ aF giving $C_\Sigma = 370$ aF. Thus, the characteristic charging energy of the transistor was ≈2.4 K. The maximum modulation amplitude measured was about 40 μV (second curve from the top). Using this curve, we estimated the charge sensitivity of the device at the lowest temperature which was equal to 70 mK. By measuring root mean square (rms) fluctuations of the output voltage V, we calculated the input charge sensitivity to be $Q_N = 1.3 \times 10^{-2}$ $e/\mathrm{Hz}^{1/2}$ at 10 Hz.

IV. DISCUSSION

We measured several transistors with a layout similar to that in Fig. 1 and with similar junction areas. The transistor described above appeared to be the best one in terms of

modulation amplitude and charge sensitivity. We observed features that are common for all the transistors studied.

(1) The $I-V$ curves show visible nonlinearity at any bias voltage in the range ± 150 mV. The nonlinearity is much stronger in comparison to Cr transistors with thermally oxidized tunnel barriers.
(2) The resistance at zero bias was at most only six times larger than the resistance at 10 mV. In fact, we could observe modulation of the tunneling current only if the resistance at 10 mV exceeded 0.5 MΩ.
(3) Sudden jumps (telegraph-like noise with various time constants) in both the $I-V$ and $V-V_g$ curves are very typical for all our samples.
(4) The transistors turned out to be much noisier in comparison to Al- and Cr-based transistors with thermally oxidized tunnel barriers. The best noise sensitivity achieved was only $Q_N = 1.3\times 10^{-2}$ $e/\mathrm{Hz}^{1/2}$.

We believe the observed behavior of the transistors can be attributed to the insulating layer forming the tunnel barrier. The nonlinearity of the $I-V$ curves can be explained by the lower barrier heights of the $Cr/CrO_x/Cr$ tunnel junctions. Indeed, our estimation[19] of the barrier height φ based on a simple model[20] yields $\varphi \approx 175$ meV. This is similar to what was reported for Cr transistors with thermally oxidized tunnel junctions and a factor of 10 lower than what is known for Al tunnel junctions. This is supported by the strong temperature dependence of the differential resistance, at any bias, when the samples are cooled from 300 K to the low temperature.

Low zero bias resistance and jumps in the $I-V$ and $V-V_g$ curves are associated with poor quality insulating layers. Intuitively, the evaporated layer is not as uniform and stable as the naturally grown oxide layer. The surface of the evaporated layer is probably rougher and also the layer may contain some normal inclusions inside. This may result in a lower zero voltage resistance.

The measured charge sensitivity for our samples is more than an order of magnitude worse than what is known for Al and Cr devices with thermal oxide barriers made on SiO_2/Si substrates. From this we can conclude that, in our samples, the charge noise was dominated by the noise coming from the evaporated CrO_x layer which forms the tunnel barrier and also underlies the island.

Further studies of evaporated chromium oxide layers including stoichiometric and microstructure measurements would be of great value. Analysis of the stoichiometry of the layer would reveal whether or not it is composed of pure Cr_2O_3. Microstructure measurements could provide important information on roughness and, possibly, the porosity of the barrier layer.

In conclusion, we demonstrated that chromium-based single-electron transistors can be fabricated using a convenient angle evaporation technique with a tunnel barrier formed by evaporation of chromium in oxygen at a relatively high pressure. Although noise properties of such transistors appear to be worse in comparison to those with native oxide, their detailed study can shed light on the nature of background charge fluctuations in SET devices.

ACKNOWLEDGMENTS

This work was supported by the Academy of Finland, Jyväskylä Science Park, and the Swedish Natural Science and Technical Research Councils. One of the authors (Yu.A.P.) acknowledges partial support from the Russian Foundation for Basic Research (Grant No. 97-02-17056) and from the Russian Program on Solid State Nanostructures.

[1] D. V. Averin and K. K. Likharev, in *Mesoscopic Phenomenal in Solids*, edited by B. L. Altshuler, P. A. Lee, and R. A. Webb (Elsevier, Amsterdam, 1991), p. 173.
[2] *Single Charge Tunneling*, edited by H. Grabert and M. H. Devoret (Plenum, New York, 1992).
[3] K. K. Likharev, IEEE Trans. Magn. **23**, 1142 (1987).
[4] T. A. Fulton and G. J. Dolan, Phys. Rev. Lett. **59**, 109 (1987).
[5] A. N. Korotkov, D. V. Averin, K. K. Likharev, and V. A. Vasenko, in *Single-Electron Tunneling and Mesoscopic Devices*, edited by H. Koch and H. Lübbig (Springer, Berlin, 1992), p. 45.
[6] L. J. Geerligs, V. F. Anderegg, and J. E. Mooij, Physica B **165&166**, 973 (1990).
[7] G. Zimmerli, T. M. Eiles, R. L. Kautz, and J. M. Martinis, Appl. Phys. Lett. **61**, 237 (1992).
[8] L. Ji, P. D. Dresselhaus, S. Han, K. Lin, W. Zheng, and J. E. Lukens, J. Vac. Sci. Technol. B **12**, 3619 (1994).
[9] E. H. Visscher, S. M. Verbrugh, J. Lindeman, P. Hadley, and J. E. Mooij, Appl. Phys. Lett. **66**, 305 (1995).
[10] S. M. Verbrugh, M. L. Benhamadi, E. H. Visscher, and J. E. Mooij, J. Appl. Phys. **78**, 2830 (1995).
[11] A. B. Zorin, F.-J. Ahlers, J. Niemeyer, T. Weimann, H. Wolf, V. A. Krupenin, and S. V. Lotkhov, Phys. Rev. B **53**, 13682 (1996).
[12] V. A. Krupenin, D. E. Presnov, M. N. Savvateev, H. Scherer, A. B. Zorin, and J. Niemeyer, J. Appl. Phys. **84**, 3212 (1998).
[13] R. J. Schoelkopf, P. Wahlgren, A. A. Kozhevnikov, P. Delsing, and D. E. Prober, Science **280**, 1238 (1998).
[14] L. S. Kuzmin, Yu. A. Pashkin, A. N. Tavkhelidze, F.-J. Ahlers, T. Weimann, D. Quenter, and J. Niemeyer, Appl. Phys. Lett. **68**, 2902 (1996).
[15] G. Dolan, Appl. Phys. Lett. **31**, 337 (1977).
[16] *The Merck Index, An Encyclopedia of Chemicals, Drugs, and Biologicals*, edited by S. Budavari (Merck, Rahway, NJ, 1989), p. 346.
[17] *Landolt-Börnstein New Series*, edited by K.-H. Hellwege and O. Madelung (Springer, Berlin, 1984), Vol. III-17g, p. 242.
[18] A. B. Zorin, Rev. Sci. Instrum. **66**, 4296 (1995).
[19] Yu. A. Pashkin, J. Pekola, and L. S. Kuzmin, Nanostructures: Physics and Technology (NANO'96), Abstracts of Invited Lectures and Contributed Papers, St. Petersburg, Russia, 24–28 June 1996, pp. 192–195.
[20] J. G. Simmons, J. Appl. Phys. **34**, 1793 (1963).

Process sensing and metrology in gate oxide growth by rapid thermal chemical vapor deposition from SiH_4 and N_2O

Guangquan Lu,[a] Laura L. Tedder,[b] and Gary W. Rubloff[c]
NSF Engineering Research Center for Advanced Electronic Materials Processing, North Carolina State University, Raleigh, North Carolina 27695-7920

(Received 7 January 1999; accepted 21 May 1999)

Active sampling mass spectrometry has been used for process sensing in gate oxide growth by rapid thermal chemical vapor deposition from SiH_4 and N_2O. Equipment and process behavior throughout the short process cycle were revealed in the detailed time-dependent changes of downstream mass spectroscopic signals. A H_2 reaction product was clearly identified during SiO_2 deposition for SiH_4/N_2O ratios of 0.5%–2.0% at 5 Torr total pressure and in the temperature range 750–850 °C. No H_2O product was observed, suggesting that the process is dominated by a two-step reaction involving SiH_4 pyrolysis and subsequent N_2O oxidation of the deposited Si to form SiO_2. The evolution of the H_2 product signal during a process was then used as a process indicator. The integrated H_2 signal was found linearly proportional to the deposited oxide thickness, providing the basis for real-time, noninvasive thickness metrology applications. This work demonstrates that properly configured real-time mass spectrometry is capable of providing not only time-dependent chemical information about system behavior, but also quantitative metrology for the film deposition process. © *1999 American Vacuum Society.* [S0734-211X(99)07704-5]

I. INTRODUCTION

With the rapid advances in semiconductor fabrication technology, there are increasing demands for new sensing techniques to achieve more reliable and informative monitoring of the equipment and process status. Mass spectrometry provides chemical information about the process environment and can be used for real-time, *in situ* diagnostics of the process effluent in semiconductor equipment. An optimal sampling system should reveal the true process gas composition under process conditions, provide a rapid response to process changes, and allow quantification of various chemical species present in the process environment. Previous work in this and other research groups has successfully applied mass spectrometry to investigate system dynamics and thickness metrology in both deposition and etch systems.[1–7] In this article, we will describe the application of a two-stage differentially pumped mass spectrometer for the deposition of ultrathin gate oxide (<100 Å) from SiH_4 and N_2O by rapid thermal chemical vapor deposition (RTCVD).

II. EXPERIMENT

Details of the RTCVD system and the mass spectrometer sampling system have been described elsewhere.[1,8] Briefly, the RTCVD chamber was part of an integrated cluster tool with two other modules for advanced wafer processing.[8] A load lock and central wafer handler allowed for automated transfer of 4 or 6 in. wafers. All experiments presented in this article were performed on 4 in. wafers. The wafer was supported on three quartz crystal rods (~1 mm outer diameter) by point contact and rotated during deposition for improved temperature uniformity across the wafer. Radiative heating of the wafer was achieved through a quartz window by an array of halogen lamps parallel to the wafer surface and located directly above the wafer outside the vacuum chamber. The chamber was water cooled to prevent deposition on the walls. A mixture of N_2O and SiH_4 was used for the deposition of thin SiO_2 film, with an excess of N_2O gas present in order to assure a stoichiometric, fully oxidized SiO_2 film. An optimal process window had been previously identified at a total pressure of 5 Torr, total flow rate of 360 sccm, 750–850 °C, and SiH_4/N_2O ratio of 0.5%–2%.[9,10] The SiH_4 source was 10% diluted in Ar.

The gate oxide film was deposited in the following operating sequence: (1) a wafer was loaded into the process module, and rotation was initiated; (2) a 5 Torr total process pressure was established and maintained by flowing SiH_4/Ar and N_2O at the appropriate flow rates to achieve the correct SiH_4/N_2O ratio and total flow rate; (3) the wafer temperature was rapidly raised to the preset value and maintained for the preset time period (based on deposition rate and total film thickness required); (4) at the end of the process the heating lamps and gas flows were turned off, the process chamber was evacuated, and a N_2 purge was turned on at 100 sccm for 10 s; and (5) the wafer was then unloaded.

The oxide deposition process was monitored with a quadrupole mass spectrometer (QMS) sampling system configured to the RTCVD module as shown in Fig. 1. A 1 mm capillary at the front of a sampling tube was placed downstream from the wafer, at the radial center of the fore line for the RTCVD module. A two-stage differential pumping system behind the sampling tube first reduced the pressure from

[a]Current address: Candescent Technologies Corporation, 6580 Via del Ora, San Jose, CA 95119; electronic mail: Brian.Lu@Candescent.Com
[b]Current address: NSF Center for Optoelectronic Computing Systems, University of Colorado, Boulder, CO 80309; electronic mail: Laura.Tedder@Colorado.Edu
[c]Current address: Institute for Systems Research, University of Maryland, College Park, MD 20742; electronic mail: rubloff@isr.umd.edu

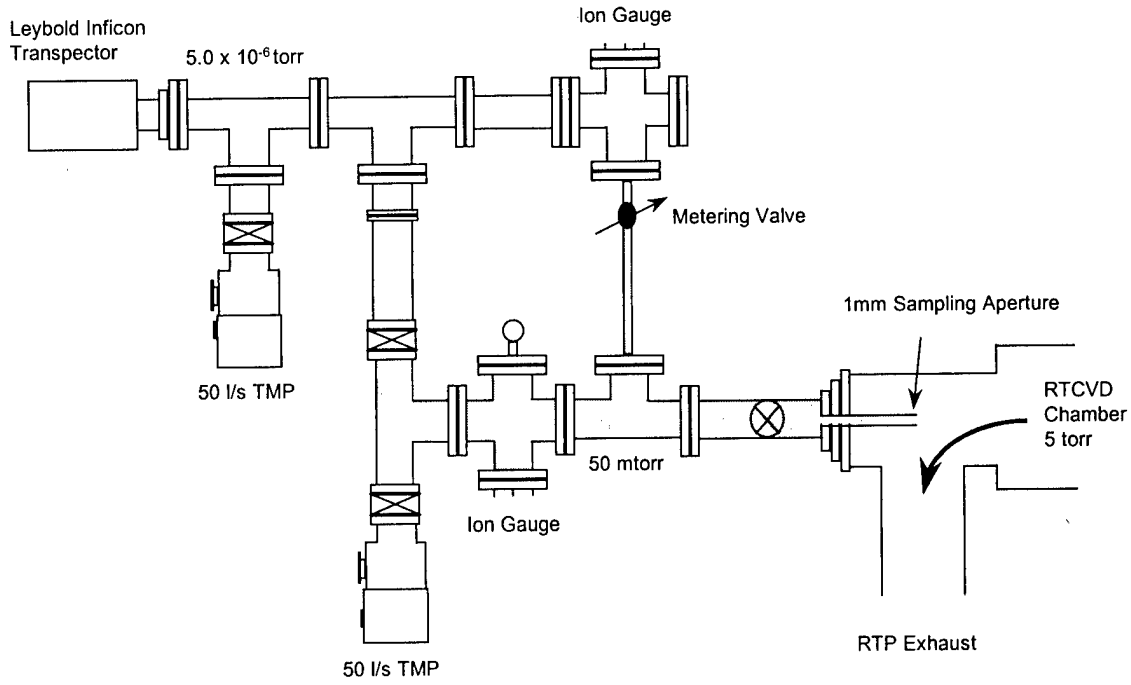


FIG. 1. Schematic diagram of the experimental setup for active sampling mass spectrometer as integrated onto a RTCVD process module. A 1 mm sampling aperture was placed in the RTCVD module downstream from the wafer and at the radial center of the fore line. The mass spectrometer was housed in the second chamber of a two-stage differential pumping system that reduced the pressure from the 5 Torr process pressure to 10^{-6} Torr operating pressure for the QMS.

the 5 Torr process pressure to ~50 mTorr in the first pumping stage, and then to $\sim 5 \times 10^{-6}$ Torr at the QMS chamber in the second pumping stage. A Leybold-Inficon mass spectrometer, configured as an open ion source, was used to sense in the mass range of 0–200 amu. This sampling system has demonstrated a response time of order ~1 s to total or partial pressure changes in the RTCVD module.[1]

The experiments were performed in two stages. First, a survey spectrum was taken over mass range 0–200 amu to identify the cracking pattern of all reactant molecules when each process gas was introduced to the process module without any wafer heating (to prevent reaction). A survey spectrum was also taken during the deposition process (i.e., with heating) to determine the reaction products. From these results, the overall process chemistry of the deposition reaction was determined. Then relevant species were selected for monitoring in trend mode (i.e., as a function of time) through the deposition process cycle. In trend mode the mass spectrometer sampled at a much faster data rate than in the survey mode due to the reduced number of masses to be scanned. Thus, the time-dependent QMS signals for the reactants and products through the process cycle provided a direct observation of the process sequence.

The SiO_2 film thickness was measured on a Rudolph AutoEl Ellipsometer. The SiO_2 uniformity from center out to a 3 cm radius was about ±10%–15%, and worse nearer the wafer edge (about ±20%–25% center to 4 cm radius). In order to obtain a first order correction for nonuniformity, an average thickness was obtained in the following way. For each wafer, the thickness was measured at nine locations equally spaced along the diameter, and an analytical expression was generated which gave a best fit to this spatial distribution of thickness. The total volume of the SiO_2 film was calculated by integrating the best-fit function over the wafer surface, assuming that the film consisted of a series of coaxial rings. The average film thickness was then determined by dividing the total volume by the wafer surface area. This uniformity calibration was essential for metrology applications since all the SiO_2 must be accounted for in order to make a valid correlation between the gas phase products and the solid phase film.

III. RESULTS

A. Time-dependent chemical sensing through the process cycle

The rapid thermal chemical vapor deposition (RTCVD) of ultrathin gate oxide was carried out as a bicomponent reaction between N_2O and SiH_4. Silane was diluted in Ar carrier gas at a 10 SiH_4/Ar ratio. As will be demonstrated, the net chemical reaction was identified as

$$SiH_4 + 2N_2O \rightarrow SiO_2 + 2N_2 + 2H_2.$$

Excessive N_2O (achieved through low SiH_4/N_2O ratios) is employed in this process to ensure full oxidation of SiH_4 to SiO_2. The partial pressure of all the gaseous reactants and

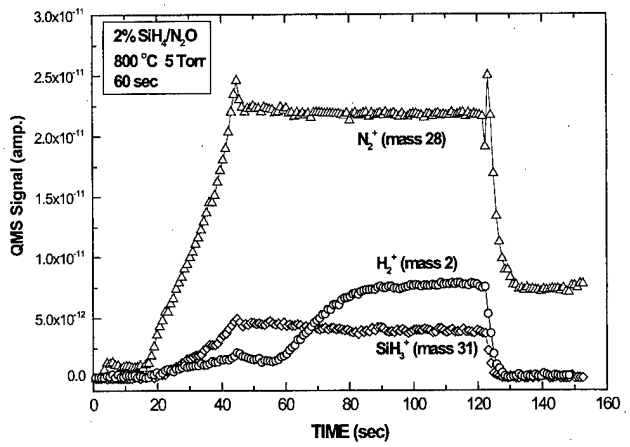


FIG. 2. Process cycle monitoring for the deposition of SiO_2 film by rapid thermal chemical vapor deposition from SiH_4 and N_2O. The QMS signals were plotted as a function of time for the reactants N_2O and SiH_4 (represented by the ionization fragments, N_2^+ and SiH_3^+, respectively), and for the reaction product H_2 (represented by H_2^+).

products in the process chamber can be monitored by QMS throughout the entire process cycle. Figure 2 shows the QMS signals for N_2^+ (an ionization fragment for N_2O), SiH_3^+ (an ionization fragment for SiH_4), and H_2^+ (for H_2) during a 60 s deposition process at 800 °C, 2% SiH_4/N_2O ratio, and 5.0 Torr process pressure. The N_2^+ fragment was selected to represent the N_2O reactant because the signal intensity of the primary ion N_2O^+ was too high to be plotted on the same scale. The signal contribution to the N_2^+ fragment arising from the reaction product N_2 was insignificant because of the large oversupply of N_2O required for material quality. The SiH_3^+ (amu 31) fragment was selected to represent SiH_4 because the primary fragment, SiH_2^+, overlaps with NO^+ (a fragment of the N_2O ionization process) at amu 30.

The three curves for reactants N_2O (N_2^+) and SiH_4 (SiH_3^+), and for product H_2 (H_2^+) in Fig. 2 clearly illustrate the partial pressure changes which occur through a deposition cycle. At time $t=3$ s, the wafer was loaded into the process chamber from the central wafer transfer module. A slight increase in the N_2^+ signal was observed because a small amount of N_2 diffused into the process chamber from the transfer module during wafer loading. At $t=15$ s, N_2O and SiH_4 were introduced to the process chamber, leading to increases in both N_2^+ and SiH_3^+ signals. Since H_2 was produced during SiH_4 ionization in the mass spectrometer, there was also a slight increase in the H_2^+ signal.

At $t=42$ s, the process pressure of 5 Torr was reached, at which a regulatory pressure control system in the process module is engaged; this caused the brief pressure overshoot seen at $t=45$ s, following which the pressure stabilized at 5.0 Torr over the next 5 s.

The heating lamp was turned on at $t=50$s, and the wafer reached 800 °C process temperature at $t=60$ s, then remaining at process temperature for the preset process time of 60 s. During this process, surface reaction leading to Si deposition on the wafer took place, with consumption of some of the SiH_4 and N_2O reactants and generated of H_2 gaseous product. The H_2^+ product signal rose over about 20 s to its steady state level, then remained relatively stable afterward until termination of the process.

At the end of the process, near $t=120$ s, the heating lamps and the SiH_4 and N_2O gas flows were turned off, leading to a rapid drop in wafer temperature, evacuation of gases from the reactor, and abrupt termination of the deposition reactor and associated H_2 product generation, leading to rapid drops in H_2^+ and SiH_3^+ signals. The N_2^+ signal (indicative of N_2O) decreased first, similar to that for SiH_3^+. Soon afterward, the N_2^+ signal increased rapidly, then decreased to 30% of the original steady state level; this behavior was associated with an intentional N_2 purge which was turned on at the end of each process cycle to flush out residual process gases. Thus the N_2^+ signal level at $t=140$ s corresponded to the N_2 pressure inside the process chamber during purge. These time-dependent observations demonstrate that the QMS sampling system reveals a variety of process cycle details of the system.

B. Rapid identification of process chemistry

The deposition of SiO_2 in this process certainly involves oxidation of SiH_4 by N_2O. While N_2O was reduced to N_2 gas, the hydrogen generated as a product from the SiH_4 reactant could be released into the gas phase as either H_2 or H_2O, such as

$$SiH_4(gas)+2N_2O(gas)\rightarrow SiO_2(solid)+2N_2(gas)+2H_2(gas)$$

or

$$SiH_4(gas)+4N_2O(gas)\rightarrow SiO_2(solid)+4N_2(gas)+2H_2O(gas).$$

To distinguish between these two pathways and to identify the primary gaseous by-products of the deposition reaction, a static deposition experiment was performed to increase the concentration of the by-products for better detection of minor reaction products. Here, instead of maintaining a steady state pressure in the process chamber, a static 5 Torr process pressure was established using a 2% SiH_4/N_2O gas mixture. The process chamber was first isolated (all mass flow controllers and gate valves were turned off), and the wafer temperature was then raised to 800 °C for 60 s. The H_2 and H_2O partial pressures were monitored using the mass spectrometer during this entire process cycle.

Figure 3 shows the QMS signals for H_2 and H_2O as a function of time during the static SiO_2 deposition. The H_2 signal displayed a shape similar to that in the regular process (Fig. 2). While this signal at amu 2 would also be produced as a fragment of H_2O ionization, H_2O would also produce substantial signal at amu 18. The measured H_2O signal, magnified by a factor of 10× in Fig. 3, shows no H_2O signal above the noise level. These data demonstrate that the H_2 formation channel is strongly favored (by something like 100×) over the H_2O formation channel for the SiO_2 deposition process.

Accordingly, the process chemistry can be considered a two-step surface reaction, involving Si deposition from SiH_4 surface decomposition, followed by N_2O oxidation of the

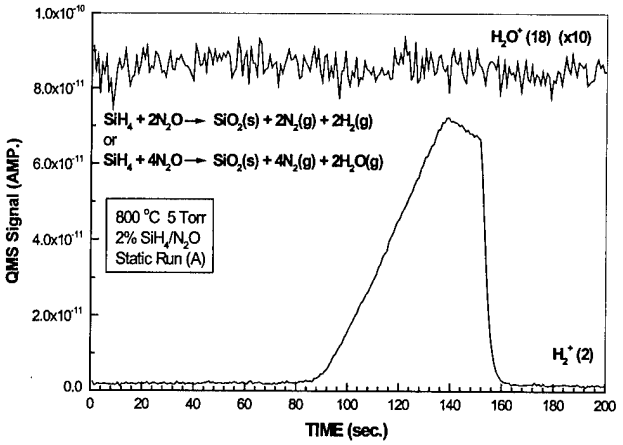


FIG. 3. Comparison of the QMS signals for H_2 and H_2O during a deposition cycle under static gas flow conditions. The process chamber was first filled with 5 Torr of 2% SiH_4/N_2O mixture, and the chamber was then isolated. The wafer temperature was rapidly raised to 800 °C for 60 s to induce surface deposition. The H_2O signal was multiplied by a factor of 10 to illustrate that the H_2O signal was below the noise level.

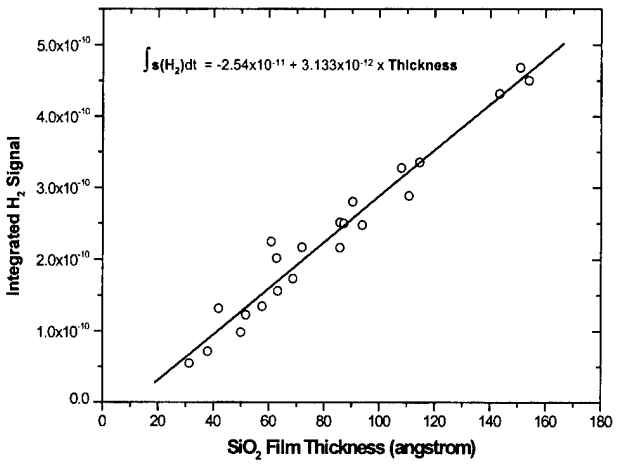


FIG. 5. Correlation of the integrated H_2 signal from the QMS and the average film thickness on the wafer for various deposition conditions within the process window for high quality SiO_2 gate oxide films. The deposition conditions (a total of 22) were selected based on a design of experiments within the temperature range of 750–850 °C and the compositional range of 0.5%–2% SiH_4/N_2O ratio. The total pressure and total gas flow rate were kept constant during these experiments.

deposited Si species. The oversupply of N_2O is needed to assure complete oxidation of the deposited Si before it is covered by more Si. This means that N_2O depletion by the reaction is difficult to measure. Furthermore, the deposition is rate limited by SiH_4 consumption/depletion, so that either H_2 product generation from SiH_4 pyrolysis or the associated SiH_4 depletion provide metrics for progress of the reaction.

C. Process metrology for oxide film thickness

The above conclusions indicate that two H_2 molecules are released to the gas phase for each SiO_2 molecular unit deposited as oxide on the wafer. Thus, the partial pressure of H_2 product in the process chamber, sensed by the QMS signal, should be directly related to the rate of reaction which generated gaseous H_2 into the process chamber and SiO_2 onto the wafer surface.

Figure 4 shows the H_2 QMS signals versus time for three

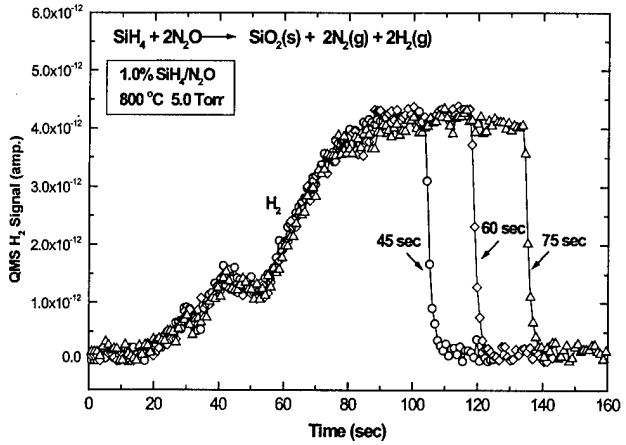


FIG. 4. Comparison of the H_2 evolution curves for three process times during the deposition of SiO_2 at 5 Torr and 800 °C using a 1% SiH_4/N_2O gas mixture.

deposition processes using 1.0% SiH_4/N_2O at 5.0 Torr, 360 sccm total flow, and 800 °C, with varying deposition time of 45, 60, and 75 s. The SiH_4 and N_2O gas flows were started at $t = 10$ s, and the process pressure was reached at $t = 45$ s. The H_2 intensity observed during this period was due to SiH_4 fragmentation during QMS ionization. There was no surface reaction yet at this point. The heating lamps were turned on, and the wafer reached the preset reaction temperature of 800 °C at $t = 55$ s. The surface reaction for SiO_2 deposition began and H_2 was produced, leading to significant rise of the H_2 signal. The wafer temperature was then maintained at 800 °C for the preset deposition time of 45, 60, and 75 s, respectively. The three lines are nearly superimposed on each other over the first 100 s of the process, illustrating the stability and repeatability of the response to process parameters by both the RTCVD module and QMS sampling system. For all three processes, the H_2 signal returned to baseline within 3 s of the end of deposition, again demonstrating the fast response time of the mass spectrometer sampling system.

Based on the process chemistry determined above, two H_2 molecules should be generated for each SiO_2 deposited. If the mass spectrometer sampling system had a linear response to the H_2 generation rate, the QMS signal for H_2 should then be proportional to the SiO_2 deposition rate. This correlation should be true as long as the reaction chemistry was preserved regardless of the process conditions (time, temperature, pressure, etc.).

The integrated H_2^+ signal is plotted against the film thickness in Fig. 5 for 22 deposition runs which produced films ranging from 30 to 155 Å. The processing conditions were randomly selected within the process window for producing high quality, stoichiometric SiO_2 film. The 22 deposition process conditions included process temperatures ranging from 750 to 850 °C, process times from 20 to 120 s, and

SiH_4/N_2O ratios from 0.5% to 2.0%. The total process pressure was kept at 5.0 Torr for all processes, and the total gas flow was kept constant at 360 sccm.

Figure 5 demonstrates an obvious linear correlation between the integrated H_2 product signal from the QMS and the average thickness of the deposited SiO_2 film. The integrated H_2 product signal has been derived by subtracting the H_2 background associated with H_2^+ fragments from the ionization of the SiH_4 reactant through calibration of the H_2^+ background before wafer heating (e.g., times $t=30-50$ s in Fig. 2). As described earlier, the average film thickness has been corrected for across-wafer thickness nonuniformity in these experiments, although manufacturing application would demand a high degree of uniformity and thereby obviate the need for such a correction. The linear correlation between the thickness and the integrated H_2 signal provides the basis for real-time *in situ* thickness metrology in CVD gate oxide deposition processes.

IV. DISCUSSION

A. Gas sampling system design

Integration of a sampling system onto a process module is key to the successful development of a capable sensor. A properly configured mass spectrometer as described in this article can sample the gas composition in the process chamber in real time, producing chemically specific signals representing true process events in the process chamber. The location of the capillary tube for gas sampling dictates whether a true representation of the reactive species (both reactants and reaction products) in the process chamber is collected for measurement. In this study, the sampling aperture was placed downstream from the wafer, at the radial center of the fore line for the process module. The sample composition collected at this location represented the gases exiting the process chamber, and therefore included changes induced by consumption of reactants and generation of gaseous products.

Rapid response of the sampling system is particularly important for monitoring processes with short process cycles, such as RTCVD. We believe that the use of a two-stage gas sampling system was crucial for achieving a short response time compared to the process cycle time. The challenge is the large pressure gradient required between the process (5 Torr) and the mass spectrometer (10^{-6} Torr). While a very small sampling aperture can be used in conjunction with a single stage of pumping to maintain this gradient, this would produce only very small gas flows through the aperture. With the process chamber in viscous flow conditions, changes in concentration caused by reaction at the wafer could diffuse only slowly to the aperture, thereby slowing response time.

The two-stage differential pumping in the gas sampling system permitted a considerably larger sampling aperture between the reactor and the first pumping stage, so that of order 1% of the total gas flow went through the sampling aperture (cf. the process pump), assuring reasonable gas flow velocities and fluxes to the sampling system. The first pump then needed to reduce pressure only to <100 mTorr, i.e., in the molecular flow regime. Once in this regime, sampling could be done with negligible time delay through a metering valve to the second pump, which achieved the pressure reduction (to the 10^{-6} Torr range) required by the mass spectrometer. The response time for the entire sampling system has been determined to be on the order of 2–3 s.[1]

With a gas sampling system optimized for rapid time response, one can understand much about the time dependencies seen in the process sensing data, e.g., in Fig. 2. Real changes in signals on 2–3 s time scales are seen in the data: in the N_2^+ data near $t=42$ s in Fig. 2 (where the pressure control system caused an overshoot); in the N_2^+ oscillation near $t=120$ s (where the N_2 purge cycle began); and in the decay of N_2^+ and H_2^+ signals near $t=125$ s (where gas flow was terminated and pumping rate enhanced by fully opening the throttle valve). In contrast, the rise time of the N_2^+ signal upon initiation of gas inlet (near $t=20$ s) was longer as a result of the relatively low inlet flow rate and significant reactor volume. The rise time of the H_2^+ product signal was also slow, dependent on the generation rate of product during heating and on the volume of the reactor.

Using mass balance principles for the process module, we have built a mathematical model and corresponding dynamic simulator which accurately represents the time dependence of the QMS signal.[11–13] This simulator provides the basis for detailed understanding of system dynamics. For example, the rise time of the H_2^+ product signal reflects system time constants arising from the time it takes for the product generation rate to establish a steady state concentration in the reactor. However, this time constant will depend on equipment specifications, such as reactor volume and pumping speed, thereby influencing the total H_2^+ product signal integrated over the process cycle. As a result, even if the mass spec sensor were itself completely linear with concentration at the sensor,[14] the integrated product signal will not be strictly a linear measure of deposition thickness; rather, for a given equipment configuration, corrections dependent on process recipe can be obtained either empirically or preferably from using physically based dynamic simulation models.

Sensitivity, dynamic range, and linearity are also important metrics for real-time process chemical sensors. A typical mass spectrometer can measure partial pressures from 10^{-5} to 10^{-12} Torr. A differentially pumped sampling system can be used to translate this dynamic range to higher pressure regimes (mTorr to Torr) as needed for CVD or etch processes. This large dynamic range assures more than adequate sensitivity to distinguish changes in process conditions, as well as to detect contaminants which may threaten process quality. Reasonable linearity of the QMS signal for process reactants and products is also very helpful for thickness metrology. Such linearity depends both on system dynamics (described above) and on sensor linearity with concentration.[14] As illustrated in Fig. 5, the two-stage gas sampling system achieves a relatively linear response.[1,15]

B. Process chemistry in CVD oxide growth

Chemical sensing using mass spectrometry has here been effective in delineating the operative process chemistry. Previous studies of SiO_2 formation from SiH_4 and N_2O using atmospheric pressure CVD (APCVD) have reported that both H_2 and H_2O are plausible reaction products at temperatures from 495 to 690 °C.[16,17] Under the low pressure RTCVD conditions employed here, the formation of H_2 is the dominant (by >100×) if not exclusive reaction channel for the deposition of SiO_2, a result confirmed under both static and regular process conditions. The lower pressures employed in the RTCVD process are expected to enhance surface processes (e.g., as compared to gas phase reactions found more commonly at APCVD conditions). The higher surface temperatures in RTCVD should accelerate pyrolysis of individual reactant species which impinge on the wafer surface, thus favoring multi-step processes in which surface decomposition of individual molecules precedes the synthesis of their products to form newly deposited material.

Identifying the operative chemistry is also an important aide to achieving effective process metrology. In the present case, the excess of N_2O reactant required for material quality rendered the monitoring of N_2O signals ineffective in assessing depletion. In contrast, with SiH_4 pyrolysis as the rate-limiting step, reaction progress could be monitored by the H_2 reaction product (as shown here) or in principle by SiH_4 depletion as well. The absence of a comparable or dominant reaction channel leading to H_2O product further simplified reaction monitoring and thickness metrology.

C. Implementation of mass spectrometry for RTCVD thickness metrology

Results here demonstrate that the integrated H_2 signal from the mass spectrometry sampling system can be linearly correlated to the film thickness on the wafer, providing an *in situ*, real-time sensing approach for thickness metrology. For a given process, a calibration experiment like that in Fig. 5 provides a model for relating the QMS signal to the thickness. The QMS signal for the H_2 product is readily integrated over time using a simple spreadsheet analysis of the data. Depending on the characteristics of the sensor interface and sensor control software, the prediction of thickness could then be exploited for run-to-run (wafer-to-wafer) or real-time process control purposes. The use of mass spectrometry for chemically based process thickness metrology assures that process chemical insights will be gained as well.

The time-dependent QMS signals (e.g., Fig. 2) illustrate areas where caution must be taken in using mass spectrometric sampling for thickness metrology. In this example, the deposition rate is quite low (~100 Å /min at 800 °C for the 0.5% SiH_4/N_2O process), and therefore the QMS signal for H_2 product during deposition is not significantly above the background H_2 level produced from SiH_4 ionization in the QMS. This contribution must be calibrated for each SiH_4 concentration and subtracted from the total H_2 signal in order to achieve accurate quantification of the actual deposition process. In addition, the dynamics of the process cycle could impact the relationship between QMS signal and deposition thickness, since the system dynamics do not permit instantaneous response of the QMS signals to equipment actuation. Finally, as indicated here and previously,[2] the accuracy of QMS-based CVD thickness metrology is not yet sufficient for manufacturing process control.

One can anticipate a broad variety of process applications for mass spectrometry sensing and metrology in semiconductor manufacturing, including plasma etching, plasma deposition, wafer cleaning, photoresist ashing, etc. At the same time, some processes will bring further technical challenge, such as processes involving gas phase nucleation steps [e.g., tetraethylorthosilicate (TEOS) processes] or wall deposition (e.g., plasma processes). While most mass spectrometry investigations of semiconductor chemical processes have addressed single-wafer, cold-wall reactor configurations, sampling from particular or individual wafer stations in a multistation CVD reactor may be more difficult.[15] Mass spectrometry does not provide information on uniformity, another key parameter for manufacturing, either for across-wafer or within-batch characteristics.

V. CONCLUSIONS

We have demonstrated the application of a two-stage differentially pumped mass spectrometer for process monitoring in rapid thermal chemical vapor deposition (RTCVD) of thin gate oxide from SiH_4 and N_2O. Detailed time-dependent changes in mass spectroscopic signals provide an accurate representation of the equipment and process variations throughout the short process cycle. A H_2 reaction product was clearly identified during SiO_2 deposition for SiH_4 concentration of 0.5%–2.0% at 5 Torr total pressure and 750–850 °C. The absence of observable H_2O product suggests that the process is dominated by separate steps of SiH_4 pyrolysis and subsequent N_2O oxidation of the deposited Si to form SiO_2. The integrated H_2 product signal was found linearly proportional to the deposited oxide thickness, providing the basis for real-time, noninvasive thickness metrology applications. These studies show that a properly configured mass spectrometer sampling system is capable of providing chemical and time-dependent information on system behavior. Thickness metrology and control applications require reasonable knowledge of the chemical process, which can be gained from mass spectrometry studies.

ACKNOWLEDGMENTS

The authors would like to thank Dr. John Hauser, Eric Vick, and John B. Flanigan, III, for the use of the RTCVD module at North Carolina State University. They acknowledge with thanks the financial support of this research project by the National Science Foundation (Grant No. CDR 8721505) and the Semiconductor Research Corporation (SRC Contract No. 94-MJ-563).

[1]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).
[2]L. L. Tedder, G. W. Rubloff, B. Conaghan, and G. N. Parsons, J. Vac. Sci. Technol. A **14**, 267 (1996).

[3]A. I. Chowdhury, W. W. Read, G. N. Parsons, G. W. Rubloff, and L. L. Tedder, J. Vac. Sci. Technol. B **15**, 127 (1997).
[4]J. J. Chambers, K. Min, and G. N. Parsons, J. Vac. Sci. Technol. B **16**, 2996 (1998).
[5]D. W. Greve, T. J. Knight, X. Cheng, B. H. Krogh, M. A. Gibson, and J. LaBrosse, J. Vac. Sci. Technol. B **14**, 489 (1996).
[6]A. I. Chowdhury, T. M. Klein, T. M. Anderson, and G. N. Parsons, J. Vac. Sci. Technol. A **16**, 1852 (1998).
[7]T. J. Knight, D. W. Greve, X. Cheng, and B. H. Krogh, IEEE Trans. Semicond. Manuf. **10**, 137 (1997).
[8]N. A. Masnari, J. R. Hauser, G. Lucovsky, D. M. Maher, R. J. Markunas, M. C. Ozturk, and J. J. Wortman, Proc. IEEE **81**, 42 (1993).
[9]X. L. Xu, R. T. Kuehn, J. J. Wortman, and M. C. Ozturk, Appl. Phys. Lett. **60**, 3063 (1992).
[10]V. Misra, X. L. Xu, B. E. Hornung, R. T. Kuehn, and J. J. Wortman, J. Electron. Mater. **25**, 527 (1996).
[11]L. L. Tedder, G. B. Lu, B. F. Conaghan, and G. W. Rubloff, Proceedings of Rapid Thermal Processing '95, Amsterdam, The Netherlands, 30 Aug.–1 Sept. 1995.
[12]G. Lu, M. Bora, and G. W. Rubloff, IEEE Trans. Semicond. Manuf. **10**, 390 (1997).
[13]G. Lu, M. Bora, L. L. Tedder, and G. W. Rubloff, IEEE Trans. Semicond. Manuf. **11**, 63 (1998).
[14]L. Lieszkovszky, A. R. Filippelli, and C. R. Tilford, J. Vac. Sci. Technol. A **8**, 3838 (1990).
[15]G. B. Lu and E. Baker, invited presentation at the Topical Conference on Partial Pressure Measurements and Process Control, American Vacuum Society, Baltimore, MD, November 1998.
[16]J. D. Chapple-Sokol, C. J. Giunta, and R. G. Gordon, J. Electrochem. Soc. **136**, 2993 (1989).
[17]C. J. Giunta, J. D. Chapple-Sokol, and R. G. Gordon, J. Electrochem. Soc. **137**, 3237 (1990).

Relaxation of strained Si layers grown on SiGe buffers

S. B. Samavedam,[a] W. J. Taylor, J. M. Grant, J. A. Smith, P. J. Tobin, A. Dip, and A. M. Phillips
Motorola Advanced Products Research and Development Laboratories, Austin, Texas 78721

R. Liu
Motorola Center for Integrated Systems Development, Mesa, Arizona 85202

(Received 28 January 1999; accepted 7 May 1999)

Thin strained Si layers grown on SiGe layers graded to 20% Ge were studied for resistance to relaxation. It was observed that in the presence of $\sim 10^5/cm^2$ threading dislocations from the underlying graded layers, the barrier to misfit dislocation formation is sufficiently reduced to induce relaxation in Si layers even when the layer thickness is less than the predicted critical thickness. Raman spectroscopy revealed that elastic strain accumulation in the uniform SiGe layers is a significant contributor to strain relaxation in the Si cap layers. Upon annealing, thermal mismatch causes the Si layers to relax further, but most of the strain relaxation is accommodated by elastic strain increase in the SiGe layers. This prevents the rampant increase in defect density that would otherwise accompany the strain relaxation. Annealing in an oxidizing ambient appears to pin pre-existing threading dislocations causing nucleation of new threading dislocations and short misfit segments to relieve the thermal mismatch stresses. © *1999 American Vacuum Society.*
[S0734-211X(99)04204-3]

I. INTRODUCTION

Biaxial tensile strain in Si splits the six-fold degeneracy in the conduction band into a lower energy double degenerate valley and a higher energy four-fold degenerate valley. The resulting increased in-plane electron mobility, due to a decrease in the average effective electron mass and decreased inter-valley carrier scattering, has been well documented.[1–5] Tensile strain in Si is also known to improve in-plane hole mobility by splitting the degeneracy between the light and heavy holes.[6,7] These effects have been successfully utilized to fabricate enhanced mobility metal–oxide–semiconductor field effect transistors[7,8] and modulation doped field effect transistors[9–11] with strained Si channels grown on relaxed SiGe layers.

The primary function of the relaxed SiGe layers on Si substrates is to serve as a larger lattice constant material to create tensile strain in the Si overlayer, and hence, is often referred to as a "virtual substrate." Si and Ge have a lattice mismatch of ~4%, which causes a high density of misfit dislocations and associated threading dislocations to form when the SiGe layer thickness on the Si substrate exceeds the critical thickness.[12–16] Currently, one of the more successful and widely accepted techniques for reducing defect densities in SiGe layers is to grow relaxed graded SiGe buffer layers.[17,18] The low grading rate (typically ~10% Ge/μm) minimizes dislocation nucleation and high growth temperature (700–900 °C) promotes dislocation glide for complete strain relaxation in the SiGe graded layers. Growth of graded SiGe layers results in the threading dislocation densities of the order of $10^5-10^6/cm^2$, compared to $10^8-10^{10}/cm^2$ for uniform SiGe layers on Si substrates. Generally a uniform SiGe buffer layer is grown on top of the graded layers to ensure complete relaxation of the graded layers and separate the active device layers from misfit dislocation strain fields in the graded region. The Si layer, which is grown on the relaxed SiGe uniform layer, is supposed to be completely strained if it is below the equilibrium critical thickness.[12–14] However, we find that in the presence of $\sim 10^5/cm^2$ threading dislocations, strained Si layers start to partially relax even when the layer thickness is well below the predicted critical thickness. Device processing subjects the strained Si layers to temperatures beyond the growth temperature for varying time periods which can relax some of the strain. Since the band engineering in these devices is almost exclusively strain induced it is crucial to retain the strain in the Si layers during the various processing steps. Strain relaxation is also accompanied by the formation of misfit and threading dislocations which can increase leakage currents, and reduce carrier mobilities by scattering. Therefore, it is essential to study how resistant the Si layers are to strain relaxation and determine the accompanying change in defect density. In this article we present the effect of growth thickness, anneal time, and temperature on strain relaxation and concomitant defect generation in strained Si layers grown on graded SiGe buffer layers.

II. EXPERIMENT

Graded SiGe layers were grown by rapid thermal chemical vapor deposition (RTCVD) with SiH_2Cl_2 and GeH_4 as the pre-cursors and H_2 carrier gas on 200 mm Si substrates. A commercially available reduced pressure RTCVD system was used to deposit the films. The system employs a N_2 purged load lock to minimize the background level of external contaminants. The process chamber is capable of depositing on one 200 mm wafer and uses wafer rotation for improved film uniformity. Heating is provided by incandescent

[a]Electronic mail: ra8799@email.sps.mot.com

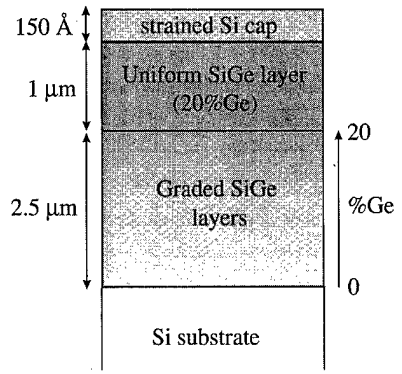


FIG. 1. Schematic of the Si/SiGe/Si structure.

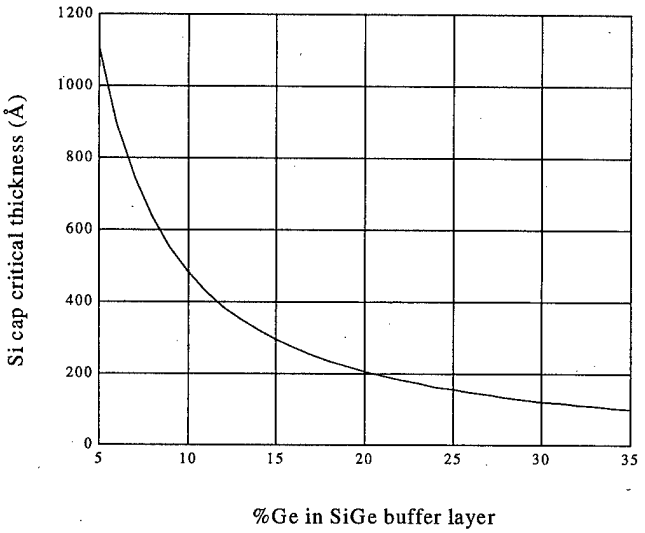


FIG. 2. Si cap critical thickness as a function of Ge% in the uniform SiGe layer.

lamp arrays located above and below the silicon carbide coated graphite susceptor with the top bank directly illuminating onto the wafer. Temperature measurements are taken with the primary pyrometer located directly beneath and facing the underside of the susceptor. The primary pyrometer is calibrated over a range of temperatures by means of a second pyrometer located above the susceptor and facing the silicon substrate. Since emissivity values for silicon are well known, calibration of the primary pyrometer to the susceptor can easily be achieved by using the second pyrometer calibrated by the silicon wafer. Temperature measurements used during processing were not directly taken from the substrate due to small changes in emissivity as the SiGe film is deposited.

The substrates were subjected to a 1080 °C *in situ* clean in an H_2 ambient to desorb the native oxide. The SiGe layers were linearly graded to 20% Ge at a Ge grading rate of 8% Ge/μm. A 1-μm-thick uniform SiGe layer was grown on top of the graded layers. The graded and the uniform SiGe layers were grown at 800 °C and 100 Torr growth pressure. The growth temperature was dropped to 700 °C for the growth of the Si cap layer. Figure 1 shows a schematic of the strained Si on the SiGe/Si structure. The Si layer thickness was varied from 108 to 190 Å. The 150 Å Si cap samples were chosen for the anneal experiments. These samples were annealed at 850 and 1025 °C for times ranging from 1 s to 3 h in 1% O_2 ambient. The anneal conditions were chosen to represent high temperature device processing steps to which the layers would eventually be subjected.

The strain relaxation in the Si layers before and after the anneals was characterized using Raman spectroscopy. The Raman spectra were obtained by using a commercially available triple spectrometer with a liquid nitrogen cooled back-thinned charge coupled device detector with a resolution of 2000×800 pixels. A high resolution triple-additive configuration of three 1800 grooves/mm gratings were used for this experiment. A 100× objective was used to focus an Ar laser beam ($\lambda = 4579$ Å) down to a 1 μm diameter spot on the sample surface. The excitation power density was about 100 kW/cm^2. The optical penetration depth at this wavelength is 4200 Å in Si and slightly lesser in the SiGe layer. This ensured that only the Si cap layer and the underlying uniform SiGe layer (1 μm) contributed to the Raman spectra. Strain-induced frequency shifts in the Si–Si longitudinal optical (LO) phonon modes in the Si cap and in the uniform SiGe layer were used to determine the residual or elastic strain (ϵ). Through peak fitting we could resolve peak shifts smaller than 0.05 cm^{-1}, which corresponds to about 10 MPa in terms of stress. The same strain-phonon coefficient, $b = -930$ cm^{-1} was used for both the Si and SiGe LO peak shift. However, a linear extrapolation based on the Ge composition was used for the conversion factor $[b(S_{11}+S_{12})]$ between the LO peak shift and stress. The elastic constants used were $(S_{11}+S_{12})^{-1} = 180$ GPa for Si and $(S_{11}+S_{12})^{-1} = 138$ GPa for Ge.

The threading and the misfit dislocation density increase due to strain relaxation was characterized using etch pit density (EPD) measurements. Dilute Schimmel etch (4:5::0.2 M CrO_3:49% HF) was used to reveal the defects. The etch depths were controlled to ensure that the delineated defects were only from the Si cap and the Si cap-uniform SiGe interface. Optical Normarski micrographs were used to estimate the misfit and threading dislocation densities. High resolution transmission electron microscopy (TEM) was used to measure the Si cap thickness. The Ge composition in uniform buffer layer was determined and verified using multiple characterization techniques such as Rutherford backscattering, Auger electron spectroscopy, and high resolution x-ray diffraction.

III. RESULTS AND DISCUSSION

A. Critical thickness and relaxation

Figure 2 is a plot of critical thickness of Si layers grown on relaxed uniform SiGe layers as a function of Ge% in the uniform SiGe layer. The assumption in deriving this curve is that the SiGe layer is completely relaxed and acts like a substrate for the thin strained Si overlayer. The curve was determined using the Matthews–Blakeslee (MB) energy minimization criterion.[12–14] Since the SiGe layer is much

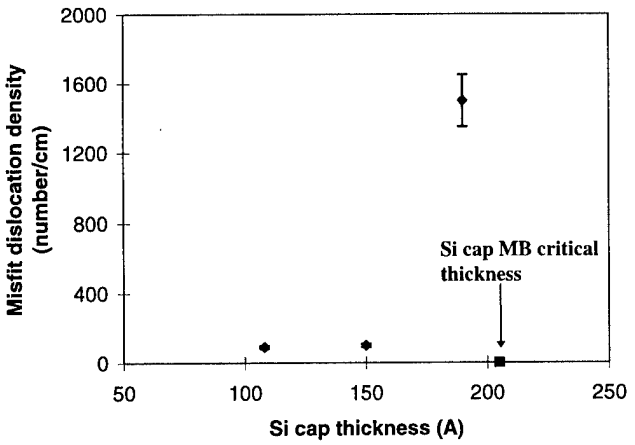


FIG. 3. Plot of misfit dislocation density as a function of Si cap thickness. The predicted MB Si cap critical thickness is also shown. The MB thickness is the thickness at which the first misfit dislocation forms. So, the theoretical misfit dislocation density at MB thickness is zero. A finite misfit dislocation density at the Si cap-SiGe interface indicates the onset of plastic relaxation of the Si cap.

thicker than the Si cap layer, it was assumed to be unstrained. The strained Si layer was assumed to be isotropic and the isotropic modulus under biaxial stress was used for computing the strain energy in the MB derivation. For the dislocation energy term, the film thickness, h, was used for the dislocation cut-off radius, since h is smaller than half the interfacial dislocation spacing at critical film thickness.[15] When the strain in the Si epilayer exceeds the critical strain, it relaxes by forming 60° misfit dislocations at the Si/SiGe interface. At any layer thickness below the critical thickness the Si layer should be completely strained since it is energetically unfavorable to form misfit dislocations at the Si/SiGe interface.[12–16] The Ge content in the uniform SiGe layer was verified to be ~20% using multiple techniques mentioned earlier. From Fig. 2, the critical thickness for a Si layer grown on a uniform $Si_{0.8}Ge_{0.2}$ layer is 205 Å. If it is assumed that the SiGe layer can be compressively strained to some degree (see Sec. III B) to accommodate the lattice mismatch, the predicted critical thickness for the Si layer will be higher than in Fig. 2.

Figure 3 is a plot of misfit dislocation density as a function of the Si cap thickness grown on uniform $Si_{0.8}Ge_{0.2}$ layers on graded SiGe layers. The cap thicknesses were accurately determined using high resolution TEM. The misfit dislocation densities were determined from Nomarski optical micrographs of large etched sample areas (typically 30 000 μm^2) for statistically accurate assessment of low defect densities. The Si cap MB critical thickness for growth on a $Si_{0.8}Ge_{0.2}$ layer, determined from Fig. 1, is also shown. The MB critical thickness is a predicted value at which the first misfit dislocations start to form at the Si/SiGe interface. Therefore, it has zero for the ordinate value and is, hence, represented as a point on the x axis. The critical thickness from the MB criterion has been experimentally verified for the cases where the growth temperature is high enough to drive the system to equilibrium. Apparent deviations from the MB criterion have been observed in SiGe/Si heterostructures.[19–21] In all cases the experimentally determined critical thickness was found to be larger than the MB prediction. This has been attributed to metastable growth and incomplete relaxation due to low growth temperatures, as well as uncertainty in the technique used for detecting relaxation.[16]

TABLE I. Strain in the Si and SiGe layers determined from shifts in the Raman spectra.

Anneal condition	Tensile elastic strain in the Si layer ($\times 10^{-3}$)	Compressive elastic strain in the $Si_{0.8}Ge_{0.2}$ layer ($\times 10^{-3}$)
As-grown	7.317	1.081
850 °C, 1 s	7.294	1.070
850 °C, 60 s	7.244	1.122
850 °C, 20 min	7.328	1.174
850 °C, 3 h	6.994	1.209
1025 °C, 1 s	6.639	1.907
1025 °C, 10 s	6.822	1.680
1025 °C, 30 s	6.572	1.988

In this case, there is an opposite deviation from the MB critical thickness. It was found that the Si layers start to relax by forming misfit dislocations at the $Si/Si_{0.8}Ge_{0.2}$ interface even when the layer thickness is significantly *less* than the critical thickness. It should be mentioned that the observed misfit densities, however, are small ($\sim 10^2$/cm) compared to expected densities for fully relaxed Si layers (10^5/cm). Nonetheless, it is significant that we observe relaxation at about half the predicted MB critical thickness. It can be explained as follows. Most experiments to verify the critical thickness are performed on substrates with very low defect densities. In this case, the "substrate" is the uniform $Si_{0.8}Ge_{0.2}$ layer that has $4\times 10^5/cm^2$ threading dislocations from the underlying graded layers. The presence of such a high defect density, sometimes in the form of dislocation pile-ups,[22] seems to lower the barrier to misfit dislocation nucleation locally. With increasing the Si cap thickness, the strain energy in the film increases and further lowers the barrier to misfit dislocation nucleation. Therefore, there is an increase in misfit dislocation density with increasing cap thickness.

B. Effect of anneals

The Raman spectra of the as-grown and annealed samples revealed that the uniform $Si_{0.8}Ge_{0.2}$ layers were not completely relaxed contrary to common expectations. Table I shows the residual elastic strains in the Si cap and the uniform $Si_{0.8}Ge_{0.2}$ layers for different anneal conditions. The strains are tensile for the Si cap and compressive for the $Si_{0.8}Ge_{0.2}$ layers. The elastic strains from the Raman measurements were used to determine the in-plane lattice constants for the Si ($a_{\|Si}$) and the $Si_{0.8}Ge_{0.2}$ ($a_{\|SiGe}$) layers. The degree of strain relaxation in the Si and $Si_{0.8}Ge_{0.2}$ was characterized using the following relaxation factors:[23]

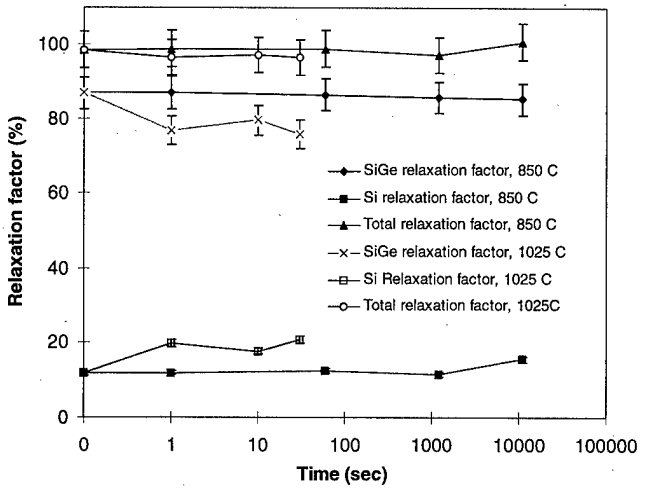


FIG. 4. Relaxation factors for the Si and the SiGe layers as a function of time at different anneal temperatures.

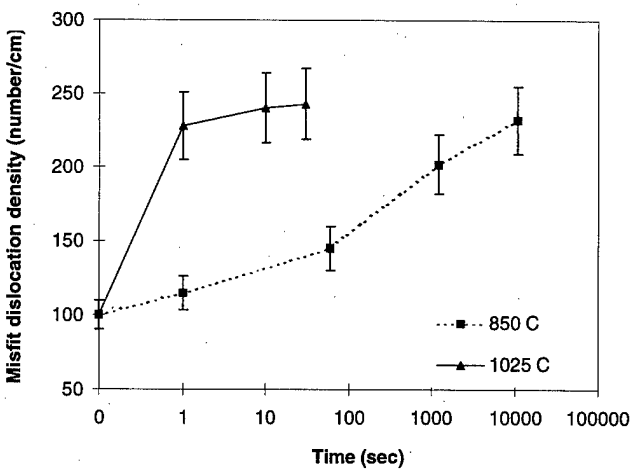


FIG. 5. Misfit dislocation density at the Si cap-SiGe interface as a function of time at the two anneal temperatures. The misfit density has a linear log (time) dependence.

$$\text{Si relaxation factor}\,(\%) = \frac{a_{\parallel \text{Si}} - a_{\text{SiGe}}}{a_{\text{Si}} - a_{\text{SiGe}}} \times 100\%, \quad (1)$$

$$\text{Si}_{0.8}\text{Ge}_{0.2}\ \text{relaxation factor}\,(\%) = \frac{a_{\parallel \text{SiGe}} - a_{\text{Si}}}{a_{\text{SiGe}} - a_{\text{Si}}} \times 100\%. \quad (2)$$

a_{Si} and a_{SiGe} refer to the fully relaxed lattice constants of Si and $\text{Si}_{0.8}\text{Ge}_{0.2}$, respectively. The relaxation factors range from 0% to 100% as the layers go from being totally strained to totally relaxed. Figure 4 is a plot of relaxation factors for the different layers as a function of anneal time at the two anneal temperatures. It is observed that in the as-grown samples the Si layer (150 Å) is already 12% relaxed and the $\text{Si}_{0.8}\text{Ge}_{0.2}$ layer is only 87% relaxed. The degree of relaxation of the Si cap layer increases with anneal time and the rate of relaxation is higher at 1025 °C than at 850 °C. There is a corresponding decrease in the $\text{Si}_{0.8}\text{Ge}_{0.2}$ layer relaxation factor indicating an accumulation of compressive elastic strain. The "total relaxation factor," which is the sum of the relaxation factors for the Si and the $\text{Si}_{0.8}\text{Ge}_{0.2}$ layers in each case is also plotted on the same graph. If the total relaxation factor adds up to 100% for each condition, it means that the strain in each $\text{Si}/\text{Si}_{0.8}\text{Ge}_{0.2}$ layer combination is completely elastic and that there is no plastic strain relaxation. The total strain relaxation factor is close to 100% for all cases, indicating that the observed strains are indeed mostly elastic in nature.

Figure 5 is a plot of misfit dislocation density at the $\text{Si}/\text{Si}_{0.8}\text{Ge}_{0.2}$ interface as a function of anneal time for the two anneal temperatures. The trends in the misfit density increase follow the trends in the Si cap relaxation at the two anneal temperatures. The rate of relaxation and misfit density increase is higher at 1025 °C, than at 850 °C as expected. The plastic strain relief (δ) due to misfit dislocation density (ρ_{md}) is given by

$$\delta = b_{\text{eff}} \times \rho_{\text{md}}. \quad (3)$$

b_{eff} is the effective Burgers vector, the component of the Burgers vector responsible for misfit strain relief. For 60° dislocations in Si, b_{eff} is $a_{\text{Si}}/2\sqrt{2}$. The misfit dislocation density increase expected for observed strain relief in the Si layers (Table I) is $\sim 10^4$/cm if the strain relief were entirely plastic, that is, due to formation of misfit dislocations. The observed misfit dislocation density increase in the Si layers is ~150/cm, indicating that only a small portion of the strain is in form of plastic strain. This agrees well with the observed elastic strains in Raman measurements. The small increase in misfit dislocation density suggests that despite the early onset of plastic strain relaxation (as evidenced by formation of misfit dislocations even in as-grown Si layers significantly below the critical thickness) the Si layers are quite resistant to subsequent relaxation upon annealing.

The observed strain increase upon annealing is due to the thermal mismatch between Si and $\text{Si}_{0.8}\text{Ge}_{0.2}$ ($\alpha_{\text{Si}} = 3.58 \times 10^{-6}$/K, $\alpha_{\text{Si}_{0.8}\text{Ge}_{0.2}} = 4.37 \times 10^{-6}$/K, at 700 °C). The difference in the coefficients of thermal expansion ($\Delta\alpha$) between the Si and the $\text{Si}_{0.8}\text{Ge}_{0.2}$ layers and the difference between the growth temperature and the anneal temperature (ΔT) results in the thermal mismatch strain ($\Delta\alpha \times \Delta T$). Table II summarizes the strain increase data.

The increase in threading dislocation density with increasing anneal time (Fig. 6) closely follows the increase in misfit dislocation density (Fig. 5) suggesting that any plastic relaxation by increase in the misfit dislocation length is accompanied by nucleation of new threading dislocations. Several short segments of misfit dislocations with a threading dislocation at each end were also observed in the annealed

TABLE II. Strain increase in the SiGe layers compared to the thermal mismatch strain.

Anneal temperature (°C)	Increase in elastaic strain in the SiGe layer ($\Delta\varepsilon$)	Thermal mismatch ($\Delta\alpha \times \Delta T$)
850	1.28×10^{-4}	1.30×10^{-4}
1025	9.07×10^{-4}	3.26×10^{-4}

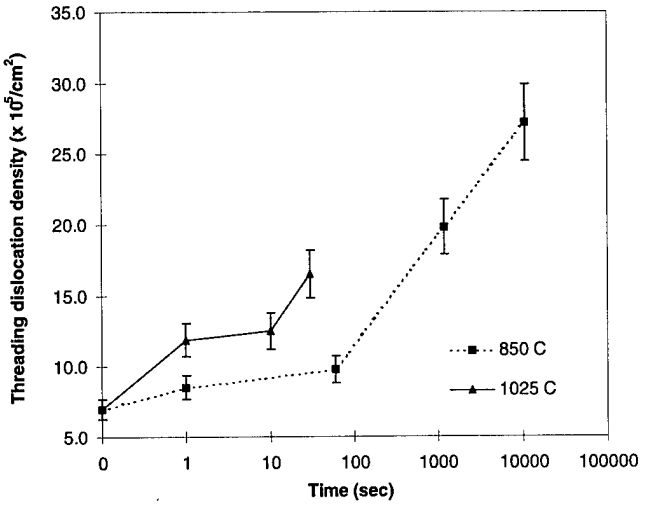


FIG. 6. Threading dislocation density through the Si cap layer as a function of time at the two anneal temperatures. The increase in threading dislocation density closely follows the increase in misfit dislocation density.

samples during the EPD characterization indicating that the freshly nucleated defects had not glided very far (Fig. 7). The misfit segments that are formed by the glide of preexisting threading dislocations would have only one threading dislocation at the leading end. The freshly nucleated misfit segments can be distinguished using this criterion. The observed plastic strain relief is small enough ($\sim 3\times 10^{-6}$) to be easily accommodated by glide of pre-existing threading dislocations ($4\times 10^5/\text{cm}^2$) from the graded layers. However, the pre-existing threading dislocations seem to be pinned, forcing the nucleation of fresh threading dislocations. Surface oxide formed due to the oxidizing annealing ambient (1% O_2) might be responsible for this. The increase in the thickness of the surface oxide with anneal time was confirmed using time-of-flight secondary ion mass spectroscopy. Recently concluded experiments indicate that changing the anneal ambient to nitrogen does indeed decrease the rate of threading dislocation increase, supporting the surface oxide pinning theory.

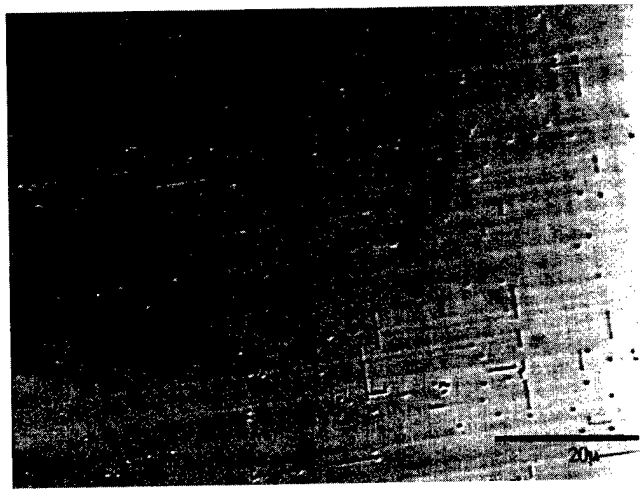


FIG. 7. Nomarski optical micrograph of an etched Si/SiGe sample that was annealed at 850 °C for 3 h. The dark dots are intersections of threading dislocations with the top surface of the film. Note the short straight segments of misfit dislocations at the Si/SiGe interface.

The increase in the threading dislocation density and the misfit density has a linear log (time) dependence at both annealing temperatures. The threading dislocation nucleation rate was determined by measuring the slope of the defect increase with time obtained from Nomarski optical micrographs of samples etched after isothermal anneals. The threading dislocation nucleation rate decreases from 4666 $\text{cm}^{-2}\text{s}^{-1}$ at short times to 162 $\text{cm}^{-2}\text{s}^{-1}$ at longer times at 850 °C. A similar decrease in nucleation rate from 4.9×10^5 to 2×10^4 $\text{cm}^{-2}\text{s}^{-1}$ is observed at 1025 °C. At shorter anneal times, the $Si_{0.8}Ge_{0.2}$ lattice response is slower than the Si lattice response, due to difference in thermal conductivities ($k_{Si}=0.564$ $\text{cm}^{-1}\text{K}^{-1}$, $k_{Ge}=0.25$ $\text{W cm}^{-1}\text{K}^{-1}$). This leads to increased stress and increased nucleation of misfit and threading dislocations in the Si cap layers at shorter anneal times. At longer anneal times, the increase in thermal stress is accommodated by the compressive elastic strain in the $Si_{0.8}Ge_{0.2}$ layer, which reduces the effective stress on the Si cap layer, and hence, the driving force for nucleation of dislocations.

IV. CONCLUSION

We have studied the strain relaxation and accompanied defect generation in thin Si layers grown on graded SiGe layers. The onset of relaxation by the formation of misfit dislocations was observed in Si layers well below the predicted critical thickness. Annealing experiments revealed that the Si layers are significantly resistant to plastic strain relief. Most of the observed strain relaxation is due to the increase in the elastic strain of the underlying SiGe layers. New threading dislocations were found to nucleate even for minimal strain relief, suggesting that the pre-existing threading dislocations are pinned.

ACKNOWLEDGMENTS

The authors would like to thank J. Conner and S. Bagchi for the high resolution TEM characterization.

[1]G. Abstreiter, H. Brugge, and T. Wolf, Phys. Rev. Lett. **54**, 2441 (1985).
[2]Y. J. Mii, Y. H. Xie, E. A. Fitzgerald, D. Monroe, F. A. Thiel, B. E. Weir, and L. C. Feldman, Appl. Phys. Lett. **59**, 1611 (1991).
[3]K. Ismail, M. Arafa, K. L. Saenger, J. O. Chu, and B. S. Meyerson, Appl. Phys. Lett. **66**, 1077 (1995).
[4]N. Sugii, K. Nakagawa, Y. Kimura, S. Yamaguchi, and M. Miyao, Semicond. Sci. Technol. **13**, A140 (1998).
[5]D. J. Paul, A. Ahmed, N. Griffin, M. Pepper, A. C. Churchill, D. J. Robbins, and D. J. Wallis, Thin Solid Films **321**, 181 (1998).
[6]D. K. Nayak and S. K. Chun, Appl. Phys. Lett. **64**, 2514 (1994).
[7]D. K. Nayak, K. Goto, A. Yutani, J. Murota, and Y. Shiraki, IEEE Trans. Electron Devices **43**, 1709 (1996).
[8]J. Wesler, J. L. Hoyt, and J. F. Gibbons, IEEE Electron Device Lett. **15**, 100 (1994).
[9]K. Ismail, Technical Digest of the IEEE International Electron Devices Meeting, 1995, p. 509.
[10]M. Gluck, T. Hackbarth, U. Konig, A. Haas, G. Hock, and E. Kohn, Electron. Lett. **33**, 335 (1997).
[11]T. E. Whall and E. H. C. Parker, J. Phys. D **31**, 1397 (1998).
[12]J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).
[13]J. W. Matthews, J. Vac. Sci. Technol. **12**, 126 (1975).

[14]J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **32**, 265 (1976).
[15]E. A. Fitzgerald, Mater. Sci. Rep. **7**, 87 (1991).
[16]R. Hull and J. C. Bean, Crit. Rev. Solid State Mater. Sci. **17**, 507 (1992).
[17]E. A. Fitzgerald, Y. H. Xie, M. L. Green, D. Brasen, A. R. Kortan, J. Michel, Y. J. Mei, and B. E. Weir, Appl. Phys. Lett. **59**, 811 (1991).
[18]F. K. LeGoues, B. S. Meyerson, J. F. Morar, and P. D. Kirchner, J. Appl. Phys. **71**, 4230 (1992).
[19]J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara, and I. K. Robinson, J. Vac. Sci. Technol. A **2**, 436 (1984).
[20]E. Kasper, H.-J. Herzog, and H. Kibbel, Appl. Phys. **8**, 199 (1975).
[21]D. C. Houghton, J. Appl. Phys. **70**, 2136 (1991).
[22]S. B. Samavedam and E. A. Fitzgerald, J. Appl. Phys. **81**, 3108 (1997).
[23]M. Fatemi and R. E. Stahlbush, Appl. Phys. Lett. **58**, 825 (1991).

High density plasma deposition of device quality silicon nitride. II. Effects of thickness on the electrical properties

M. C. Hugon, F. Delmotte,[a] and B. Agius
Laboratoire Charles Fabry de l'Institut d'Optique (URA D0014), 91403 Orsay, France

E. A. Irene
Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290

(Received 28 October 1998; accepted 9 April 1999)

Dielectric behavior of SiN_x films, fabricated by microwave electron cyclotron resonance discharge, has been studied as a function of film thickness on the basis of the current–voltage and the capacitance–voltage characteristics. In the thickness range ($20\,nm<d<80\,nm$), the resistivity and the critical field for SiN_x were found not to be sensitive to the film thickness (d) and which was opposite to strong dependence of the dynamic dielectric constant ϵ_d on thickness. To explain the ϵ_d behavior as a function of d, a model based on trapped space charge effects is proposed. The dominant mode of electronic conduction, determined from $J-E^{1/2}$ curves and Arrhenius plots of leakage current, appears to be Poole–Frenkel emission only for thicker films ($d>20\,nm$). Finally, the spatial profile of fixed charges reveals that SiN_x/Si interface has a much greater concentration of defects than the bulk film. © *1999 American Vacuum Society.* [S0734-211X(99)00704-0]

I. INTRODUCTION

SiN_x thin films have important applications in microelectronics, optoelectronics, optics, and hard surface coatings. Presently, in the ultralarge scale integration (ULSI) circuit applications, as device dimension shrink into the submicron range, a proportional decrease in the gate dielectric thickness is indispensable.[1] For the development of InP based microelectronics, the formation of a high quality dielectric insulator is key to technology. As a consequence, low temperature deposition is required to prevent thermal decomposition of the InP surface.[2–5] These different requirements have stimulated considerable research on the plasma deposition of dielectric thin films in particular the use of high density plasmas (HDPs).[6]

With film thicknesses below 20 nm now commonplace, control of the film physical and electrical properties becomes very critical. Only a few papers dealing with the electrical characteristics of SiN_x/Si structures have been published,[4,7] even though the behavior of SiN_x devices strongly depends on the film's electrical properties, especially when the film thickness is less than 20 nm. For example, Poole–Frenkel emission is known to be the dominant conduction mechanism for silicon nitride thin films deposited at high temperatures (800–1000 °C) by low pressure chemical vapor deposition (LPCVD)[8] or by radio frequency plasma enhanced CVD (PECVD).[9,10] However, only a few papers[11] have described the electrical properties of SiN_x films deposited by electron cyclotron resonance (ECR)-PECVD, and it is not particularly obvious whether the conduction mechanism will be the same when the thickness varies from nanometers to hundreds of nanometers. Therefore, understanding of the conduction mechanism for such thin films is necessary.

In this article, we report the study of the electrical properties of SiN_x films deposited using the distributed ECR (DECR) plasma method at floating temperature, on silicon substrates, as a function of the film thickness in the 7–80 nm range. This DECR plasma process has proved to be one of the most suitable "soft" deposition techniques to obtain device quality insulator thin films with good electrical properties at very low substrate temperatures (below 150 °C). In view of the increasing interest in SiN_x interlayers in gate dielectric structures, the aim of this article is to study the electrical properties of SiN_x films by means of current–voltage ($I-V$), current–temperature ($I-T$), and capacitance–voltage ($C-V$) measurements. We focus essentially on the positive fixed charge distribution and on the determination of dominant conduction mechanisms.

II. EXPERIMENT

The DECR equipment used in this study was described previously.[12,13] Briefly, it consists of a 400 mm diam chamber equipped with a loadlock; the base pressure in both chambers is below 5×10^{-5} Pa. The microwave power (2.45 GHz) is coupled to the plasma through 14 antennas evenly distributed at the periphery of the chamber and permanent magnets (875 G) are placed near each antenna outside the chamber.

SiN_x films were deposited on (100) n-type ($5-6\times10^{15}\,cm^{-3}$) silicon substrates. Before loading into the chamber, the substrates were subjected to conventional RCA cleaning.[14] This cleaning procedure ended with a 1 min rinse in diluted HF, a 10 min rinse in de-ionized (DI) water and being blown dry with nitrogen.

The deposition process parameters were the same throughout this study: the pressure and microwave power were 0.3 Pa and 1500 W, respectively, and the N_2/SiH_4 flow ratio was fixed to 19 (19 sccm N_2 and 1 sccm SiH_4). From our previous results, these process parameters allow us to optimize the physical and electrical properties of SiN_x thin

[a]Electronic mail: franck.delmotte@iota.u.psud.fr

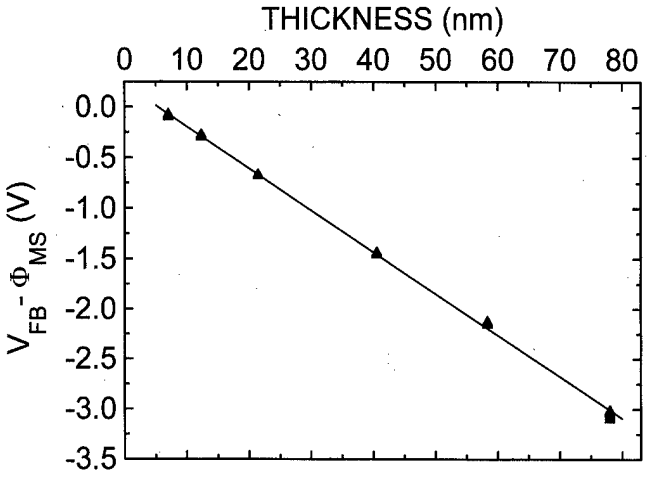


FIG. 1. $\Delta V_{FB} - \Phi_{MS}$ as a function of film thickness ($P = 1500$ W, $N_2/SiH_4 = 19$, $p = 0.3$ Pa).

films.[5,13,15] These films display a homogeneous composition over a large thickness range (7–100 nm). However, the film stoichiometry near the SiN_x/Si interface (N/Si = 1.52 ± 0.05) is higher than in the film bulk. From a previous spectroscopic ellipsometry study, an interface layer with a lower index than that in the bulk film is necessary for modeling the SiN_x/Si structure.[16]

For the electrical measurements, Al dots of 0.152 mm^2 were evaporated through a shadow mask. The metal–insulator–semiconductor (MIS) structures were then annealed for 30 min at 450 °C in forming gas ($Ar + 10\%H_2$). The electrical properties were characterized with ($I-V$) measurements carried out with a HP 4140B picoammeter at various temperatures within the range of 150–500 K, the structure being biased in the accumulation regime. High frequency (1 MHz) $C-V$ characteristics, performed at room temperature with an impedance analyzer (HP 4192 A), were used to estimate the density of the fixed charges in the film.

III. RESULTS AND DISCUSSION

A. Fixed charge distribution

In our previous study of SiN_x/Si structures,[16] we showed evidence for the presence of an interface layer with a higher N/Si ratio than the bulk film and having a lower refractive index than the bulk SiN_x film. In the present study our $C-V$ characteristics show a large flat band voltage shift (ΔV_{FB}) toward negative values which is indicative of fixed positive charges in the SiN_x films. The ΔV_{FB} values indicate about $10^{12}\,cm^{-2}$ positive charges assuming that the charges are at the interface of the SiN_x/Si capacitor or in the bulk of the SiN_x layer.[13,15] This high level of positive charges has been previously observed in similar films.[9,10] Figure 1 displays $\Delta V_{FB} - \Phi_{MS}$ as a function of film thickness (d) where Φ_{MS} is the work function difference between the Al gate and the semiconductor. If ΔV_{FB} is due to an uniform fixed charge distribution, we would expect that ΔV_{FB} varies like d^2. Since Fig. 1 shows a linear relationship, we conclude that the fixed charges are not distributed uniformly in the film. If we

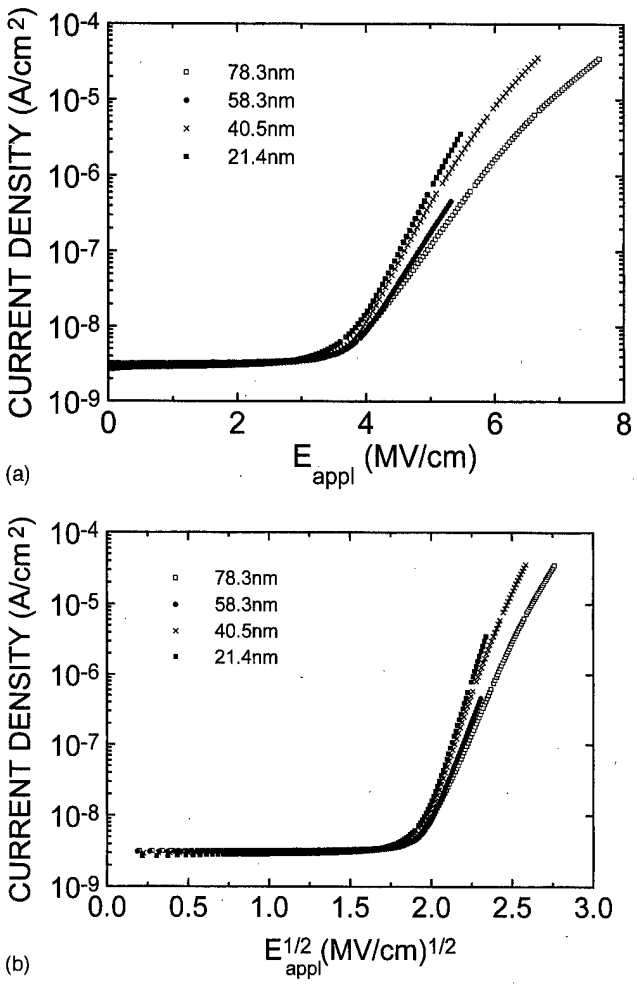


FIG. 2. Current density vs electric field ($J-E_{appl}$) (a) and square root of the electric field ($J-E_{appl}^{1/2}$) (b) for different film thicknesses [78.3, 58.3, 40.5, and 21.4 nm (electric field ramp 5×10^{-3} MV/cm s)] ($P = 1500$ W, $N_2/SiH_4 = 19$, $p = 0.3$ Pa).

assume that the fixed charges are mainly located at the SiN_x/Si interface where the film is nitrogen rich, the charge distribution [$Q_f(x)$] could be described by the following:

$$Q_f(x) = \rho_0 \quad \text{for } 0 < x < d_1, \tag{1}$$

$$Q_f(x) = 0 \quad \text{for } x > d_1, \tag{2}$$

where d_1 is the characteristic length that describes the fixed charge location. Thus, ΔV_{FB} varies linearly with d^2 only for $0 < x < d_1$ [Eq. (1)] and ΔV_{FB} should display a linear dependence with d for $x > d_1$ [Eq. (2)]. This charge profile is in better agreement with the results in Fig. 1 and provides further evidence that the SiN_x/Si interface and SiN_x bulk film are different.

B. Thick films

The current density versus applied electric field strength ($J-E_{appl}$) characteristics for thick SiN_x films (78.3, 58.3, 40.5, and 21.4 nm) are shown in Fig. 2(a), and these data were obtained using an electric field ramp rate (r) of 5×10^{-3} MV/cm s. For low electric fields (<4 MV/cm), the

TABLE I. Electrical properties (ρ, Ec, and ϵ_d) as a function of film thickness ($P=1500$ W, $N_2/SiH_4=19$, $p=0.3$ Pa).

Electric field ramp (MV/cm s)	5×10^{-3}	5×10^{-3}	5×10^{-3}	5×10^{-3}	8×10^{-3}	1.4×10^{-2}
Thickness (nm)	78.3±0.5	58.3±0.5	40.5±0.5	21.4±0.5	12.3±1	7±1
ρ (10^{16} Ω cm)	1.7±0.2	1.3±0.2	1±0.2	0.7±0.1	0.25±0.5	0.12±0.05
Ec (MV/cm)	3.5±0.1	3.5±0.1	3.5±0.1	3.2±0.1	2.44±0.5	1.45±0.5
ϵ_d	7	5.4	3.7	3.2	...	...

different $J-E_{appl}$ characteristics are similar, and from this initial (low field) region of the $J-E_{appl}$ data, we determine the film resistivity (ρ) from the slope of the $J-E_{appl}$ curve at low voltages as well as the critical field (Ec) which is defined as the field required for a 1 nA cm^{-2} increase of the current density through the Al gate. The values of ρ and Ec for different thicknesses are reported in Table I. For films with thickness in the range of 20–78 nm, only small variations of ρ and Ec are observed and values of 0.7×10^{16}–1.7×10^{16} Ω cm and 3.2–3.5 MV/cm, respectively, are obtained for ρ and Ec, which are two electrical properties that describe the bulk behavior of SiN_x films. As was the case for the film composition,[5] ρ and Ec are independent of the film thickness in the film thickness range of 20–78 nm.

For high electric fields (>4 MV/cm), the shape of the $J-E_{appl}$ data is different, as is shown in Figs. 2(a) and 2(b), and is directly related to the conduction mechanism in SiN_x films. Figure 2(b) shows that $\ln J$ is linear with $E_{appl}^{1/2}$ which suggests that the carrier transport process in SiN_x films is via Poole–Frenkel emission, as is commonly observed,[8–10] and the current–voltage characteristics can be expressed as

$$J=J_0\exp\left(\frac{-q}{kT}[\phi-\beta\sqrt{E_{appl}}]\right), \tag{3}$$

where J, ϕ, and J_0 are, respectively, the current density, the Poole–Frenkel barrier height, and a constant that depends on trap density. E_{appl} is equal to V/d where V and d are the applied voltage and the dielectric film thickness, respectively. β is given by

$$\beta=\left(\frac{q}{\pi\epsilon_0\epsilon_d}\right)^{1/2}, \tag{4}$$

where ϵ_0 and ϵ_d are, respectively, the dielectric constant of free space and the dynamic dielectric constant of SiN_x films. Figure 2(b) shows that for all the film thickness studies there is a linear relationship between $\ln J$ and $E_{appl}^{1/2}$. The dynamic dielectric constant is determined directly from the slope of the characteristic $\ln J$ vs $E_{appl}^{1/2}$ and Eqs. (3) and (4). The results are reported in Table I. These results show that ϵ_d is a strong function of film thickness while ρ, Ec, and the atomic ratio N/Si remain constant in this thickness range (20–78 nm). This suggests that the conduction mechanism changes with the thickness. For the determination of ϵ_d, it has been incorrectly assumed that the electric field is constant in SiN_x films and equal to E_{appl}. In fact, it is well known that injected carriers (electrons in our case) from Si are trapped in the nitride film and a space charge is built up.[17] Therefore, the electric field is different at the SiN_x/Si interface and in the bulk nitride film. The distribution of occupied electronic traps, N_T, is the result of the trapping and emission rates for electrons. Andrews *et al.*[18] assumed that N_T is an exponentially decreasing function of distance from the injecting interface and given as

$$N_T=N_0\exp\left(-\frac{x}{x_0}\right), \tag{5}$$

where N_0 is the trap density and x_0 is the characteristic length that describes electronic occupation traps. The electric field distribution $E(x)$ is deduced by integration of the Poisson equation:

$$E=E_0+E_T\left[1-\exp\left(-\frac{x}{x_0}\right)\right], \quad \text{where } E_T=\frac{qN_0x_0}{\epsilon_0\epsilon_S}, \tag{6}$$

where E_0 represents the electric field without charge in SiN_x film, and ϵ_S is the static dielectric constant of the nitride film. Integrating the Poisson equation a second time yields the total voltage through the structure:

$$V=\int_0^d E(x)dx=(E_0+E_T)d+E_Tx_0\left[\exp\left(-\frac{d}{x_0}\right)-1\right]. \tag{7}$$

From Eq. (7), the applied field can then be expressed as

$$E_{appl}=\frac{V}{d}=(E_0+E_T)+\frac{E_Tx_0}{d}\left[\exp\left(-\frac{d}{x_0}\right)-1\right]. \tag{8}$$

Close to the SiN_x film/Si substrate interface, the traps are filled and a high negative space charge is present. Further into the dielectric, a point is reached where the traps are emptied ($N_T\approx0$) due to field assisted emission [the Poole–Frenkel (PF) mechanism]. This occurs at a field E_{PF} equals to

$$E_{PF}=E_0+E_T=E_{appl}-\frac{E_Tx_0}{d}\left[\exp\left(-\frac{d}{x_0}\right)-1\right]. \tag{9}$$

If we assume that $d\gg x_0$, E_{PF} can be written as

$$E_{PF}=E_{appl}+\frac{E_Tx_0}{d}. \tag{10}$$

Finally Eq. (3) should be written as

$$J=J_0\exp\left(\frac{-q}{kT}[\phi-\beta_{eff}\sqrt{E_{PF}}]\right). \tag{11}$$

β_{eff} is given by an equation similar to Eq. (4):

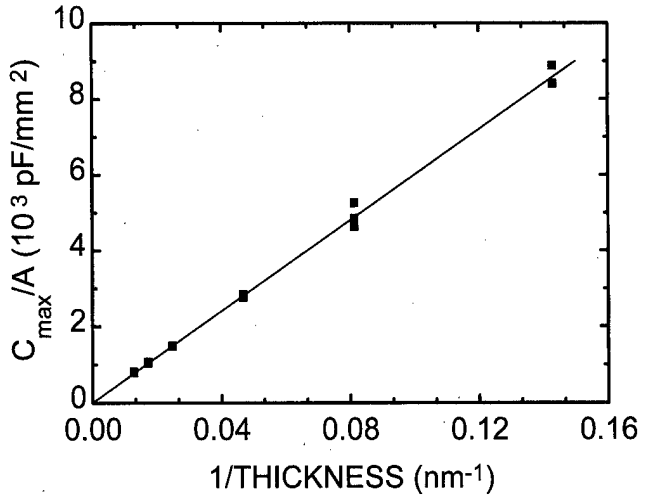


FIG. 3. C_{max}/A ratio as a function of $1/d$ ($P=1500$ W, $N_2/SiH_4=19$, $p=0.3$ Pa).

$$\beta_{\text{eff}}=\left(\frac{q}{\pi\epsilon_0(\epsilon_d)_{\text{eff}}}\right)^{1/2}, \tag{12}$$

where $(\epsilon_d)_{\text{eff}}$ is the effective dynamic dielectric constant of the film. From Eq. (10), E_{PF} decreases with increasing thickness. Therefore, for a given E_{appl}, the current density is lower for thick films. This result is confirmed by the $J-E_{\text{appl}}$ and $J-E_{\text{appl}}^{1/2}$ characteristics in Figs. 2(a) and 2(b), respectively. From Eq. (11), $\ln J$ is proportional to $E_{\text{PF}}^{1/2}$. Since we observe a linear relationship between $\ln J$ and $E_{\text{appl}}^{1/2}$, we deduce that E_{PF} is proportional to E_{appl} ($E_{\text{PF}}=\alpha E_{\text{appl}}$, α being a constant that decreases with increasing d). Using Eqs. (3) and (11), we find that the apparent β depends on nitride thickness:

$$\frac{d\ln J}{d(E_{\text{appl}}^{1/2})}=\frac{q\beta}{kT}=\frac{q\beta_{\text{eff}}}{kT}\alpha^{1/2}. \tag{13}$$

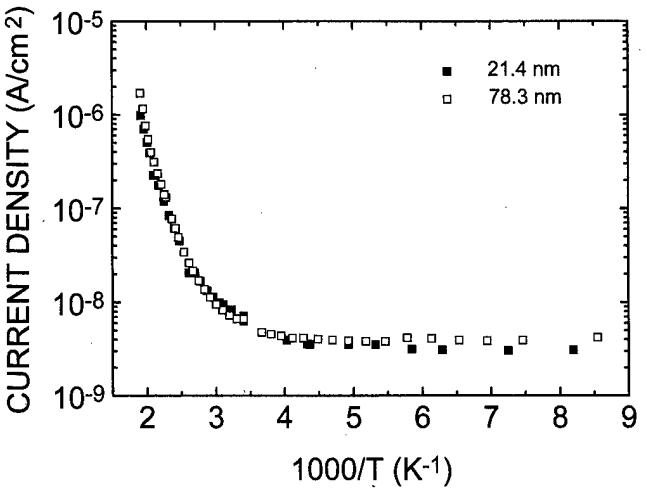


FIG. 4. Temperature dependence of the current density for two film thicknesses (78.3 and 21.4 nm). $\log J$ is represented vs $1/T$ ($P=1500$ W, $N_2/SiH_4=19$, $p=0.3$ Pa).

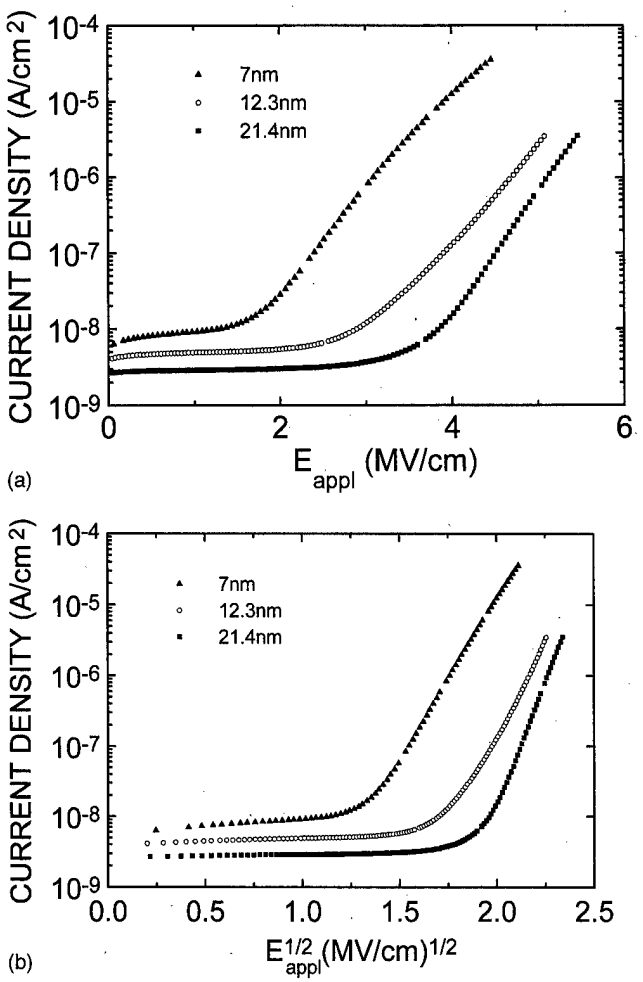


FIG. 5. Current density vs electric field ($J-E_{\text{appl}}$) (a) and square root of the electric field ($J-E_{\text{appl}}^{1/2}$) (b) for different film thicknesses [21.4, 12.3, and 7 nm (for the electric field ramp; see Table I)] ($P=1500$ W, $N_2/SiH_4=19$, $p=0.3$ Pa).

Figure 3 shows a linear relationship between the C_{max}/A ratio and $1/d$ where C_{max} and A are, respectively, the capacitance in the accumulation regime and the Al dot area. Thus we can assume that the static dielectric constant ϵ_S does not depend on film thickness. Furthermore, we can assume that β_{eff} is also independent of thickness (20–78 nm), since we observe no significant changes in ρ, Ec, and ϵ_S. Using Eq. (13), it is found that β decreases with increasing d and therefore the apparent ϵ_d increases with d. Table I confirms this conclusion for film thicknesses in the range of 20–78 nm.

The value of the dynamic dielectric constant is quite high, especially for large film thicknesses ($\epsilon_d=7$ for 78.3 nm). Therefore, to confirm the Poole–Frenkel mechanism, it is necessary to examine the Arrhenius plots of leakage current for a constant electric field. Since the purpose of this article is not to find the charge distribution in SiN_x films, the same current density, measured at room temperature, was chosen. Figure 4 displays Arrhenius plots of leakage current for two different film thicknesses (78.3 and 21.4 nm). At high temperature, the currents yield straight lines which strongly suggests that the electronic conduction corresponds to the Poole–Frenkel emission mechanism.

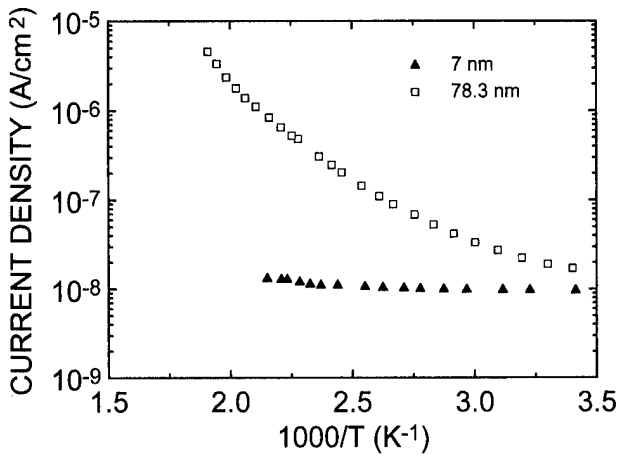


FIG. 6. Temperature dependence of the current density for two film thicknesses (7 and 21.4 nm). $\log J$ is represented vs $1/T$ ($P=1500$ W, $N_2/SiH_4=19$, $p=0.3$ Pa).

From the dependence of $\ln J$ vs $1/T$ in Fig. 4, the Poole–Frenkel barrier height Φ is found to be 1.2 ± 0.2 eV. In many silicon nitride films, the trap level Φ is in the range of 1–1.5 eV.[19] Since evaluation of Φ by Poole–Frenkel emission has an error of ±0.2 eV, it is difficult to compare the conductivity of different silicon nitride films just by comparing the values of Φ.

C. Thin films

Figure 5(a) and Table I show that, for thin films (<21 nm), there is a degradation of electrical properties. Contrary to Fig. 2(a), Fig. 5 shows that in the low field region the film characteristics are different, but this difference is essentialy due to the electric field ramp rate (r) which is not the same (see Table I). In the low field region, the current strongly depends on the displacement current $[I=C(dV/dt)]$ which is correlated to r and depends slightly on the leakage current. As can be seen in Table I, the resistivity and critical field decrease with film thickness. This evolution could be correlated to the SiN_x/Si interface which plays a more important role for thin films. However, we observe a linear relationship between $\ln J$ and $E_{\text{appl}}^{1/2}$ which could mean that the current density is governed by Poole–Frenkel emission. The value of the dynamic dielectric constant determined from the slope of the characteristic $\ln J$ vs $E_{\text{appl}}^{1/2}$ is too high ($\epsilon_d=8.4$ for a thickness of 7 nm) to attribute the conduction mechanism to Poole–Frenkel emission. The current is not or not only limited by the bulk conduction through the SiN_x film. An Arrhenius plot (Fig. 6) confirms this conclusion, and, contrary to Fig. 4, we observe a slight change of the current density as a function of the measurement temperature for thin films. Although the behavior is very similar to the Fowler–Nordheim emission process, the film is too thick (7 nm) to consider this conduction mechanism. The SiN_x/Si interface, which is more nitrogen rich and where the fixed charges are mainly located, now has more influence than the SiN_x bulk film.

IV. CONCLUSION

In this article, a DECR plasma source has been used to deposit SiN_x films, at floating temperature, with thicknesses in the range of 7–80 nm. For films thicker than 20 nm the influence of film thickness is minimal on the resistivity ($0.7-1.7\times10^{16}\,\Omega$ cm) and on the critical field (3.2–3.5 MV/cm). For these films, the dominant conduction mechanism is Poole–Frenkel emission. The trap level, as determined from the temperature dependence of current density, is equal to 1.2 ± 0.2 eV. The assumption of a constant electric field throughout the structure is responsible for an apparent ϵ_d variation. For thinner films (<20 nm), the electrical properties become degraded, particularly the critical field ($Ec=1.45$ MV/cm for a 7 nm thick film). The conduction mechanism is not only determined by the bulk conduction; the SiN_x/Si interface also plays an important role. This interface presents properties which differ from those for the SiN_x bulk film. We found, from $C-V$ characterization, that the fixed charges are mainly located at the SiN_x/Si interface.

In conclusion, the DECR process appears to be a promising technique by which to prepare gate dielectric films in a wide range of thicknesses.

ACKNOWLEDGMENTS

This work was supported by the Centre National de la Recherche Scientifique (GDR86) and France Telecom. One of the authors (E.A.I.) gratefully acknowledges the support of the National Science Foundation (NSF) Division of Materials Research.

[1]A. Izumi and H. Matsumura, Appl. Phys. Lett. **71**, 1371 (1997).
[2]R. W. M. Kwok, W. M. Lau, D. Landheer, and S. Ingrey, J. Vac. Sci. Technol. A **11**, 990 (1993).
[3]D. Landheer, G. H. Yousefi, J. B. Webb, R. W. M. Kwok, and W. M. Lau, J. Appl. Phys. **75**, 3516 (1994).
[4]S. Garcia, I. Martil, G. Gonzalez Diaz, E. Castan, S. Duesñas, and F. Fernandez, J. Appl. Phys. **83**, 600 (1998).
[5]M. C. Hugon, F. Delmotte, B. Agius, and J. L. Courant, J. Vac. Sci. Technol. A **15**, 3143 (1997).
[6]M. A. Lieberman and R. A. Gottscho, Phys. Thin Films **18**, 1 (1994).
[7]Y. C. Park, W. B. Jackson, N. M. Johnson, and S. B. Hagstrom, J. Appl. Phys. **68**, 5212 (1990).
[8]S. M. Sze, J. Appl. Phys. **38**, 2951 (1967).
[9]S. V. Hattangady, G. G. Fountain, R. A. Rudder, and R. J. Markunas, J. Vac. Sci. Technol. A **7**, 570 (1989).
[10]S. W. Hsieh, C. Y. Chang, Y. S. Lee, C. W. Lin, and S. C. Hsu, J. Appl. Phys. **76**, 3645 (1994).
[11]Y. C. Jeon, H. Y. Lee, and S. K. Joo, J. Appl. Phys. **75**, 979 (1994).
[12]F. Plais, B. Agius, F. Abel, J. Siejka, M. Puech, G. Ravel, P. Alnot, and N. Proust, J. Electrochem. Soc. **139**, 1489 (1992).
[13]S. Sitbon, M. C. Hugon, B. Agius, F. Abel, J. L. Courant, and M. Puech, J. Vac. Sci. Technol. A **13**, 2900 (1995).
[14]W. Kern and D. A. Puotinen, RCA Rev. **6**, 187 (1970).
[15]F. Delmotte, M. C. Hugon, B. Agius, and J. L. Courant, J. Vac. Sci. Technol. B **15**, 1919 (1997).
[16]F. Delmotte, M. C. Hugon, B. Agius, and E. A. Irene, J. Vac. Technol. B (to be published).
[17]P. C. Arnett and D. J. DiMaria, J. Appl. Phys. **27**, 34 (1975).
[18]J. M. Andrews, B. G. Jackson, and W. J. Polito, J. Appl. Phys. **51**, 495 (1980).
[19]K. Watanabe and S. Wakayama, J. Appl. Phys. **53**, 568 (1982).

Characterization of Al, Cu, and TiN surface cleaning following a low-*K* dielectric etch

P. J. Matsuo,[a] T. E. F. M. Standaert, S. D. Allen, and G. S. Oehrlein[b]
Department of Physics, University at Albany, State University of New York, Albany, New York 12222

T. J. Dalton[c),d)]
IBM Microelectronics, Hopewell Junction, New York 12533

(Received 29 May 1998; accepted 5 April 1999)

The cleaning of Al, TiN, and Cu blanket samples was investigated in a high density inductively coupled plasma reactor, and compared with results for silicon. After simulating the dielectric overetch exposure of these substrates to a CHF_3 discharge, an *in situ* O_2 plasma clean and subsequent Ar^+ premetal sputter clean were performed and evaluated using ellipsometry and x-ray photoelectron spectroscopy. Following the fluorocarbon exposure, significant C and F residues were observed. Exposure to a O_2 plasma clean greatly reduced this contamination. Subsequent treatment with an Ar^+ sputter further reduced the thickness of the modified surface layer. Comparisons of the cleaning results with silicon suggest an efficient cleaning procedure, especially in the cases of copper and titanium nitride. The response of several blanket, oxide-like low-*K* dielectrics to the O_2 plasma treatment were also studied and compared to SiO_2. While a fluorinated SiO_2(SiOF) exhibited SiO_2-like stability, deep modifications were observed in both hydrogen silsesquioxane and methyl silsesquioxane, consistent with the removal of hydrogen and carbon from these films. These results were compared to a dedicated chamber design, where no fluorocarbons contaminate the reactor. The dedicated chamber methodology offered no significant advantage. © *1999 American Vacuum Society.* [S0734-211X(99)00204-8]

I. INTRODUCTION

With the evolution of feature sizes to submicron dimensions, RC delay becomes a critical factor in the overall performance of logic chips based on multilevel interconnect designs. There are two conventional ways of reducing the RC delay, one of which is to reduce the resistivity in the metal lines. To achieve this, interconnect metals with lower resistivity can be implemented. The current methodology for producing such integrated circuits involves a damascene approach wherein the features are first etched in a dielectric. Then, after a premetal clean, these trenches and vias are filled with the appropriate liner and metallization. This procedure is repeated for each metallization layer. Producing a clean interface between these metallization layers reduces the contact resistance of the plugs, and hence the overall resistivity of the line. In this work, we will address both aluminum and copper thin films, which are important interconnect metals. We will then compare the results with silicon, for which several evaluations of O_2 plasma and Ar^+ sputter cleaning have been made.[1–3] Diffusion of Cu or Al into the Si or SiO_2 makes it imperative to use a diffusion barrier. TiN has proven effective in preventing this diffusion.[4–7] We therefore include TiN films in our study.

The other conventional way of reducing RC delay is to use a dielectric material of a lower dielectric constant (*K*), hence reducing the capacitance between the metal lines. We have investigated several oxide-like dielectrics: a fluorinated SiO_2 (SiOF), hydrogen silsesquioxane (HSQ), and methyl silsesquioxane (MSQ). We have tested their stability under exposure to a O_2 plasma clean and compared these results to a conventional thermal SiO_2. Each of these dielectrics offers a dielectric constant considerably lower than the $K=3.9$ exhibited by the conventional thermal oxide. More information of these dielectrics is available in a companion publication.[8]

A schematic of the processing steps required for this work is shown in Fig. 1. In step 1 the contact hole is etched in the dielectric material using a fluorocarbon plasma. Following this etch step, there is a CF_x residue on the structure bottom. This residue increases the contact resistance between layers and hence the overall connection resistivity. This etching process is addressed by Standaert *et al.*[8] Step 2 involves removing this residue with a O_2 plasma. This, in turn leaves an oxidized surface. Although the oxygen contamination is more inert than the fluorine, which can lead to chemical degradation of the metal and dielectric,[9] this oxidation still creates a high contact resistance and thus, must be removed.[10] In step 3, the final step, this oxidation is removed using an Ar^+ premetal sputter clean.

The right side of the illustration depicts the steps involved in cleaning up after the etch of an organic class of dielectric. Although quite similar to the steps 2 and 3 on the left, this process will be addressed in a future work. This article focuses on the response of the Si, Al, Cu, and TiN surfaces to each step in the cleaning process, after simulating the dielectric overetch on blanket surfaces. Selected conditions were applied to actual structures, where the O_2 induced modifications of the dielectric were observed.

[a]Electronic mail: pm1747@csc.albany.edu
[b]Electronic mail: oehrlein@csc.albany.edu
[c]Electronic mail: tjdalton@us.ibm.com
[d]Work done while at Digital Semiconductor, Hudson, MA 01749.

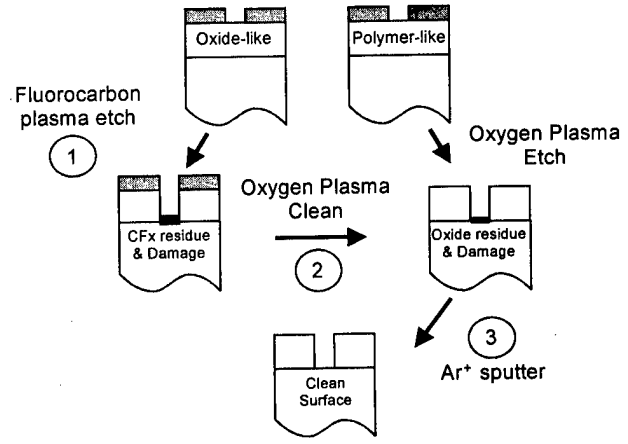


FIG. 1. Schematic of the processing steps involved in this study.

II. EXPERIMENTAL SETUP AND PROCEDURE

The cleaning of the residual CF_x and oxides was carried out in a planar coil inductively coupled plasma (ICP) reactor. The specifics of this reactor are explained elsewhere.[8,11,12] A schematic of this tool is shown in Fig. 2.

During processing, the surface modifications were monitored in real time using a SOFIE rotating compensator ellipsometer in the polarizer-compensator-sample-analyzer (PCSA) configuration with a 638.2 nm He/Ne laser source beam. After processing, samples were transferred under ultrahigh vacuum (UHV) to a multitechnique surface analysis vessel for x-ray photoelectron spectroscopy (XPS) measurements. Some samples were selected for secondary electron microscopy (SEM) measurements after atmospheric transport.

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. Metal and barrier layers

Between metallization layers, a clean interface ensures a low contact resistance. Two relevant interconnect metals, aluminum and copper, were investigated after a cleaning procedure designed to remove the CF_x residue left from the dielectric etch step. Also of related importance is the cleaning efficiency of titanium nitride, a commonly used diffusion barrier. Silicon is of importance both at the lowest level of metallization and as a reference for comparing the cleaning efficiency of the other materials.

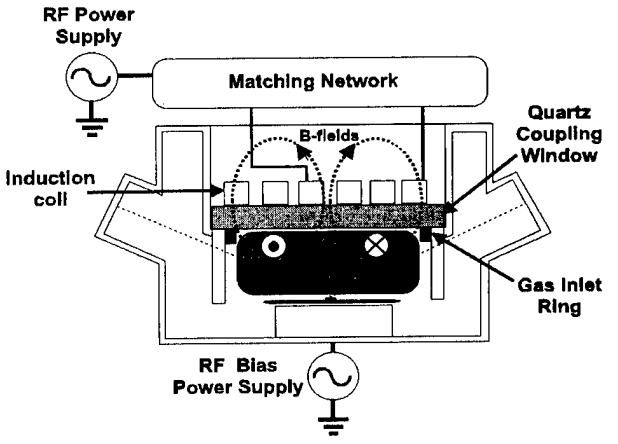


FIG. 2. Schematic of the tool used in this work.

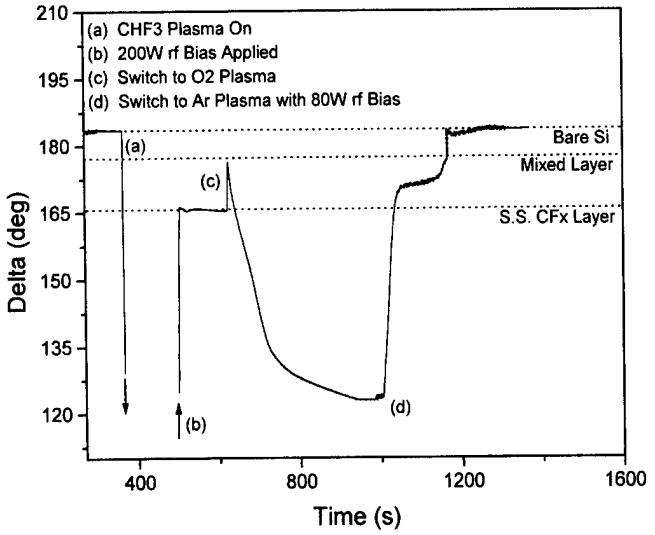


FIG. 3. Time evolution of the ellipsometric variable delta monitored in real time during the CF_x contamination of c-Si, the O_2 cleaning step, and the final Ar sputter.

1. Silicon

The modifications on the silicon surface were monitored in real time with ellipsometry during each phase of the processing steps depicted in Fig. 1. These results are presented in Fig. 3 in terms of the ellipsometric variable delta as a function of time. In this case, a decrease in delta represents an increase in the thickness of a modified surface layer. The conductivity of the c-Si results in its semi-infinite appearance to the 632.8 nm source beam, and insures that all observed modifications are taking place on the Si surface, since the silicon thickness change cannot be detected. At time (a), a 1400 W 6 mTorr 40 sccm CHF_3 discharge is ignited. Immediately, a thick fluorocarbon layer is formed on the Si surface. At time (b), a 200 W bias is applied to the sample. The induced ion bombardment quickly etches away the CF_x film until only a thin steady state CF_x layer and a thin reacted Si layer remain.[12] The plasma is then extinguished and the chamber evacuated until at time (c) a 1000 W 6 mTorr 40 sccm O_2 plasma is started. At this time, the fluorocarbon layer is quickly removed from the reacted Si surface and an oxide layer begins to grow. At time (d), an Ar plasma is ignited and 80 W radio frequency (rf) bias (corresponding to 25 V bias voltage) is applied to the sample. This ion bombardment removes the oxide and reacted layer until the bare Si surface is recovered.

We observe that the oxide layer is considerably thicker than the CF_x mixed layer. A previous work describes the CF_x layer on Si in great detail.[12] XPS data were taken at each step of this process, and these results are shown in Figs. 4 and 5.

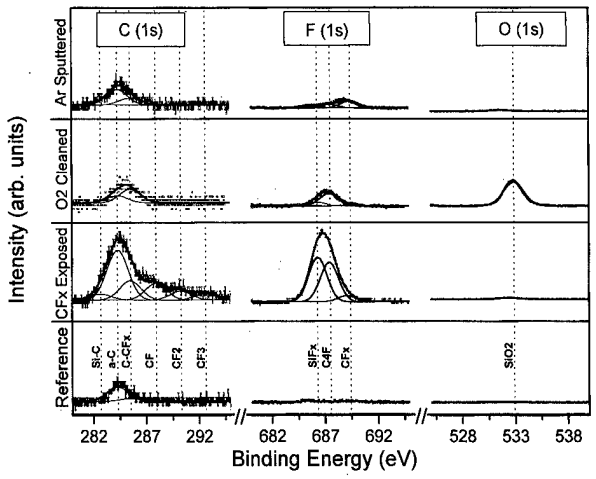


FIG. 4. XPS spectra representing the important stages of the Si cleaning process. Relevant binding energies are listed in the Appendix. The O(1*s*) spectra was reduced in intensity by a factor of 8.

The lowest curves in Fig. 4 depict the state of the sample after a HF dip, i.e., with the native oxide removed. Steps 1–3, as depicted in Fig. 1, are then presented with each successively higher curve. Reference binding energies for the curve fits are marked by the vertical dotted lines, and are listed in the Appendix. Fits were obtained by constraining the full width at half maximum (FWHM) and binding energies. Comparing the post-Ar^+ sputter stage with the HF dipped reference sample, we see that only a slight contamination, attributed to CF_x related species, remains. More resolved Si(2*p*) spectra are shown in Fig. 5. Again, panel (i) represents the reference sample, which in this case has been HF dipped to remove the native oxide. Panel (ii) illustrates the Si surface after exposure to the CHF_3 plasma, panel (iii) after the O_2 plasma clean, and panel (iv) after the final Ar^+ sputter step. The elemental Si peak at 99.15 eV is in good

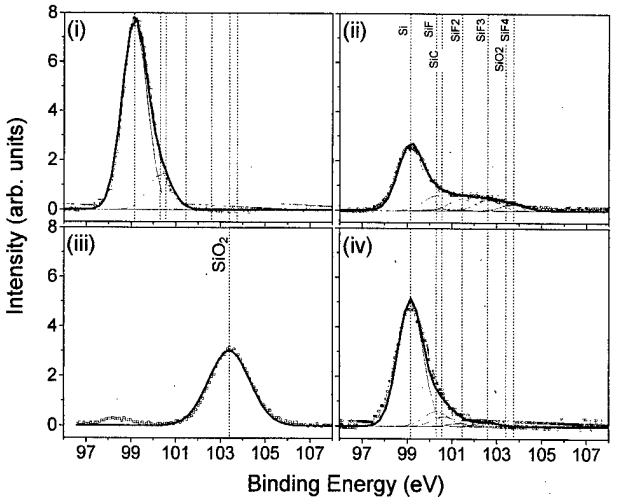


FIG. 5. Si(2*p*) XPS spectra and chemical shifts. Curve (i) depicts the state of the surface as received, (ii) after CF_x exposure, (iii) following the O_2 plasma clean, and (iv) after the Ar^+ sputter. Chemical shifts were fitted with a constant FWHM of 1.36 eV except in panel (v) where no constraint was imposed. The Appendix lists the relevant binding energies.

agreement with values in the literature.[13] We have fitted the silicon-fluorine bonded species with a +1.15 eV shift from the elemental peak for each fluorine bonded to the silicon, i.e., 100.3 eV for SiF, 101.45 eV for SiF_2, 102.6 eV for SiF_3, and 103.75 eV for SiF_4. These choices give a satisfactory fit and are in agreement with several previous works.[14–16] Because we are using the fitted chemical shifts for a qualitative understanding, we have ignored the spin-orbit separation in the fitting of the elemental peak. Evidence for a Si–C bonded species at 282.7 eV for the fluorocarbon plasma treated surface was found (see Fig. 4),[17] but ignored in the fit of the Si(2*p*) spectra in panel (ii). Although this compound exists, it would overlap the Si–F bonded species, making its deconvolution difficult.

The diminished intensities in panels (ii) and (iv) suggest the presence of a CF_x overlayer. Although most significant after the CHF_3 exposure, such a residue is suggested by all the spectra, aside from the HF dipped reference.

The thicknesses of the modified surface layers on Si were also determined using ellipsometry. The refractive index used in the fit was $n+ik=3.866+i0.028$ for the Si, $n=1.48$ for the steady state CF_x layer, and $n=1.46$ for the SiO_2-like overlayer. Such fitting yielded a steady state CF_x thickness of 3 ± 1.0 nm. The thickness of the oxide layer was determined to be 16 ± 1.0 nm. The error reported for these measurements are primarily due to the quality of the fit, with intentional misfits used as reference. In general, the refractive index of the CF_x overlayer will represent the F:C ratio of the film, i.e., more fluorinated films will exhibit a lower refractive index.[18] In this case, $n=1.48$ provided the best fit to the data.

Using photoelectron escape depth considerations, the apparent thicknesses of the modified Si surface were calculated by comparing the reacted Si(2*p*) peak areas with the area of the elemental Si peak, as described elsewhere.[16] This yielded a thickness of 1.1 ± 0.3 nm for the HF dipped reference sample, 3.0 ± 0.6 nm for the CHF_3 treated sample, greater than 10 nm for the O_2 plasma treated sample (determined from the complete loss of the elemental peak), and 1.5 ± 0.4 nm for the Ar^+ treated sample. Within the error, primarily determined by the peak fits, these thicknesses are in agreement with the ellipsometrically determined thicknesses. In the case of the CHF_3 plasma exposed surface, the agreement between ellipsometric and XPS determined thicknesses suggests little if any nonreacted CF_x residue on the sample surface. It is possible that some adventitious carbon does indeed lie on the surface, but we cannot support this conclusion beyond the 1 nm error in the ellipsometric calculation.

The XPS spectra presented in Figs. 4 and 5 indicate a successful clean following the final Ar^+ sputter step.

2. Aluminum

Figure 6 shows the time evolution of the ellipsometric variable delta for an Al film throughout the processing sequence as was shown for silicon in Fig. 3. Again, the thickness of the Al layer was sufficient for it to be semi-infinite to the 632.8 nm source beam, therefore all changes in delta are

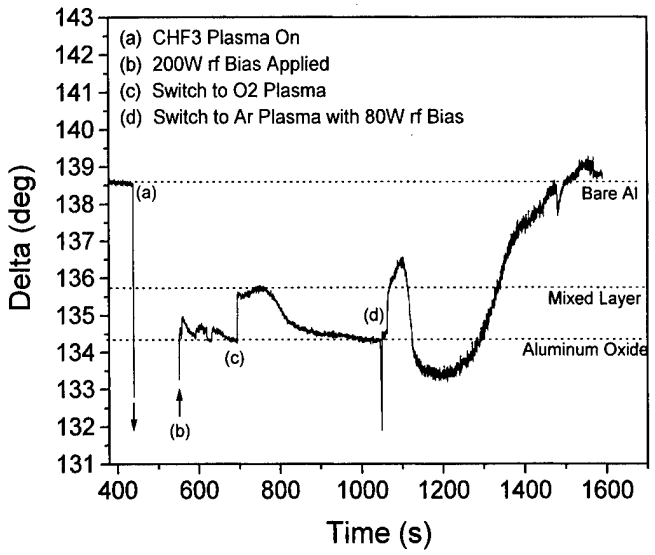


FIG. 6. Time evolution of the ellipsometric variable delta monitored in real time during the CF_x contamination of Al, the O_2 cleaning step, and the final Ar sputter.

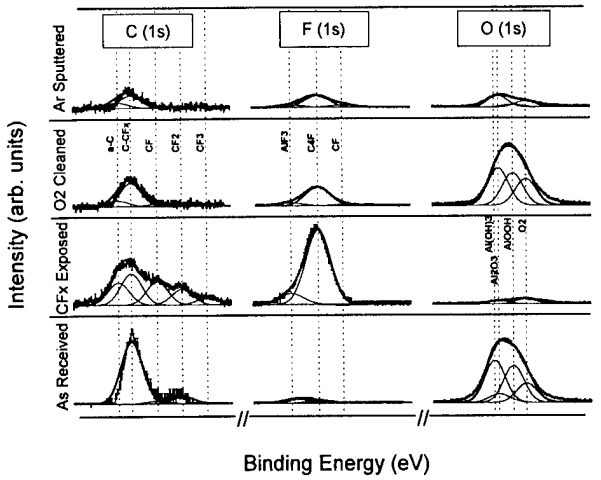


FIG. 7. XPS spectra representing the important stages of the Al cleaning process. The spectra were fit using binding energies as listed in the Appendix with FWHM of 2.36 eV for the C(1*s*) and F(1*s*) spectra and 2.3 eV for the O(1*s*) spectra. The total area of all the F(1*s*) and O(1*s*) peaks were cut in half with respect to the C(1*s*) peaks.

resulting from the surface modification. The Al was first sputtered with Ar^+ to remove the native oxide. At time (d), when the Ar plasma is ignited and 80 W rf bias is applied to the sample, there is an initial period of optical response of the ellipsometer to the untuned discharge, after which the oxide is sputtered away to the bare Al point. This optical response varies with the chamber pressure as the system relaxes from the plasma ignition. We observe that the steady state CF_x layer and the oxide layer are of roughly the same thickness (~5 nm), as opposed to the silicon processing. This can be attributed to the oxidation kinetics of Al_2O_3 taking place at the surface thus leading to the formation of a self-passivating oxide, while for SiO_2, diffusion of oxygen through the SiO_2 interface supports oxidation at the Si/SiO_2 interface.[19]

XPS data were taken at each of step of this process, and these results are shown in Fig. 7. The lowest curves depict the state of the sample as they were received, i.e., with native oxide and carbonaceous contamination. Steps 1–3, as depicted in Fig. 1, are again presented with each successively higher curve. Reference binding energies for the fitted curves are marked by the vertical dotted lines, and are listed in the Appendix. The reader is also directed to Fig. 8, where the Al(2*p*) region has been magnified and fitted with Gaussians representing individual chemical shifts. Panel (i) represents the "as received" reference spectrum, panel (ii) is after the Al has been exposed to the dielectric etch environment, panel (iii) shows the effect of the O_2 plasma clean, and panel (iv) shows the state of the Al surface after the Ar^+ sputter step. The Al peak at 72.65 ± 0.5 eV is consistent with several values found in the literature.[13,20–23] The spin-orbit separation is too small for us to resolve (~0.42 eV),[27] so we fitted the Al(2*p*) doublet with one Gaussian peak. The contribution of the oxides at 73.95 ± 0.2 and 75.25 ± 0.2 eV are attributed to Al–O and Al_2O_3 oxidation states,[21,24] although contributions from $Al(OH)_3$ and AlOOH are supported by the O(1*s*) spectra,[25] and assumed present. Both these peak positions are in excellent agreement with the observed energies reported by both Carley and Roberts[21] and by Flodstrom *et al.*[24] The 76.45 ± 0.5 eV peak is attributed to the Al–F_x bonded species.[13] In order to fit the CHF_3 treated spectrum properly, an additional peak at ~77.75 eV was also necessary and we also attribute this to a fluoride contribution.[20]

In the reference sample, strong oxidation and carbonaceous residues are observed, both in terms of the O(1*s*) and C(1*s*) spectra seen in Fig. 7, and the reacted Al(2*p*) peak components outlined above and illustrated in Fig. 8. Only a small fluorine signal is observed, and of this, the majority results from Al–F bonded species as seen from the 685.19

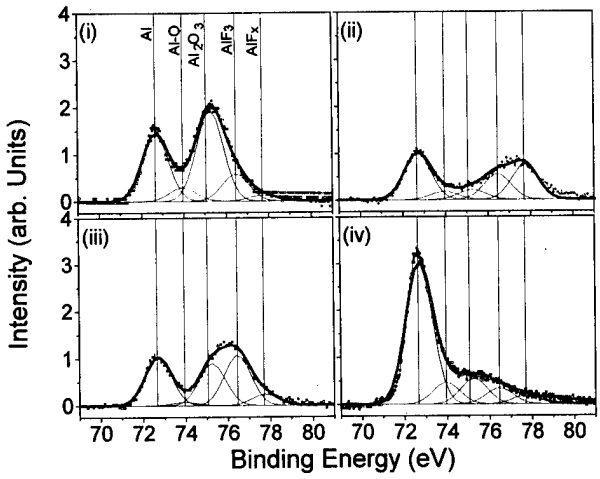


FIG. 8. Al(2*p*) XPS spectra and fitted chemical shifts for (i) the as received reference sample, (ii) after the CF_x exposure, (iii) following the O_2 plasma clean, and (iv) after the Ar^+ sputter. The binding energies used for the fits are listed in the Appendix.

eV Al–F component of the F($1s$) peak and the high energy side of the Al($2p$) spectrum.

In the CHF_3 treated samples, there is a significant CF_x overlayer. Although most of this layer is deposited on the Al surface, there are still Al–F bonds visible suggested by the low binding energy shoulder (~685.19 eV) of the F($1s$) spectrum and more easily seen in the Al($2p$) spectra of Fig. 8. The strong reduction in the observed O($1s$) intensity suggests the efficient removal of the oxide, and this is supported by the diminished Al_2O_3 and/or $Al(OH)_3$ peak at 75.25 eV in the Al($2p$) spectrum.

After the O_2 plasma cleaning step, a significant level of the Al_2O_3 returns. There is substantially less fluorine in the form of C–F_x bonded species, although the Al–F_x component of the Al($2p$) spectrum is still prominent. Further removal of both fluorine and carbon is found after the Ar^+ sputter clean. There is also substantially less Al–F bonded species as seen in the Al($2p$) spectrum of Fig. 8.

Thicknesses of the modified Al surface can be obtained by fitting the ellipsometric data presented in Fig. 7. Considering the refractive index of the Al substrate to be $n+ik = 1.60-i7.53$[26–28] and of the deposited CF_x film to be $n = 1.5$ we obtain a steady state thickness of the fluorocarbon layer to be 2 ± 1.0 nm. For the oxidation of the Al, we used a refractive index of $n=1.62$.[29] This gave a consistent fit in delta with a $<2°$ shift in Ψ and yielded a thickness of 2.7 ± 1.5 nm.

The XPS spectra as seen in Fig. 8 were used to determine an apparent thickness of the modified Al surface. As was done in the case of silicon, the ratio of the reacted Al to the elemental Al peak areas was used for this calculation. The density coefficient k was assumed to lie between 0.2 and 0.5 and the mean free path of the Al($2p$) photoelectrons through this modified layer was assumed to be 2.5 nm.[30] Such a range of density coefficients allow a mixed layer stoichiometry from AlX to AlX_3, where X can represent an oxygen or fluorine. These assumptions are the major source of the error in these values. For the untreated reference sample, a modified layer thickness of 5.12 ± 1.0 nm was found. After exposure to the fluorocarbon plasma, we calculated a thickness of 4.77 ± 1.0 nm, while the O_2 plasma treated sample yielded 5.32 ± 1.1 nm. Finally, after the Ar^+ treatment, the thickness was determined to be 2.55 ± 0.8 nm. The reason for the discrepancy in thicknesses measured by ellipsometry and XPS is attributed to the bare Al point considered by the ellipsometry calculations to actually correspond to a modified surface of 1–1.5 nm.

Throughout the F($1s$) XPS spectra presented in Fig. 7, Al–F bonding is visible. The clean aluminum surface is extremely reactive, so that trace fluorine contamination in the chamber could account for such signals, especially following the final Ar^+ sputter step. This surface contamination is also evident in the return of the bulk plasmon peak seen in Fig. 9, but the absence of the surface plasmon at lower binding energy.[31,32] In Fig. 9, each data point has been averaged with the nearest two neighboring data points. Such peaks are attributed to discrete energy losses via the collective excitation of valence electrons.[24,33]

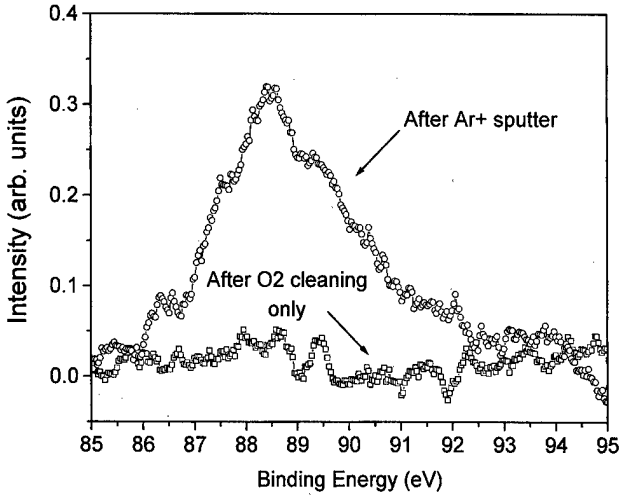


FIG. 9. XPS spectra of the Al($2p$) plasmon at various stages in the cleaning process.

3. Copper

In Fig. 10 we present the time evolution of the ellipsometric response of the Cu surface throughout the processing sequence as was shown for Si in Fig. 3. Consistent with the Si processing, yet as opposed to Al, the oxide layer thickness is almost three times greater than the CF_x damage layer thickness. It is important to note, that although the thicknesses of the modified CF_x and oxide layers are both proportional to the change in delta in this region of the Δ/Ψ plane, the constants of proportionality differ, i.e., for a given delta shift, the corresponding CF_x thickness is greater than that of the oxide. The XPS spectra in Fig. 11 also indicate a much more efficient removal of the residual F and O contamination. Here, the chemical shifts are marked and fitted according to the Appendix. Several of the O($1s$) spectral data points have

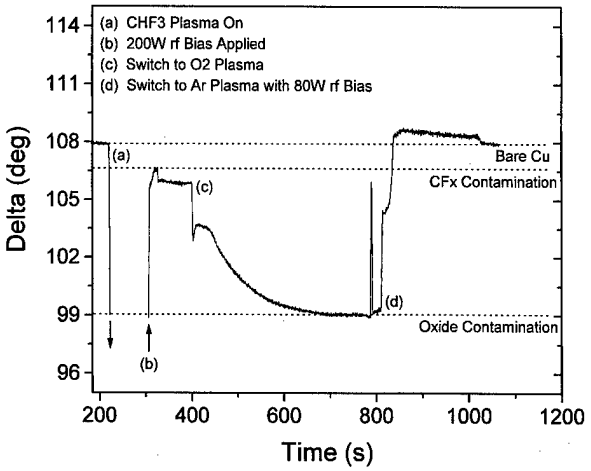


FIG. 10. Time evolution of the ellipsometric variable delta monitored in real time during the CF_x contamination of Cu, the O_2 cleaning step, and the final Ar sputter.

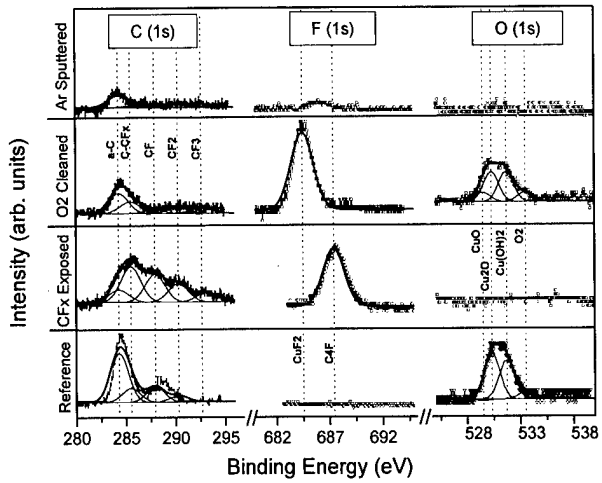


FIG. 11. XPS spectra representing the important stages of the Cu cleaning process. The spectra were fit using binding energies as listed in the Appendix with FWHM of 2.36 eV for the C(1*s*) and F(1*s*) spectra and 2.3 eV for the O(1*s*) spectra. The total area of all the F(1*s*) and O(1*s*) peaks were cut in half with respect to the O(1*s*) peaks. Note that the $Cu(OH)_2$ and $CuCO_3$ bonded species are not resolved and therefore may both exist on the Cu surface.

been averaged with their nearest neighbors. In the case of the Cu(2*p*) spectra, the spin-orbit separation is much greater than the chemical shifts, and due to the cutoff at 960 eV, the reacted $Cu(2p)_{1/2}$ peaks (binding energy >960) are not seen. In Fig. 12, the $Cu(2p)_{3/2}$ regions have been analyzed with fits of the chemical shifts. The Cu peak at 932.4 eV is in good agreement with values reported in the literature.[13,34,35] Moving to higher binding energies, we have fit a CuO peak at 933.45 ± 0.5 eV, which again agrees with values reported by Pan *et al.*[35] and others.[13] At 934.75 eV, we have included a $Cu(OH)_2$ peak, although we cannot distinguish this from $CuCO_3$ at 934.8 ± 0.5 eV.[13] We assume, based on literature,[25,32] that the reference sample is indeed passivated with $Cu(OH)_2$, while after fluorocarbon plasma exposure, the $CuCO_3$ may be present. The well known Cu(*I*) oxidation state of Cu_2O is shifted only slightly from the elemental Cu(2*p*) binding energy (~0.1 eV), and we have therefore left this unresolved. However, Auger analysis taken on the reference sample has clearly resolved this oxidation state.[36] The CuF_2 is fitted with two peaks, as reported in Ref. 13, and supported by trends in the F(1*s*) and O(1*s*) spectra. They are located at 935.83 and 936.75 eV and should be considered within ±0.5 eV. Finally, the shakeup satellite characteristic of Cu(II) oxidation state (~942 eV)[37] was fitted using three weakly constrained peaks. These peaks should therefore only be interpreted as a general indication of the CuO concentration. Although visible in all spectra, it is predominant after the O_2 plasma cleaning step.

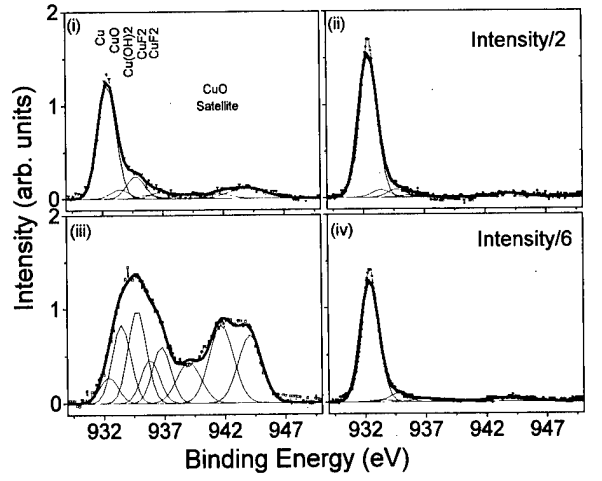


FIG. 12. XPS spectra and fitted chemical shifts for the $Cu2p_{3/2}$ spectra. The binding energies used for the fits can be found in the Appendix. A FWHM of 1.53 eV was used in fitting all the chemical shifts. The intensities in panels (ii) and (iv) were reduced by factors 2 and 6, respectively.

The lack of significant reactive components to the $Cu(2p)_{3/2}$ peak after CHF_3 plasma exposure, coupled with the reduced Cu(2*p*) emission intensity suggests that the carbon and fluorine residue is primarily an overlayer rather than chemical reaction with the Cu. This is also supported by the F(1*s*) spectra, where the 687.4 eV peak suggests F–C bonded species.[13] The lack of an O(1*s*) intensity signifies the removal of most of the Cu–O bonded species, and this is supported by the reduced Cu(II) oxidation peaks in Fig. 12.

After the O_2 plasma treatment, the shift of the F(1*s*) spectrum indicates a change from F–C bonded species to F–Cu bonds (684.5 eV). There is a strong Cu(II) oxidation signal, as well as what is consistent with the $CuCO_3$ patina at 934.8 eV. Again, we note that the $Cu(OH)_2$ peak at 934.75 eV could also be present.

There is no observable F(1*s*) or O(1*s*) intensity after the Ar^+ sputter step, as seen in Fig. 11, although there is still a small $CuCO_3$ peak [seen in Fig. 12, panel (iv)]. As with the Al, the ellipsometry data presented in Fig. 10 were fit in order to get an idea of the modified layer thicknesses. In this region of the Δ/Ψ plane, the thickness is proportional to the change in delta. Using a refractive index of the Cu substrate of $n+ik=0.17-i21.5$,[38] and of $n=1.5$ for the modified CF_x layer, a thickness of 1.2 ± 0.2 nm was obtained for the steady state CF_x layer. As seen in Fig. 10, after the CHF_3 discharge was extinguished (the plateau before point c), there was film growth attributed to postplasma fluorocarbon species depositing on the sample. The growth of this layer can be prevented by switching off the inductive power before the rf bias of the sample, and for the XPS analysis, this was done.

With a complex refractive index of $n+ik=1.5-i0.3$ for the oxide film,[38] an oxide thickness of 8 ± 2.0 nm was obtained after the O_2 plasma exposure. Although the reference referring to the oxide film did not identify the oxidation state of the copper, this value gave an acceptable fit to what our surface analysis has shown to be both CuO and $Cu(OH)_2$. Although we could not attain a perfect fit, by acceptable, we refer to a fit identical to the data in shape and delta variations with a <3° shift in psi. Forcing the simulated psi values to err both higher and lower than the data by 3° allowed us to assess the error in film thickness induced by the poor fit in psi.

Again, apparent thicknesses of the modified copper were calculated from the Cu($2p$) XPS spectra. For the as received reference sample, a thickness of 2.41 ± 0.55 nm was found. After the CHF_3 exposure, the thickness was determined to be 1.35 ± 0.38 nm. After the O_2 plasma step, we calculated the thickness as 9.05 ± 0.85 nm, and after the Ar^+ sputter, 1.28 ± 0.36 nm.

Referring again to Fig. 12, it is clear that the Ar^+ treated surface [panel (iv)] is the least contaminated, however, the CF_x treated surface [panel (ii)] is nearly free from modified Cu residues. In panel (i), the as received reference sample is shown. There is only a small satellite at ~942 eV suggesting only a low level of Cu(II) oxidation. However, the O($1s$) spectra seen in Fig. 11 suggest significant oxidation. We therefore conclude that the primary oxidation states are Cu(I), $Cu(OH)_2$, and/or $CuCO_3$. This is supported by the $Cu(OH)_2$ peak at 934.75 eV and by native oxidation models found in the literature.[25,32] The presence of carbon is also clear in the C($1s$) panel of Fig. 11. Panel (ii) of Fig. 12 represents the state of the copper surface after exposure to the CHF_3 plasma with ion bombardment. Note that we have reduced the intensity of the displayed spectrum by a factor of 2. Again, the characteristic Cu(II) oxidation satellite at 942 eV is extremely small. Here, we also observe that the oxygen contributions at 933.5 (CuO) and 934.8 [$Cu(OH)_2$ and/or $CuCO_3$] have become significantly diminished. A return to Fig. 11 will verify that the O($1s$) peak is absent. Again, note the significant CF_x overlayer in the absence of any significant Cu–F or Cu–C bonded species. After the copper surface, previously exposed to the CHF_3 plasma, has been exposed to an oxygen plasma, one immediately notices the significant characteristic Cu(II) oxidation satellite. Also, the ratio of the 932.4 eV Cu peak to the 933.5 CuO peak has been greatly reduced. This state of the surface is supported by the O($1s$) peaks seen in Fig. 11.[13,34,35] Also of interest is the return of the Cu–F bonded species after the O_2 plasma treatment. This is apparent in the CuF_2 peaks seen in panel (iii), as well as the F($1s$) spectrum in Fig. 11. Possibly, this is due to the removal of carbon by the oxygen plasma and the reaction of the residual fluorine with the copper. Illustrated in panel (iv) is the final stage of the cleaning process studied here. The Cu sample has undergone treatments as in panel (iii), and then is subjected to an Ar^+ sputter at 80 W. Note that the displayed intensity is only one sixth the measured intensity. This is clearly the cleanest sample, with a strong intensity and low levels of reacted Cu. There does, however, remain a small amount of carbon contamination.

4. Titanium nitride

Unlike the other films used in this part of our study, the TiN was thin enough (~50 nm) so that changes in its thickness could be observed with the ellipsometer. It is the quasi-infinite response, at 632.8 nm, of the other three substrates investigated that allowed for a more accurate determination of the surface modifications during each stage in the processing sequence. Therefore these data must be interpreted differently than the analogous figures for Si, Al, and Cu, i.e., in such a way as to account for the etching of the TiN as well as the surface modification. The real-time evolution of the TiN in response to the processing is depicted in Fig. 13. At time (a), a thick fluorocarbon layer is deposited in response to the ignition of a CHF_3 discharge. The removal of this layer then takes place upon applying the 200 W rf bias to the sample. Once this passively deposited CF_x film is removed, the TiN begins to etch through a thin steady state reaction layer. At this point the plasma is extinguished. At time (c), a 1000 W O_2 plasma is ignited. A portion of the CF_x damage layer is removed before the TiN begins to oxidize. The oxide layer is quite thin, as is the modified CF_x layer. Finally, at time (d), with the O_2 plasma extinguished, an Ar plasma at 440 W is turned on and a rf bias of 80 W is quickly applied to the sample. After an initial optical response of the ellipsometer to the high pressure Ar discharge, the TiN surface is recovered, though there is an increase in delta due to the etching of the film. Each stage in this evolution is depicted in Fig. 14 in

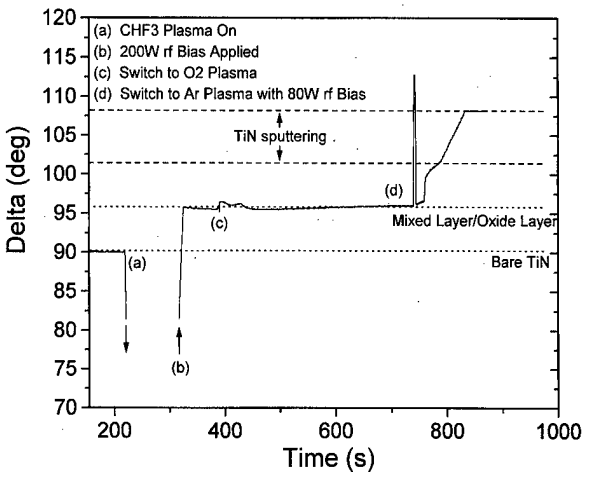


FIG. 13. Time evolution of the ellipsometric variable delta monitored in real time during the CF_x contamination of TiN, the O_2 cleaning step, and the final Ar sputter.

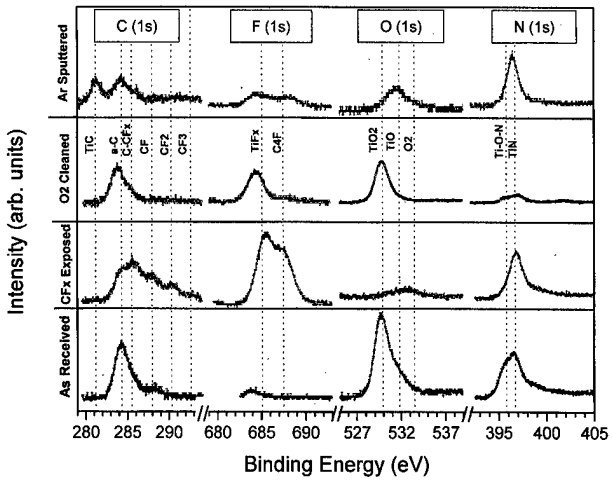


FIG. 14. XPS spectra representing the important stages of the TiN cleaning process. Specific binding energies for the chemical shifts can be found in the Appendix.

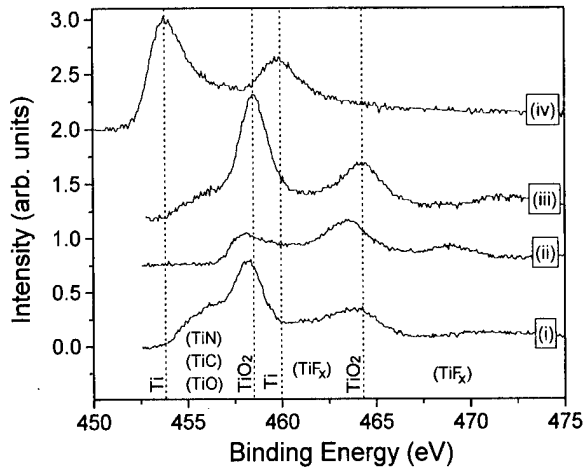


FIG. 15. Ti($2p$) XPS spectra and chemical shifts. Curve (i) depicts the state of the surface as received, (ii) after CF_x exposure, (iii) following the O_2 plasma clean, and (iv) after the Ar^+ sputter. Only the Ti and TiO_2 binding energies are lined in due to the overlap of other chemical shifts. Other bonds are suggested and listed in the Appendix.

terms of XPS spectra. The Appendix lists the relevant chemical shifts used in the fitting.

It is very difficult to resolve the chemical shifts in the $Ti(2p)_{3/2}$ spectrum due to the overlap of the TiO, TiN, and TiC binding energies.[13] Nevertheless, the Ti($2p$) spectra have been magnified and are presented in Fig. 15. Certain chemical shifts are identified in Fig. 15, and the Appendix presents some values from literature. A linear background subtraction was used which has been shown to compare well with an integrated background correction.[39]

In the case of the as received reference spectra, we see a strong TiO_2 component at 458.5 eV in addition to the TiN component at about 455.5 eV [the low binding energy shoulder of the $Ti(2p)_{3/2}$ peak]. This is supported by the observed O($1s$) and N($1s$) emission intensities. In both of the latter peak shapes, an asymmetry is observed. In the case of the O($1s$) spectra, the main peak at 529.9 eV is due to oxygen from Ti–O bonded species.[40] A slight shoulder on the high binding energy side of this main peak is observed, though we do not speculate as to its origin. We acknowledge the probable occurrence of oxynitride phases, but have not attempted to resolve their contributions. Also observed in the reference spectra are a C($1s$) peak and a small F($1s$) peak. Close inspection of the O($1s$) spectrum reveals at least two main features: (1) a graphitic contribution at 284.3 eV, and (2) a CF_x related peak at 287.95 eV. No C–Ti bonds are visible. In most of the N($1s$) spectra, high binding energy asymmetries are observed. Other investigators have attributed these to absorbed N_2 and nitrosyl species.[41–43]

After the TiN sample has been exposed to the ion bombardment of a CHF_3 discharge, several changes have taken place. There is clearly a CF_x overlayer, as is seen in the strong CF_x related species found in the C($1s$) spectrum and supported by the F–C bonded species found in the F($1s$) spectrum. The diminished intensity of the Ti($2p$) spectrum also suggests the presence of an overlayer. The F($1s$) spectrum is clearly made up of two features: (1) a 685.1 eV peak originating from Ti_xF_y species, and (2) the C_4F 687.4 eV peak. There appears to be no carbon that is chemically bonded to the titanium. The strength of the Ti–F bonded species in the F($1s$) spectrum coupled with the low O($1s$) intensity suggests the Ti($2p$) peaks near the characteristic TiO_2 binding energies are originating from TiF_x contributions. This is consistent with values reported in the Appendix.

Following the O_2 plasma treatment of the CF_x contaminated TiN sample, the surface is mostly TiO_2, although the low binding energy shoulder of the TiN is still visible. This is reflected nicely in the heightened O($1s$) and diminished N($1s$) emission intensities. The fluorine is mainly due to Ti_xF_y species, although a small C–F related peak is observed. Again, we see no contribution of Ti–C bonds to the C($1s$) spectra, but mainly a graphitic component.

In the final step, after the sample has been sputter cleaned with Ar^+, there is a return to the TiN surface. There is a clear Ti peak, as well as the TiN peak at its high energy shoulder. In this case, we assume that a portion of this intraspin-orbit region is also due to Ti–C bonds. This is supported by the O($1s$) spectrum where there is clearly a TiC contribution. An oxide contribution can be seen on the low energy shoulder of the $2p_{1/2}$ state.

An accurate determination of the modified surface layer thickness was not attempted in the case of TiN processing, due to the overlap of the etching dynamics observed with the ellipsometer. However, we estimate this layer to be quite thin, i.e., <2 nm. We have based this assumption on a simulation involving the etching of TiN either through a CF_x layer or through a modified TiO_2 surface. The refractive index of the TiN used in the fitting was $n+ik=1.5-i2.00$.[44,45] The fluorocarbon overlayer is again $n=1.5$ and the TiO_2 was fitted using $n=2.2$.[45] When various thicknesses for the steady state modified CF_x layer or the TiO_2 were introduced into the fitting routine as the top layer in a three layer model of Al/TiN/(modified TiN), the correlation between the delta values at a given psi between the ideal TiN and overlayer and the dataset of the TiN etching and overlayer suggested such a thickness. Due to the complex chemical shifts and satellite presence in the Ti($2p$) spectra, apparent thicknesses of the modified surface layers were not calculated from the XPS data.

The contamination due to atmospheric exposure of TiN is well documented.[7,43,46] The spectra obtain in our work are in good agreement for such a reference sample. After the sample has been exposed to the fluorocarbon plasma, a CF_x overlayer is observed. Although quite a bit of fluorine has bonded to the titanium, the carbon has not. Rather, the carbon seems to be graphitic or part of the physisorbed CF_x overlayer. The O_2 plasma treatment is quite efficient at removing the C–F bonded species, however, the surface TiN has been converted to TiO_2. A significant amount of the fluorine incorporated in Ti_xF_y bonds has been removed, though a substantial amount remains. After the Ar^+ sputter step, and only in this stage, the C($1s$) spectrum reveals the

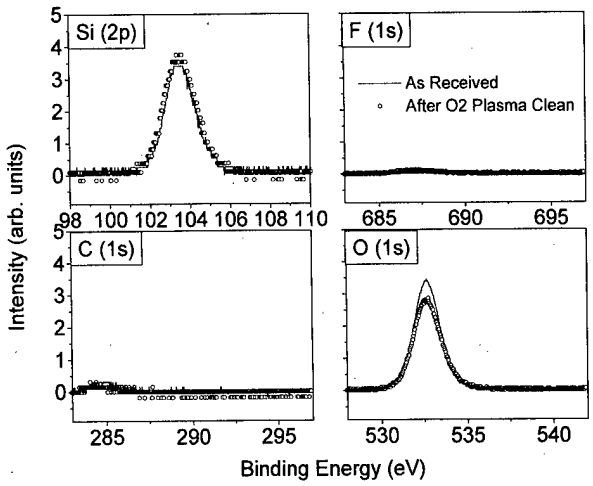


FIG. 16. XPS spectra illustrating the stability of SiO_2 during O_2 plasma exposure.

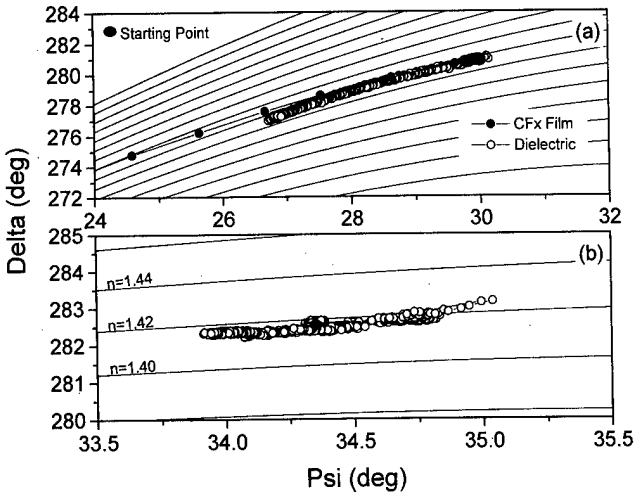


FIG. 17. Real-time evolution of the SiOF surface during O_2 plasma exposure as monitored by the ellipsometric variables psi and delta. Contours of constant refractive index are separated by $\Delta n = 0.02$. Panel (a) represents *in situ* cleaning following CF_x exposure, while panel (b) is for a reactor free of CF_x contaminants.

emergence of TiC species. Also the fluorine incorporation in the film has been significantly reduced. The TiO_2 has been sputtered away and the surface has returned to primarily TiN.

B. Effect of cleaning on low-*K* dielectrics

In a successful cleaning procedure, the dielectric will be exposed to the O_2 plasma. Hence, the stability of the dielectric in this environment is a critical issue. To retain the etched profile, the dielectric must not etch significantly, nor should the dielectric constant increase. We investigated the stability of three oxide-like dielectrics in response to O_2 plasma exposure and compared the results with a conventional oxide.

1. SiO_2

The stability of thermally grown SiO_2 in an O_2 plasma environment is well known and we have confirmed this ellipsometrically. No change in refractive index was observed, even in the presence of residual CF_x contaminants as found in an *in situ* cleaning procedure. In Fig. 16, XPS spectra, taken at normal electron emission, are presented which also support the stability of the dielectric. The different curves represent the unprocessed reference oxide and the state of the SiO_2 after O_2 plasma exposure. The small $F(1s)$ signal visible after processing is attributed to the residual fluorine in the reactor.

Stoichiometries of the SiO_2 were referenced to a Si:O ratio of 1.0:2.0 for the observed normal emission intensities of the $Si(2p)$ and $O(1s)$ peaks. After processing in the O_2 plasma, the stoichiometric ratio of silicon to oxygen remained SiO_2.

The reference spectrum had a $Si(2p)$ FWHM of 1.8 eV, while the processed oxide was 1.9 eV. This is also consistent with the uptake of residual fluorine from the reactor.

2. SiOF

The ellipsometric response of the SiOF during O_2 plasma exposure is shown in Fig. 17. The top panel displays data obtained during an *in situ* plasma clean, i.e., immediately following the fluorocarbon plasma exposure. Initially, the sample is exposed to a 1400 W CHF_3 plasma. This results in the deposition of a thick CF_x layer. After considerable growth (>200 nm), the discharge is extinguished and the system evacuated. Then a 1000 W O_2 plasma is ignited, and the CF_x layer etches back, ultimately reaching the starting point. Once the SiOF surface has been recovered, slight deposition is observed, but this deposition is consistent with the formation of an oxide. We attribute this to the erosion of the quartz coupling window that separates the discharge region from the inductive coil. In any case, the refractive index, and hence the dielectric constant, remains constant during this exposure. In a dedicated chamber environment (reactor free from CF_x contamination), as seen in the bottom panel of Fig. 17, the refractive index again remains constant throughout the O_2 plasma exposure. Noting the magnified ordinate, we also see the sputter deposition from the coupling window is eliminated. XPS measure O_2 plasma exposure. These results are presented in Fig. 18. Consistent with the ellipsometrically observed stability of these films, no modification is seen. Characteristic in both samples is the SiO_2 peak from the $Si(2p)$ core level at 103.4 eV.

The stability of the SiOF under exposure to a O_2 plasma environment can be attributed to the fact that these films are compositionally close to SiO_2. Fits of the XPS spectra as seen in Fig. 18 yield a stoichiometry of $SiO_{1.9}$ for the reference sample (FWHM=2.1 eV) and the SiO_2 for the O_2 plasma exposed (FWHM=2.1 eV).

3. HSQ

When HSQ is exposed to the O_2 plasma environment, an increase in refractive index is observed. This increase should correspond to an increase in the dielectric constant. The increase in refractive index is consistent with the loss of hy-

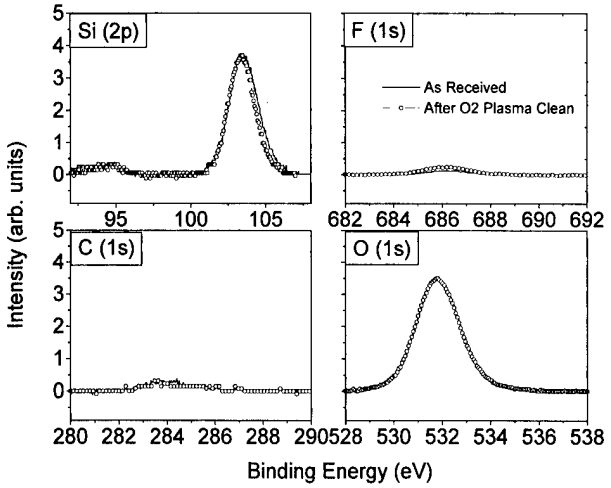


FIG. 18. XPS spectra illustrating the stability of SiOF during O_2 plasma exposure.

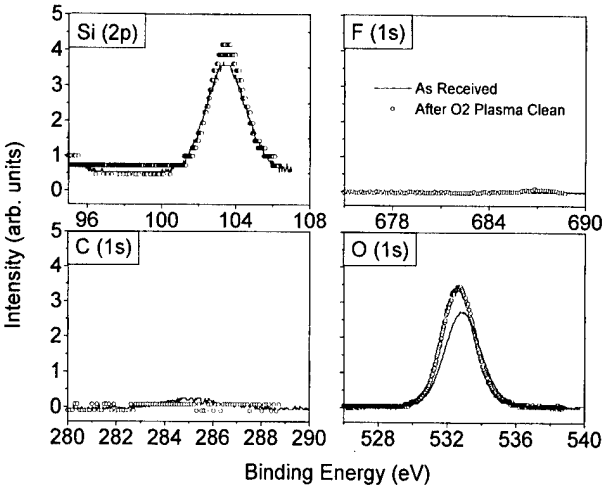


FIG. 20. XPS spectra illustrating the instability of HSQ during O_2 plasma exposure.

drogen from the film and is observed both for an *in situ* clean and in a dedicated chamber configuration. These results are illustrated in Figs. 19(a) and 19(b), respectively. After the *in situ* clean, a significant etch of the dielectric takes place through the fluorocarbon film, before the bare HSQ is recovered. The duration of the O_2 plasma exposure was 300 s, and in this time, the refractive index increased by $\Delta n = \sim 0.1$. Figure 20 shows the surface analysis of the HSQ before and after a dedicated chamber O_2 plasma clean. The increase of the 103.4 eV SiO_2 peak in the Si($2p$) spectrum and O($1s$) peak area suggest the increase of Si–O bonding at the expense of the Si–H bonds, and is therefore consistent with the hydrogen loss model. Further support for this notion is the modification in stoichiometry of the films as calculated from the XPS spectra presented in Fig. 20. The reference sample was observed to have a stoichiometry of $SiO_{1.7}$, while after processing it was measured to be SiO_2. Also the fit of the Si($2p$) spectra after processing had a FWHM 0.2 eV less than the reference sample (FWHM=2.6 eV). This suggests that the unresolved Si–H bonds, which were contributing to the FWHM, have been broken. These trends all suggests that during processing, hydrogen is lost from the dielectric.

In order to determine the depth of this modification, we sputtered a modified HSQ sample with Ar^+, and noted the recovery of the refractive index versus depth. After the O_2 plasma modification, the refractive index increased by $\Delta n = +0.11$. After 75 nm the HSQ was removed by the ion bombardment, the index was observed to increase by an additional 0.1. After approximately 300 nm of the dielectric was sputtered away, a change of $\Delta n = +0.05$ from the initial refractive index was observed. From this we conclude that the modification is extremely deep into the HSQ.

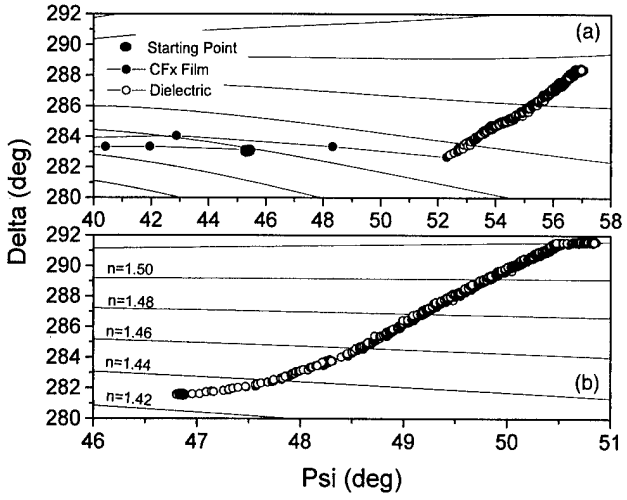


FIG. 19. Real-time evolution of the HSQ surface during O_2 plasma exposure as monitored by the ellipsometric variables psi and delta. Contours of constant refractive index are separated by $\Delta n = 0.02$. Panel (a) represents *in situ* cleaning following CF_x exposure, while panel (b) is for a reactor free of CF_x contaminants.

4. MSQ

We also observed an instability of MSQ in the O_2 plasma cleaning environment. Ellipsometric evidence of this is presented in Fig. 21, where again, panel (a) represents the *in situ* clean and panel (b) the dedicated chamber configuration. Etching through the fluorocarbon film is again observed in panel (a). During an *in situ* resist strip and via clean, this etching results in a tapering of the sidewall profile at the dielectric–photoresist interface. This degradation, as well as the residual sidewall veils, is illustrated in Fig. 22. Panels (a) and (b) show structures etched in SiOF before and after the O_2 cleaning step, respectively. The stability of the SiOF as presented in Sec. III B 2 is manifested in the unmodified sidewalls. Panels (c) and (d) show the results of the same process for MSQ. Although the veils along the sidewalls are clearly visible in the MSQ etched structure after the O_2 plasma clean, they also exist along the patterned SiOF. This is shown in Fig. 23. Here, the dielectric was removed after the cleaning treatment using a diluted HF solution. All that remains on the sample surface are the insoluble veils, which for SiOF are depicted in panel (a) and for MSQ in panel (b).

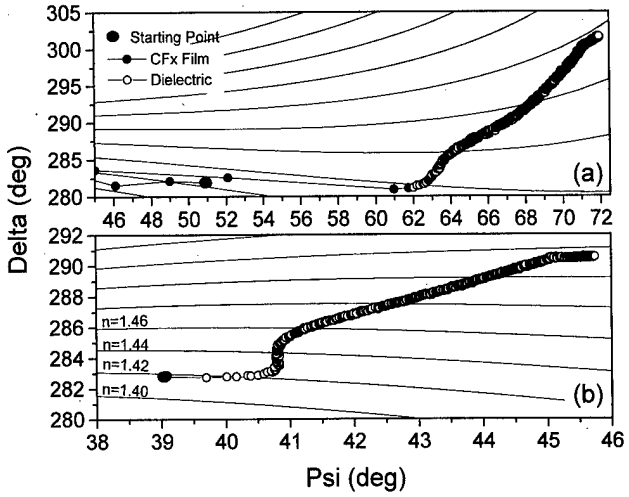


FIG. 21. Real-time evolution of the MSQ surface during O_2 plasma exposure as monitored by the ellipsometric variables psi and delta. Contours of constant refractive index are separated by $\Delta n = 0.02$. Panel (a) represents *in situ* cleaning following CF_x exposure, while panel (b) is for a reactor free of CF_x contaminants.

Referring back to Fig. 21 and the response of the blanket samples, we see that during the 300 s O_2 plasma exposure, the refractive index has increased by $\Delta n = \sim 0.1$. XPS spectra illustrating this modification are presented in Fig. 24. A strong reduction in the C($1s$) spectra and increase in the observed SiO_2 is consistent with the increase in refractive index observed in Fig. 21. The small F($1s$) signal visible after processing is again attributable to the residual fluorine in the reactor. Fits of the MSQ Si($2p$) and O($1s$) spectra both before and after processing yielded stoichiometries consistent with the evolution towards a more SiO_2-like film. For the unprocessed reference sample, we observed a

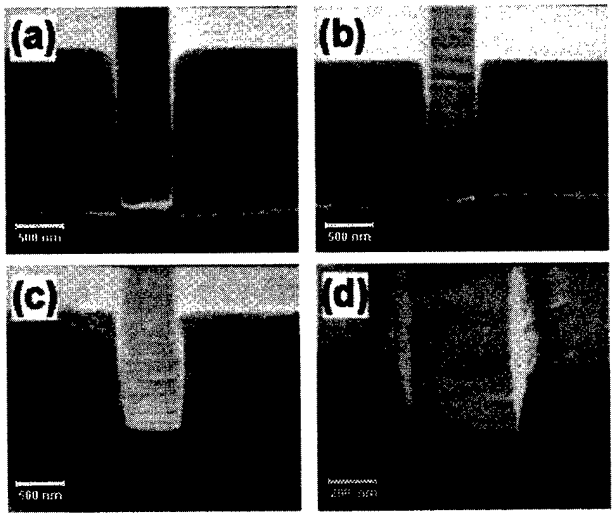


FIG. 22. SEM of 0.5 μm trenches in SiOF and MSQ after the dielectric etch both before and after the O_2 plasma cleaning step. Panels (a) and (b) represent the SiOF before and after the O_2 plasma exposure, respectively. Panels (c) and (d) represent the MSQ before and after the O_2 plasma exposure, respectively.

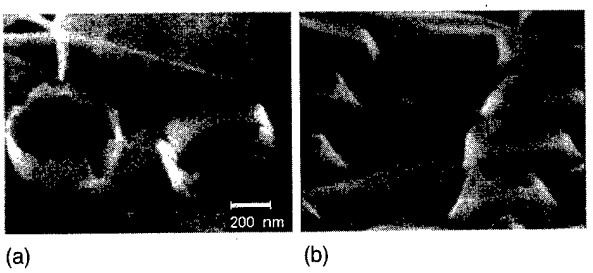


FIG. 23. SEM of (a) SiOF and (b) MSQ veils remaining after the O_2 plasma clean. The veils were exposed after the treated samples were HF dipped to remove the dielectric material.

Si($2p$):O($1s$) peak area ratio of 1.0:1.7 with a FWHM of 2.1 eV. After O_2 plasma exposure, the ratio became 1.0:2.0, while the FWHM slightly lessened to 2.0 eV. An Ar^+ sputter was used to determine the depth of the MSQ modification. In this case, the O_2 plasma exposure also resulted in an increase of the refractive index of $\Delta n = +0.11$. After about 75 nm of sputtering, we see a slight return towards the original refractive index ($\Delta n = +0.10$). After almost 300 nm removal, the refractive index is within $\Delta n = +0.04$ of the initial value. From this result and the analogous result for HSQ, we conclude that the depth of the HSQ modification is noticeably greater than that of MSQ. This is consistent with expected diffusion coefficients for hydrogen and carbon through an oxide-like film.

IV. CONCLUSIONS

The contamination after a simulated fluorocarbon overetch of various low-K oxide-like dielectrics has been characterized and the efficiency of a two step cleaning process has been evaluated for various substrates.

Using silicon as a comparison, we have found the efficiency of the cleaning process to be quite good, especially in the cases of TiN and Cu. In the case of copper, there is a lack

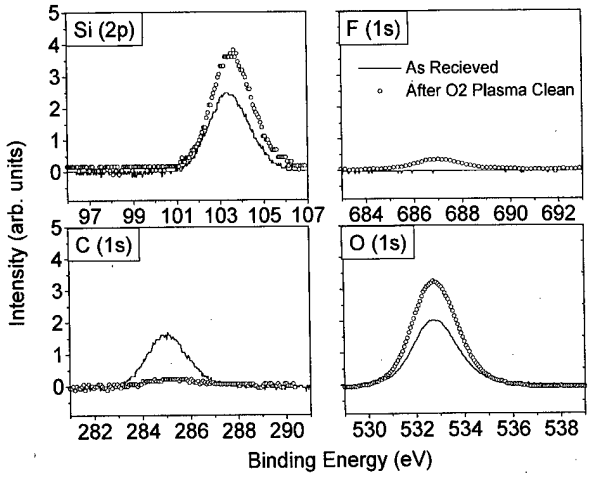


FIG. 24. XPS spectra illustrating the instability of MSQ during O_2 plasma exposure.

of reacted metal after fluorocarbon plasma exposure. Only after the subsequent O_2 plasma treatment are Cu–F bonds observed.

The stability of the dielectrics to the O_2 plasma environment is dependent on their chemical makeup. SiOF, as SiO_2, exhibited good stability while HSQ and MSQ were significantly modified. We have presented evidence that this modification is due to the removal of the hydrogen or the carbon atoms. We have also provided evidence that the depths of these modifications are extremely deep into the dielectric. We are currently investigating various cleaning chemistries that can maintain the integrity of the HSQ and MSQ, as well as other novel dielectrics.

ACKNOWLEDGMENTS

This work was sponsored in part by the Semiconductor Research Corporation. The authors acknowledge Digital Equipment Co. for supplying them with the wafers used in this work. M. Schaepkens, B. Kastenmeier, and P. Wrschka are thanked for their ideas. T. Yagi is thanked for his help with the data taking.

APPENDIX: RELEVANT BINDING ENERGIES FOR XPS SPECTRA AND CHEMICAL SHIFTS USED IN THIS WORK

Material	Core level	Compound	Binding energy (eV)	Reference
Generic	C(1*s*)	C[a]	284.3	13, 20
		C–CF_x	285.5	12
		CF	287.95	12, 47
		CF_2	290.3	12, 47
		CF_3	292.65	12, 47
	F(1*s*)	C_4F	687.4	13
		CF_x	689.4	48
	O(1*s*)	O_2	533.5	30
	N(1*s*)	N_2	399.8	43
Silicon	C(1*s*)	Si–C	282.7	17
	F(1*s*)	SiF_x	686.3	49
	O(1*s*)	SiO_2	532.9	13, 12
	Si(2*p*)	Si	99.15	13
		SiF	100.3	14, 15, 16
		SiC	100.65	17
		SiF_2	101.45	14, 15, 16
		SiF_3	102.6	14, 15, 16
		SiF_4	103.75	14, 15, 16
		SiO_2	103.4	13
Aluminum	F(1*s*)	AlF_3	685.19	20
	O(1*s*)	$Al(OH)_3$	531.3	25
		Al_2O_3	531.6	13, 19
		AlOOH	532.6	25
	Al(2*p*)	Al	72.65	13, 20, 19, 21, 22
		Al–O	73.95	19, 23
		Al_2O_3	75.25	13, 19, 21, 23
		$Al(OH)_3$	75.25	25
		AlOOH	75.45	25
		AlF_3	76.45	13
		AlF_x	77.7	20
Copper	F(1*s*)	CuF_2	684.5	13
	O(1*s*)	CuO	529.5	32
		Cu_2O	530.3	32
	Cu(2*p*)$_{3/2}$	Cu	932.4	13, 33, 35
		CuO	933.45	13, 35
		$CuCO_3$	934.8	13
		CuF_2	935.8	13
		CuF_2	936.75	13
		CuO[b]	942.5	13, 34
	Cu(2*p*)$_{1/2}$	Cu	952.2	13
Titanium nitride	C(1*s*)	TiC	281.2	13
	F(1*s*)	TiF_x	685.1	50
	O(1*s*)	TiO_2	529.9	51
		TiO	531.8	51
	N(1*s*)	Ti-O-N	395.8	51
		TiN	396.7	40, 41, 43
	Ti(2*p*)$_{3/2}$	Ti	453.8	13
			453.9	52, 53
		TiN	454.8	39, 41
			455.2	52, 46
			454.5	43, 54
		TiO	454.7	55
			455	56
		TiO_2	458.5	13
			458.3	52
			459	39, 57
		TiC	454.7	52
			454.8	58
		TiF_x	461	59
		TiF_3	461	50
			462	41
		TiF_4	463.5	50

[a]Graphitic carbon.
[b]CuO satellite.

[1]H.-H. Park, K.-H. Kwon, J.-L. Lee, K.-S. Suh, O.-J. Kwon, K.-I. Cho, and S.-C. Park, J. Appl. Phys. **76**, 4596 (1994).
[2]M. Delfino, S. Salimian, and D. Hodul, J. Appl. Phys. **70**, 1712 (1991).
[3]Y. Z. Hu, P. P. Buaud, Y. Wang, L. Spanos, and E. A. Irene, Appl. Phys. Lett. **64**, 1233 (1994).
[4]A. Sherman, J. Electrochem. Soc. **137**, 1982 (1990).
[5]D.-H. Kim, S.-L. Cho, K.-B. Kim, J. J. Kim, J. W. Park, and J. J. Kim, Appl. Phys. Lett. **69**, 4182 (1996).
[6]W. Sinke, G. P. A. Frijlink, and F. W. Saris, Appl. Phys. Lett. **47**, 471 (1985).
[7]M. Mandl, H. Hoffman, and P. Kucher, J. Appl. Phys. **68**, 2127 (1990).

[8]T. E. F. M. Standaert, P. J. Matsuo, S. D. Allen, and G. S. Oehrlein, J. Vac. Sci. Technol. A **17**, 741 (1999).
[9]K. Sakuma, K. Machida, K. Kamoshida, Y. Sato, K. Imai, and E. Arai, J. Vac. Sci. Technol. B **13**, 902 (1995).
[10]C. Ernsberger, J. Nickerson, and T. Smith, J. Vac. Sci. Technol. A **4**, 2784 (1986).
[11]N. R. Rueger, J. J. Beulens, M. Schaepkens, M. F. Doemling, J. M. Mirza, T. E. F. M. Standaert, and G. S. Oehrlein, J. Vac. Sci. Technol. A **15**, 1881 (1997).
[12]T. E. F. M. Standaert, M. Schaepkens, N. R. Rueger, P. G. M. Sebel, G. S. Oehrlein, and J. M. Cook, J. Vac. Sci. Technol. A **16**, 239 (1997).
[13]C. D. Wagner, W. M. Riggs, L. E. Davis, and J. F. Moulder, *Handbook of X-ray Photoelectron Spectroscopy*, edited by G. E. Muilenberg (Perkin-Elmer, Eden Prairie, MN, 1979).
[14]B. Roop, S. Joyce, C. Schultz, and J. I. Steinfeld, Surf. Sci. **173**, 455 (1986).
[15]M. Delfino, B-C. Chung, W. Tsai, and S. Salimian, J. Appl. Phys. **72**, 3718 (1992).
[16]G. S. Oehrlein, J. Vac. Sci. Technol. A **11**, 34 (1993).
[17]G. J. Coyle, Jr. and G. S. Oehrlein, Appl. Surf. Sci. **25**, 423 (1986).
[18]J. M. Cook, O. Turmel, T. Wicker, and J. Winniczek, Technical Proceedings, Semicon Japan, SEMI, Chiba, 1993, p. 414.
[19]I. P. Batra, J. Electron Spectrosc. Relat. Phenom. **33**, 175 (1984).
[20]A. C. Miller and F. P. McCluskey, J. Vac. Sci. Technol. A **9**, 1461 (1991).
[21]A. F. Carley and M. W. Roberts, Proc. R. Soc. London, Ser. A **363**, 403 (1978).
[22]R. A. Rosenberg, M. W. McDowell, and J. R. Noonan, J. Vac. Sci. Technol. A **12**, 1755 (1994).
[23]A. Arranz and C. Palacio, Surf. Sci. **355**, 203 (1996).
[24]S. A. Flodstrom, R. Z. Bachrach, R. S. Bauer, and S. B. M. Hagstrom, Phys. Rev. Lett. **37**, 1282 (1976).
[25]T. L. Barr, J. Phys. Chem. **82**, 1801 (1978).
[26]A. C. Nyce and L. P. Skolnick, J. Opt. Soc. Am. **65**, 792 (1975).
[27]C. J. Dell'Oca, Thin Solid Films **26**, 371 (1975).
[28]R. Greef and C. F. W. Norman, J. Electrochem. Soc. **132**, 2362 (1985).
[29]V. N. Egorov, I. L. Chistyl, and V. F. Gusynin, Sov. J. Opt. Technol. **53**, 440 (1986).
[30]D. Briggs and M. P. Seah, *Practical Surface Analysis* (Wiley, New York, 1990).
[31]S. A. Flodstrom, R. Z. Bachrach, R. S. Bauer, J. C. McMenamin, and S. B. M. Hagstrom, J. Vac. Sci. Technol. A **14**, 303 (1977).
[32]T. L. Barr, *Modern ESCA* (Chemical Rubber, Boca Raton, 1994).
[33]C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1986).
[34]J. Bloch, D. J. Bottomley, S. Janz, H. M. Driel, and R. S. Timsit, J. Chem. Phys. **98**, 9167 (1993).
[35]F. M. Pan, S. R. Horng, T. D. Yang, and V. Tang, J. Vac. Sci. Technol. A **8**, 4074 (1990).
[36]J. Hernandez and P. Wrschka (private communication).
[37]B. Wallbank, I. G. Main, and C. E. Johnson, J. Electron Spectrosc. Relat. Phenom. **5**, 259 (1974).
[38]T. Smith, J. Opt. Soc. Am. **67**, 48 (1977).
[39]M. J. Vasile, A. B. Emerson, and F. A. Baiocchi, J. Vac. Sci. Technol. A **8**, 99 (1990).
[40]N. Kaufherr and D. Lichtman, J. Vac. Sci. Technol. A **3**, 1969 (1985).
[41]F. Fracassi, R. d'Agostino, R. Lamendola, and I. Mangieri, J. Vac. Sci. Technol. A **13**, 335 (1995).
[42]N. D. Shinn and K. L. Tsang, J. Vac. Sci. Technol. A **9**, 1558 (1991).
[43]N. C. Saha and H. G. Tompkins, J. Appl. Phys. **72**, 3072 (1992).
[44]H. G. Tompkins, J. Appl. Phys. **70**, 3876 (1991).
[45]H. G. Tompkins and J. A. Sellers, J. Vac. Sci. Technol. A **12**, 2446 (1994).
[46] N. Van Hieu and D. Lichtman, Appl. Surf. Sci. **20**, 186 (1984).
[47]T. A. Schreyer, A. J. Bariya, J. P. McVittie, and K. C. Saraswat, J. Vac. Sci. Technol. A **6**, 1402 (1988).
[48]G. S. Oehrlein, J. G. Clabes, and P. Spirito, J. Electrochem. Soc. **133**, 1002 (1986).
[49]J. P. Simko, G. S. Oehrlein, and T. M. Mayer, J. Electrochem. Soc. **138**, 277 (1991).
[50]R. d'Agostino, F. Fracassi, and C. Pacifico, J. Appl. Phys. **72**, 4351 (1992).
[51]C. Ernsberger, J. Nickerson, A. E. Miller, and J. Moulder, J. Vac. Sci. Technol. A **3**, 2415 (1985).
[52]L. Ramquist, K. Hamrin, G. Johansson, A. Fahlmann, and C. Nordling, J. Phys. Chem. Solids **31**, 2669 (1969).
[53]A. Platau, L. I. Johansson, A. L. Hagstrom, S. E. Karlsson, and S. B. M. Hagstrom, Surf. Sci. **63**, 153 (1977).
[54]S. Hofmann, J. Vac. Sci. Technol. A **4**, 2789 (1986).
[55]H. F. Franzen, M. X. Umana, J. R. McCreary, and R. J. Thorn, J. Solid State Chem. **18**, 363 (1976).
[56]P. T. Dawson and K. K. Tzatzov, Surf. Sci. **149**, 105 (1985).
[57]M. Murata, K. Wakino, and S. Ikeda, J. Electron. Spectros. **6**, 459 (1975).
[58]H. Ihara, Y. Kumashiro, A. Itoh, and K. Madea, Jpn. J. Appl. Phys. **12**, 1462 (1973).
[59]S. W. Robey, M. A. Jaso, and G. S. Oehrlein, J. Appl. Phys. **65**, 2951 (1988).

Reaction and thermal stability of cobalt disilicide on polysilicon resulting from a Si/Ti/Co multilayer system

A. Alberti
INFM and Physics Department, Catania University, Corso Italia 57, Catania, Italy

F. La Via[a)]
CNR–IMETEM, Stradale Primosole 50, Catania, Italy

F. Rimini
INFM and Physics Department, Catania University, Corso Italia 57, Catania, Italy

(Received 7 August 1998; accepted 16 April 1999)

The formation and thermal stability of $CoSi_2$ has been investigated using a poly-Si/Ti/Co multilayer structure. The presence of the Ti layer allows the phase formation by a single step with an optimum reaction temperature in the range between 750 and 850 °C. The layer agglomerates at temperatures higher than 950 °C. Sheet resistance measurements have been used to monitor the kinetics of the agglomeration process. The dependence of both the sheet resistance and the thermal stability on the film thickness has been studied in detail. The agglomeration process is thermally activated with an energy of ~3.9 eV, lower than the value 5.6 eV measured in $CoSi_2$ layers formed by direct reaction of Co with polycrystalline silicon. These activation energies have been correlated to the different layer structures. In the silicide formed without the Ti thin film, the grains are columnar, while the presence of Ti causes a random distribution in the grain shape and smaller size. The low thermal stability of the $CoSi_2$ layer on polysilicon can be a problem to use the Ti/Co bilayer process in self-aligned complementary metal–oxide–semiconductor technology. © *1999 American Vacuum Society.* [S0734-211X(99)01404-3]

I. INTRODUCTION

Cobalt disilicide has received a great deal of attention for microelectronics applications. As devices are scaled to deep submicron dimensions, thermally stable, highly conductive interconnects and silicided contacts to ultrashallow junctions are required. The stable $CoSi_2$ phase has been extensively investigated as a material for both interconnects on *polycrystalline* silicon and contacts on *crystalline* Si in integrated circuit (IC) (Refs. 1 and 2) fabrication. The main reasons for the interest in this silicide is its low resistivity (15–20 $\mu\Omega$ cm at room temperature) and its good thermal stability.

On silicon (001) substrates, $CoSi_2$ is a favorite candidate for epitaxial growth because it has the same face-centered-cubic CaF_2 structure and a small lattice mismatch with silicon (−1.2% at room temperature). Epitaxial $CoSi_2$ has been grown on Si using a wide variety of methods, such as the reaction of the pure cobalt atoms deposited by sputtering or evaporation onto heated silicon,[3] codeposition of Co and Si,[4] high-dose implantation of Co in Si,[5] and thermal reaction starting from a $Co_{75}W_{25}$ amorphous layer.[6] In another approach, a thin titanium layer has been interposed between the (001) silicon substrate and the Co thin film.[7] The use of a Ti/Co bilayer leads to the growth of an epitaxial silicide with a very smooth interface.[8,9]

The thermodynamic and the kinetic factors involved in the two reactions Co:Si and Ti:Si have been studied.[10] Increasing the Ti thickness, increases the distance x crossed by the Co to reach the Si interface, then the Co diffusion rate v slows down ($v=D/2x$ and $d\Delta G/dt=Fv$, where D is the diffusion coefficient for the cobalt atoms through the Ti layer, ΔG, the free-energy change, and F is the driving force of the reaction). At a Ti thickness of 20 nm or higher $(d\Delta G/dt)_{\mathrm{Co-Si}} \leqslant (d\Delta G/dt)_{\mathrm{Ti-Si}}$: in this case, the Ti–Si reaction occurs first, while at lower thickness the cobalt diffusion is retarded and the $CoSi_2$ epitaxial growth is promoted. A similar investigation using a polycrystalline silicon substrate has not been reported in the literature.

The thermal stability of the silicide film is very important in device fabrication because it has to sustain a set of annealing steps without degradation to maintain the device characteristics. There are many theoretical approaches to the thermal stability of thin films.[11–13] These models are based on the assumption that agglomeration in thin films is predominantly caused by thermal grooving at grain boundaries via matter diffusion away from the high-energy interfaces. The main parameters that control the agglomeration are the surface (γ_s), the interface (γ_i), and grain boundaries (γ_b) energies. Also, the film thickness and the grain dimension should have a great influence on the deterioration process. From a microscopic model, the agglomeration requires several steps, i.e., breaking of the $CoSi_2$ molecules close to the grain boundaries, migration of cobalt and silicon atoms, reaction of the cobalt with the silicon substrate, and epitaxial growth of silicon at the grain-boundary intersections. Several investigations on the thermal stability[14] of epitaxial $CoSi_2$ layers resulting from the reaction of a Si/Ti/Co multilayer were performed, while the same kind of study has not been reported for the polysilicon substrate.

Then, for a better characterization of the silicide process using the Ti/Co bilayer, a study of the formation process and the $CoSi_2$ thermal stability on a *polysilicon* substrate was

[a)]Electronic mail: lavia@imetem.ct.cnr.it

performed. Furthermore, we focused on the dependence of the thermal stability on the film thickness and, for a fixed $CoSi_2$ thickness, the influence of the intermediate titanium layer.

II. EXPERIMENTAL PROCEDURES

An amorphous silicon layer, ~200 nm thick, was deposited by low-pressure chemical-vapor deposition (CVD) at 500 °C on a thermally oxidized (001) Si substrate. At this point two different sets of samples were prepared.

In the first set, Ti and Co thin films of different thickness were sequentially deposited without breaking the vacuum in an ultra-high-vacuum (UHV) chamber by an electron gun system. The deposition rate for both materials was 0.1 nm/s and the base pressure in the chamber was 2×10^{-9} Torr. The following samples were prepared: (a) 4 nm titanium film and 5 nm thin cobalt layer, (b) 10 nm titanium film and 15 nm cobalt layer, and (c) 10 nm titanium film and 30 nm Co layer. These samples were annealed in the rapid thermal annealer (RTA) Heatpulse 610 under a nitrogen ambient in the temperature range between 650 and 950 °C for 60 s. After the reaction the samples were etched in a $NH_4OH:H_2O_2$ (1:1) solution at 60 °C for 10 min to remove as much as possible the titanium-rich surface layer and the unreacted cobalt. The final $CoSi_2$ layers had the following thicknesses: (a) 15 nm, (b) 48 nm, and (c) 100 nm, respectively. At the end of the thermal processes the unreacted silicon layer has become polycrystalline, as evidenced by transmission electron microscopy (TEM) and transmission electron diffraction (TED) analysis.

Another set of cobalt disilicide samples, to be used as a reference, was prepared without the intermediate titanium layer (d). In this set, 30 nm of cobalt were deposited on the amorphous silicon layer and the same type of substrate as previously used. The silicidation was performed with a standard two-step RTA process. The first anneal is at 540 °C for 30 s in forming gas flux ($N_2:H_2$ 90:10) to reduce the cobalt oxidation. After 2 min etch in a 4% HNO_3 solution to remove the unreacted cobalt, the sample was annealed at 800 °C for 40 s in a N_2 ambient. The reacted cobalt disilicide is about 100 nm thick. In spite of the different adopted reaction procedures, samples (c) and (d), i.e., with and without the titanium layer, have the same thickness and resistivity (~25 $\mu\Omega$ cm).

All the samples were analyzed with 2 MeV He^+ Rutherford backscattering spectroscopy (RBS) both in normal and glancing geometry to follow the silicide reaction and to measure the final $CoSi_2$ thickness. The interface morphology has been investigated by cross-sectional TEM and the $CoSi_2$ surface microstructure has been analyzed by a Nanoscope 3 atomic force microscope (AFM). Sheet resistance measurements were performed at room temperature with a four-point probe. With this technique, both the silicide reaction and the effect of the morphological evolution on the average resistivity of the layer upon annealing have been experimentally determined.

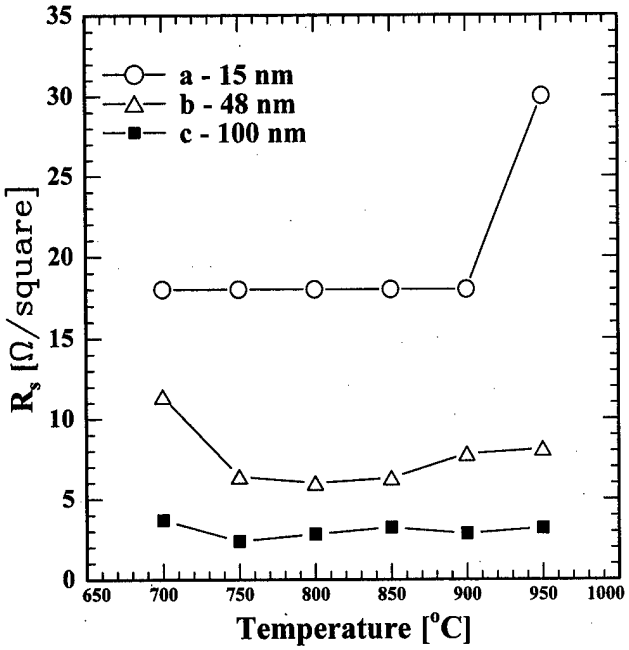


FIG. 1. Sheet resistance vs reaction temperature for $CoSi_2$ samples of different layer thickness (a) 15 nm, (b) 48 nm, and (c) 100 nm. The silicide reaction is complete at 750 °C 60 s.

III. RESULTS AND DISCUSSION

A. Reaction temperature

As a first step we have investigated the formation of cobalt silicide by the Ti/Co bilayer on amorphous silicon looking for the best reaction temperature in terms of sheet resistance. The sheet resistance as a function of the annealing temperature, for the three different samples (a, b, and c) is reported in Fig. 1. The annealing time was fixed at 60 s. A fast reduction of the sheet resistance between 700 and 750 °C has been observed for sample (b). Raising the anneal temperature, the sheet resistance remains constant up to 850 °C, and increases at 900 °C. The lowest value is about 6Ω/□. In sample (a) the sheet resistance reaches the lowest values to 18 Ω/□, a factor 3 higher than in sample (b), at 700 °C. The sheet resistance increases after the 900 °C anneal and reaches 30 Ω/□ at 950 °C. In sample (c) the sheet resistance at 750 °C is 3 Ω/□ and remains constant over the range studied (<950 °C).

The decrease of the sheet resistance up to 750 °C observed in samples (b) and (c), is due to the sequence of phases formed between cobalt and silicon. During the annealing, at the interface with the silicon substrate, the CoSi and $CoSi_2$ phases nucleate sequentially.[10] The resistivity of the CoSi phase is 150 $\mu\Omega$ cm, (in the bulk system), while for the $CoSi_2$ phase it is ~25 $\mu\Omega$ cm. Increasing the temperature to 750 °C, the CoSi layer is consumed at the expense of the $CoSi_2$ layer. At 750 °C the reaction of the $CoSi_2$ phase is complete and the sheet resistance saturation at the lowest value differs for the three samples because of the different thickness.

In order to investigate the kinetics, the mass transport, and the Co–Si reaction in the Si/Ti/Co multilayer system, we

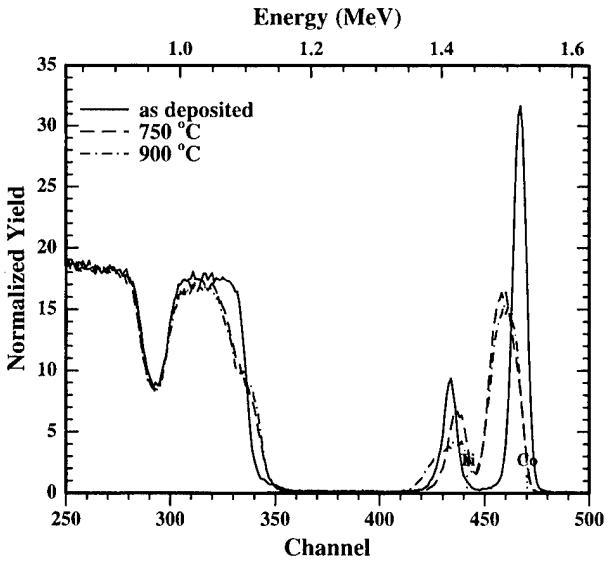


FIG. 2. 2 MeV He^+ Rutherford backscattering spectra of Si/Ti/Co (10 nm Ti and 15 nm Co) as deposited, and reacted at 750 and 900 °C for 60 s. In the last case the titanium atoms diffuse inside the $CoSi_2$ layer.

have analyzed the samples annealed at different temperature by RBS. In Fig. 2 we report the spectra for sample (b) in three different cases: the as deposited, and the annealed at 750 and at 900 °C for 60 s, respectively. In the as-deposited sample, the cobalt layer is at the surface and the titanium at the interface with the amorphous silicon substrate. After the 750 °C thermal process, the cobalt signal shifts to lower He backscattered energy, while the titanium signal to higher energy and the silicon signal show a small shoulder at the high-energy edge that is due to the Co/Si reaction. At 900 °C, the cobalt signal remains unchanged while the titanium peak broadens towards the low-energy region of the RBS spectrum. These spectra indicate that during the anneal at 750 °C the cobalt atoms diffuse through the titanium layer and react with the silicon substrate to form the $CoSi_2$ phase (also detected by x-ray diffraction). In the 900 °C annealed sample the broadening of the Ti signal indicates that a small amount of Ti atoms remains in the $CoSi_2$ layer, probably segregated at the grain boundaries. The titanium presence inside the $CoSi_2$ thin film might be responsible for the increase in the sheet resistance reported in Fig. 1. Being diffused inside, the titanium atoms cannot be removed by the subsequent etch, so that a suitable temperature range has to be chosen in order to avoid this in-diffusion. The comparison of the RBS analysis spectra with the sheet resistance measurements indicates in 750 °C a suitable temperature for the reaction. The temperature range (750–850 °C) over which the reaction of the Ti/Co bilayer on polysilicon should take place is narrower than that for the same kind of reaction on (001) silicon substrates. Previous works[7,15] have reported that the optimum anneal temperature in the last case is between 800 and 1100 °C. In both cases, the reaction temperatures are higher than that one used in the standard two-step process (540 °C for the first step and 800 °C for the second one). The presence of the interfacial titanium layer reduces the cobalt flux to the interface in such a way that, for the same annealing time, higher temperatures are needed to complete the silicide reaction.

According to our experimental results, we choose 750 °C for the reaction process. The $CoSi_2$ reaction is completed and the diffusion of titanium in the cobalt disilicide layer is negligible. After etching in $NH_4OH:H_2O_2$ solution, the residual titanium atoms in the top layer overlying the $CoSi_2$, measured by RBS, are about 1×10^{15} atoms/cm^2 for samples (a) and (b). In sample (c) this evaluation is difficult because of the superimposition of the Co and Ti signals.

B. Grain growth

When a polycrystalline silicide layer is annealed at high temperature two phenomena occur: grain growth and agglomeration. The driving force for both processes is the reduction of the surface energy of the layer.

In Fig. 3 we show the surface microstructure of the silicide in samples (b), for two different cases: the as-reacted sample [Fig. 3(a)] and the sample annealed at 950 °C [Fig. 3(b)]. The annealing process increases the average grain diameter of the silicide from 45 to 90 nm. Moreover, the spread in grain size is noteworthy. Large grains up to 100–150 nm surrounded by small grains with the same diameter of the layer thickness are present in the annealed samples. The grain diameter distribution is reported in Fig. 3(c) and the change with temperature indicates the occurrence of an Ostwald ripening process. The trend is typical of grain growth, the larger grains increase at the expense of the smaller ones. While the grains grow, they also tend to become spherical (agglomeration) reducing in this way the total surface energy. This process deteriorates the resistivity of the film,[16] so it has to be studied in detail to fix an upper limit to the thermal process that the silicide layer can tolerate.

C. Thermal stability versus $CoSi_2$ layer thickness

A previous study[17] showed a dependence of the silicide thermal stability on the layer thickness. Nolan, Sinclair, and Beyers have explained this dependence theoretically[11] and they have derived the relationship between the silicide thickness and the critical grain size, i.e., the size for which the grain-boundary (GB) area becomes zero. From this expression it is clear that thicker silicide layers should have higher thermal stability, as observed[17] in platinum silicide by sheet resistance measurements. No data are available, to our knowledge, for cobalt silicide.

To investigate the dependence of the thermal stability on the film thickness, the sheet resistance (R_s) of samples (a), (b), and (c), annealed for 30 s between 800 and 1050 °C, was measured by a four-point probe at room temperature. The ratio of R_s to the initial value R_{s0} of the as-reacted sample, is reported versus the annealing temperature in Fig. 4. R_{s0} is 18 Ω/□ for sample (a), 6 Ω/□ for sample (b), and 2.5 Ω/□ for sample (c), respectively. The sheet resistance of sample (a) remains constant up to 950 °C and it increases quickly to

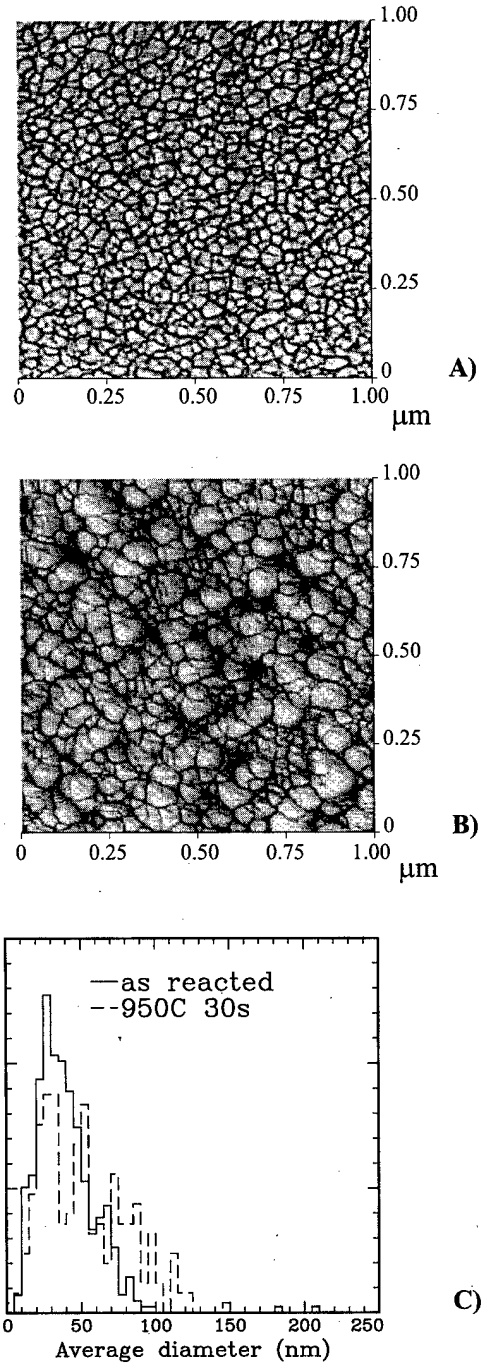


FIG. 3. Surface atomic force microscopy of sample (b) as reacted (A) and annealed at 950 °C for 60 s (B). During the annealing the average silicide grain diameter increases from 45 to 90 nm. the histograms of the silicide grain diameter distribution are reported in (C).

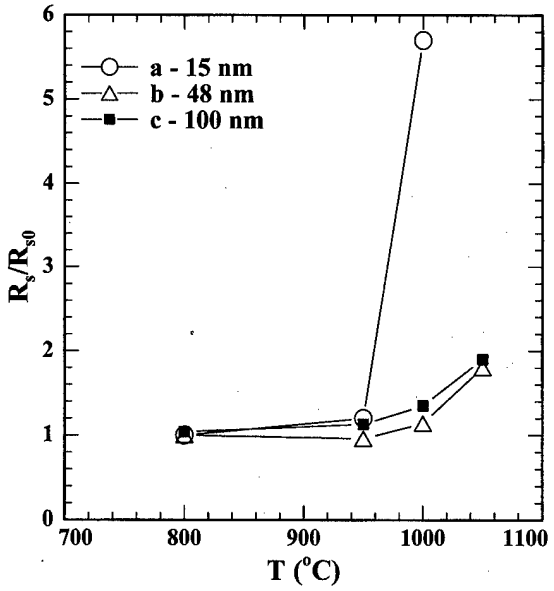


FIG. 4. Sheet resistance normalized to the R_{S_0} value of the as-reacted sample vs the annealing temperature.

$6R_{s0}$ at 1000 °C. Instead, the sheet resistance of both samples (b) and (c) increases slowly starting from 1000 °C.

The data show a steep increase of the sheet resistance on the annealing temperature for the 15 nm thick silicide sample. For thicker samples (48 nm and higher) the ratio follows the same curve: sample (c) (100 nm) has the same values of sample (b) (48 nm $CoSi_2$). This behavior is similar to that reported[17] for platinum silicide: the dependence of the sheet resistance on the silicide thickness was strong for thin PtSi films and almost negligible for films thickness higher than about 60 nm.

To understand better this phenomenon we have analyzed by transmission electron microscopy the morphology of the silicide layers. In all the samples (a, b, and c), the interface between the as-reacted silicide and the polycrystalline silicon is almost flat and the layers are uniform in thickness. After the 1000 °C 30 s annealing, the thinnest silicide (a) becomes discontinuous, and some silicon grains appear at the surface [Fig. 5(a)]. The TEM cross sections of samples (b) and (c) [Figs. 5(b) and 5(c)] show a thermal grooving process at the interface close to the groove root, while the $CoSi_2$ surface remains almost flat. For both samples, the film roughness increases mainly at the silicide–silicon interface. This behavior can be explained using the Nolan model.[11] The equilibrium groove angle θ_s at the silicide surface is given by

$$\vartheta_s = \arcsin\left(\frac{\gamma_b}{2\gamma_s}\right),$$

and at the silicide/substrate interface by

$$\vartheta_i = \arcsin\left(\frac{\gamma_b}{2\gamma_i}\right),$$

where γ_b, γ_s, and γ_i are the grain boundary, the surface, and the interface free-energy, respectively (see the inset in Fig. 5).

The larger roughness of the silicide/silicon interface is due to the larger value of the surface free-energy γ_s with respect to the interface one γ_i ($\gamma_b \sim \gamma_i \sim \gamma_s/3$).[18] The data of Fig. 5 indicate also that the agglomeration process depends on the film thickness, the thinner film, the more likely it will break into islands. For thicker layers, the holes in the silicide decrease and the interface roughness increases. This agglomeration process produces a deterioration of the sheet resistance as reported in Fig. 4. In fact, as previously observed,[16]

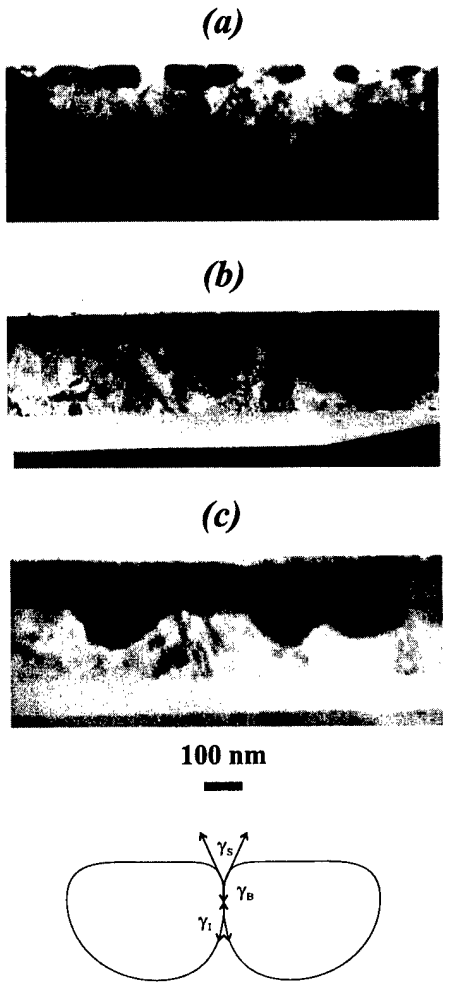


FIG. 5. Cross section (TEM) of samples (a), (b), and (c) after 1000 °C for 30 s anneal. In the inset the equilibrium shape of the grains is shown: γ_b, γ_s, and γ_i are the grain boundary, the surface, and the interface free-energy, respectively.

a strong correlation exists between the shape of the samples and the electrical features. When the silicide starts to agglomerate, the film thickness decreases drastically close to the grain boundary and becomes comparable to the mean-free path of the electrical carriers. This thickness reduction produces a "*local*" resistivity increase and the deterioration of the "*average*" resistivity of the film.

To compare the resistivity behavior of films with different thicknesses careful analysis is necessary. In our previous work,[16] we reported a model that correlates the average resistivity value with both the thickness and the roughness of the layer. Using this model, we can calculate the roughness corresponding to a fixed average resistivity for different silicide thicknesses (solid line in Fig. 6). The resistivity has been chosen equal to the average value of samples (b) and (c) annealed at 1000 °C for 30 s, i.e., 33 $\mu\Omega$ cm (see Fig. 4). From the model results (see Fig. 6) it is clear that, for a fixed resistivity, thicker films should have a larger layer roughness. In the same figure we have also reported the measured values for the samples (b) and (c) annealed at 1000 °C. Thickness and roughness are obtained from the TEM micrographs of Fig. 5. In sample (b) (48 nm thick) the observed sheet resistance and the film roughness of 20 nm agree quite well with the calculated solid line. A reasonable agreement is also found for sample (c) (100 nm thick) where a roughness of 60 nm is measured for the same sheet resistance value. The error bar on the reported experimental points is due to the experimental evaluation of the layer roughness on the micrographs.

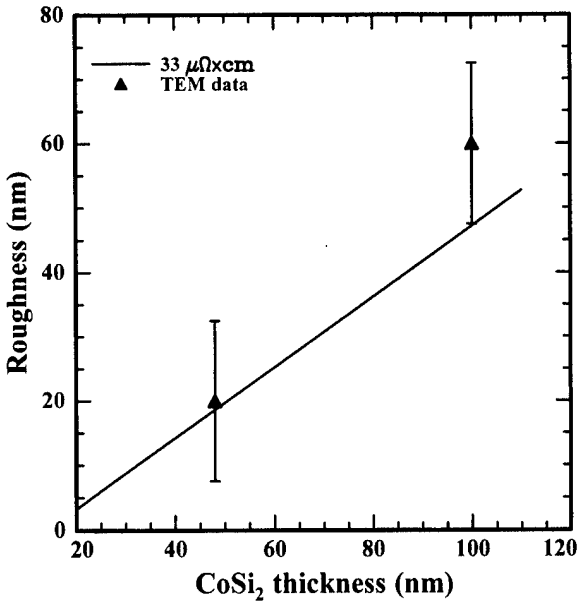


FIG. 6. $CoSi_2$ roughness vs layer thickness for a fixed resistivity: the solid line represents the model; the experimental data, obtained from Figs. 5(b) and 5(c), are reported, too.

The previous TEM results (Fig. 5) indicate that there is a critical thickness (around 50 nm) above which the layer roughness increases faster in thicker than in thinner samples. But increasing the film thickness, the changes of the sheet resistance caused by the agglomeration is reduced (Fig. 6). In fact, in thick films the "*local*" thickness reduction to achieve the same resistance increase corresponds to a stronger interface roughness than in thin films. Then, considering only the analysis of the electrical measurements (Fig. 4), one can erroneously argue that thick films have a better thermal stability.

D. Thermal stability for the systems Si/Ti/Co and Si/Co

The aim of the present section is to compare the thermal stability of the $CoSi_2$ thin film, obtained with or without the intermediate Ti layer [samples (c) and (d)]. The normalized sheet resistance data are plotted in Fig. 7 as a function of the annealing time for different temperatures. The experimental points are fitted by straight lines and for all the samples the slope increases with the annealing temperature. As we have shown in a previous work,[16] this value is proportional to the agglomeration rate of the $CoSi_2$ thin film and it is then a reliable parameter in the study of the silicide thermal stability. At each temperature, the slope of sample (c) is steeper

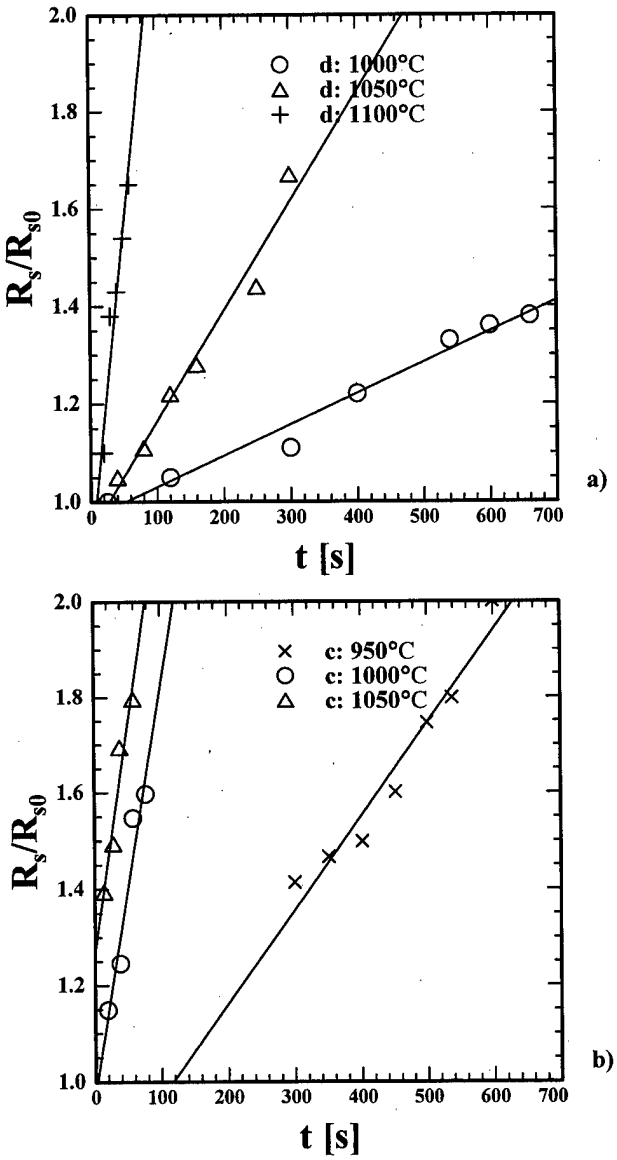


FIG. 7. Agglomeration rate at different annealing temperature vs annealing time for the silicide reacted with (c) and without (d) titanium layer. Each temperature gives rise to an agglomeration process faster in sample (c) then in sample (d).

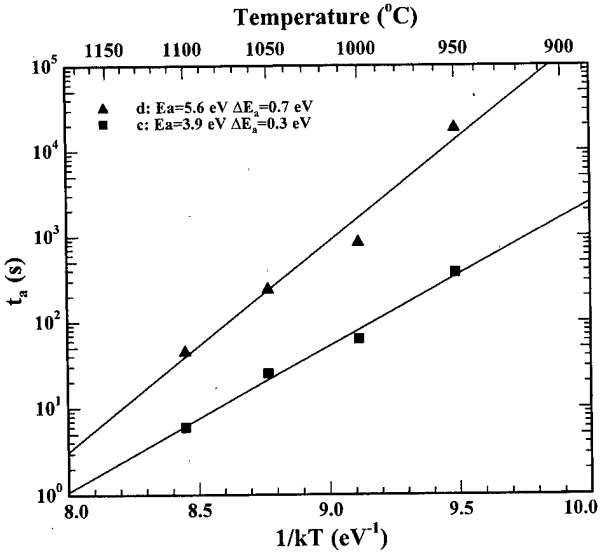


FIG. 8. Arrhenius plot for the agglomeration process. Sample (d) shows the highest activation energy.

than that of the (d) sample. Then the silicide formed by the direct reaction of Co and Si is more stable with respect to that grown in the presence of a thin titanium layer at the cobalt/silicon interface.

The different behavior between the two samples can be more clearly evidenced if we increase the annealing time t_a to reach a fixed value of the normalized sheet resistance. We have chosen a 50% increase. The time t_a is reported in an Arrhenius plot in Fig. 8. Clearly, the agglomeration process is thermally activated and two different energies are extracted. The choice of the sheet resistance increase does not influence the obtained activation energies. In fact, if we select an increase of the sheet resistance of 20%, and if we report the corresponding agglomeration times in the Arrhenius plot, the same activation energies, within the experimental errors, are obtained.

The activation energies are 5.6 eV for the $CoSi_2$ layer grown by the direct reaction (Co/Si) and 3.9 eV for the film reacted starting from the Ti/Co bilayer. The first value is close to that reported[19] for the agglomeration of a 50 nm polycrystalline $CoSi_2$ film grown on Si (001). In the case of $CoSi_2$ formed on a polysilicon substrate, the reported activation energies ranged between 3.8 and 5.4 eV.[2,16] In Ref. 2, the low activation energy (3.8 eV) was associated with the underlying polysilicon grain growth, to which the deterioration of the film has been attributed. Instead, when the substrate did not take part in the process,[16] the activation energy is 5.6 eV. This value was related, instead, to the break of the cobalt silicide molecule close to the grain boundary or at the reaction between cobalt and silicon at the $CoSi_2$/Si interface.

The activation energy of the silicide layer formed from the Si/Ti/Co multilayer is close to the low value reported[2] in the case of the polysilicon substrate grain growth. We have analyzed the substrate after etching the silicide layer but we have observed no large differences in the silicon grain size distribution before and after the thermal treatment at high temperature. Consequently, the low value of the activation energy cannot be related to the structure of the underlying substrate. With the exclusion of the substrate, it can be argued that either the surface or the silicide morphology comes into play.

The TEM cross sections of Fig. 9 show the presence of a thin surface layer on top of the silicide reacted in the presence of Ti [Fig. 9(a)]. This layer is 30 nm thick and it remains after the standard etch solution for the TiN (NH_4OH:H_2O_2). According to previous works,[8,14] there are only two possible compositions: TiN(O), or Ti–Co–Si (1:1:2) ternary alloy. From RBS analysis in glancing configuration (not shown), the average composition of the thin surface layer resulted very close to the second one previously reported. Nevertheless, we can exclude that it affects the silicide thermal stability for the two following reasons. First,

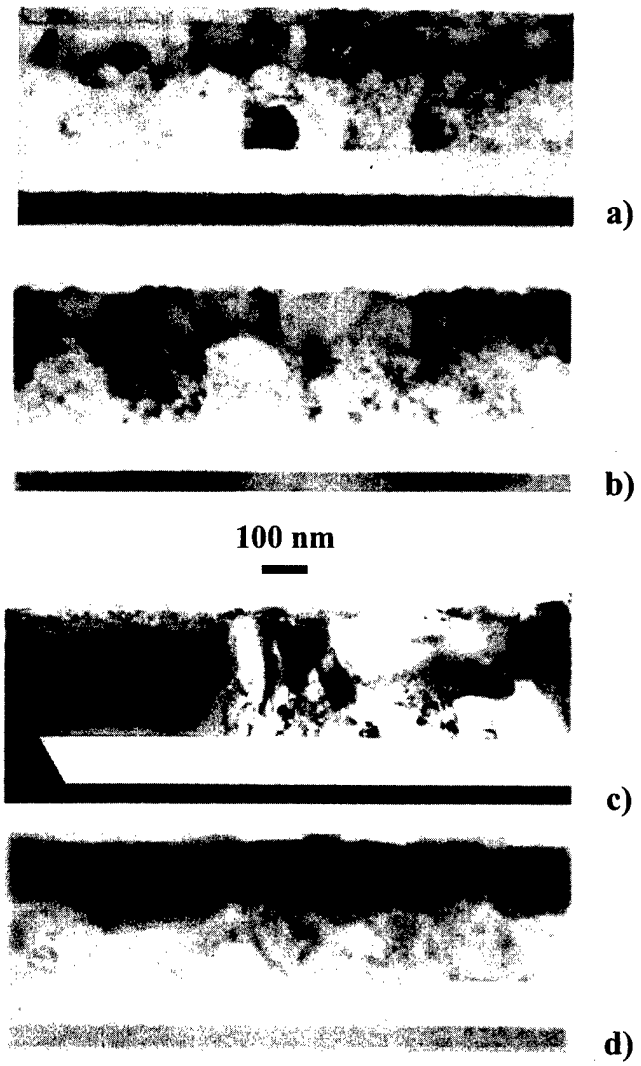


FIG. 9. TEM cross sections of sample (d) and (c). Sample (d): (b) as reacted and (d) after 1050 °C 40 s anneal. Sample (c): (a) as reacted and (c) after the same thermal process of sample (d).

the agglomeration process starts at the silicide/silicone interface so that the surface layer does not affect considerably the surface silicide energy. Second, several experimental data[8,20] show an unchanged or, somewhere else, an increase of the silicide thermal stability in the presence of these thin surface cap layers.

The low value of the activation energy found in our experiment might be related only to the morphology and to the microstructure of the silicide layer. The as-reacted sample (d) [Fig. 9(b)] shows a columnar structure of the $CoSi_2$ grains with an average grain size of 150 nm, while sample (c) [Fig. 9(a)] is characterized by several noncolumnar grains and a lower average grain dimension (50–100 nm). After a process at 1050 °C for 30 s, the structure of sample (d) remains almost unchanged [Fig. 9(d)], while for sample (c) a large increase in the silicide grain size occurs and several holes appear in the layer [Fig. 9(c)]. The agglomeration rate is enhanced in this last sample probably because of the presence of small and noncolumnar grains compared to those of the first sample. This particular structure is due to the high temperature used in the reaction that produces a higher nucleation rate with respect to sample (d). Therefore, in the case of the Ti/Co bilayer the reacted film has a high surface–volume ratio, i.e., the density of grain boundaries, where the break of Co–Si bonds occurs, increases. The direct consequence is a lower agglomeration time in sample (c) (Fig. 8), because of the quicker grooving and deterioration of the silicide. Furthermore, the curvature radius of the grain boundaries in the noncolumnar grains is reduced, so this can produce a decrease of the energy required to break the Co/Si bonds with respect to the columnar grains. Then, the different activation energy for agglomeration between the silicide layers obtained with or without the Ti thin film at the cobalt/silicon interface can be explained in terms of the different film structure.

IV. CONCLUSIONS

The growth and the thermal stability of $CoSi_2$ thin films obtained by the reaction of a poly Si/Ti/Co multilayer was studied.

The optimal reaction temperature to form the $CoSi_2$ phase was found in the range between 750 and 850 °C. At higher temperatures the titanium atoms diffuse in the $CoSi_2$ film increasing the silicide resistivity.

The sheet resistance after the high-temperature process increases as a function of the silicide thickness and it saturates for a thickness larger than about 48 nm. This behavior does not correspond to a saturation process in the thermal groove at the GB, but it can be explained in terms of the roughness/thickness ratio. In fact, thicker layers have a higher interface roughness under the same anneal conditions.

The high-temperature anneals also produce a $CoSi_2$ grain growth of the largest grain at the expense of the smallest ones. Then, an Ostwald ripening process is present. During the agglomeration process the morphology of the silicide layer plays a relevant role. It is, in fact, responsible for the reduction of the agglomeration activation energy observed for the films obtained by the titanium-mediated reaction with respect to the usual $CoSi_2$ film. This low thermal stability of the $CoSi_2$ layer on polysilicon can be a problem for the use of the Ti/Co bilayer process in self-aligned complementary metal–oxide–semiconductor technology.

ACKNOWLEDGMENTS

The authors wish to thank V. Raineri for AFM measurements, M. Furnari for technical assistance, and S. Pannitteri for TEM sample preparation and analysis. This work has been partially supported by the Progetto Finalizzato MADESS II.

[1]B. Shing and M. C. Chen, J. Appl. Phys. **72**, 4619 (1992).
[2]Q. Z. Hong, S. Q. Hong, F. M. D'Heurle, and J. M. E. Harper, Thin Solid Films **253**, 479 (1994).
[3]A. H. Reader, J. P. W. B. Duchateau, and J. E. Crombeen, Semicond. Sci. Technol. **8**, 1204 (1993).
[4]S. M. Yalisove, R. T. Tung, and D. Loreto, J. Vac. Sci. Technol. A **7**, 1472 (1988).

[5]A. E. White, Appl. Phys. Lett. **50**, 95 (1987).
[6]F. La Via, A. H. Reader, J. P. W. B. Ducheteau, R. A. Hakvoort, and A. van Veen, J. Vac. Sci. Technol. B **11**, 1807 (1993).
[7]A. Vantomme, M. A. Nicolet, and N. D. Theodore, J. Appl. Phys. **75**, 3882 (1994).
[8]A. Lauwers, R. J. Schreutelkamp, B. Brijs, H. Bender, and K. Maex, Appl. Surf. Sci. **73**, 19 (1993).
[9]F. Hong, G. A. Rozgonyi, and B. Patnaik, Appl. Phys. Lett. **61**, 1519 (1992).
[10]F. Hong and G. A. Rozgonyi, J. Electrochem. Soc. **141**, 3480 (1994).
[11]T. P. Nolan, R. Sinclair, and R. Beyers, J. Appl. Phys. **71**, 720 (1992).
[12]D. J. Srolovotz and S. A. Safran, J. Appl. Phys. **60**, 255 (1986).
[13]K. T. Miller, F. F. Lange, and D. B. Marshall, J. Mater. Res. **5**, 151 (1990).
[14]S. L. Hsia, T. Y. Tan, P. Smith, and G. E. McGuire, J. Appl. Phys. **72**, 1864 (1992).
[15]S. L. Hsia, T. Y. Tau, P. Smith, and G. E. McGuire, J. Appl. Phys. **70**, 7579 (1991).
[16]F. La Via, A. Alberti, V. Raineri, S. Ravesi, and E. Rimini, J. Vac. Sci. Technol. B **16**, 1129 (1998).
[17]S. R. Das, K. Sheergar, D. X. Xu, and A. Naem, Thin Solid Films **253**, 467 (1994).
[18]T. P. Nolan, R. Sinclair, and R. Beyers, J. Appl. Phys. **71**, 720 (1992).
[19]H. Jiang, C. M. Osburn, Z. G. Xiao, G. Mcguire, and G. A. Rozgonyi, J. Electrochem. Soc. **139**, 211 (1992).
[20]R. T. Tung and F. Schrey, Appl. Phys. Lett. **67**, 2164 (1995).

Impact of plasma-enhanced chemical vapor deposited oxide characteristics on interconnect via resistance and device performance of four-transistor static random access memory with polysilicon load resistors

C.-F. Lin
Institute of Materials Science and Engineering, Chiao-Tung University, Hsinchu 300, Taiwan and Thin Film Engineering Department, FAB-2, Winbond Electronics Corporation, Hsinchu 300, Taiwan

W. T. Tseng[a)]
Department of Materials and Engineering, Cheng-Kung University, Tainan 701, Taiwan

M. S. Feng
Institute of Materials Science and Engineering, Chiao-Tung University, Hsinchu 300, Taiwan

Y. L. Wang
Engineering Department-2 FAB-4, Taiwan Semiconductor Manufacturing Company, Limited, Hsinchu 300, Taiwan

(Received 21 January 1999; accepted 14 May 1999)

By varying the gas ratio during the process of plasma-enhanced chemical vapor deposited (PECVD) SiO_x, the composition of the oxide used as the intermetal dielectric (IMD) in the device is modified, and its impact on the integrity of the interconnect metal via hole and the stability of four-transistor (4-T) static random access memory (SRAM) with poly-Si load resistors is investigated. PECVD oxides using precursors of SiH_4 and N_2O under various gas ratios are adopted to manipulate the Si:O atomic ratio and other dielectric characteristics. An increase in the Si atomic percent in IMD film would induce a higher via resistance. Cross-sectional scanning electron microscopy of the post-etching via holes in a Si-rich IMD sample reveals a high amount of plasma-induced polymer formation around the via holes, which is perceived as the root cause of the dimensional decrease in via hole size and a corresponding increase in via resistance. As the gas ratio of SiH_4/N_2O increases, the IMD films become more Si rich with a higher refractive index and an *a*-Si-like dangling bond ($\cdot Si{\equiv}Si_3$) density. The *a*-Si-like dangling bonds ($\cdot Si{\equiv}Si_3$) in IMD films serve as effective trapping centers for hydrogen or moisture above the second polysilicon load resistor and hence protect them from attack by back-end process-induced mobile charges. The resistance of these poly-Si load resistors is maintained at a high level and device performance is secured. Thus, the high quality of a 4-T SRAM device with stable load resistance, could be realized, while maintaining a low interconnect metal via resistance. © *1999 American Vacuum Society.*
[S0734-211X(99)06404-5]

I. INTRODUCTION

During the past few years, increasing attention has been paid to the impact of material issues, such as interlayer dielectric (ILD), intermetal dielectric (IMD) and passivation layers, rather than the active gate area alone, on semiconductor device performance and reliability. For four-transistor (4-T) cache static random access memory (SRAM) devices, reliability and performance issues such as hot-carrier-induced degradation,[1,2] time-dependent dielectric breakdown,[3] and poly-Si load resistors[4,5] have all been influenced by the dielectric material characteristics. The resistance of the poly-Si load resistor in 4-T SRAM is one of the critical electrical characteristics for device performance and reliability. The resistance should be kept as low as possible during the programming cycle when a pull-up voltage passes through the polyload resistor. Conversely, as the other extreme, the resistance of a poly-Si load resistor should be as high as possible to prevent excessive power consumption over concern for storage data security. Therefore, the resistance shift of the poly-Si load resistor induced, for example, by mobile charges released from the surrounding dielectrics, will lead to fluctuation and degradation in device performance, manifested by seed slowdown, high power consumption, heat dissipation and data loss. Ion contamination[5] and humidity[6] have been held responsible as the main charge loss mechanisms.

The deposition of plasma-enhanced chemical vapor deposited (PECVD) oxide is a common process for an IMD, which is followed sequentially by the planarization process and interconnect via etching for the realization of multilevel interconnects. All of these back-end-of-line (BEOL) plasma-assisted process steps and materials used would have certain impact on the interconnect via through hole resistance and the resistivity of lightly doped poly-Si load resistor in 4-T SRAM. Previous studies indicate that the composition of the PECVD silicon nitride passivation layer and reactive-ion

a)Electronic mail: wttsen@mail.ncku.edu.tw

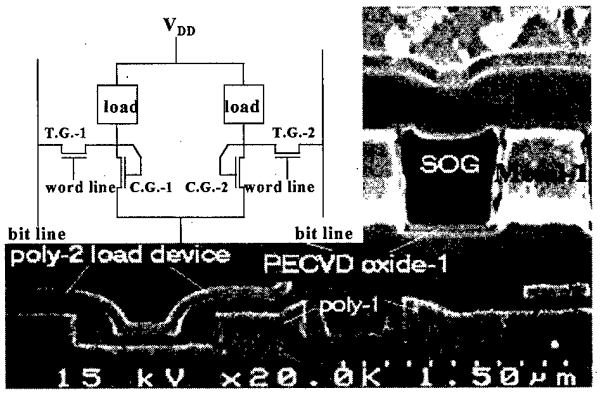


FIG. 1. Schematic and SEM cross-sectional view of a 0.4 mm flip–flop double-polysilicon double-metal 4-T SRAM device using the second polysilicon layer stack above the control gate as the load resistor.

etching (RIE) of the metal interconnect line have a major influence on hot electron degradation of the device.[1,7] PECVD oxides with a high Si content were proposed to improve the hot carrier reliability. This is achieved through the incorporation of a high density of Si dangling bonds, which act as trap centers for the hydrogen that evolves from the plasma nitride passivation, or from the moisture from silicon oxide. Additionally, the silicon-rich oxide (SRO) also forms a charge shield that prevents damage induced degradation in device reliability.[8]

In this investigation, the stoichiometry, refractive index, stress and other characteristics of a PECVD SiO_x IMD were modified to optimize memory device performance. The impact of PECVD silicon oxide stoichiometry on resistivity degradation and the variation of the lightly doped second polysilicon load are evaluated. The trap centers in the oxides will be identified and their trapping mechanisms will be elucidated. However, modifications in the PECVD oxide composition also lead to variation in the metal via dimension and resistance. This abnormal rise in via resistance due to the compositional changes in the IMD is believed to be the direct consequence of via hole etching, which also depends on the composition of the IMD. The mechanism between the compositional change of IMD material characteristics and the interconnect metal via plasma etching ambient which induce higher via resistance will be investigated. An optimization scheme will be proposed in this study to enhance the device reliability of 4-T SRAM, while maintaining a low enough via resistance over concern for device speed.

II. EXPERIMENT

A. Film deposition and characterization

The test vehicle used for this study is a 0.4 μm 4 Mb *N*-metal–oxide–semiconductor (*N*-MOS) 4-T SRAM with two *N*-MOS transfer gates, two *N*-MOS storage gates and two second level thin polysilicon load resistors, completed by double level polysilicon and double level Al interconnects. The cross-sectional scanning electron microscopy (SEM) micrograph in Fig. 1 reveals details of the device. The IMD layers consisted of a 2000 Å PECVD oxide followed by a siloxane spin-on-glass (SOG) planarization coating and furnace curing. The SOG thus coated was etched back partially in CHF_3/CF_4 plasma ambient for planarization, followed by deposition of a 5000 Å PECVD oxide cap layer. The via formation process includes first a 40 s wet etching step with 10:1 NH_4F buffered-HF oxide etching (BOE) solution. This is followed by CHF_3/CF_4 plasma dry etching conducted to define the taper metal via through hole with a size 0.55 μm in order to maintain a reasonable interconnect wire step coverage and chip size.

A commercially available PECVD reactor was used for oxide deposition with a gas mixture of SiH_4, N_2O, and N_2 at a deposition temperature of 400 °C under a pressure of 2 Torr and a rf power of around 150 W. The composition of the PECVD oxide film was modified by varying the SiH_4 to N_2O flow ratio (i.e., the Si:O atomic ratio).

For material characterization, PECVD oxide films were deposited on 150 mm bare silicon wafers. The refractive index (RI) and film thickness were measured by an ellipsometer and an interferometer, respectively. The thin film stress and bulk density were measured by a flatness gauge and by a microbalance. PECVD oxide wet etching was performed in 10:1 BOE for 1 min. Fourier transform infrared (FTIR) spectra of the PECVD oxide films were taken to provide bonding configuration information. The Si:O atomic ratio was determined by an electron probe x-ray microanalyzer (EPMA), a JEOL model JXA-8800M analyzer. The Si dangling bond ($\cdot Si{\equiv}Si_3$) density was measured by an electron paramagnetic resonance (EPR) spectrometer to characterize the mobile ion trapping density of the as-deposited oxide films. The EPR analysis was conducted using a Brüker EMX-10 spectrometer under a center field of 3489.9 G with a width of 50 G. The microwave frequency was set at 9.779 GHz with a power of 19.971 mW. The receiver modulation frequency and its amplitude were 100 kHz and 1.6 G, respectively.

B. Processing monitoring and device characterization

The dimension of the metal via through hole was measured during after-via photoresist development inspection (ADI) and after-via etching inspection (AEI) by top view SEM. The via dimension was determined from the average of five measurements taken across the wafer. Direct observation of the postetching metal via profile and of plasma-induced polymer formation on the via sidewall was conducted by field emission SEM before photoresist removal.

The via resistance and its variation were measured at the end of device processing by probing the via chain test patterns with five site probings per wafer. The resistance of the second polysilicon load resistor was determined by monitoring the voltage drop at constant current. The average load resistance and its variation were taken from five site measurements per wafer and the result is expressed in units of gigaohm (GΩ).

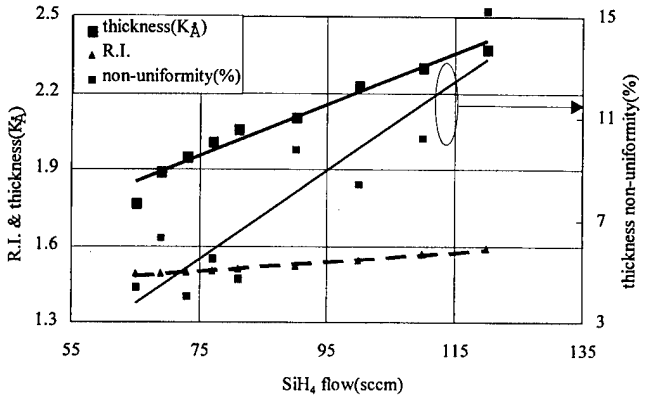


FIG. 2. Experimental results showing PECVD oxide thickness, thickness nonuniformity and reflective index as a function of the silane flow rate used during deposition.

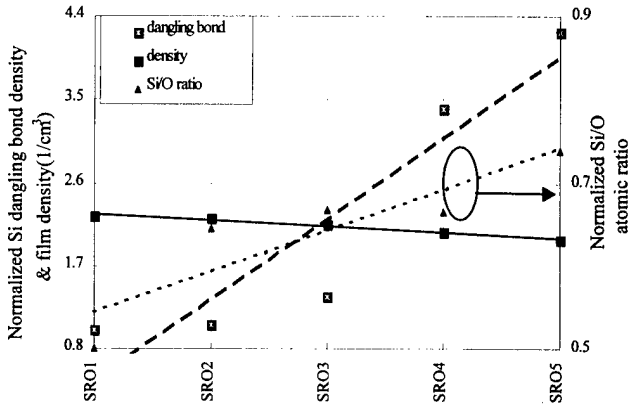


FIG. 3. Normalized Si to O atomic ratio, film density and RI of the as-deposited SRO films.

III. RESULTS AND DISCUSSION

A. Characteristics of PECVD dielectric materials

Figure 2 shows how the increase in the SiH_4 flow rate (i.e., increasing Si content of the oxide) during the PECVD process would affect the film thickness, thickness nonuniformity and reflective index (RI) of the oxide films deposited. For process integration concerns, a target thickness of 2000 Å is desired for the first layer PECVD oxide film. This was achieved on the basis of the deposition conditions provided in Fig. 2. Since the decomposition reaction of SiH_4 in plasma ambient is more efficient and quick than N_2O, the increase in the SiH_4 gas flow rate would result in an enhanced deposition rate and decreased thickness uniformity. Oversaturation of SiH_4 gas flow in the plasma decomposition process would result in Si richness in the as-deposited oxide films. A higher Si content in the film also corresponds to a higher RI. The above result agrees well with at least one previous report that an oxygen-deficient glassy film will exhibit a higher refractive index compared with the stoichiometric silicon dioxide.[9] The extra Si content, in the form of dangling bonds, enhances the electronic polarizability of the films, leading to a higher refractive index. This will be elaborated on later. The results of material characterization for PECVD oxides used in this study are listed in Table I. The results seem to suggest that the modification of oxide characteristics is dominated by the changing SiH_4 gas flow rate.

Figure 3 exhibits the changes in normalized film density, Si dangling bond density and Si/O atomic ratio with the Si content of the as-deposited oxide films. The Si to O atomic ratio, film density and *a*-Si-like dangling bond ($\cdot Si{\equiv}Si_3$) density are detected by EPMA, the microbalance and EPR, respectively. The decrease in oxygen content would introduce more vacancies into the oxide network since, in this case, fewer oxygen atoms are bonded to silicon thus the thin film bulk density would be decreased. In addition, the enhancement in the deposition rate by increasing SiH_4 flow would degrade growth of the silicon oxide network and hence increase the porosity.

The introduction of more Si atoms to the silicon oxide network also alters the optical and dielectric characteristics. When an external electric field E causes a slight displacement x of an electron cloud with a charge q on an atom, the induced dipole moment m is simply qx. The electronic polarizability, α_e can then be expressed as

$$\alpha_e = \frac{qx}{E}. \tag{1}$$

Assuming that the electrons are elastically bound to the nucleus,[10] the charge displacement is related to the interaction force between electrons and the positive nucleus so that

$$qE = kx, \tag{2}$$

where k is the force constant which represents the force per unit length needed to displace the electrons from the nucleus. Substituting Eq. (3) into Eq. (2) yields an alternative expression for electron polarizability:

TABLE I. Summary of the PECVD oxide film characteristics used in this study.

Experimental condition	SiH_4 flow (sccm)	RI	Stress (MPa)	Wet etch rate (nm/min)	Atomic ratio (Si/O)	Normalized dangling bond	Density (g/cm³)
SRO1	65	1.492	−72	362	0.5	1	2.23
SRO2	73	1.499	−34	387	0.64	1.06	2.22
SRO3	81	1.509	−19	390	0.66	1.37	2.14
SRO4	100	1.542	−12	341	0.66	3.4	2.07
SRO5	120	1.589	−34	285	0.73	4.22	1.99

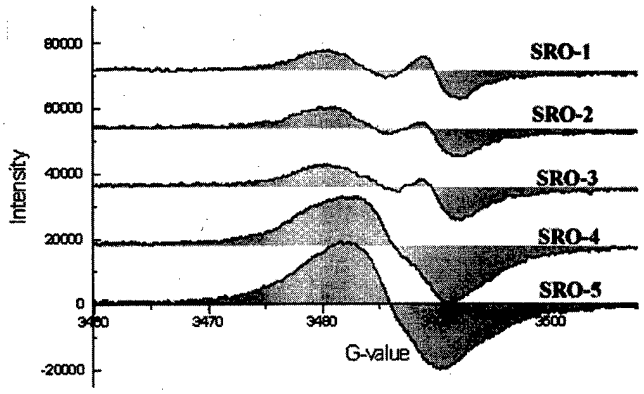


FIG. 4. EPR spectra of SRO films showing the increase in electron spin density and shift in the g factor with increasing Si content from SRO1 to SRO5.

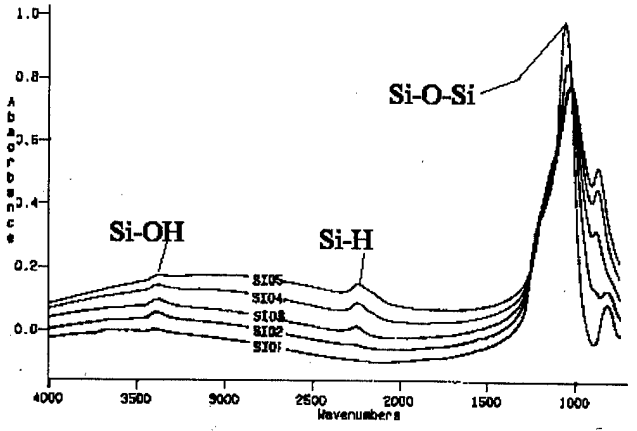


FIG. 5. FTIR spectra showing that the bonding density of Si–OH and Si–O–Si decreases while that of Si–H increases. The peak position shifts to lower wave numbers for Si–O–Si (stretch), Si–H and Si–O–H bands as the SiH_4 gas flow rate increases.

$$\alpha_e = \frac{q^2}{k}. \tag{3}$$

The electronegativities of O, Si and H are 3.6, 1.8 and 2.2, respectively. An atom with a low electronegativity (e.g., Si) will also have a small force constant k, since the nucleus has a weak affinity for electrons. As a consequence, as more Si is incorporated into the oxide network, the electron polarizability is enhanced according to Eq. (4). The electronic polarization k_e, and hence the refractive index, increase correspondingly.

Based on the perspectives above, the continuous rise in RI from that near the stoichiometric oxide (SRO1, RI=1.492) to 1.589 of SRO5 should be the direct consequence of increased silicon content, instead of a densification effect in the deposited oxide films. In fact, the density decreases, *not* increases, with increasing Si content in the oxide. This finding is similar to one in a previous study,[11] which attributed the change of RI in annealed PECVD oxides to alternation in the Si–O bond chemistry rather than densification in the oxide network.

The increase in the number of nonbonded Si atoms would introduce a higher a-Si-like dangling bond ($\cdot Si{\equiv}Si_3$) density in the oxide. This is detected by EPR and the results are shown in Fig. 4. As the silicon content increases from SRO1 to SRO5, the type of dominant defect species present in the oxides also changes. For SRO1, SRO2 and SRO3, a peak with a G value equal to 3489 can be clearly observed. This corresponds to the characteristic spin of E' centers ($\cdot Si{\equiv}O_3$), which are the dominant defect species in these oxides. For SRO4 and SRO5, however, the E' centers disappear and a new path with a G value of 3482.5 emerges. This coincides with the peak of a-Si spin centers ($\cdot Si{\equiv}Si_3$) identified by Kamigaki *et al.*[12] This new peak can be attributed to the existence of Si nanocrystals generated during the deposition. Overall, as the Si to O atomic ratio increases from the stoichiometric value of 0.5 to 0.73 for SRO5, the dominant defect spin centers shift from the E' centers to a-Si spin centers as extra silicon atoms are incorporated into the oxide network. In fact, the oxide in this case can be perceived as a mixture of two phases: the oxide plus the Si nanocrystals.[13] The former contains E' centers as the dominant traps, while the latter comprises the a-Si centers as the major and more effective hydrogen traps than $\cdot Si{\equiv}O_3$. The presence of these Si nanocrystals, 30–50 nm in size,[14] significantly changes the characteristics of the silicon-rich oxides. Therefore, they can be distinguished as a separate phase from the oxide. Further identification and characterization of these Si nanocrystals can be found in another study.[14]

Figure 5 shows the as-deposited FTIR spectra of the oxides used for this study. As more SiH_4 gas was added to the deposition process, both the Si–O–Si and Si–O–H bond densities decrease while the Si–H bond concentration increases gradually. This result suggests that at least part of the additional Si atoms is incorporated in the form of Si–H bonds. The existence of a higher Si–H bond density also corresponds to a higher dangling bond density in the oxide.[11] The peak positions also shift due to this change in chemical composition, i.e., Si–O–Si (stretch), Si–H, and Si–O–H bands all move to lower wave numbers as more Si is present. Note that for stoichiometric SiO_2, the Si–O–Si stretch band should be positioned around 1067 1/cm, corresponding to a RI of 1.452. The lower wave numbers for the Si–O–Si stretch band in the silicon-rich oxides correlate well with their higher RI, as seen in Fig. 2. The shift in Si–O stretching frequency to lower wave numbers with increasing silicon richness also agrees with the observation of Lucovsky *et al.*[15]

The change in atomic composition of the PECVD oxide films also results in different wet etching rates as shown in Fig. 6. The results indicate that the etching rate increases slightly from IMD1 to IMD3 and then decreases to that of IMD5. The excess Si in the first three films may have relieved the compressive strain, which leads to an increase in the etching rate. As reported previously,[16] the relief in compressive stress will result in a corresponding increase in the water diffusion rate in the glassy film and hence a higher dissolution or hydration rate (wet etching rate) in the BOE

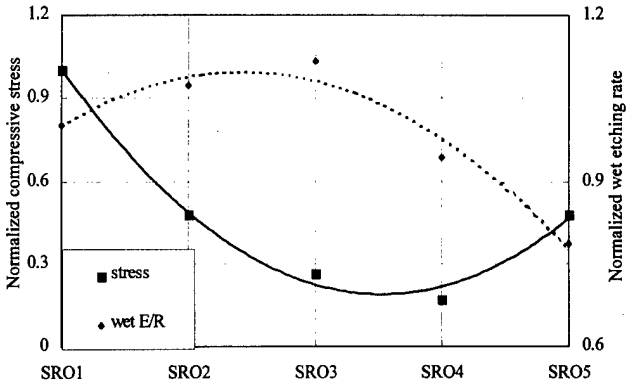


FIG. 6. Normalized wet etching rate and film stress of various SRO-IMD materials.

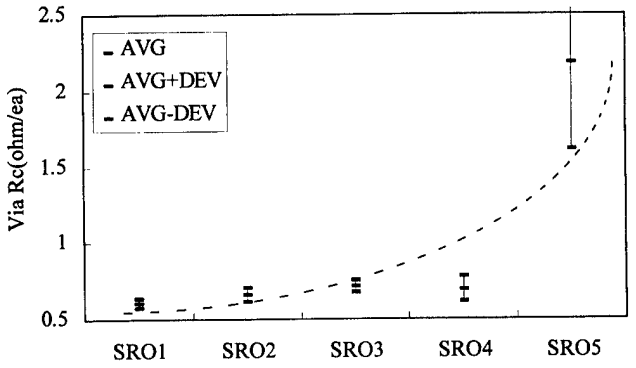


FIG. 7. Metal via resistance and its deviation increase as more Si is incorporated into the IMD oxides.

solution. The incorporation of additional adsorbed Si phases into the film tends to inhibit perfect growth locally, leading to elongated bonds and/or extra silicon dangling bonds and hence a tensile stress component[17] and enhanced chemical activity. The etch rate, in turn, increases correspondingly. Besides, the increase in the Si/O atomic ratio is accompanied by a continuous fall in film density as shown in Fig. 3. This lowering density or increased porosity of silicon oxide will also result in high BOE wet etching rates due to an enlarged reactive contact area, As the Si contents increases further to that of IMD4 and IMD5, the existence of extra Si atoms eventually reduces the reactivity and dissolution rate of the oxide in BOE etchant, leading to a drop in the etching rate again. Overall, the trends in stress and wet etching rate shown in Fig. 3 are consistent with the above postulations. Besides, as the IMD oxide becomes more silicon rich as in SRO4 and SRO5, the film's compressive stress increases again. This can be attributed to the existence of Si microcrystals generated during deposition, and identified previously by the EPR analysis. However, further experiments including high-resolution transmission electron microscopy (TEM) observation are underway to verify this point.

B. Effects of IMD materials on interconnect metal via integrity

An unexpected rise in interconnect metal via resistance with a large variation was detected when the Si-rich oxide films were implemented as the IMD layer for device reliability improvement. This is shown in Fig. 7. To determine the root cause of such a variation in metal via resistance, the in-line metal via related process data were monitored and the results are shown in Fig. 8. According to the results, there appears to be no significant difference in metal via hole critical dimension (CD) for any of the IMD materials from results of ADI. However, the CD from AEI exhibits a decreasing trend with increasing Si content in the films.

To investigate the underlying mechanism leading to the variation in metal via AEI-CD, cross-sectional SEM (XSEM) observation was performed on the postetching via hole and the results are shown in Figs. 9(a) and 9(b). Close inspection of the two micrographs reveals the significant differences in the wet-and-dry etch sidewall profile and the polymer formation in the via sidewall in Figs. 9(a) and 9(b).

The difference in the wet etching profile between Figs. 9(a) and 9(b) might be associated with the difference in the wet etching rate between PECVD IMD materials as discussed and shown in Fig. 6. The sharp metal via profile shown in Fig. 9(b) would contribute in part to the increase in metal via resistance of the IMD5 sample, since the sharp via profile would degrade the sidewall step coverage of the sputter deposited Al interconnect wire. This leads to a higher average via Rc with a large variation as shown in Fig. 7. Based on the results, fine-tuning of the via wet etching recipe according to the different characteristics between dielectrics would be required to achieve a tapered via profile.

It is believed that, as the composition changes in the IMD materials, the interconnect via etching process also needs to be optimized in order to compensate for the change in the characteristics of the complex metallic polymers (metal-$C_xF_yO_z$) formed during the plasma etching process. The change in IMD composition would change the complex metallic polymer (metal-$C_xF_yO_z$) formation behavior on the bottoms and sidewalls of via holes, leading to different via size and via resistance.[18,19]

According to the glow discharge mechanism[20] of halocarbon plasma etching ambient, the fluorine and carbon in the

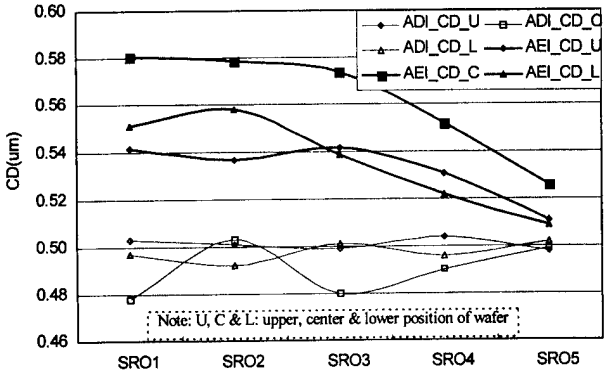
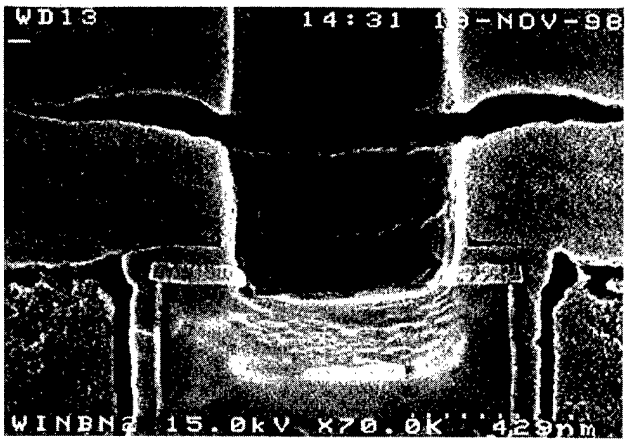


FIG. 8. Metal via AI and ADI critical dimension vs PECVD deposition conditions. Refer to the text for details.



(a)



(b)

FIG. 9. (a) XSEM picture of a via with SRO1 as the IMD. (b) XSEM picture of a via with SRO5 as the IMD.

etch gas would decompose into F* or C*, respectively, which are then converted into parts of etching and polymerization reactions with the reactive species that came from the ambient. As observed earlier from FTIR and EPMA spectra, the IMD materials contain different atomic percentages of Si, O and H. The composition of IMD materials seems to be a dominant factor in the deviation of metal via size at the same etching ambient.

The most important oxide etching mechanism is probably the breaking and reforming of bonds within and on the surface of the $SiC_xF_yO_z$ layer due to the collision cascade produced when an energetic ion hits and penetrates the surface.[20] This procedure easily desorbable etch products that are weakly bonded to the surface, such as SiF_4, SiF_2, CO, $SiOF_2$ and possibly, O_2. The adsorbed carbon atoms can act as etchants here, removing oxygen from the surface by reactions such as

$$\text{ion}+\text{C(s)}+\text{SiO}_2\text{(s)}\rightarrow\text{CO(g)}+\text{SiO(s)}+\text{ion}$$
$$\rightarrow\text{CO}_2\text{(g)}+\text{Si(s)}+\text{ion}. \quad (4)$$

Hence the presence of oxygen in the lattice impedes carbon buildup, allowing the surface to be etched. For the same

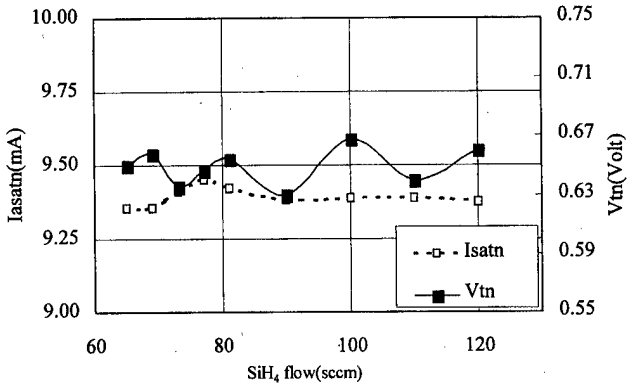


FIG. 10. Changes in N-MOS threshold voltage (V_{tn}) and saturation current (I_{satn}) of the 4-T SRAM vs the SiH_4 flow rate used for IMD deposition.

conditions, carbon-containing polymer films as thick as 100–200 Å are observed to form on silicon as well as on a non-active surface.[20] The film on silicon surface inhibits the etching reaction there, leading to the overall lower plasma etching rate for silicon-rich oxides and polymer formation on their surface.

Therefore, for oxides with more Si and H, more F ions are required to form volatile gas species such as SiF_x or HF fluorides, described above, during plasma etching. That is, more reactive F* ions are consumed so as to lower the F to C ratio and hence an enhanced polymerization reaction during plasma etching process. Meanwhile, if an insufficient amount of oxygen is fed into the system for the silicon-rich SRO5, it will be more difficult for the active carbon radicals to be oxidized to form a volatile gas such as CO_x, thereby leading to a higher carbon concentration and again a relatively lower F to C ratio, in favor of the polymerization reaction.

The mechanisms described above would account for the results of decrease AEI-CD shown in Fig. 8. The higher polymer content on the via sidewall shown in Fig. 9(b) for a Si-richer oxide compared with the one shown in Fig. 9(a) confirms the arguments above. The results suggest that, to alleviate the side effect of introducing a high Si content to IMD materials, proper modifications in the extent of via wet etching and plasma etching gas ratio are required. These issues will be pursued in future work.

C. Effects of IMD materials on 4-T SRAM device characteristics

As the SiH_4 gas flow rate increases during the deposition of the IMD oxide films, the N-MOS threshold voltage (V_{tn}) and saturation current (I_{satn}) remain virtually at the same level, as shown in Fig. 10. This implies that none of the IMD dielectric materials would impact N-MOS device performance. However, the second polysilicon load resistor shows an increase in resistance and a decrease in its variation with increasing Si content in the IMD, as shown in Fig. 11.

According to the previous study of Kamigaki *et al.*,[12] Si dangling bonds in PECVD oxide film may act as trapping centers for mobile charges induced during the back-end process. This is also one way to improve the hot carrier reliabil-

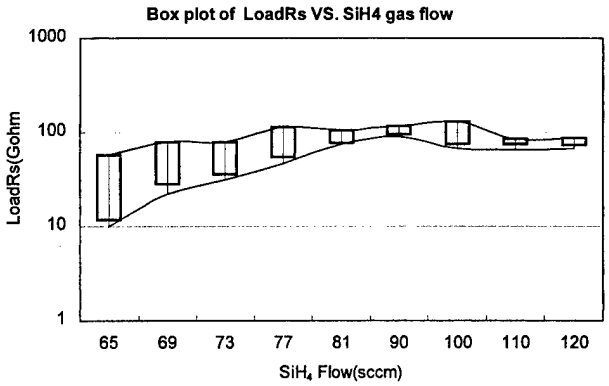


FIG. 11. Changes in load resistance of the poly-Si resistor vs the SiH_4 flow rate used for IMD deposition.

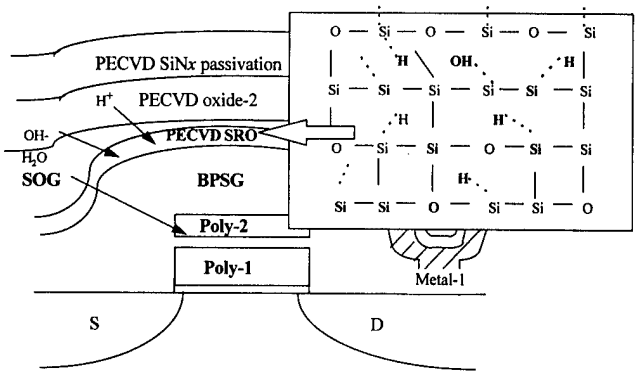


FIG. 12. Schematic of 4-T SRAM showing the impact of mobile ions on a poly-2 load resistor and the trapping mechanisms of silicon-rich oxides containing E' centers ($\cdot Si{\equiv}O_3$) and a-Si centers ($\cdot Si{\equiv}Si_3$).

ity of MOS devices. The mobile charges here usually originate from the plasma, or from the H^0, H^+ or OH^- species released from SOG, PECVD oxide, or the SiN_x passivation layer, respectively. It is reported that the mean diffusion length, $2\sqrt{Dt}$, for H^0, H_2 and H_2O in silica at 400 °C for 5 min is 740, 174 and 0.4 mm, respectively.[21] The fast diffusion of hydrogen (H^0 or H^+) can certainly be held responsible in part for the fluctuation in electrical performance. The high dangling bond density in, for example, SRO5-IMD effectively traps the hydrogen released from the oxide or nitride layers above, protecting the poly-Si load resistor from "poisoning," and maintaining its resistance at a high enough level. Here, the most probable species responsible for the resistance fall in the poly-Si resistors is the positively charged H^+ species, since they counterbalance the negatively charged poly-Si resistors and reduce its resistance thereafter during normal device operation. The trapping mechanism for these H^+ species by the silicon dangling bonds can be expressed as follows:

$$\equiv Si\cdot + H^+ \leftrightarrow \equiv Si{-}H^+. \tag{5}$$

Alternatively, if molecular hydrogen is present in the oxide or nitride layers, the trapping reaction may proceed as

$$Si{-}H + H \leftrightarrow \equiv Si\cdot + H_2. \tag{6}$$

The driving force for reaction (6) is the large difference between the bond energies of Si–H, 3.4 eV, and H–H, 4.5 eV. The Si–H bond that is attacked can be at the interface that has at least one Si-atom neighbor, on in the oxide that has three O-atom neighbors.

The diffusion paths of hydrogen or moisture that cause resistance degradation are shown schematically in Fig. 12. Since oxide is inserted beneath the nitride layer, the silicon-rich oxide, with its abundant number of dangling bonds, can effectively trap the mobile hydrogen or moisture released from the top nitride, forming additional Si–H bonds along the oxide/nitride interface. At the other extreme, the incorporation of a stoichiometric oxide layer provides poor trapping capability for hydrogen, which imposes a negative impact on the device characteristics. The experimental results in Fig. 11 suggest that the diffusion of mobile charges may have been suppressed by the a-Si-like dangling bonds ($\cdot Si{\equiv}Si_3$) traps. The existence of these trap centers effectively improves the durability of the lightly doped polysilicon load resistor against moisture and mobile ions to achieve higher and more stable load device resistivity.[22]

Since $\cdot Si{\equiv}Si_3$ defects are found to be more effective hydrogen traps than $\cdot Si{\equiv}O_3$, the SRO5 oxide, when incorporated as an IMD layer, would be more efficient in trapping the mobile ions released from the post-IMD process, protecting the device below. Once the moisture or hydrogen radicals reach the vicinity of the thin and very lightly doped second polysilicon resistor, defect reactions occur and interfacial carrier traps are generated at grain boundaries of the thin polysilicon, hence reducing its resistance. Therefore, the device performance of 4-T SRAM with poly-Si load resistors is improved.

Also worth noting is the fact that the MOS device characteristics have not deteriorated in the presence of SRO1, with low mobile ion trapping efficiency (see Fig. 10). The situation may be different when aggregated and prolonged electrical stressing (e.g., hot-carrier stressing) is conducted. Under this circumstance, the hydrogen or moisture that has made its way through the SRO1 and borophosphosilicate glass (BPSG) layers down to the oxide/Si interface may be converted into charged defects there, under the bombardment of hot carriers.[23] Then, the impact of the hydrogen or moisture on the device surface and the electrical characteristics of the device (e.g., the drain current) will be affected.

IV. CONCLUSIONS

By tailoring the composition of PECVD oxide films incorporated as IMD layers, we developed a process to improve the reliability of a 4-T SRAM device with polysilicon load resistors. Oxide films with more Si content, by increasing the SiH_4 gas flow rate during deposition, contain a higher a-Si-like dangling bond ($\cdot Si{\equiv}Si_3$) density. The presence of these a-Si dangling bonds in silicon-rich oxide serves as effective traps for hydrogen and other mobile ions, hence providing more robust protection for the device underneath with a more stable poly-Si load resistance of the 4-T SRAM. These modifications in the IMD material characteristics also

impose an unexpected impact on the metal via geometry and resistance through via wet etching and plasma dry etching processes. Under the same etching ambient, oxide films with more silicon and hydrogen but less oxygen will contribute to the smaller metal via size after etching due to polymer formation on the via sidewall. In general, fine-tuning of the IMD oxide characteristics during 4-T SRAM fabrication can improve device performance with reasonable interconnect metal via integrity.

ACKNOWLEDGMENT

The authors acknowledge financial support from the National Science Council of the Republic of China under Project No. NSC 88-2215-E-009-057.

[1]M.-L. Chen, C.-W. Leung, W.-T. Cochran, W. Jüngling, C. Dziuba, and T. Yang, IEEE Trans. Electron Devices **ED-35**, 2210 (1988).
[2]N. Lifshitz and G. Smolinsky, IEEE Trans. Electron Device Lett. **EDL-12**, 140 (1991).
[3]A. T. Wu, V. Murali, J. Nulman, B. Triplett, D. B. Fraser, and M. Garner, IEEE Electron Device Lett. **EDL-10**, 443 (1989).
[4]K. Wu, C. Pan, J. J. Shaw, P. Freiberger, and G. Sery, Proceedings of the 28th International Reliability Physics Symposium, 1990, p. 145.
[5]N. R. Mielke, Proceedings of the 21st IEEE International Reliability Physics Symposium, 1983, p. 106
[6]R. E. Shiner, J. M. Caywood, and B. L. Euzent, Proceedings of the 18th IEEE International Reliability Physics Symposium, 1980, p. 238.
[7]S. Yoshida, K. Okuyama, F. Kanai, Y. Kawate, M. Motoyoshi, and H. Katto, Proc. Int. Electron Devices Meet. 22 (1988)
[8]V. Jain, D. Pramanik, S. R. Nariani, and K. Y. Chang, Proceedings of the 8th VLSI Multilevel Interconnection Conference, 1991, p. 272.
[9]W. A. Pliakin and H. S. Lehman, J. Electrochem. Soc. **112**, 1013 (1965).
[10]H. A. Lorentz, Ann. Phys. (Leipzig) **9**, 641 (1880).
[11]C. Jiang, C. Hu, C. H. Chen, and P. N. Tseng, Proceedings of the 30th IEEE International Reliability Physics Symposium, 1992, p. 121.
[12]Y. Kamigakai *et al.*, Proceedings of the 33rd IEEE International Reliability Physics Symposium, 1995, p. 12.
[13]M. Hamasaki, T. Adachi, S. Wakaymaa, and M. Kikuchi, J. Appl. Phys. **49**, 3987 (1978).
[14]C.-F. Lin, W.-T. Tseng, and M.-S. Feng, J. Appl. Phys. (to be published).
[15]G. Lucovsky, S. Y. Lin, P. D. Richard, S. S. Chao, Y. Takagi, P. Pai, J. E. Keem, and J. E. Tyler, J. Non-Cryst. Solids **75**, 429 (1985).
[16]M. Nogami and M. Tomozawa, J. Am. Ceram. Soc. **67**, 151 (1984).
[17]M. Stadtmüeller, J. Electrochem. Soc. **139**, 3669 (1992).
[18]P. T. Chu, S. F. Chen, C. C. Hung, and Y. C. Chao, Proceedings of the Dielectrics for ULSI Multilevel Interconnection Conference, 1997, p. 179.
[19]Y. P. Lee, S. T. Loong and Y. F. Lu, J. Electrochem. Soc. **145**, 3966 (1998).
[20]J. W. Coburn and H. F. Winters, J. Vac. Sci. Technol. **16**, 397 (1979).
[21]S. Shuto, M. Tanaka, M. Sonoda, T. Idaka, K. Sasaki, and S. Mori, Proceedings of the 35th International Reliability Physics Symposium, 1997, p. 17.
[22]K. Machida, N. Shimoyama, J. Takahashi, E. Arai, and N. Yabumoto, Proceedings of the 11th VLSI Multilevel Interconnection Conference, 1994, p. 299.
[23]C.-F. Lin, W.-T. Tseng, and M.-S. Feng, J. Electrochem. Soc. **146**, 1510 (1999).

Study of crystal orientation in Cu film on TiN layered structures

Kazuhide Abe,[a] Yusuke Harada,[b] and Hiroshi Onoda[b]
VLSI R&D Center, Oki Electric Industry Co., Ltd., 550-1 Higashiasakawa, Hachioji, Tokyo 193-8550, Japan

(Received 25 November 1998; accepted 16 April 1999)

The effect of underlayer texture on Cu film orientation has been studied. Cu texture correlates well with the underlayer texture. The Cu⟨111⟩ crystallographic orientation is enhanced on TiN film with a strong TiN⟨111⟩ orientation. Cu preferred orientation can be controlled by choosing an appropriate underlayer. Cross-sectional transmission electron microscope observation revealed that the Cu(111) plane grows epitaxially on the TiN(111) plane in spite of the large lattice mismatch between Cu and TiN. The atomic arrangement between Cu(111) and TiN(111) planes in TiN film deposited on Cu film has a rotational angle within ± 10° around the ⟨111⟩ axis. On the other hand, in Cu film deposited on TiN film, there are two rotational angles: a rotational angle of 25°–30° in large Cu grain region and little rotation with rotational angle distributions less than 10° in a small Cu grain region. These results can be explained by consideration of superlattice mismatch. Cu(111) plane rotation occurs so as to have more energy stable position at coalescence stage. © *1999 American Vacuum Society.* [S0734-211X(99)01304-9]

I. INTRODUCTION

Al-based metal lines have been widely used as interconnects for Si large scale integrated circuits, because of its low resistivity and fabrication feasibility. In miniaturized device scales, however, the increasing trend in the current density makes it difficult to continue to use Al-based metal lines, because Al alloys are susceptible to electromigration failures. Other metals which have lower resistivity and higher electromigration endurance than Al alloys are necessary for future devices. Cu has been studied extensively to replace the Al-based metal lines for this purpose. Cu is expected to have a better electromigration performance than Al,[1,2] however, the electromigration performance of the Cu metal line depends largely on its fabrication methods [chemical vapor deposition (CVD), plasma vapor deposition (PVD), electroplating, and so on], adhesion layer materials and metal line structures (damascene method or reactive ion etching). The choice of these conditions will affect the reliability of Cu metal lines, therefore, great efforts are necessary to ensure the reliability of Cu interconnects for future devices demanding a high current density.

In the history of improving reliability of Al-based metal lines, various approaches have been performed by a lot of research institutes. Enhancing the preference of the crystallographic orientation has been known as one of the most effective methods to improve the electromigration performance of Al-based metal lines.[3–7] Vaidya *et al.* suggested that median time to failure (MTF) of an Al metal line is improved in Al metal lines having a large grain size, a small grain-size standard deviation, and a strong ⟨111⟩ orientation. This is the result of single layer Al interconnects, and a lot of efforts have been made to enhance the Al texture by changing the underlying materials in layered interconnects. We have reported that a highly ⟨111⟩-oriented Al metal line formed by controlling the underlying TiN preferred orientation has an excellent electromigration performance.[8] Other researchers also show the crystallographic effects of undermetal planes on Al⟨111⟩ orientation in Al/Ti and Al/TiN/Ti layered structures changing the pretreatment and Ti thickness.[5] Those studies use the crystallographic epitaxial relations with underlayer metals.

This method is also effective in Cu interconnects. It has been shown that the Cu⟨111⟩ crystallographic orientation is enhanced on a TiN film having a strong ⟨111⟩ orientation, similar to that of Al interconnects.[9] Electromigration performance has been improved by the realization of damascene Cu interconnects with a strong Cu⟨111⟩ orientation.[10]

In this article, Cu crystallographic orientation control by underlayer materials has been studied in detail. First, it will be shown that Cu crystallographic orientation can be controlled easily by choosing appropriate underlayer materials. Here, the epitaxial growth of the Cu(111) plane occurs on the TiN plane which has extremely ⟨111⟩ preferred orientation. Second, the plane rotation around the ⟨111⟩ axis in the Cu(111) plane on TiN(111) plane, and the TiN(111) plane on the Cu(111) plane will be studied intensively. It will be shown that the different rotation results were obtained according to the layered structures. The rotation mechanism will be explained by superlattice mismatch consideration.

II. EXPERIMENT

TiN (50 nm), TiN (40 nm)/Ti (10 nm) layered, and rapid thermal nitrided-TiN (50 nm) (RTN-TiN) films were studied as underlayers for Cu film. Each underlayer was deposited on a SiO_2/Si substrate by direct current (dc) magnetron sputtering systems with a base pressure less than 10^{-7} Torr. The Ti film was made by sputtering using a Ti target in conditions of a dc power of 1 kW and a total pressure of 3 mTorr in Ar gas. The TiN film was deposited by reactive sputtering

[a]Electronic mail: abe749@oki.co.jp
[b]Present address: LSI Production Division, Oki Electric Industry Co., Ltd., 550-1 Higashi-asakawa, Hachioji, Tokyo 193-8550, Japan.

using a Ti target in conditions of a dc power of 5 kW and a total pressure of 7 mTorr in N_2 gas. The substrate temperature during deposition is 100 °C both for the Ti and TiN films. The distance between target and substrate is 45 mm. These conditions resulted in the same deposition rates of approximately 76 nm/min for both the films. In the TiN/Ti layered structure, the Ti and reactive sputtered TiN films were deposited successively in the same vacuum. The RTN-TiN film was formed by rapid thermal annealing of a sputter-deposited Ti film at 760 °C for 30 s in N_2 ambient. After these underlayers were exposed to air, Cu films (390 nm) were deposited by collimated sputtering (the aspect ratio of the collimator is 1) without substrate heating in Ar gas. The dc power and the total pressure for Cu deposition were 5 kW and 2 mTorr, respectively. The deposition rate was approximately 110 nm/min. The samples were exposed to air again and then were annealed at 450 °C for 15 min in a vacuum of 10^{-6} Torr range. The crystallographic textures of the underlayers and Cu films were evaluated by x-ray diffraction (XRD), and the Cu/RTN-TiN interface was observed using cross-sectional transmission electron microscopy (TEM). In addition, TiN/Cu/RTN-TiN and Cu/TiN/Cu/RTN-TiN layered structures were fabricated. Here, Cu (390 nm)/RTN-TiN (50 nm) layered films were formed on SiO_2/Si substrates, and annealed at 450 °C for 30 min in a vacuum to enhance the Cu grains. A TiN film (100 nm) were then deposited on the samples after breaking the vacuum. The second Cu film (390 nm) was deposited again on the TiN/Cu/RTN-TiN layered structure and annealed at 450 °C for 30 min. The microstructures of the Cu and TiN films were observed using TEM, and the orientational plane rotations between the first Cu(111) and TiN(111) planes, and the second Cu(111) and TiN(111) planes were also evaluated by using selected-area electron diffraction (SAED).

III. RESULTS AND DISCUSSION

A. Cu epitaxial growth on TiN underlayer

XRD spectra of three underlayers deposited on SiO_2/Si substrates are shown in Fig. 1. In the figure, the results of (a) RTN-TiN (50 nm), (b) TiN (40 nm)/Ti (10 nm), and (c) TiN (50 nm) films are shown. TiN(111) and TiN(200) peaks are observed mainly from TiN film in each sample. The values in the parentheses after TiN(111) notation in the figures show TiN(111) peak intensities. TiN(111) peak intensity is largely different due to the changing film formation method or film structure. RTN-TiN film provides a remarkably strong TiN(111) peak intensity about two orders higher than TiN film. The TiN(111)/TiN(200) peak intensity ratios calculated from the XRD results are 33.06 for RTN-TiN, 7.14 for TiN/Ti, and 0.77 for TiN films. Considering the ratio from powdered TiN (0.75: referred from JCPDS card), the TiN film has a random orientation whereas TiN/Ti layered film and RTN-TiN film have a ⟨111⟩ preferred orientation.

Cu (390 nm) film was deposited on underlayers described above and studied by XRD. XRD spectra show only Cu(111) and Cu(200) peaks. Cu(111) peak intensity depends strongly on TiN(111) peak intensities. In Table I, XRD results of

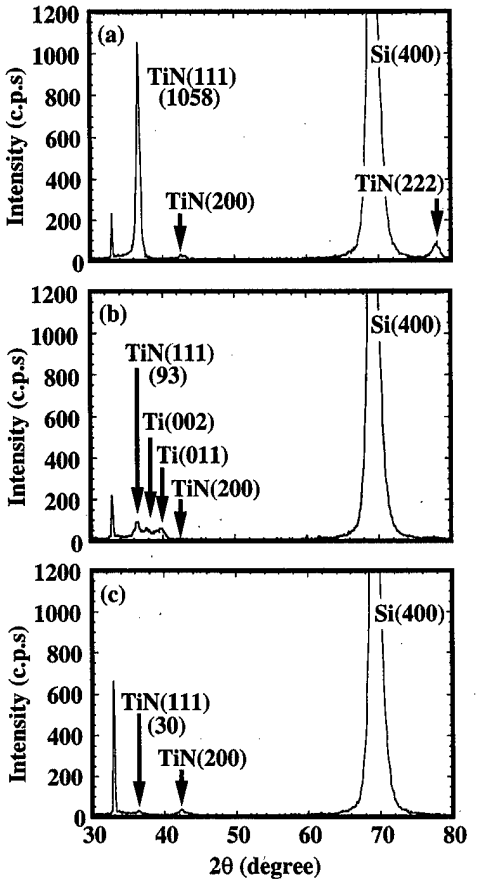


FIG. 1. XRD spectra of various underlayers deposited on SiO_2/Si substrates: (a) RTN-TiN (50 nm); (b) TiN (40 nm)/Ti (10 nm); and (c) TiN (50 nm).

as-deposited and 450 °C annealed Cu films deposited on TiN-based underlayers are summarized. In as-deposited Cu film, the peak ratio of $I(111)/I(200)$ on a RTN-TiN underlayer with a strong TiN(111) peak intensity is about 60 times larger than that on a TiN underlayer with weak TiN(111) peak intensity. The Cu⟨111⟩ crystallographic orientation is enhanced on TiN film having a strong ⟨111⟩ orientation. Furthermore, Table I indicates that 450 °C annealing improves the Cu⟨111⟩ texture except for the case of the TiN underlayer.

Cu(111) rocking curves of Cu films deposited on various underlayers are shown in Fig. 2. The full width at half maximum (FWHM) value indicates the level of texture distribu-

TABLE I. XRD results of as-deposited and 450 °C annealed Cu films deposited on TiN based underlayers.

Samples		XRD peak intensity (cps) $I(111)$	$I(200)$	Peak ratio of $I(111)/I(200)$
Cu/RTN-TiN	as depo (RT)	18 157	17	1160.0
	450 °C 15 min	62 750	20	3137.5
Cu/TiN/Ti	as depo (RT)	6 320	336	19.6
	450 °C 15 min	15 206	538	28.2
Cu/TiN	as depo (RT)	4 356	227	19.5
	450 °C 15 min	7 482	719	10.4

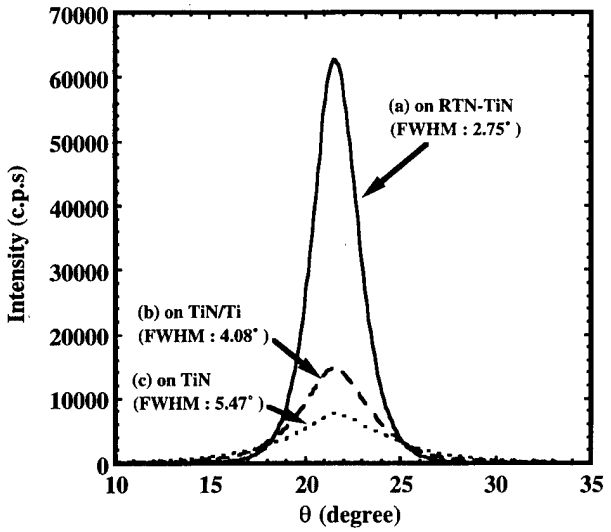


FIG. 2. Cu(111) rocking curves of Cu films deposited on (a) RTN-TiN (50 nm), (b) TiN (40 nm)/Ti (10 nm), (c) TiN (50 nm) films after annealing at 450 °C for 15 min.

tion around the Cu⟨111⟩ axis, that is, the smaller value means tighter texture distribution. The results indicate that the distribution of the Cu⟨111⟩ texture is improved by using highly textured TiN film as an underlayer. Cu preferred orientation can be controlled by choosing an appropriate underlayer.

In order to observe the interface of Cu/TiN layered structure, a cross-sectional TEM image was obtained from a Cu/RTN-TiN layered structure, which has the most ⟨111⟩ preferred orientation (Fig. 3). The TiN grains are much smaller than the Cu grains (average TiN grain size: 10 nm, average Cu grain size: 840 nm), however, the crystal continuity is clearly observed at some parts of a Cu/RTN-TiN interface. Each lattice-fringe spacing of the Cu and TiN layers in the TEM image corresponds to the plane distances of Cu(111) and TiN(111) planes (Cu: 2.008 Å, TiN: 2.44 Å). It can be said that the Cu(111) plane grows epitaxially on the TiN(111) plane.

Here, the lattice mismatch between TiN and Cu should be considered. Cu has a face-centered-cubic (fcc) structure with a lattice constant of 3.615 Å, while the TiN crystal has the

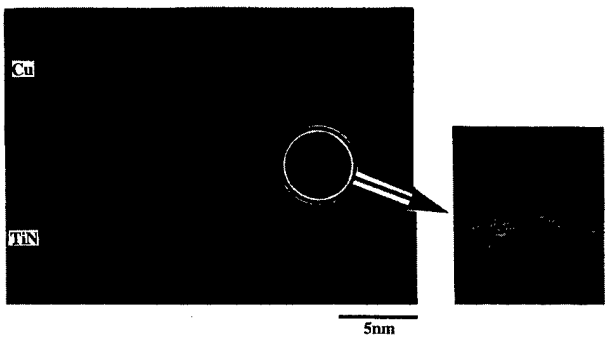


FIG. 3. Cross-sectional TEM image obtained from a Cu/RTN-TiN layered structure after annealing at 450 °C for 15 min.

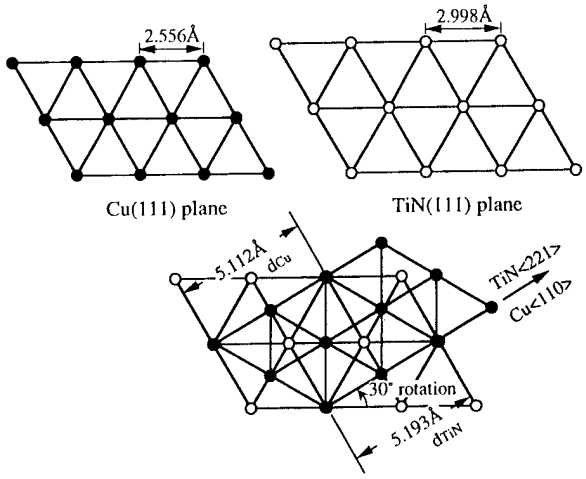


FIG. 4. Schematic diagram showing a Cu(111) plane rotated 30° counterclockwise around the ⟨111⟩ axis on TiN(111) plane.

same cubic NaCl type structure with a lattice constant of 4.24 Å. Therefore, the lattice mismatch between Cu(111) and TiN(111) planes is about 14.7%. In spite of the large lattice mismatch, Cu can be grown epitaxially on TiN film. If it happens, the lattice mismatch seems to be relaxed by the defects formed at the Cu/TiN interface as shown in the circle in Fig. 3.

B. Cu(111) plane rotation around ⟨111⟩ axis

Another consideration, however, is possible to explain the epitaxial growth of Cu deposited on TiN in spite of the large lattice mismatch. Figure 4 shows a schematic diagram showing the Cu(111) plane rotated 30° counterclockwise around the ⟨111⟩ axis on TiN(111) plane. On the hypothesis that the Cu film grows on the TiN film with the Cu(111) plane rotating 30° around the ⟨111⟩ axis on the TiN(111) plane as shown in Fig. 4, two Cu atomic periods along the ⟨110⟩ direction nearly correspond to the TiN atomic periodicity along the ⟨221⟩ direction: the second nearest atomic distance of 5.112 Å along the ⟨110⟩ direction in the Cu(111) plane matches to the nearest atomic distance of 5.193 Å along the ⟨221⟩ direction in the TiN(111) plane with an effective lattice mismatch of about 1.6%. Therefore, there is a possibility that Cu(111) plane grows aligned on the TiN(111) plane with a rotation angle of 30°. It will provide the energy stable position.

In order to clarify the rotational relationship between Cu(111) and TiN(111) planes around the ⟨111⟩ axis, TiN/Cu/RTN-TiN and Cu/TiN/Cu/RTN-TiN layered structures were evaluated by XRD measurement and TEM observation. Underlayer RTN-TiN was used to enhance the preference of overlying Cu and TiN. Figure 5 shows XRD spectra for (a) the TiN (100 nm)/Cu (390 nm)/RTN-TiN (50 nm)/SiO_2 layered structure and (b) the TiN (100 nm)/Cu (390 nm)/SiO_2 layered structure. The values in the parentheses after Cu(111) and TiN(111) notation in the figures show Cu(111) and TiN(111) peak intensities. The (111) peak intensity of Cu film with the RTN-TiN underlayer (a) is eight times

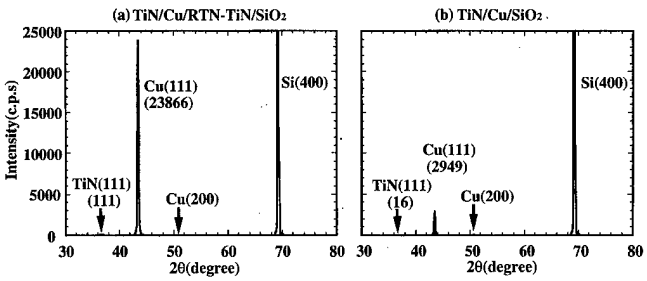


FIG. 5. XRD spectra for (a) the TiN (100 nm)/Cu (390 nm)/SiO_2 layered structure and (b) the TiN (100 nm)/Cu (390 nm)/RTN-TiN (50 nm)/SiO_2 layered structure.

higher than that without the RTN-TiN underlayer (b). The difference of Cu texture also affects the overlying TiN film texture. By depositing TiN film on a Cu/RTN-TiN layered structure with a strong Cu⟨111⟩ orientation, a strong ⟨111⟩-oriented overlying TiN film was obtained. Figure 6(a) is a plan-view TEM image showing typical grains for TiN film deposited on a Cu/RTN-TiN layered structure. The TiN film has grains of about 10 nm in size. A lot of small TiN grains exist on a single Cu grain having median grain size of approximately 0.9 μm. Figure 6(b) is a SAED pattern obtained from the TiN/Cu layered structure. The probe beam diameter is about 0.4 μm. The SAED pattern consists of two types of diffraction spots from the overlying TiN and Cu films, since the sample was thinned down from the sample bottom, and the overlying TiN and Cu remained to be observed. The external spots coincide well with those for the ⟨111⟩ zone axis of cubic Cu from one grain, while the internal arc-like spots are from the ⟨111⟩ zone axis of cubic TiN, since the spots from TiN include crystallographic information from a lot of TiN grains. The directions of the internal spots and the external spots coincide well within the rotation angle distribution of $\pm 10°$ around the ⟨111⟩ axis. Therefore, in TiN film deposited on Cu film, it is considered that little rotation of the TiN(111) plane occurs on the Cu(111) plane within $\pm 10°$ distribution around the ⟨111⟩ axis.

Next, Cu film deposited on TiN film has been studied. Figure 7 shows XRD spectrum for a Cu (390 nm)/TiN (100nm)/Cu (390 nm)/RTN-TiN (50 nm)/SiO_2 layered structure. Figure 7 was obtained with a same condition of XRD measurement as Fig. 5. The Cu(111) peak intensity is almost double that of Fig. 5(a). This result indicates that the second Cu layer also grows epitaxially on a TiN/Cu layered structure, and the preferred orientation of the second Cu layer is also enhanced. Figure 8 is a plan-view TEM image showing typical grains for the Cu film deposited on a TiN/Cu/RTN-TiN layered structure. Two featured regions in the image can be seen: a region which has large Cu grains about 2 μm in size and a region which has small Cu grains about 0.2 μm in size. Small Cu grains always exist as colonies in a total size of approximately 2 μm. In a colony, small Cu grains have relatively close orientation and form small angle tilted grain boundaries as described in Fig. 9. This observation indicates that a small Cu grain colony could form one large Cu grain if some more energetic stimulation occurs.

Figure 9 shows SAED patterns from a sample prepared from a Cu (390 nm)/TiN (100 nm)/Cu (390 nm)/RTN-TiN/SiO_2 structure; a large Cu grain region (a) and a small Cu grain region (b), (c). The SAED pattern from large Cu grain region (a) consists of two types of diffraction spots from Cu and underlying TiN films as shown in Figs. 6(a) and 6(b). The external spots are those for the ⟨111⟩ zone axis of cubic Cu from one grain. The internal arc-like spots are from the ⟨111⟩ zone axis of cubic TiN. The difference from Fig. 6(b) is that the Cu{110} spots have a rotational angle of 25°–30° compared with the arc-like TiN{110} spots around the ⟨111⟩ axis. On the other hand, the SAED patterns from the small Cu grain region [Figs. 9(b) and 9(c)] are different. Figure 9(b) shows the SAED pattern from the small Cu grain region without spots from underlying TiN. The underlying TiN film was removed through the sample preparing process. Rela-

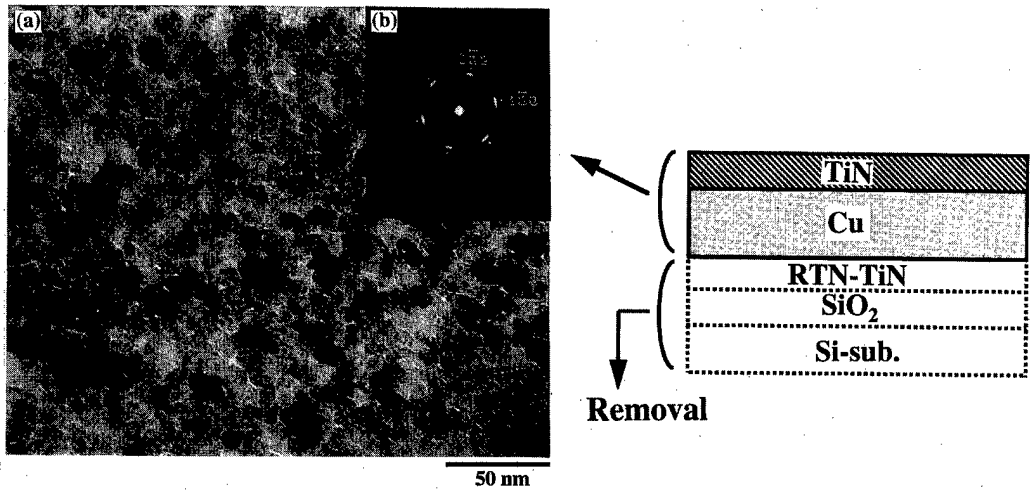


FIG. 6. Plan-view TEM image showing typical grains for the TiN film deposited on (a) the Cu/RTN-TiN layered structure, and a SAED pattern obtained from (b) the TiN/Cu layered structure.

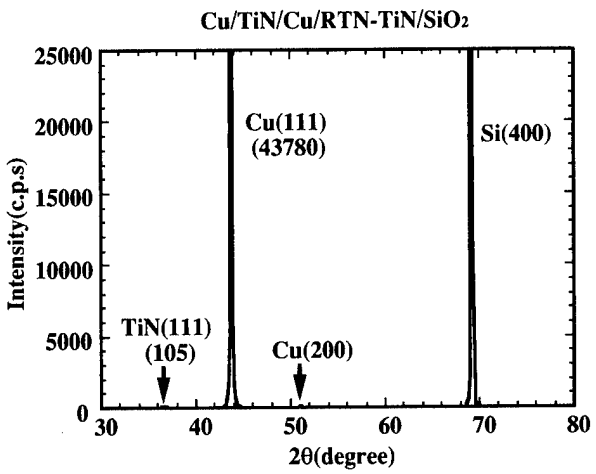


FIG. 7. XRD spectrum for the Cu (390 nm)/TiN (100 nm)/Cu (390 nm)/RTN-TiN (50 nm)/SiO_2 layered structure.



FIG. 8. Plan-view TEM image showing typical grains for the Cu film deposited on the TiN/Cu/RTN-TiN layered structure.

tively coherent ⟨111⟩ preferred orientation with a little rotational distribution less than 10° can be seen. Figure 9(c) is the SAED pattern from the small Cu grain region with spots from underlying TiN. In this case, both the internal spots from TiN and the external spots from Cu have rotational angle distributions less than 10°, and the directions of the internal spots and the external spots coincide well.

The plane rotational results will be summarized here. In TiN film deposited on Cu film, little rotation of the TiN(111) plane occurs on the Cu(111) plane within ±10° distribution around the ⟨111⟩ axis. On the other hand, in Cu film deposited on TiN film, there are two rotational angles: a rotational angle of 25°–30° in the large Cu grain region and little rotation with rotational angle distributions less than 10° in the small Cu grain region. In both cases, it is believed that epitaxial growth occurs.

The difference in plane rotation results according to layered structure and grain size can be explained as follows. At the initial stage of Cu deposition, ⟨111⟩-oriented Cu nucleates on ⟨111⟩-oriented TiN small grains, and the deposited Cu is aligned to the underlayer TiN orientation. As Cu film grows thicker during deposition, grain coalescence occurs and Cu grains become larger keeping highly ⟨111⟩ preferred orientation on the TiN film with a strong ⟨111⟩ orientation. It seems that there is little plane rotation at the initial stage of Cu deposition, however, in the course of further Cu deposition, coalescence of Cu grains induces the plane rotation to a more energy stable position where the effective lattice mismatch is 1.6% as shown in Fig. 4. On the other hand, in TiN film deposited on Cu film, TiN seems to be more persistent to the underlayer crystallographic orientation.

IV. CONCLUSION

The effect of underlayer texture on Cu film orientation has been studied. It was found that Cu texture correlates well with the underlayer texture. The Cu⟨111⟩ crystallographic orientation is enhanced on TiN film with a strong TiN⟨111⟩ orientation. Cross-sectional TEM observation showed the Cu(111) plane grows epitaxially on the TiN(111) plane. The atomic arrangement between Cu(111) and TiN(111) planes in a TiN/Cu/RTN-TiN layered structure has a rotational angle within ±10° around the ⟨111⟩ axis. On the other hand, in Cu film deposited on TiN film, there are two rotational angles: a rotational angle of 25°–30° in the large Cu grain region and little rotation with rotational angle distributions less than 10° in the small Cu grain region. These results can be explained by consideration of superlattice mismatch. On

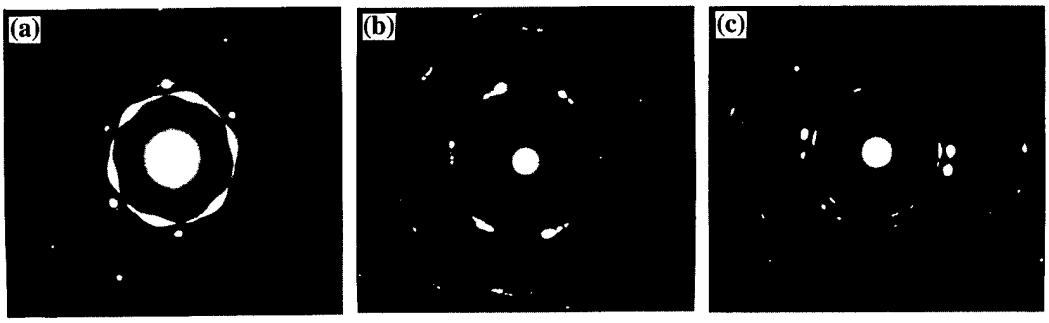


FIG. 9. Selected area electron diffraction patterns from a sample prepared from Cu (390 nm)/TiN (100 nm)/Cu (390 nm)/SiO_2 structure; (a) large Cu grain area and (b), (c) small Cu grain area.

the hypothesis that the Cu film grows on the TiN film with the Cu(111) plane rotating 30° around the ⟨111⟩ axis on the TiN(111) plane, two Cu atomic periods along the ⟨110⟩ direction nearly correspond to the TiN atomic periodicity along the ⟨221⟩ direction: the second nearest atomic distance of 5.112 Å along the ⟨110⟩ direction in the Cu(111) plane matches to the nearest atomic distance of 5.193 Å along the ⟨221⟩ direction in the TiN(111) plane with an effective lattice mismatch of about 1.6%. The large lattice mismatch of 14.7% between Cu and TiN reduces to 1.6% by above consideration. Cu(111) plane rotation seems to occur so as to have a more energy stable position at coalescence stage of deposition.

[1]H-K. Kang, J. S. H. Cho, I. Asano, and S. S. Wong, Proceedings of the VLSI Multilevel Interconnection Conference, Santa Clara, 1992, pp. 337–343.

[2]J. Tao, N. W. Cheung, and C. Hu, IEEE Electron Device Lett. **14**, 249 (1993).

[3]S. Vaidya and A. K. Sinha, Thin Solid Films **75**, 253 (1981).

[4]K. Hinode and Y. Homma, Proceedings of the 28th IEEE International Reliability Physics Symposium, Las Vegas, 1990, pp. 25–30.

[5]H. Shibata, M. Murota, and K. Hashimoto, Jpn. J. Appl. Phys., Part 1 **32**, 4479 (1993).

[6]H. Toyoda, T. Kawanoue, M. Hasunuma, H. Kaneko, and M. Miyauchi, Proceedings of the 32nd IEEE International Reliability Physics Symposium, San Jose, 1994, pp. 178–184.

[7]H. Onoda, T. Narita, K. Touchi, and K. Hashimoto, Proceedings of the 34th IEEE International Reliability Physics Symposium, Dallas, 1996, pp. 139–147.

[8]M. Kageyama, K. Hashimoto, and H. Onoda, Proceedings of the 29th IEEE International Reliability Physics Symposium, Las Vegas, 1991, pp. 97–102.

[9]K. Abe, Y. Harada, and H. Onoda, Appl. Phys. Lett. **71**, 2782 (1997).

[10]K. Abe, Y. Harada, and H. Onoda, Proceedings of the 36th IEEE International Reliability Physics Symposium, Reno, 1998, pp. 342–347.

Barrier properties of Ta–RuO_2 diffusion barrier for dynamic random access memory capacitor bottom electrodes

Dong-Soo Yoon and Hong Koo Baik[a)]
Department of Metallurgical Engineering, Yonsei University, Seoul 120-749, Korea

Sung-Man Lee
Department of Materials Engineering, Kangwon National University, Chuncheon, Kangwon-Do 200-701, Korea

Sang-In Lee
Development of Technology, Semiconductor R&D Center, Samsung Electronics, Co. San#24, Nongseo-Lee, Kiheung-Eup, Yongin-City, Kyungki-Do 449-900, Korea

(Received 22 December 1998; accepted 8 May 1999)

We proposed the Ta–RuO_2 diffusion barrier for oxygen in the dynamic random access memory capacitor bottom electrode, and investigated the barrier and electrical properties of the developed diffusion barrier. The Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system deposited with and without the SiO_2 capping layer showed the lower total resistance and ohmic characteristics up to 800 °C. For the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system, no other phases observed except for the formation of conductive RuO_2 phase in the barrier film by reaction with the indiffused oxygen after annealing in air, but the thin oxidized layer at the Ta–RuO_2/$TiSi_2$ interface was formed by external oxygen. However, a large number of the crystallites in the annealed samples compared to that of as-deposited film were observed even after depth profile. The crystallites consisted of Ru and O containing a small amount of Ta. In addition, the embedded RuO_2 crystalline phase was observed in the thin oxidized $TiSi_2$ surface layer. Correspondingly, we suggest that the ohmic mechanism of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system is an embedded RuO_2 crystalline phase involving a small amount of Ta in a Ta amorphous structure. © *1999 American Vacuum Society.* [S0734-211X(99)06004-7]

I. INTRODUCTION

Perovskite oxides-lead zirconium titanate ($PbZr_xTi_{1-x}O_3$)(PZT),[1] strontium bismuth tantalum oxide ($SrBi_2Ta_2O_9$) (SBT),[2] and barium strontium titanate ($Ba_{1-x}Sr_xTiO_3$) (BST)[3] have been studied for dynamic random access memory (DRAM)/ferroelectric random access memory (FRAM) devices because of high dielectric constant, good paraelectric and ferroelectric properties. In particular, BST, as a dielectric thin film for DRAM capacitor, is required because the data storage in the DRAM use the paraelectric properties. The high dielectric oxides during the DRAM/FRAM capacitor fabrication processes usually require high temperature process and oxidizing ambient to improve the chemical and electrical properties. Platinum (Pt)[4] is widely used as a bottom electrode in the high dielectric thin film capacitors. Under these conditions, the oxygen rapidly diffuses at more than 600 °C through grain boundaries of a polycrystalline Pt film and thus oxidizes the underlying layer, such as the ohmic contact layer and poly-Si plug. This results in the contact resistance degradation of capacitor layer structure. Therefore, a diffusion barrier for oxygen is needed and should be developed for high density DRAMs beyond 64 MDRAMs. For COB memory cell structure, both bottom electrode and diffusion barrier are patterned before BST deposition. This implies that both sides of a diffusion barrier directly contact with the BST film and thus a diffusion barrier requires high oxidation resistance during deposition or postannealing in an oxidizing ambient. Therefore, a diffusion barrier for DRAM capacitor bottom electrode should not only prevent the oxygen indiffusion up to high temperatures but also require the high oxidation resistance.

Currently, many efforts have been performed to apply an oxygen diffusion barrier for capacitor bottom electrodes, including nitrides and amorphous ternary compound barriers developed for Al or Cu metallization.[5–9] These are oxidized under BST integration process condition. However, RuO_2 or CeO_2-incorporated Ta barrier has been investigated as a diffusion barrier for Cu and Al metallization, and dynamic random access memory and ferroelectric random access memory capacitor bottom electrode, and reported to perform good diffusion barrier properties for both purposes.[10–17]

In this article, we focused on the development of new diffusion barrier for oxygen, and characterized the barrier and electrical properties of the developed diffusion barrier. In order to develop the diffusion barrier for oxygen, first of all, two properties of oxygen should be considered. Oxygen not only reacts easily to oxidize the diffusion barrier, but also diffuses rapidly through diffusion barrier and then oxidizes the underlying layer. That is, oxygen for most elements in the periodic table has good reactivity as well as fast diffusivity. In order to retard fast diffusion of oxygen at more than 650 °C, first, careful control of the microstructural characteristics of diffusion barrier itself is required. The microstruc-

a)Electronic mail: thinfilm@bubble.yonsei.ac.kr

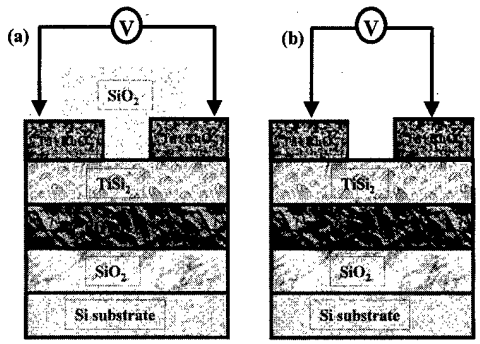


FIG. 1. Cross section of samples for $I-V$ measurement (a) with (b) without SiO_2 deposition.

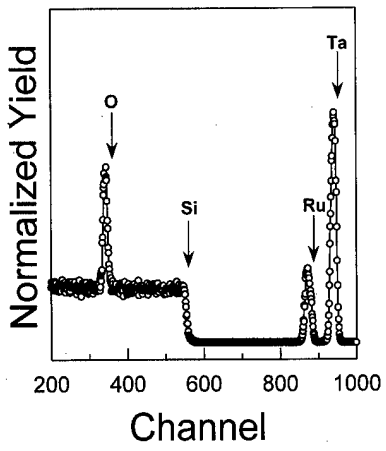


FIG. 2. Rutherford backscattering spectrum of Ta–RuO_2 film prepared with 100 W rf power.

ture of barrier itself must be amorphized by the incorporation of a material in a matrix barrier layer and thus it can prevent the indiffusion of oxygen because of no grain boundaries, which can act fast diffusion path of oxygen. On the other hand, to prevent the oxidation of diffusion barrier itself, the binding force formed between matrix metal and the added material should be chemically strong in the as-deposited state. To satisfy these requirements, the diffusion barrier needs the added material into the matrix metal. The added material should be conductive oxide because the electrical properties should remain intact after annealing in oxygen atmosphere. In this work, ruthenium dioxide (RuO_2), as an added material into the Ta matrix metal, is chosen because of electrically conductive oxide. The incorporation of RuO_2 into a Ta barrier layer will amorphize the microstructure of barrier itself and form the strong bonds with Ta–O or Ta–Ru–O. Thus, it is expected to increase the oxidation resistance and prevent the indiffusion of oxygen up to high temperatures.

II. EXPERIMENT

The Ta–RuO_2 films were deposited on the $TiSi_2$/poly-Si/SiO_2/Si(100) substrate in a radio frequency magnetron sputtering system with a base pressure of 2.0×10^{-6} Torr. The substrates were ultrasonically cleaned in T.C.E., acetone, alcohol and rinsed in deionized water, and then dried with nitrogen. RuO_2 mosaic targets were prepared and rf power of Ta–RuO_2 target was 100 W. Samples were annealed between 650 and 800 °C in air for 30 min.

Figure 1 showed schematically the cross section of samples for $I-V$ measurement. In order to investigate the ohmic behavior for the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact structure with and without the SiO_2 capping layer, the $I-V$ measurement was carried out between the diffusion barriers. All diffusion barriers were patterned during deposition using a stainless steel shadow mask and patterns were 2×2 mm in size. The area between barriers was deposited by the SiO_2 capping layer for preventing the oxidation of $TiSi_2$ channel layer.

Total resistance and the $I-V$ curve were measured by HP 4145B semiconductor parameter analyzer. The thickness and composition of barrier film were evaluated by Rutherford backscattering spectroscopy (RBS). Auger electron microscopy (AES) was used for depth profiling of the layer structures. The microstructures, crystal structures, and phase identification were analyzed by x-ray diffraction (XRD) and transmission electron microscopy (TEM).

III. RESULTS AND DISCUSSION

The thickness and composition ratio of the deposited Ta–RuO_2 (100 W) barrier film was analyzed by RBS and oxygen resonance (3.06 MeV, H^{++} ion) was used to obtain the amount of oxygen in the Ta–RuO_2 film, and the result was shown in Fig. 2. From the rump simulation, the thickness and composition ratio (Ta:RuO_2) was found to be 100 nm and 50:50, respectively.

The total resistance for Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system prepared with RuO_2 addition was measured at $V=0$ as a function of post-annealing temperatures in air atmosphere for 30 min. The result was shown in Fig. 3. When the Ta–RuO_2 film was prepared with 100 W rf power, the total resistance showed the higher value in the as-deposited state compared to the annealed samples, as shown in Fig. 3. After annealing at 650 °C, the total resistance decreased initially and gradually increased up to 800 °C. Since good electrical conductivity between bottom electrode, diffusion barrier, poly-Si plug, and transistor is responsible for the capacitor performance, it is thought that the lower total resistance of the annealed sample is desirable feature for the DRAM capacitor bottom electrode.

In order to investigate the ohmic characteristics, $I-V$ measurement between -4 and 4 V was carried out for Ta–RuO_2 (100 W)/$TiSi_2$/poly-Si/SiO_2/Si contact system before and after annealing at various temperatures. The result was shown in Fig. 4. The $I-V$ curve of the Ta–RuO_2 (100 W)/$TiSi_2$/poly-Si/SiO_2/Si contact system showed the nonohmic behavior in the as-deposited state. This result should be thought for the Ta–RuO_2/$TiSi_2$ and $TiSi_2$/poly-Si interfaces, respectively. First, the formation of the ohmic contact layer ($TiSi_2$) was grown by rapid thermal annealing of the Ti/poly-Si/SiO_2/Si sample in nitrogen ambient and

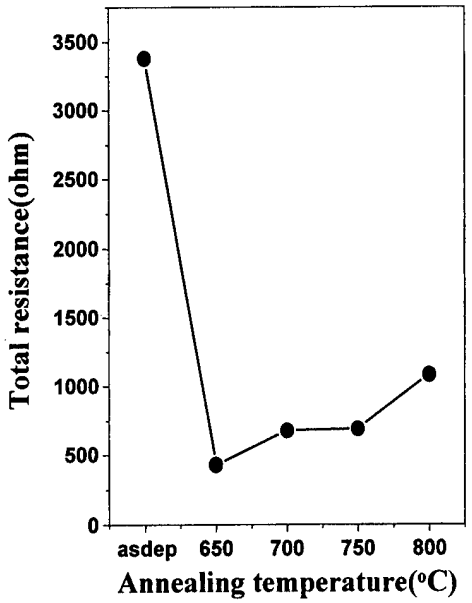


FIG. 3. Total resistance for Ta layer with RuO_2 addition depending on various post-annealing temperatures in air atmosphere.

thus the $TiSi_2$/poly-Si interface does not affect the ohmic behavior. The surface of the $TiSi_2$ ohmic contact layer is thinly oxidized by oxygen ions in the plasma during the deposition of the Ta–RuO_2 film. It is thought that the nonohmic behavior in the as-deposited state is attributed to the formation of thin oxidized layer at the interface between Ta–RuO_2 barrier film and the $TiSi_2$ ohmic layer. In the case of the Ta film prepared without RuO_2 addition, the $I-V$ curve of the Ta/$TiSi_2$/poly-Si/SiO_2/Si contact system showed the linear characteristics in the as-deposited state. This result is evidence for nonohmic behavior of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system in the as-deposited state because of the formation of thin oxide layer at the $TiSi_2$ surface. In addition, the thin amorphous layer at

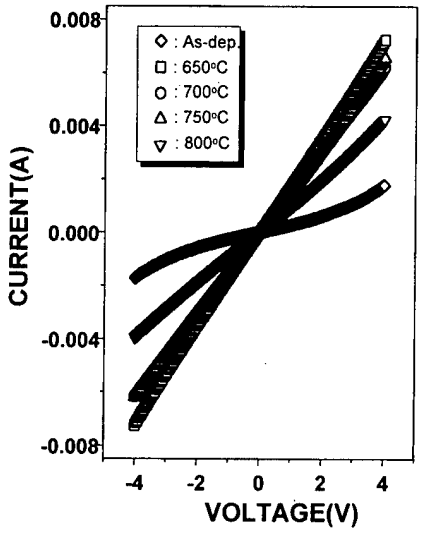


FIG. 4. $I-V$ curve for Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system depending on various post-annealing temperatures in air atmosphere.

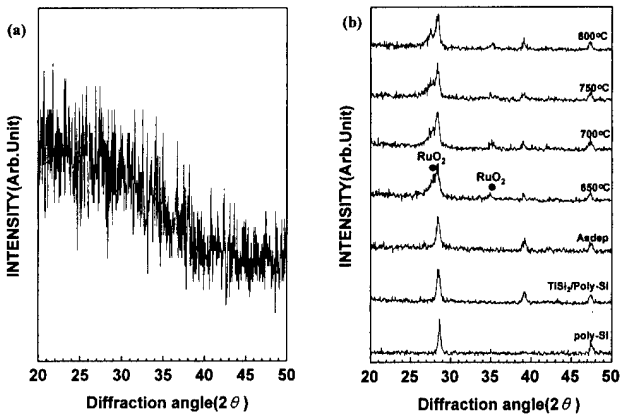


FIG. 5. XRD patterns of (a) Ta–RuO_2/Si contact system in the as-deposited state (b) Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system after annealing at various temperatures for 30 min in air atmosphere.

the Ta–CeO_2/$TiSi_2$ interface for the Pt/Ta–CeO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system before annealing was observed by cross-sectional TEM. The Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system showed the perfect ohmic behavior with increasing annealing temperature. Although the total resistance slightly increased, in particular, the $I-V$ curve showed the linear characteristics even after annealing at 800 °C as shown in Fig. 4.

In order to investigate the reaction products for Ta–RuO_2 (100 W)/$TiSi_2$/poly-Si/SiO_2/Si contact system with annealing temperature, the x-ray diffraction analysis was performed and the result was shown in Fig. 5. First, to find out the microstructure of the Ta–RuO_2 diffusion barrier, the Ta–RuO_2 film was deposited on the Si(100) substrate and the glancing angle XRD (Cu K_α radiation) result was shown in Fig. 5(a). It clearly showed that the microstructure of the Ta–RuO_2 diffusion barrier was amorphized by RuO_2 addition during the Ta deposition in the as-deposited state. It was found that the chemically strong Ta–O bond was formed during the deposition at room temperature and thus it resulted in an amorphous microstructure of the Ta–RuO_2 diffusion barrier because of the surface mobility limitation of adatoms. Although the amount of RuO_2 is sufficient for detecting by XRD from RBS data shown in Fig. 2, RuO_2 peaks were not detected in XRD patterns in the as-deposited state. This is because the crystallite size of RuO_2 is very small. That is, the nanocrystal size of RuO_2 was embedded in a Ta amorphous structure. As shown in Fig. 5(b), an amorphous Ta–RuO_2 diffusion barrier has lower intensity than other elements (Si) and compounds ($TiSi_2$), and thus it did not appear in the XRD pattern in the as-deposited state. When the Ta/$TiSi_2$/poly-Si/SiO_2/Si contact system was deposited with RuO_2 addition, the RuO_2 peaks in the Ta–RuO_2 film annealed at 650 °C were observed and no other phases were observed except for the formation of the RuO_2 phase increased at the annealing temperature. In particular, Ta_2O_5 was not observed even after annealing at 800 °C, as shown in Fig. 5(b). Both the formation of Ta_2O_5 in the Ta–RuO_2 film and oxidation of the $TiSi_2$ layer by reaction with external

oxygen from air after annealing would have increased the total resistance and not appeared the ohmic behavior because the oxidized layer has a dielectric property. This indicates that the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si layer structure was not oxidized by the reaction with the indiffused oxygen from air atmosphere. In the present study, the lower total resistance and linear $I-V$ curve of Ta–RuO_2 films after annealing compared to that of as-deposited film is attributed to the formation of the conductive RuO_2 phase in the barrier film by reaction with oxygen after annealing in air.

From total resistance, the $I-V$ curve and glancing angle XRD results, both barrier properties and oxidation resistance of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system was better than that of various diffusion barriers previously reported by others. For instance, an amorphous $W_{80}N_{20}$ film[8] was oxidized by dry and wet oxidizing ambient in the temperature range of 450–575 °C. In the case of TiN film,[7] although the Pt bottom electrode is deposited and annealing duration is short, the electrical resistivity between Pt, TiN, and the poly-Si substrate at 650 °C is too high. The amorphous, Ta–Si–N barrier,[9] has been reported that the barrier surface is oxidized. The formation of thin oxidized layer, such as Ta_2O_5 and SiO_2 at the surface of the Ta–Si–N film terminates the electrical contact from the bottom electrode to transistor. Therefore, since the various barriers developed for metallization, as an oxygen diffusion barrier for a DRAM capacitor bottom electrode, are oxidized under BST integration processes, it is expected that they could not be applied to high density memory devices. In the present study, the Ta–RuO_2 diffusion barrier prepared with 100 W rf power showed the lower total resistance and ohmic characteristics and its surface oxidation was prevented even after annealing at 800 °C.

In order to investigate the diffusion behavior of the oxygen and other elements, AES depth profile was carried out for the Ta–RuO_2 (100 W)/$TiSi_2$/poly-Si/SiO_2/Si sample at no SiO_2 capped layer before and after annealing at various temperatures and the results were shown in Fig. 6. We first noted that the sputtering depth by the Ar^+ ion was performed at the surface of poly-Si for the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample. The layer structure of the Ta–RuO_2/$TiSi_2$/poly-Si was well defined in the as-deposited state. After annealing at 700 °C, the surface of the $TiSi_2$ ohmic contact layer was oxidized by reaction with the indiffused oxygen from air atmosphere. However, the oxidized thickness of the $TiSi_2$ layer was little changed with the increasing annealing temperature. As shown in Fig. 6, the Si atoms of $TiSi_2$ ohmic layer were oxidized by external oxygen. This is supported by the fact that the Si atom is a dominant moving species during the silicide formation and oxidation.[18] However, the oxygen did not diffuse to the $TiSi_2$/poly-Si interface even after annealing at 800 °C. Note that the amount of Ru atoms in the Ta–RuO_2 film was gradually decreased with increasing annealing temperature.

Figure 7 showed the depth profile of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample at the SiO_2 capped layer after annealing at 700 °C. This result is similar to the data of the

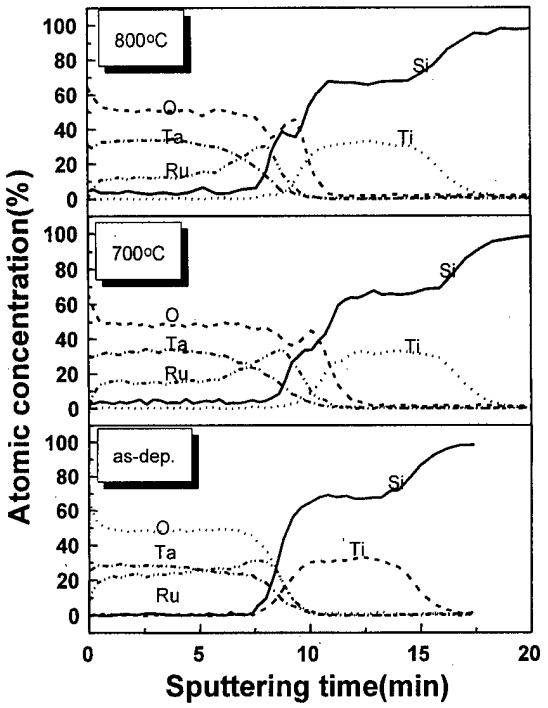


FIG. 6. AES depth profile for Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system at no SiO_2 capped layer depending on various post-annealing temperatures in air.

Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample at no SiO_2 capped layer after annealing at 700 °C, as shown in Fig. 6. This implies that the SiO_2 capping layer does not play a role in the prevention of oxygen diffusion.

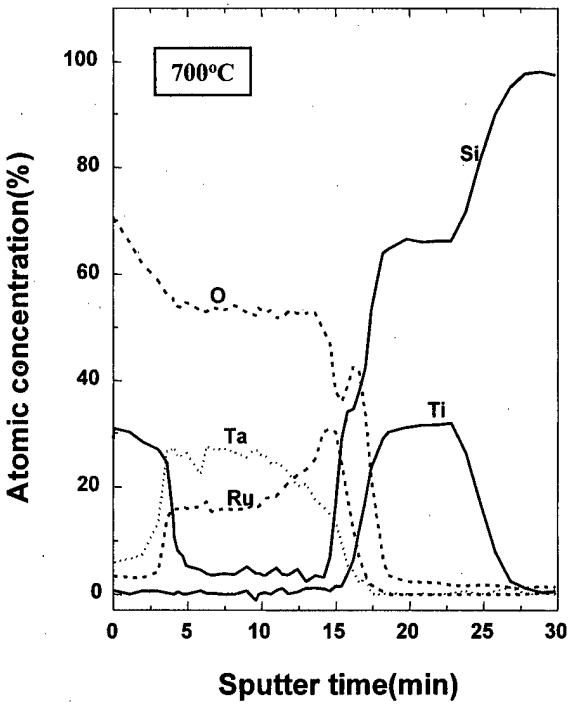


FIG. 7. AES depth profile for Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system at the SiO_2 capped layer after annealing at 700 °C in air.

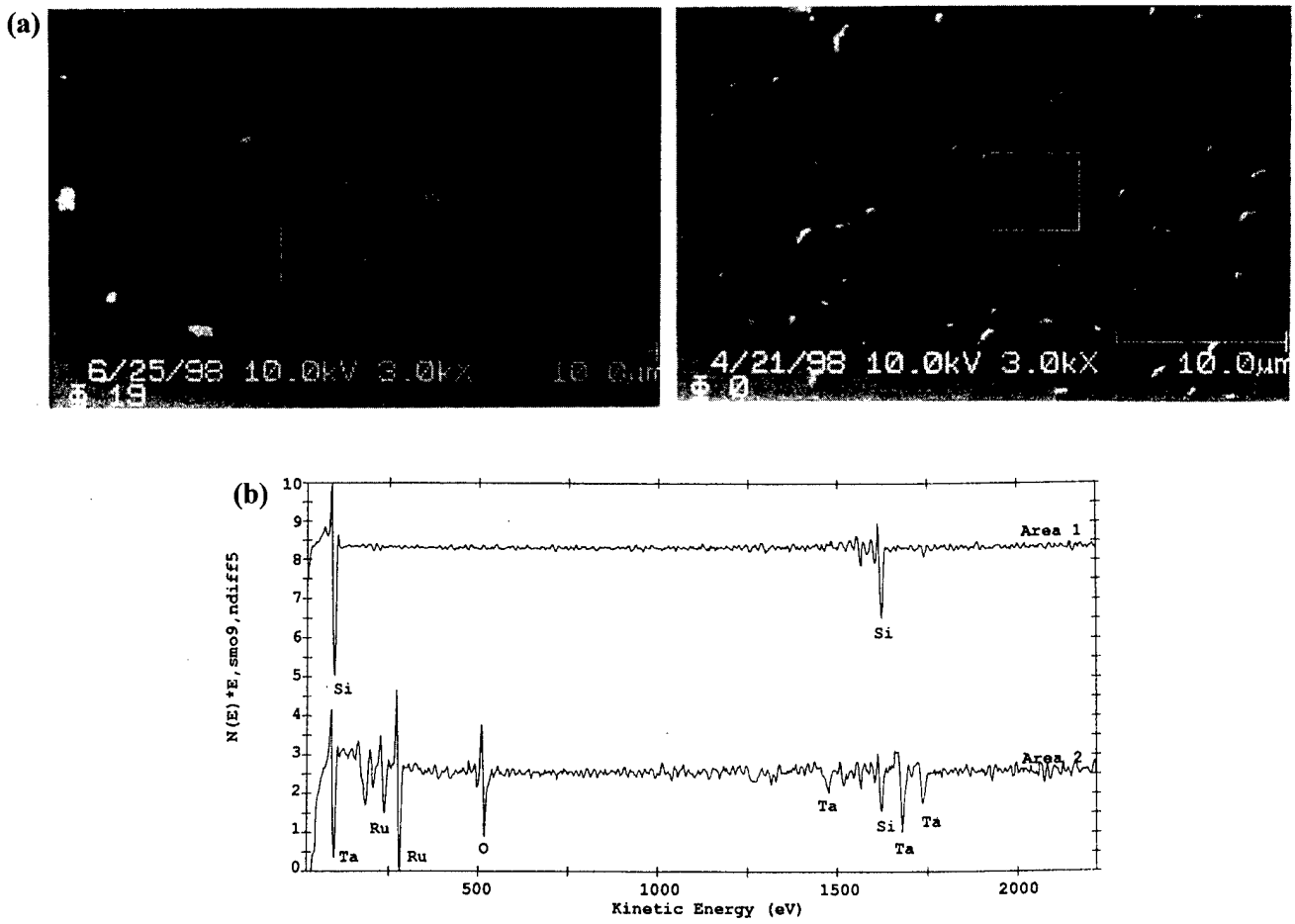


FIG. 8. SAM micrographs and Auger spectrum for Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system at no SiO_2 capped layer before and after annealing at various temperatures in air.

Figures 6 and 7 showed the diffusion behavior of each element for the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system with and without the SiO_2 capped layer part depending on the annealing temperature in air. For the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample, that the oxidized layer at the Ta–RuO_2/$TiSi_2$ interface was observed, irrespective of no SiO_2 and SiO_2 deposition layer. This indicates that the electrical contact between the Ta–RuO_2 barrier layer and the $TiSi_2$ ohmic layer is broken. Therefore, the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system should show the higher total resistance and non-ohmic behavior with increasing annealing temperature. As shown in Figs. 3 and 4, however, the lower total resistance and ohmic characteristics for the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system appears after annealing at various temperatures.

To find out the mechanism of ohmic behavior, the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact systems with different annealing temperatures were analyzed by scanning Auger microscopy (SAM) micrographs and AES spectra, and the results are shown in Fig. 8. In SAM images of Fig. 8, the depth profiles for the annealed samples were obtained from area 1, the crystallite-free zone in Fig. 6. In addition, AES spectra in Fig. 8 were obtained by Auger electron point analysis from area 2, the crystallites. A few of the crystallites were distributed in the as-received state of as-deposited film and remained after depth profile. After annealing at 700 and 800 °C, however, a large number of the crystallites compared to that of as-deposited film were observed even after depth profile. This indicates that the crystallites exist at the surface of poly-Si. However, it is thought that the crystallites do not diffuse to the surface of poly-Si during the annealing because they are observed in the as-deposited state. Therefore, the difference of the sputtering rate by the Ar^+ ion between crystallites and crystallite-free zone during *in situ* ion beam etching would take place. The surface element constituents of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si samples at 700 and 800 °C were characterized by AES. The spectra taken from the crystallites consisted of Ru and O containing a small amount of Ta, but only Si peaks were detected for the crystallite-free zone. Although the surface of $TiSi_2$ ohmic contact layer was oxidized, the ohmic behavior of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system is attributed to an embedded RuO_2 crystallite in the thin oxidized $TiSi_2$ surface layer.

To find out the reaction products of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample after annealing at 800 °C, cross-sectional TEM analysis was carried out and the results were shown in Fig. 9. Figure 9(a) showed the full layer structure at low magnification. The thin oxidized layer at the

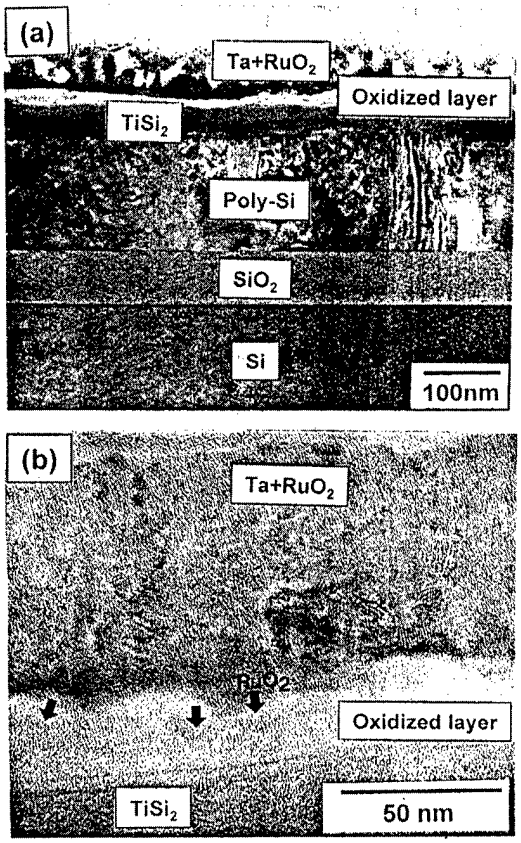


FIG. 9. Cross-sectional TEM micrographs for Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system after annealing at 800 °C in air. (a) Low magnification; (b) high magnification.

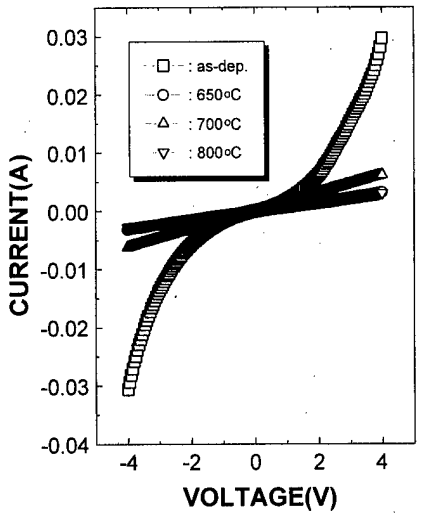


FIG. 10. $I-V$ curve for Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system prepared without the SiO_2 capping layer depending on various post-annealing temperatures in air.

Ta–RuO_2/$TiSi_2$ interface was observed. This result was consistent with AES depth profile data shown in Fig. 6. In order to clearly investigate the Ta–RuO_2/oxidized layer/$TiSi_2$ interfaces, high-resolution TEM was performed and the result was shown in Fig. 9(b). The invaded crystalline phases were observed in the thin oxidized layer. This was found to be RuO_2 containing a small amount of Ta by d spacing the value.

From the $I-V$ curve, AES and TEM results, although the surface of the $TiSi_2$ layer is oxidized by the reaction through indiffused oxygen from air, the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system showed the linear $I-V$ curve. The reason why ohmic behavior is exhibited by the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system is attributed to an embedded RuO_2 crystalline phase involving a small amount of Ta as shown by SAM, AES point analysis, and TEM data. If the invaded RuO_2 crystalline phase is the ohmic mechanism for the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system after annealing at different temperatures, the $I-V$ curve of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample prepared without the SiO_2 protection layer should show the linear behavior. That is, if the $TiSi_2$ channel layer was not fully oxidized by the direct exposure in air and the $TiSi_2$ layer on both sides of the Ta–RuO_2/$TiSi_2$ interface was only partially oxidized by the lateral diffusion of oxygen, and if the invaded RuO_2 crystalline phase, which is the current path, did exist, then the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system should show an ohmic behavior after annealing in air. Figure 10 showed the $I-V$ curve of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample prepared without the SiO_2 capping layer after annealing at different temperatures in air. This result is similar to the data of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si sample prepared with the SiO_2 capping layer shown in Fig. 3. Therefore, we suggest that the ohmic mechanism of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system is an embedded RuO_2 crystalline phase in a Ta amorphous structure.

IV. CONCLUSION

The Ta–RuO_2 diffusion barrier for oxygen in the DRAM capacitor bottom electrode is proposed, and the barrier and electrical properties of the developed diffusion barrier is characterized. The Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system deposited with and without the SiO_2 capping layer showed the lower total resistance and ohmic characteristics up to 800 °C. For the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system, no other phases observed except for the formation of the conductive RuO_2 phase in the barrier film by reaction with the indiffused oxygen after annealing in air, but the thin oxidized layer at the Ta–RuO_2/$TiSi_2$ interface was formed by external oxygen. However, a large number of the crystallites in the annealed samples compared to that of as-deposited film were observed even after depth profile. The crystallites consisted of Ru and O containing a small amount of Ta. In addition, an embedded RuO_2 crystalline phase was observed in the thin oxidized $TiSi_2$ surface layer. Correspondingly, we suggest that the ohmic mechanism of the Ta–RuO_2/$TiSi_2$/poly-Si/SiO_2/Si contact system is an embedded RuO_2 crystalline phase involving a small amount of Ta in a Ta amorphous structure.

[1]H. N. Al-Shareef, A. I. Kingon, X. Chen, and K. R. Bellur, and O. Auciello, J. Mater. Res. **9**, 2968 (1994).
[2]H. N. Al-Shareef, D. Dimos, T. J. Boyle, W. L. Warren, and B. A. Tuttle, Appl. Phys. Lett. **68**, 690 (1996).
[3]D. Roy and S. B. Krupanidhi, Appl. Phys. Lett. **62**, 1056 (1993).
[4]Y. Shimada, A. Inoue, T. Nasu, K. Arita, Y. Nagano, A. Matsuda, Y. Uemoto, E. Fujii, M. Azuma, Y. Oishi, S. Hayashi, and T. Otsuki, Jpn. J. Appl. Phys., Part 1 **35**, 140 (1996).
[5]D.-S. Yoon, B.-S. Kang, S.-M. Lee, and H. K. Baik, J. Appl. Phys. **80**, 6550 (1996).
[6]J. H. Kim, D.-S. Yoon, J. S. Kwak, S.-M. Lee, and H. K. Baik, J. Electron. Mater. **28**, 6 (1999).
[7]K. Kushida-Abdelghafar, M. Hiratani, K. Torii, and Y. Fujusaki, Integr. Ferroelectr. **13**, 113 (1996).
[8]Q. T. Vu, P. J. Pokela, C. L. Garden, E. Kolawa, S. Raud, and M. A. Nicolet, J. Appl. Phys. **68**, 6420 (1991).
[9]T. Hara, T. Kitamura, M. Tanaka, T. Kobayashi, K. Sakiyama, S. Onishi, K. Ishihara, J. Kudo, Y. Kino, and N. Yamashita, J. Electrochem. Soc. **143**, L264 (1996).
[10]D.-S. Yoon, S.-M. Lee, and H. K. Baik, J. Appl. Phys. **83**, 1333 (1998).
[11]D.-S. Yoon, S.-M. Lee, and H. K. Baik, J. Appl. Phys. **83**, 8074 (1998).
[12]D.-S. Yoon, S.-M. Lee, and H. K. Baik, J. Vac. Sci. Technol. B **17**, 174 (1999).
[13]J. H. Kim, D.-S. Yoon, J. S. Kwak, S.-M. Lee, and H. K. Baik, J. Appl. Phys. **85**, 2170 (1999).
[14]D.-S. Yoon, S.-M. Lee, and H. K. Baik, J. Vac. Sci. Technol. A **15**, 2781 (1997).
[15]D.-S. Yoon, S.-I. Lee, S.-M. Lee, and H. K. Baik, J. Vac. Sci. Technol. B **16**, 1137 (1998).
[16]D.-S. Yoon, S.-I. Lee, S.-M. Lee, and H. K. Baik, Appl. Phys. Lett. **73**, 324 (1998).
[17]D.-S. Yoon, S.-I. Lee, S.-M. Lee, and H. K. Baik, J. Vac. Sci. Technol. B **16**, 3059 (1998).
[18]H. Jiang, C. S. Petersson, and M. A. Nicolet, Thin Solid Films **140**, 115 (1986).

Superlatticed negative differential-resistance heterojunction bipolar transistor

Shiou-Ying Cheng
Department of Electrical Engineering, National Cheng-Kung University, 1 University Road, Tainan, Taiwan, Republic of China

Jung-Hui Tsai
Department of Electronic Engineering, Chien-Kuo College, Changhwa, Taiwan, Republic of China

Po-Hung Lin and Wen-Chau Liu[a)]
Department of Electrical Engineering, National Cheng-Kung University, 1 University Road, Tainan, Taiwan, Republic of China

(Received 23 October 1997; accepted 23 April 1999)

In this article, we report the fabrication of a new superlatticed negative-differential-resistance heterojunction bipolar transistor (SNDR-HBT) based on a lattice-matched $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ material system. A 20-period AlInAs/GaInAs superlattice is used to serve the resonant tunneling route and the confinement barrier for minority carriers. From experimental results, an extremely low offset voltage of 61 mV and a common-emitter current gain of about 25 has been obtained at room temperature for the non-optimized device, simultaneously. In addition, we also observe the interesting *N*-shape negative-differential-resistance (NDR) phenomena in both two- and three-terminal operations at 77 K. It is believed that the *N*-shaped NDR results mainly from resonant tunneling within the 20-period AlInAs/GaInAs superlattice. © *1999 American Vacuum Society.* [S0734-211X(99)02104-6]

I. INTRODUCTION

Recently, owing to rapid advances in growth technologies, such as molecular beam epitaxy (MBE) and metal organic chemical vapor deposition (MOCVD), etc., it has become easy to control material composition and doping profiles nearly to the atomic scale. Due to these remarkable advances, these techniques have led to an increased interest in the fabrication and development of quantum-dot, quantum-wire, quantum-well, and superlattice related devices. The resonant-tunneling diode (RTD) and resonant-tunneling transistor (RTT), due to their negative-differential-resistance (NDR) and high-speed properties for integrated circuit applications, have been applied widely in analog-to-digital (A/D) converters, parity bit generator, frequency multipliers, and multiple-valued logic circuits.[1–3] Recently, Liu *et al.* have reported a superlattice-emitter resonant tunneling bipolar transistor (SE-RTBT) based on the AlGaAs/GaAs material system. A five-period AlGaAs/GaAs superlattice was used as a confinement layer and RT route through which the RT action was developed. Interesting transistor action and NDR phenomenon were observed in this device.[4] On the other hand, the InP-based crystal system is suitable for opto-electronic integrated circuits (OEICs) because this material system is sensitive to the $1.3\sim1.55\ \mu m$ wavelength range. In addition, AlInAs/GaInAs heterojunctions are suitable for RT devices due to their superior material properties such as higher direct conduction-band discontinuity ΔE_C ($Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$: 0.53 eV, $Al_xGa_{1-x}As/GaAs$: 0.30 eV), smaller turn-on voltage of base-emitter junction, and lower surface recombination velocity ($Ga_{0.47}In_{0.53}As$: $\sim10^3$ cm/s, GaAs: $\sim10^7$ cm/s).[5] In this article, we demonstrate a new superlatticed negative-differential-resistance heterojunction bipolar transistors (SNDR-HBT) with a long-period-superlattice emitter based on the lattice-matched $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ material system. A 20-period $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ superlattice was employed as the confinement and RT route layer.

II. EXPERIMENTS

The SNDR-HBT device was grown by metal organic chemical vapor deposition (MOCVD) on a (100) oriented n^+-InP substrate. The epitaxial layers included a 0.2 μm $n^+=3\times10^{18}\,cm^{-3}$ $Ga_{0.47}In_{0.53}As$ buffer, a 0.5 μm $n=5\times10^{16}\,cm^{-3}$ $Ga_{0.47}In_{0.53}As$ collector, a 0.1 μm $p^+=1\times10^{19}\,cm^{-3}$ $Ga_{0.47}In_{0.53}As$ base, a 500 Å $n=5\times10^{17}\,cm^{-3}$ $Ga_{0.47}In_{0.53}As$ emitter, a 20-period i-$Al_{0.48}In_{0.52}As$/n-$Ga_{0.47}In_{0.53}As$ superlattice and a 0.3 gmm $n^+=5\times10^{18}\,cm^{-3}$ $Ga_{0.47}In_{0.53}As$ cap layer. In the superlattice, the 50-Å-thick $Al_{0.48}In_{0.52}As$ barriers were undoped while the 50-Å-thick $Ga_{0.47}In_{0.53}As$ wells were doped to $n=5\times10^{17}\,cm^{-3}$ to reduce the series resistance and increase the tunneling current. After the epitaxial growth, the conventional photolithography, chemical wet etching, and vacuum evaporation techniques were used to fabricate the mesa-type devices. The ohmic contact metals were Au/Ge and AuZn for the *n*-type emitter, collector, and the *p*-type base, respectively. The emitter area is $60\times60\ \mu m^2$ and the collector-to-emitter area ratio is 10. The common-emitter $I-V$ characteristics of the device were measured at room temperature by an HP-4156A precision semiconductor parameter analyzer.

[a)]Electronic mail: wcliu@mail.ncku.edu.tw

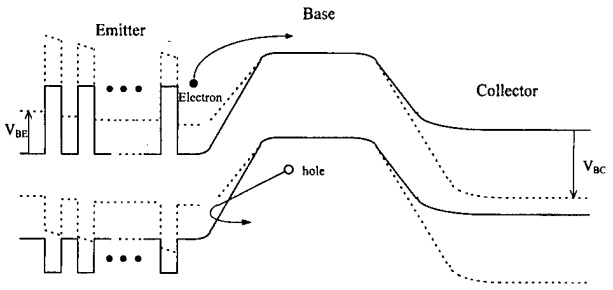


FIG. 1. Energy band diagrams of the studied structure at thermal equilibrium (solid lines) and under forward-biased operation (dashed lines).

III. RESULTS AND DISCUSSION

The corresponding energy band diagram for the device operated at thermal equilibrium and forward-active mode are illustrated in Fig. 1. In the studied structure, the emitter-base (E-B) structure contains an effective p-n homojunction and a hetero-confinement layer. The homojunction controls the electron injection and the heterojunction is used for hole confinement. The measured typical common-emitter current-voltage (I-V) characteristics of the studied device at room temperature are shown in Fig. 2. An expanded view near the origin of Fig. 2(a) is shown in Fig. 2(b). It is worthy to notice that the collector-emitter offset voltage ΔV_{CE} is as low as 61 mV. This ΔV_{CE} is smaller than that of the abrupt and graded AlInAs/GaInAs HBT (350 mV and 120 mV) and superior to that of InGaP/GaAs and AlGaAs/GaAs HEBTs (100 mV and 80 mV).[6,7] The small ΔV_{CE} is mainly caused by the disappearance of potential spike at the effective E-B junction. The ΔV_{CE} leading to the power consumption is also a significant factor for the transistor performance especially in digital applications. The ΔV_{CE} value ($\Delta V_{CE}=V_{BE}-V_{BC}$ when $I_C=0$) could be expressed as[8]

$$\Delta V_{CE}=I_B R_E+\frac{kT}{q}\ln\left(\frac{A_C}{A_E}\right)+\frac{kT}{q}\cdot\left(\frac{J_{CS}}{\alpha_F\cdot J_{ES}}\right), \tag{1}$$

where R_E is the emitter series resistance. A_C and A_E are the base-collector (B-C) and base-emitter (B-E) junction area, respectively. α_F is the forward common-base current gain. J_{ES} and J_{CS} are the reverse saturation current density of the emitter and collector junctions, respectively. For small values of I_B and R_E, there are two terms which determine the offset voltage. One is due to the differences between the emitter and collector geometry (A_C and A_E), and the other one is due to the electrical differences between each injection (J_{CS} and J_{ES}).

Figure 3 reveals the common-emitter current gain versus the collector current density. The common-emitter gain is 25 at a collector current density of 80 A/cm^2 with a high base-to-emitter doping ratio of 20. This demonstrates that the 20-period superlattice performs as an excellent minority carrier confinement layer. This behavior is similar to the N-AlInAs bulk layer in an heterostructure-emitter bipolar transistor (HEBT) device. As compared with other structures, the relatively small current gain is believed to be due to (i) the larger emitter thickness which introduces the extra neutral emitter

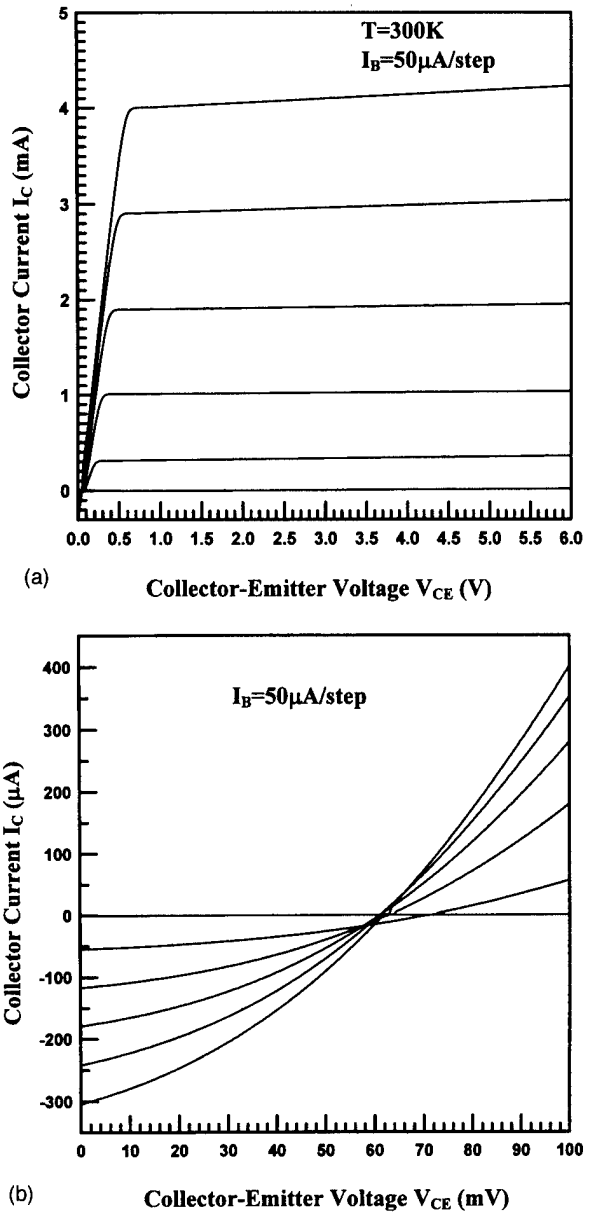


FIG. 2. Common-emitter current-voltage (I-V) characteristics of the studied device (a) and (b) the enlarged view near the origin.

recombination, (ii) the interface charge trap associated with the quality of epitaxy layer, and (iii) the hole tunneling current through the superlattice confinement layer.

The relationship between the current gain and collector current is illustrated in Fig. 3. The common-emitter current gain increases monotonically with the increase of collector current I_C. As stated above, the studied device is similar to a conventional HEBT. If the n-GaInAs emitter layer is too thin, a potential spike ΔV_E exists and the offset voltage ΔV_{CE} increases. So the device performs like a conventional HBT. On the other hand, if the emitter layer is too thick, it will result in degraded current gain due to charge storage and bulk recombination. Hence the n-GaInAs emitter layer thickness is an important factor for device performance. Chen *et al.* reported that the minimum value of emitter layer thickness is required to eliminate the potential spike and maintain

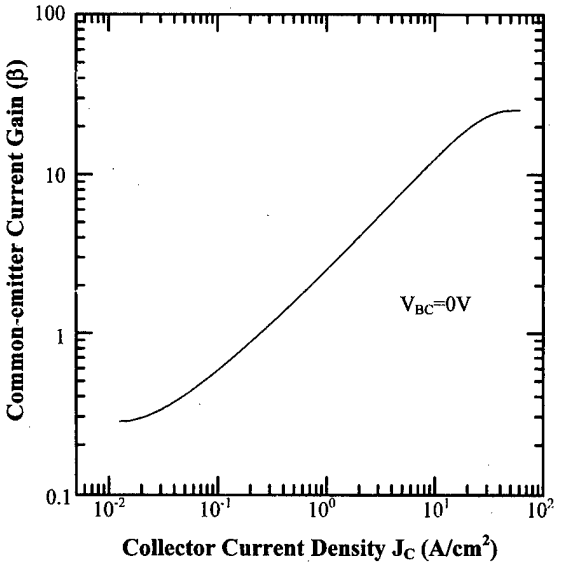


FIG. 3. Common-emitter current gain as a function of collector current. The base-collector bias V_{BC} is maintained at 0 V.

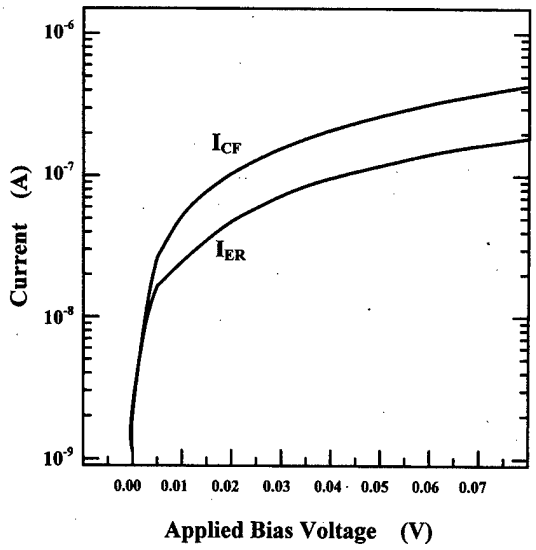


FIG. 4. Current-voltage transfer characteristics of the studied device.

the considerable current gain.[9] The minimum emitter layer thickness $W_{E\,\min}$ can be obtained as

$$W_{E\,\min} \geq \sqrt{\frac{2\epsilon_S \Delta E_C}{q^2 N_D}\left(\frac{N_A}{N_D+N_A}\right)}, \tag{2}$$

where N_D and N_A are the doping concentration of the emitter and base layer. ΔE_C and ϵ_s are the conduction band discontinuity and the dielectric permittivity of GaInAs. From theoretical calculations, the $W_{E\,\min}$ is about 390 Å for the studied device. Then we could choose an adequate emitter layer thickness to eliminate the potential spike ΔV_E. If ΔV_E is close to zero, the offset voltage is decided by the aspect ratio of the collector and emitter junctions. From Eq. (1), when $(A_C/A_E)=10$ is chosen, the ΔV_C is about 59.5 mV at room temperature. This also reveals that the studied device structure exhibits a nearly zero potential spike at E-B heterojunction.

The determination of the potential spike energy, by using a direct measurement technique, was employed to examine the existence of potential spike.[10,11] The forward transfer current I_{CF} was measured under the biased emitter-base (E-B) junction while the base-collector (B-C) junction is shorted. Next, the B-E junction was shorted while the B-C junction was biased to measure the reverse transfer collector current I_{ER}. Lin *et al.* have shown that the potential spike energy ΔV_E can be determined as[10]

$$\Delta V_E = kT \cdot \ln\left(\frac{I_{ER}}{I_{CF}}\right), \tag{3}$$

where k is the Boltzmann constant. If I_{CF} and I_{ER} are equal under the same applied voltage until the series resistance effect appears, the potential spike at E-B junction does not exist. From Fig. 4, it was found that I_{ER} and I_{CF} overlap each other until the series resistance effect causes the curves to deviate from the linear regime. Figure 5 shows the plots of the collector current I_C and the base current I_B as a function of base-emitter voltage V_{BE} under the applied zero base-collector voltage V_{BC}. The ideality factor n of I_C is 1.01 and that for I_B is 1.51. The 1 kT collector current indicates that the potential spike was not found between the E-B junction. On the other hand, the 1.5 kT base current means that the neutral base recombination and space charge recombination dominate the total base current.[12]

A transfer matrix method and envelope function approximation are employed to analyze the transmissivity of the 20-period superlattice structure. From the theoretical calculation, the quantized miniband energies of the 20-period superlattice were obtained as a ground band at $E_1=123$ meV and a first excited band at $E_2=480$ meV calculated from the bottom of the conduction band in the strongly coupled superlattice regime. The band widths of E_1 and E_2 are 4 meV and

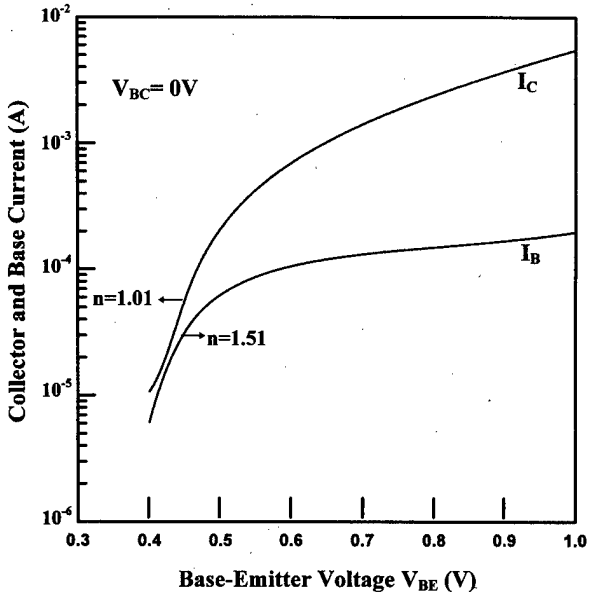


FIG. 5. Gummel plot of the studied device. The base-collector bias V_{BC} is maintained at 0 V.

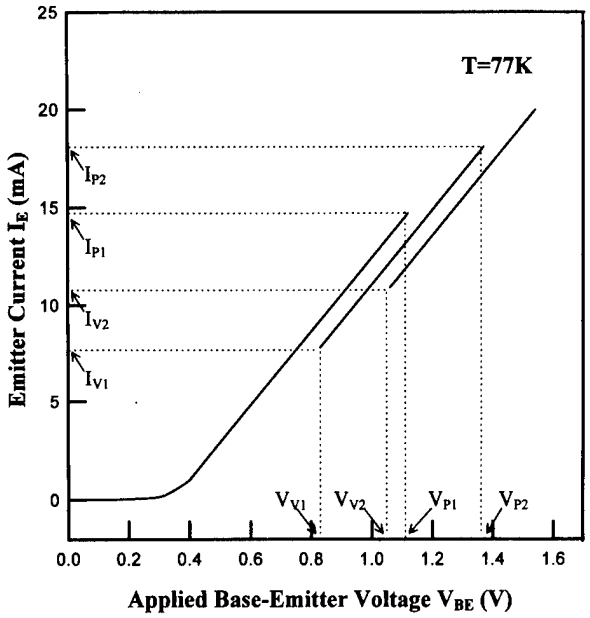


FIG. 6. Experimental current-voltage characteristics between base and emitter of the studied device at 77 K. The I_{P1}, I_{V1}, I_{P2}, and I_{V2} are 14.7, 7.8, 10.9, and 18.1 mA with the V_{P1}, V_{V1}, V_{P2}, and V_{V2} of 1.12, 0.84, 1.36, and 1.07 V, respectively. The peak-to-valley current ratio of first and second NDR are 1.89 and 1.66, respectively.

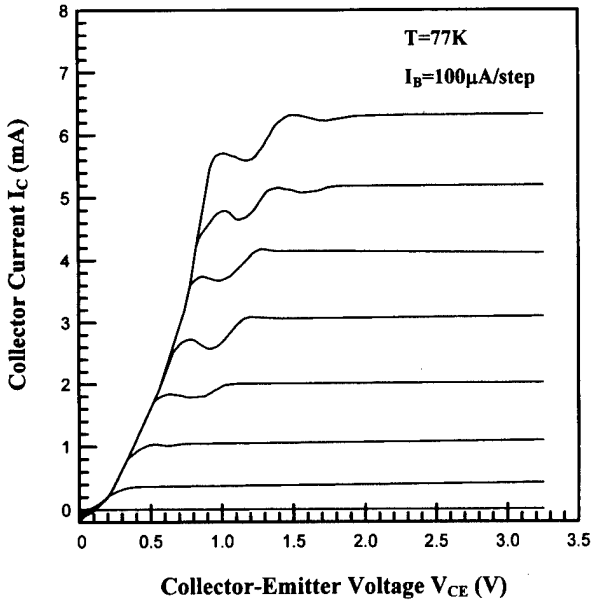


FIG. 7. Measured common-emitter current-voltage characteristics of the studied device at liquid nitrogen temperature (77 K).

60 meV, respectively. The measured base-emitter junction current-voltage (I-V) characteristics of the studied device at liquid nitrogen temperature (77 K) is shown in Fig. 6. Significantly, the double N-shape NDR phenomena were observed. The first peak current I_{P1}, first valley current I_{V1}, second peak current I_{P2}, and second valley current I_{V2} are 14.7, 7.8, 10.9, and 18.1 mA with the corresponding first peak voltage V_{P1}, first valley voltage V_{V1}, second peak voltage V_{P2}, and second valley voltage V_{V2} are 1.12, 0.84, 1.36, and 1.07 V, respectively. We find that the peak-to-valley current ratios of first and second NDR are 1.89 and 1.66.

Figure 7 shows the common-emitter current-voltage (I-V) characteristics of the studied device at 77 K. The interesting N-shaped NDR phenomena are found in the I-V curves. The transport mechanism at 77 K may be described as follows. For a low value of base current I_B (e.g., $I_B < 0.1$ mA), the base-emitter voltage V_{BE} is smaller than the p-n junction built-in-voltage V_{bi}. Thus the applied V_{BE} voltage is essentially across the E-b p-n junction until the flat-band condition is established. Within this regime, the studied device performs as a bipolar transistor. Beyond the flat-band condition, the additional increase in V_{BE} voltage will mostly fall across the superlattice region. For some V_{BE} voltage ($0.2\,\text{mA} \leq I_B \leq 1.0\,\text{mA}$), the Fermi level E_F on the emitter side aligns with the miniband within the superlattice. This increases the transmissivity and current through the superlattice structure by RT action. When the collector-emitter voltage V_{CE} is increased, for fixed I_B, the band structure near the emitter side is elevated quickly. This will cause the misalignment between the Fermi energy and the ground miniband within the superlattice. Hence, a quenching of RT through the superlattice and a sudden decrease in the conduction current were found. This certainly exhibits an N-shape NDR phenomenon in the I-V curves. When the V_{CE} voltage is further increased, a second NDR will be observed once the Fermi level aligns with the upper miniband within the superlattice. Experimentally, the second NDR phenomena was observed in the higher I_B regime ($I_B \geq 0.6\,\text{mA}$) as shown in Fig. 7. Because the upper miniband approaches the top surface of the superlattice region when the Fermi level on the emitter side reaches the upper miniband level, many electrons will travel directly over the superlattice rather than tunnel through the miniband. Thus, the second NDR behavior is relatively insignificant when compared with the first NDR phenomena. On the other hand, at room temperature (300 K), a sufficient amount of thermal energy causes most of the electrons to transport directly over the superlattice region. Therefore, the RT action and NDR phenomena are not observed.

IV. CONCLUSIONS

In conclusions, we have successfully fabricated and demonstrated a new SNDR-HBT with a long-period-superlattice emitter based on lattice-matched $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ material system. We observed a collector-emitter offset voltage of 61 mV and common-emitter current gain of 25 when the studied device operated at room temperature. The extremely small collector-emitter offset voltage is very important for digital circuit applications because it can significantly reduce unnecessary power consumption. By employing a direct measurement technique, it was revealed that the potential spike at the effective E-B junction of the studied device is almost negligible. Furthermore, the relatively low common-emitter current gain can be improved by a decrease of the emitter and base layer thickness in order to

degrade the neutral emitter and base bulk recombination current. In addition, the interesting *N*-shape NDR phenomena introduced by RT both in two- and three-terminal operations at 77 K was found. The NDR characteristic is important to multiple-valued logic applications. Consequently, the studied device provides the promise for high speed, low power consumption digital circuit applications.

ACKNOWLEDGMENT

This study was supported by the National Science Council of the Republic of China under Contract No. NSC.88-2215-E-006-012.

[1]F. Capasso, S. Sen, A. Y. Cho, and D. Sivco, IEEE Electron Device Lett. **8**, 297 (1987).
[2]F. Capasso, S. Sen, F. Beltram, L. M. Lunardi, A. S. Vengurlekar, P. R. Smith, N. J. Shah, R. J. Malik, and A. Y. Cho, IEEE Trans. Electron Devices **36**, 2065 (1989).
[3]F. Capasso, A. S. Vengurlekar, A. Hutchinson, and W. T. Tsang, Electron. Lett. **25**, 1117 (1989).
[4]W. C. Liu, W. S. Lour, and Y. H. Wang, IEEE Trans. Electron Devices **39**, 2214 (1992).
[5]S. M. Sze, *High-Speed Semiconductor Device* (Wiley, New York, 1990).
[6]T. Won and H. Morkoc, IEEE Trans. Electron Devices **10**, 138 (1989).
[7]K. B. Thei, J. H. Tsai, W. C. Liu, and W. S. Lour, Solid-State Electron. **39**, 1137 (1996).
[8]S. C. Lee, J. N. Kau, and H. H. Lin, Appl. Phys. Lett. **45**, 1114 (1984).
[9]H. R. Chen, C. Y. Chang, C. P. Lee, C. H. Huang, J. S. Jang, and K. L. Tsai, IEEE Electron Device Lett. **15**, 336 (1994).
[10]S. C. Lee, J. N. Kau, and H. H. Lin, J. Appl. Phys. **58**, 890 (1985).
[11]H. H. Lin and S. C. Lee, IEEE Electron Device Lett. **6**, 431 (1985).
[12]S. S. Lu and C. C. Huang, IEEE Electron Device Lett. **13**, 214 (1992).

High-quality InAlAs layers grown on (411)A-oriented InP substrates by molecular beam epitaxy

Takahiro Kitada,[a] Keisuke Nii, Tetsuya Hiraoka, Satoshi Shimomura, and Satoshi Hiyamizu
Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

(Received 25 November 1998; accepted 16 April 1999)

High-quality InAlAs layers lattice matched to InP were successfully grown on (411)A-oriented InP substrates by molecular beam epitaxy (MBE). High-resolution x-ray diffraction and photoluminescence (PL) measurements of InAlAs layers grown on (411)A InP substrates revealed that crystalline quality of the (411)A InAlAs layer strongly depends on the substrate temperature (T_s), and much improved crystalline quality of (411)A InAlAs layer was achieved at a high T_s (570 °C). The linewidth of the PL (12 K) peak from the best (411)A InAlAs layer is 10.7 meV which is 16%–29% smaller than those (12.8–15 meV) of InAlAs layers grown on conventional (100) InP substrates by MBE. © *1999 American Vacuum Society.* [S0734-211X(99)01204-4]

I. INTRODUCTION

In recent years, effectively atomically flat heterointerfaces over a wafer-size area ["(411)A super-flat interfaces"] have been reported in GaAs/AlGaAs quantum wells (QWs) grown on (411)A-oriented GaAs substrates by molecular beam epitaxy (MBE).[1,2] Photoluminescence (PL) linewidths (4.2 K) of (411)A GaAs/AlGaAs QWs are much narrower than those of QWs grown on conventional (100) substrates, and they are comparable to those of split peaks of (100) growth-interrupted QWs.[3] In our recent studies, the (411)A super-flat interfaces have been also achieved in MBE grown pseudomorphic $In_xGa_{1-x}As/Al_{0.3}Ga_{0.7}As$ QWs with $x\sim0.08$[4] and $x\sim0.15$[5] and $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ QWs lattice matched to InP substrates,[6] which are important materials for ultrahigh speed devices such as high electron mobility transistors (HEMTs), resonant tunneling diodes (RTDs), and long-wavelength optical devices. In order to achieve high-quality InGaAs/InAlAs heterostructures with the (411)A super-flat interfaces, MBE growth of a high-quality InAlAs barrier layer is essential. In this article, we investigated the MBE growth conditions of high-quality InAlAs layers on (411)A-oriented InP substrates with the use of high-resolution x-ray diffraction (HRXRD) and low-temperature (12 K) PL measurements.

II. MBE GROWTH

The (411)A-oriented InP substrates were degreased and etched by Br_2/CH_3OH to remove residual impurities on the substrate surfaces, and they were mounted side by side with (100) InP substrates on a 2 in. Mo holder with In solder. After loading into the growth chamber of a Nissin RB-2001G MBE system, native oxides of InP were desorbed by raising substrate temperature (T_s) up to 520 °C under As_4 beam flux. 800-nm-thick $In_xAl_{1-x}As$ layers ($x=0.49-0.52$) were grown on these substrates under two growth conditions of (1) $T_s=520$ °C, V/III [As_4/(In+Al)] $=16$ (in pressure) and (2) $T_s=570$ °C, V/III=9. The substrates were rotated at 30 rpm during MBE growth. The growth rate of the InAlAs layers was about 1 μm/h. Beam flux intensities of In and Al were adjusted with the use of a beam flux monitor, and the In contents (x) of $In_xAl_{1-x}As$ layers were checked by HRXRD measurements.

III. CRYSTALLINE PROPERTIES

Figure 1 shows surface morphologies of the InAlAs layers simultaneously grown on (411)A- [(a) and (b)] and (100)-oriented [(c) and (d)] InP substrates observed by a Nomarski interference microscope. Figures 1(a) and 1(c) are surface morphologies of the InAlAs layers grown at $T_s=520$ °C, and Figs. 1(b) and 1(d) are those of the InAlAs grown at $T_s=570$ °C. Surface morphologies of the (100) InAlAs layers are quite featureless even for two different growth conditions. In the case of the (411)A InAlAs layers, a featureless surface was observed for $T_s=570$ °C, while a rough surface for $T_s=520$ °C, indicating that enhancement of surface migration of group III atoms (In, Al) by a high substrate temperature is necessary to obtain flat surfaces of InAlAs layers grown on (411)A InP substrates. It is reasonable that a high substrate temperature is favorable for MBE growth on a (411)A substrate, because an epitaxial growth on the (411)A substrate is made in the step-flow growth mode, which was confirmed by reflection high-energy electron diffraction (RHEED) measurements.

Figure 2 shows HRXRD rocking curves around low-incidence (511) reflection of the InAlAs layers grown on (411)A InP substrates. An incident x-ray beam with a wavelength ($\lambda=0.15406$ nm) of Cu $K\alpha_1$ was obtained by a (220)-oriented Ge asymmetric four-crystal monochromator. InAlAs samples were mounted on a computer-controlled goniometer stage, and the diffracted x-ray intensity was counted as a function of the incident beam angle. A sharp peak [full width at half maximum (FWHM) of 34 s] from the (411)A InAlAs layer grown at $T_s=570$ °C is clearly seen at 104 s apart from the (511) reflection peak of the InP substrate in Fig. 2(a), while a broadened peak (FWHM of 450 s)

[a]Electronic mail: kitada@d310.mp.es.osaka-u.ac.jp

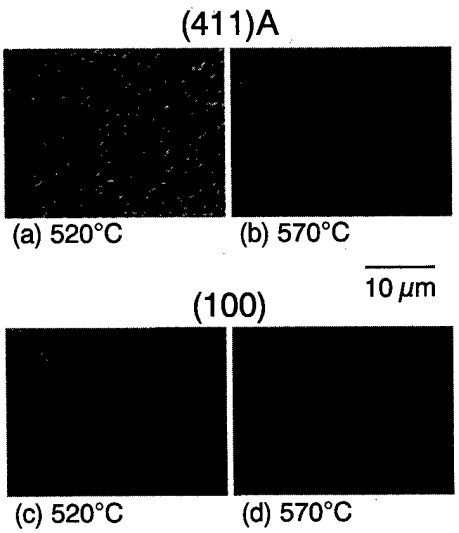


FIG. 1. Surface morphologies of the InAlAs layers simultaneously grown at T_s=520 and 570 °C on (411)A- [(a) and (b), respectively] and (100)-oriented [(c) and (d)] InP substrates by MBE.

from the (411)A InAlAs grown at T_s=520 °C is observed at 775 s apart from the substrate peak in Fig. 2(b). By comparing these reflection angles with calculated ones taking into account of the strain component of the InAlAs layers on (411)A InP substrates,[7] the In contents (x) of the (411)A $In_xAl_{1-x}As$ layers grown at T_s=520 and 570 °C were determined to be 0.49 and 0.52, respectively. The difference of the In content between these two samples is due to different beam flux intensity ratio [In/(In+Al)]. Relaxation of lattice mismatch did not occur in these two samples, which was confirmed by HRXRD measurements of high- and low-incidence asymmetric (511) reflections for these two samples. The InAlAs layers simultaneously grown on

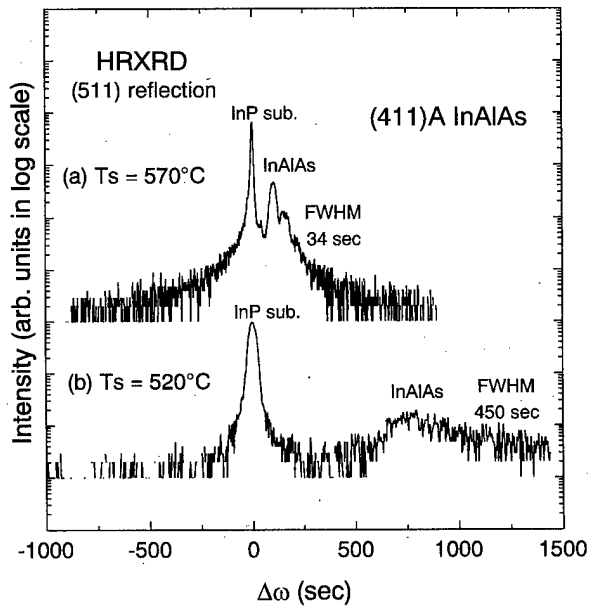


FIG. 2. HRXRD rocking curves around (511) reflections of the InAlAs layers grown at (a) T_s=570 °C and (b) T_s=520 °C on (411)A InP substrates.

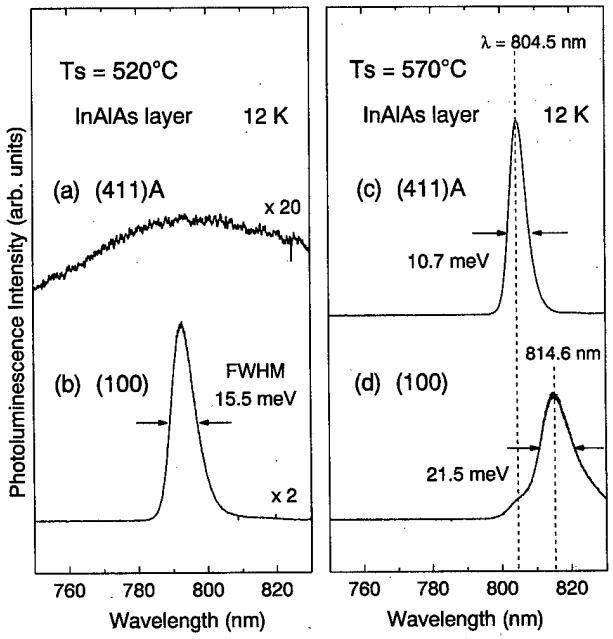


FIG. 3. PL spectra (12 K) from InAlAs layers simultaneously grown at T_s =520 °C and T_s=570 °C on (411)A- [(a) and (c), respectively] and (100)-oriented [(b) and (d)] InP substrates.

(100) InP substrates showed the same In contents as those of the (411)A InAlAs layers. HRXRD measurements around (400) reflections were used for the (100) samples. A most prominent feature in HRXRD rocking curves in Fig. 2 is that the linewidth of the (411)A InAlAs peak is much improved by increasing T_s from 520 to 570 °C. These results indicate that the high substrate temperature (T_s=570 °C) is very important for MBE growth of InAlAs layers with featureless surface morphology and high crystalline quality.

IV. OPTICAL PROPERTIES

Figure 3 shows 12 K PL spectra from InAlAs layers simultaneously grown on (411)A and (100) InP substrates. The excitation laser was a He–Cd laser (λ=325 nm) with an excitation power of 2 mW. The excitation beam was focused on an area of about 200 μm in diameter on the substrate surface. A much broadened PL peak was observed from the (411)A InAlAs layer grown at T_s=520 °C, as shown in Fig. 3(a), while a sharp peak at 792.8 nm (FWHM=15.5 meV) from the (100) InAlAs is clearly seen in Fig. 3(b). A broadened PL peak from the (411)A InAlAs grown at T_s =520 °C is due to poor crystalline quality of the InAlAs layer, which is consistent with results of crystalline characterization by HRXRD. In contrast with this, a very sharp PL peak was observed at 804.5 nm (FWHM=10.7 meV) from the (411)A InAlAs layer grown at T_s=570 °C [Fig. 3(c)], while a broadened PL peak was observed at 814.6 nm (FWHM=21.5 meV) from the (100) InAlAs [Fig. 3(d)]. It is worth to note that when the substrate temperature was increased from 520 to 570 °C, PL properties of the (411)A InAlAs layer were drastically improved, while those of the (100) InAlAs were slightly degraded. This result implies that

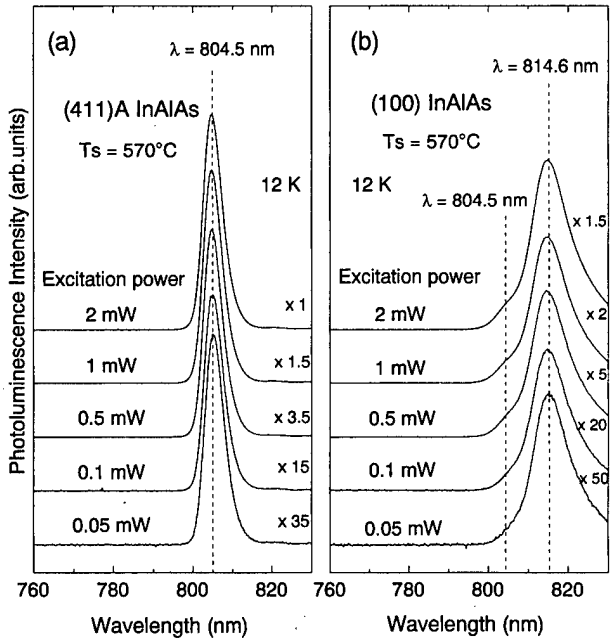


FIG. 4. Excitation power dependence of PL spectra (12 K) from the InAlAs layers simultaneously grown on (a) (411)A- and (b) (100)-oriented InP substrates.

MBE growth of the InAlAs layer at the high T_s on the (411)A face is more stable than that on (100). The linewidth of the PL peak from the (411)A InAlAs layer is narrower than not only those of the (100) samples in this study but also those of (100) InAlAs layers grown by MBE reported so far (FWHM=12.8 meV at 15 K,[8] 15 meV at 4 K[9]). Therefore, the (411)A InAlAs layer grown at the high substrate temperature of 570 °C has better optical quality compared with any conventional (100) InAlAs layers. Furthermore, the line shape of the (411)A InAlAs is a single peak at 804.5 nm [Fig. 3(c)], while that of the (100) InAlAs has a double-peak character, which consists of a main peak at 814.6 nm with a high-energy side shoulder around 804.5 nm [Fig. 3(d)].

Figure 4 shows excitation power dependence of PL spectra from the (411)A and (100) InAlAs layers simultaneously grown at T_s=570 °C. The line shape of the PL peak from the (411)A InAlAs layer almost does not depend on the excitation power in the range of 0.05–2 mW as shown in Fig. 4(a). For the (100) InAlAs layer [Fig. 4(b)], the relative PL intensity of the high-energy side peak (λ = 804.5 nm) against the main peak (λ = 814.6 nm) decreases with decreasing excitation power. The main PL peak at 814.6 nm of the (100) InAlAs layer is probably due to As-vacancy-related luminescence because of the high substrate temperature of 570 °C, and the high-energy side shoulder around 804.5 nm should be due to band-to-band luminescence. The PL peak from the (411)A InAlAs is due to band-to-band luminescence, because (1) the line shape almost does not depend on the excitation power, and (2) the wavelength of the PL peak from the (411)A InAlAs layer (λ = 804.5 nm) is the same as the high-energy side shoulder of the (100) InAlAs layer, which has the same In content (0.52) as that of the (411)A InAlAs layer. This result indicates that the incorporation of defects such as As vacancies in the InAlAs layer grown on a (411)A InP substrate is much suppressed compared with (100) InAlAs even at a high substrate temperature of 570 °C. (411)A growing surface consists of high-density microsteps, which are responsible for the incorporation of As atoms, therefore, the incorporation of As vacancies on the (411)A substrate is suppressed.

V. CONCLUSION

High-quality InAlAs layers were successfully grown on (411)A-oriented InP substrates by MBE. Crystalline and optical properties of the (411)A InAlAs layers grown under different conditions revealed that a high substrate temperature of 570 °C is very important for (411)A InAlAs layers with featureless surface morphology and good crystalline quality. From PL measurements, the incorporation of defects such as As vacancies in the (411)A InAlAs layer grown at a high substrate temperature (570 °C), is much suppressed compared with the simultaneously grown (100) InAlAs layer, which was characterized by PL measurements. Furthermore, the linewidth of the PL peak from the (411)A InAlAs layer (10.7 meV) is 16%–29% narrower than those of conventional (100) InAlAs layers reported so far.[8,9] Hence, MBE grown (411)A InGaAs/InAlAs heterostructures should be suitable for novel quantum devices such as HEMTs.

[1] S. Shimomura, A. Wakejima, A. Adachi, Y. Okamoto, N. Sano, K. Murase, and S. Hiyamizu, Jpn. J. Appl. Phys., Part 2 **32**, L1728 (1993).
[2] S. Hiyamizu, S. Shimomura, A. Wakejima, S. Kaneko, A. Adachi, Y. Okamoto, N. Sano, and K. Murase, J. Vac. Sci. Technol. B **12**, 1043 (1994).
[3] M. Tanaka, H. Sakaki, J. Yoshino, and T. Furusawa, Surf. Sci. **174**, 65 (1986).
[4] T. Saeki, T. Motokawa, T. Kitada, S. Shimomura, A. Adachi, Y. Okamoto, N. Sano, and S. Hiyamizu, Jpn. J. Appl. Phys., Part 1 **36**, 1786 (1997).
[5] M. Ohashi, T. Saeki, T. Kitada, S. Shimomura, Y. Okamoto, and S. Hiyamizu, Jpn. J. Appl. Phys., Part 1 **37**, 4515 (1998).
[6] T. Kitada, T. Saeki, M. Ohashi, S. Shimomura, A. Adachi, Y. Okamoto, N. Sano, and S. Hiyamizu, J. Electron. Mater. **27**, 1043 (1998).
[7] K. Yang, T. Anan, and L. J. Schowalter, Appl. Phys. Lett. **65**, 2789 (1994).
[8] A. Chin, P. Bhattacharya, W. P. Hong, and W. Q. Li, J. Vac. Sci. Technol. B **6**, 665 (1988).
[9] S. Monéger, A. Tabata, C. Bru, G. Guillot, A. Georgakilas, K. Zekentes, and G. Halkias, Appl. Phys. Lett. **63**, 1654 (1993).

Evolution of microstructure and dislocation dynamics in $In_xGa_{1-x}P$ graded buffers grown on GaP by metalorganic vapor phase epitaxy: Engineering device-quality substrate materials

Andrew Y. Kim,[a] Wendy S. McCullough, and Eugene A. Fitzgerald
Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 2 December 1998; accepted 5 April 1999)

This study explores the dislocation dynamics of strain relaxation in graded composition buffers of $In_xGa_{1-x}P$ grown on GaP ($In_xGa_{1-x}P$/GaP) by metalorganic vapor phase epitaxy. Transmission electron microscopy, cathodoluminescence imaging, atomic force microscopy, and triple-axis x-ray diffraction are applied to the characterization of $In_xGa_{1-x}P$/GaP with final compositions ranging from $x=0.09$ to $x=0.39$ and growth temperatures ranging from 650 to 810 °C. The previously reported escalation of defect density with continued grading of $In_xGa_{1-x}P$/GaP beyond $x\sim0.3$ is discovered to be due to the formation of dislocation pileups. A new defect microstructure with a branching morphology and featuring sharp local strain fields, hereafter referred to as branch defects, is observed to pin dislocations and cause the dislocation pileups. Branch defect morphology varies strongly with growth temperature, becoming significantly stronger with increasing growth temperature and causing severe material degradation above 700 °C. Further experiments show that branch defects evolve during growth and that the onset of branch defect formation is delayed by increasing growth temperature. Comparison with the literature suggests that the evolution of branch defects may control the microstructure of indium-bearing phosphides and arsenides over a very wide range of conditions. In the absence of branch defects at high growth temperatures and low indium compositions near $x\sim0.1$, nearly ideal dislocation dynamics dominated by dislocation glide kinetics are recovered, providing the first experimental proof of a kinetic model for graded buffer relaxation. This new understanding of the evolution of microstructure and dislocation dynamics in $In_xGa_{1-x}P$/GaP suggests that growth temperature must be optimized as a function of composition for optimal material quality. A simple process optimization in $In_xGa_{1-x}P$/GaP graded to $x=0.39$ results in an overall threading dislocation density of 4.7×10^6 cm^{-2}, which is the lowest reported value to date for $x>0.3$. Combining the new observations with earlier findings, we present three basic design rules for producing practical, device-quality graded buffers: branch defects must be avoided or suppressed, growth temperature must be maximized, and surface roughness must be minimized. Using these design rules, we also present optimization strategies for achieving device-quality substrate materials. Applying these design rules and optimization strategies, we hope to achieve threading dislocation densities of $<10^6$ cm^{-2} in $In_xGa_{1-x}P$/GaP over the full range of useful compositions. © *1999 American Vacuum Society.* [S0734-211X(99)00404-7]

I. INTRODUCTION

Epitaxial graded composition buffers (graded buffers) of $In_xGa_{1-x}P$ on GaP substrates ($In_xGa_{1-x}P$/GaP) are promising materials for high-performance optoelectronic devices. $In_xGa_{1-x}P$ alloys with band gaps that are difficult or impossible to achieve lattice-matched to GaAs substrates can be grown on graded buffers,[1] expanding the range of emission wavelengths into the infrared and providing direct band gap emission of the critical amber to infrared wavelengths inaccessible to GaN-based light emitting diode (LED) and laser diode technologies. GaP also has nearly twice the thermal conductivity of GaAs for high power applications.[2] $In_xGa_{1-x}P$/GaP substrates are also inherently transparent to devices grown on them, which roughly doubles light extraction efficiency in LEDs compared to those grown on absorbing substrates such as GaAs, and offer significant processing advantages over the current transparent-substrate LED technology.[3,4] The transparency of $In_xGa_{1-x}P$/GaP has also been used to produce negative electron affinity GaAs[5] and InGaAs[6] photocathodes that operate in transmission mode, and a variety of other optoelectronic detectors and modulators can be envisioned to take advantage of a transparent semiconductor substrate. Furthermore, GaP is nearly lattice-matched to Si, so $In_xGa_{1-x}P$/GaP is one natural choice for integrating compound semiconductor devices on Si substrates.[7]

Graded buffers are grown to efficiently relieve lattice-mismatch strain between substrates and films of differing lattice constants. For most optoelectronic device applications, direct band gap compositions of $In_xGa_{1-x}P$ are desired. The >2% lattice mismatch between GaP and direct band gap compositions of $In_xGa_{1-x}P$ results in heavily defective single heterostructures, due to the large and abrupt

[a]Electronic mail: andy@mtl.mit.edu

introduction of strain at one interface.[8,9] A graded buffer of $In_xGa_{1-x}P$ on GaP slowly introduces strain over many interfaces, which minimizes dislocation interactions, maintains a low state of strain, and minimizes dislocation nucleation during growth. Consequently, graded buffers typically have orders of magnitude lower threading dislocation densities than single heterostructures.[8]

The growth of $In_xGa_{1-x}P$/GaP has been studied for decades using a variety of growth techniques, including hydride vapor phase epitaxy (HVPE), gas source molecular beam epitaxy (GSMBE), and metalorganic vapor phase epitaxy (MOVPE). Early HVPE experiments with $In_xGa_{1-x}P$/GaP[10] and $GaAs_xP_{1-x}$/GaAs[11] by Abrahams and co-workers established some of the basic principles of dislocation dynamics in graded buffers. Since then, visible LEDs have been demonstrated on $In_xGa_{1-x}P$/GaP. Stinson *et al.* used HVPE to produce LEDs operating at wavelengths from 565 to 650 nm and found that device efficiency decreased dramatically when $In_xGa_{1-x}P$/GaP was graded beyond $x \sim 0.35$.[1] Chin *et al.* used GSMBE to grow $In_xGa_{1-x}P$/GaP with photoluminescence (PL) ranging from 560 to 600 nm and observed decreasing PL intensity in buffers graded beyond $x \sim 0.32$.[12]

The agreement of results showing degradation beyond $x \sim 0.3$ with two very different growth techniques is striking. Both groups concluded that material degradation is a natural consequence of increasing lattice mismatch, presumably through increasing defect density. This intuitive picture is inconsistent with the earlier work of Abrahams *et al.*, who concluded from experimental and theoretical considerations that strain relaxation in graded buffers is a steady-state process, hence defect density should be constant.[10,11]

Developments in the Ge_xSi_{1-x}/Si system have provided new insights into dislocation dynamics in graded buffers that can aid in understanding $In_xGa_{1-x}P$/GaP. Samavedam and Fitzgerald demonstrated that the formation of dislocation pileups is the primary cause of material degradation with continued grading in Ge_xSi_{1-x}/Si.[13] Since dislocations immobilized in pileups can no longer glide to relieve strain, pileups force the nucleation of additional dislocations to continue the relaxation process. Samavedam and Fitzgerald proposed that pileups are caused by an interaction between dislocations and surface roughness. Misfit dislocation strain fields produce surface undulations during growth and gliding dislocations can be pinned between the strain fields and undulations, resulting in pileups. The pinned dislocations increase local surface roughness, resulting in a recursive and escalating cycle of dislocation pinning and surface roughening. Currie *et al.* then showed that controlling surface roughness by periodic planarization can suppress pileup formation in Ge_xSi_{1-x}/Si and recover a steady-state dislocation density between $x=0.3$ to $x=1$.[14] The recovery of steady-state dislocation dynamics through planarization is compelling evidence that pileup formation due to the interaction of dislocations and surface roughness is responsible for material degradation with continued grading in Ge_xSi_{1-x}/Si.

Recent work with $In_xGa_{1-x}P$/GaP grown by MOVPE also showed a strong correlation between surface roughness and the density of dislocations and pileups.[9] Pileup formation was tentatively attributed to the mechanism proposed by Samavedam and Fitzgerald, but comparison with Ge_xSi_{1-x}/Si results suggests that the sensitivity of defect density to surface roughness in $In_xGa_{1-x}P$/GaP is much greater than expected. Related work with $In_xGa_{1-x}As$/GaAs noted the presence of "high-energy boundaries" that appeared to pin dislocations, although their overall impact on relaxation was unclear.[15] Defects similar to the high-energy boundaries have not been observed in Ge_xSi_{1-x}/Si, so defects of this type may account for the difference in pileup behavior noticed between $In_xGa_{1-x}P$/GaP and Ge_xSi_{1-x}/Si.

In this study, we explore the evolution of dislocation dynamics in $In_xGa_{1-x}P$/GaP grown by MOVPE. Starting with the question of escalating defect density, we show by microscopic characterization and macroscopic modeling that a previously unrecognized defect microstructure causes pileups and dominates dislocation dynamics in $In_xGa_{1-x}P$/GaP. We map the evolution of microstructure in graded buffers and model its effects on dislocation dynamics. The evolution of the new defect microstructure is used to explain the microstructure of both indium-bearing phosphides and arsenides over a wide range of conditions. By controlling the new defect microstructure, we are also able to observe nearly ideal relaxation behavior dominated by dislocation kinetics. Through analysis and modeling, we confirm and extend a proposed kinetic model for relaxation in graded buffers. The new understanding and control of dislocation dynamics and microstructure are used to derive a set of design rules and an optimization strategy for high-quality graded buffer growth. A simple process optimization results in material with a dislocation density of $4.7\times10^6\ cm^{-2}$ at $x=0.39$.

II. EXPERIMENT

A. MOVPE film growth

Film growth was performed in a Thomas Swan atmospheric pressure MOVPE reactor. Trimethylgallium (TMGa) and solution trimethylindium (TMIn) were used as group III sources, and PH_3 and AsH_3 were used as group V sources. Source gases were carried in Pd-purified H_2 flowing at 5 slpm through a fast-switching manifold with parallel, pressure-balanced vent and reactor lines. Substrates were placed on a graphite susceptor in a horizontal, rectangular quartz reactor. Growth temperature was controlled through a halogen lamp under the reactor and a thermocouple inside the susceptor.

The substrates were S-doped (001) GaP wafers off-cut 10° towards an orthogonal {001}. To prepare substrates for growth, wafers were cleaved into approximately 5 cm^2 pieces, etched for 1 min in 25:1:1 H_2SO_4:H_2O_2:H_2O, rinsed with deionized water, and dried with N_2. The substrates were then loaded into the reactor and baked for 2 min at 150 °C, 5 min at 200 °C, and 15 min at 800 °C under a flow of H_2 and PH_3. The desired growth temperature was then set and allowed to stabilize for 5 min.

TABLE I. General growth conditions for all $In_xGa_{1-x}P$/GaP samples. All graded buffers were grown with an average growth rate of 5 μm/h, an average grading rate of $-0.4\%/\mu$m, and incorporated a 0.5 μm homoepitaxial GaP buffer and a 4 μm uniform composition cap layer.

Series	I	I	II	II	II	III	III	III	III	IV
Temperature (°C)	760	760	700	760	810	650	700	760	800	Variable
PH_3 pressure (Pa)	800	800	2600	800	2600	2600	2600	800	800	Variable
Indium composition (nominal/measured)	0.27/ 0.26	0.32/ 0.31	0.33/ 0.34	0.33/ 0.31	0.33/ 0.36	0.10/ 0.15	0.10/ 0.09	0.10/ 0.11	0.10/ 0.12	0.40/ 0.39
Grading rate	Variable	Variable	Constant	Constant	Constant	Constant	Constant	Constant	Constant	Constant

Film growth began with a 0.5 μm GaP buffer layer grown at 3 μm/h. Graded buffers were then grown by setting a constant TMGa flow of 0.235 sccm and stepping TMIn flow up in 2 min intervals. The group III flows corresponded to growth rates ranging from 3 to 6 μm/h as a function of temperature and composition. For changes in growth temperature, growth was interrupted and the system was equilibrated at the new temperature for 5 min. Growth was completed with a 4 μm uniform composition cap layer and samples were cooled under a flow of H_2 and PH_3 down to 150 °C. Throughout the growth process, the PH_3 flow was adjusted to at least twice the level previously found necessary for the high-quality growth of GaP/GaP and $In_{0.31}Ga_{0.69}P/GaAs_{0.64}P_{0.36}$ in our reactor.

All graded buffers were grown with a nominal mean grading rate of -0.4% strain/μm, which corresponds to 5% In/μm. To maintain a constant grading rate profile, TMIn flow step size was increased continuously. To produce a variable grading rate profile that begins at $-0.7\%/\mu$m and decreases to $-0.2\%/\mu$m, TMIn flow step size was kept constant.

A 4 μm film of $In_{0.33}Ga_{0.67}P$ was grown directly on GaP at 700 °C as a reference. Five series of graded buffer experiments were performed and are summarized in Table I. Series I was grown to investigate the degradation with continued grading beyond $x\sim0.3$ in $In_xGa_{1-x}P$/GaP reported by Stinson *et al.*[1] The growth temperature of 760 °C was chosen to be within the range typical of HVPE. Series II was grown to study the effect of growth temperature on buffers graded up to $x\sim0.33$. Series III was grown to study the effect of growth temperature on buffers graded up to $x\sim0.1$. Series IV was grown to optimize growth temperature as a function of composition: the sample was graded up to $x=0.1$ at 760 °C, up to $x=0.2$ at 700 °C, and up to $x=0.4$ at 650 °C.

B. Characterization

The microstructure of the graded buffers was examined by transmission electron microscopy in both plan-view (PVTEM) and cross-sectional (XTEM) configurations with a JEOL 2000FX operating at 200 kV. Panchromatic cathodoluminescence (CL) images were obtained with a JEOL 5300 scanning electron microscope fitted with a GaAs photomultiplier tube detector. Surface morphology was characterized by tapping-mode atomic force microscopy (AFM) using a Digital Instruments Dimension 3000 Nanoscope IIIa system and also by differential interference contrast microscopy using a Zeiss Axioplan microscope. Composition and Bragg peak width were determined by triple-axis x-ray diffraction (TAXRD) using a Bede D^3 diffractometer with a Rigaku rotating anode x-ray generator. Compositional grading was verified by cross-sectional Auger electron spectroscopy (AES) using a Physical Electronics Model 660 scanning Auger microprobe.

C. Defect density

Accurately determining relatively low dislocation densities is the primary challenge in evaluating process optimization and making comparisons with the literature. The common practice of quoting detection limits for PVTEM is inadequate because it is neither accurate nor standardized: various authors have reported detection limits ranging from 10^5 to 10^7 cm^{-2}. Also, the sampling area of PVTEM is far too small to accurately measure the heterogeneous dislocation distributions that result from dislocation pileups. Large area techniques, such as CL or etch-pit density, can be used to evaluate low defect densities, but unfortunately both techniques have relatively large features sizes that limit resolution to approximately 2×10^7 cm^{-2}.

PVTEM and CL characterization and analysis were combined in series I, II, and IV to accurately measure the density of threading dislocations in the field areas between pileups, ρ_{field}, the dislocation pileup density, ρ_{pileup}, and the linear density of dislocations in pileups, ρ_{linear}. The indirect band gap material in series III was difficult to characterize with CL, so only ρ_{field} was measured by PVTEM. The overall threading dislocation density, $\rho_{overall}$, was calculated by

$$\rho_{overall}=\rho_{field}+\rho_{pileup}\rho_{linear}. \quad (1)$$

The sample grown at 700 °C in series II and the optimized sample of series IV both have very low and similar overall defect densities, so extensive characterization was performed to accurately determine and compare the defect densities. For each sample, approximately 50 CL images were taken at 600× from different areas of the sample and approximately 100 PVTEM images were taken at 10 kX from three TEM foils prepared from different areas of the sample. The total sampling areas for each sample were approximately 7×10^3 μm^2 in PVTEM and 2.5×10^5 μm^2 in CL.

III. RESULTS

A. Reference samples

The surface roughness and defect density of GaP substrates and the $In_{0.33}Ga_{0.67}P$ film grown directly on GaP were characterized as references. For GaP, the surface root-mean-

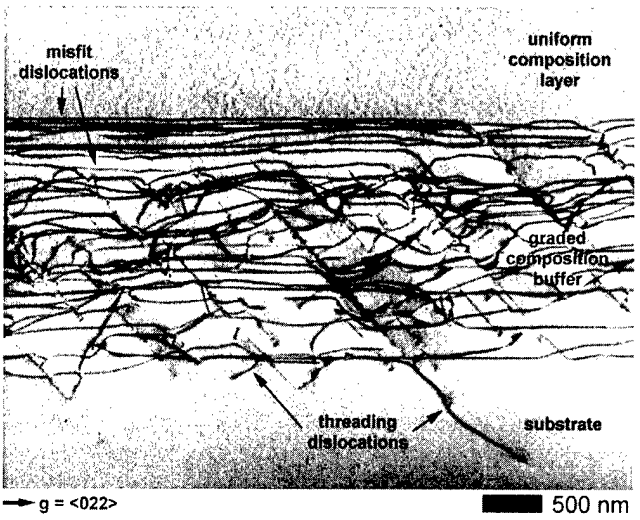


FIG. 1. XTEM of $In_xGa_{1-x}P$/GaP graded to $x=0.11$ at 760 °C. The dark lines confined primarily to the graded buffer are dislocations. The speckle contrast seen in the uniform composition layer and in the top half of the graded buffer is characteristic of indium-bearing III–V compound semiconductor alloys.

square (rms) roughness, R_q, is 5 Å and CL imaging shows $\rho_{overall}=3\times10^4\,cm^{-2}$. For $In_{0.33}Ga_{0.67}P$ grown directly on GaP, $R_q>250$ nm and PVTEM shows $\rho_{overall}=5\times10^9\,cm^{-2}$.

B. Step graded buffers

All of the graded buffers exhibit the crosshatch surface morphology characteristic of efficient strain relaxation by dislocation glide in low-mismatch systems.[16] Cross-sectional AES quantitatively verifies that compositional grading was achieved consistent with our growth process designs. XTEM with $\mathbf{g}=\langle004\rangle$ reveals sharp, regular step-grade interfaces with misfit and threading dislocation segments. XTEM with $\mathbf{g}=\langle022\rangle$ shows the expected networks of dislocations confined primarily to the graded buffers. The speckle contrast characteristic of indium-bearing III–V compound semiconductor alloys[17] is seen in the uniform composition layers and in the graded buffers for $x>0.05$. Figure 1 shows the substrate, graded buffer, and uniform composition layer in a typical sample.

C. Series I: Effects of continued grading beyond the indirect-direct band gap transition

To investigate material degradation with continued grading, samples were graded to different final compositions under otherwise identical conditions. The first sample was graded to the indirect-direct band gap transition at $x\sim0.27$,[18] while the second sample was graded an additional $\Delta x=0.05$. Both samples in this series have specular surfaces with easily visible crosshatch. All quantitative indications of material quality show significant degradation between $x=0.26$ and $x=0.31$. The most marked change is an increase in $\rho_{overall}$ from 1.6×10^7 to $2.5\times10^8\,cm^{-2}$, which correlates with increases in ρ_{pileup} and R_q. Relevant quantitative data are summarized in Table II.

TABLE II. Materials characterization results for series I, $In_xGa_{1-x}P$/GaP grown at 760 °C with variable grading rates.

Temperature (°C)	760	760
x–indium composition	0.26	0.31
R_q–(10 μm)^2 scan (nm)	25±6	45±25
$\beta_{(004)}$–Bragg peak FWHM (arc sec)	170±8	303±10
ρ_{field}–PVTEM (cm^{-2})	6.3×10^6 $\pm2.4\times10^6$	1.1×10^8 $\pm0.2\times10^8$
ρ_{pileup}–CL (cm^{-1})	377	1128
ρ_{linear}–PVTEM (cm^{-1})	2.7×10^4	1.2×10^5
$\rho_{overall}$–PVTEM+CL (cm^{-2})	1.6×10^7	2.5×10^8
ρ_{branch} (transverse)–PVTEM (cm^{-1})	6000±496	4773±932
ρ_{branch} (axial)–PVTEM (cm^{-1})	939±108	832±151

PVTEM reveals defects with a branching structure, hereafter referred to as branch defects, textured in one of the $\langle011\rangle$ directions, as shown in Fig. 2. The density of branch defects, ρ_{branch}, is the same in both samples, approximately $5500\,cm^{-1}$ transverse and $900\,cm^{-1}$ axial, but the branch defects themselves are narrower and sharper in the sample graded to $x=0.31$. The branch defects show strong contrast for $\mathbf{g}=\langle022\rangle$, which suggests that they are crystallographic. Weaker contrast is also seen for $\mathbf{g}=\langle004\rangle$, indicating that the local strain fields in branch defects are strong enough to be visible under other diffraction conditions. Bend contours turn sharply where they intersect the branches, which also indicates significant local strain in branch defects. The appearance of branch defects does not change in PVTEM foils prepared from material where the growth surface is polished down to $R_q=5$ Å, so branch defects are not merely topographic features. No strong topography is resolved in AFM that can be correlated with the branch defects, so any surface features caused by branch defects must be gentler than the general surface roughness of these samples.

Fine $\langle001\rangle$ ridges are also seen intersecting the branch defects in PVTEM. Ridges with the same density and orien-

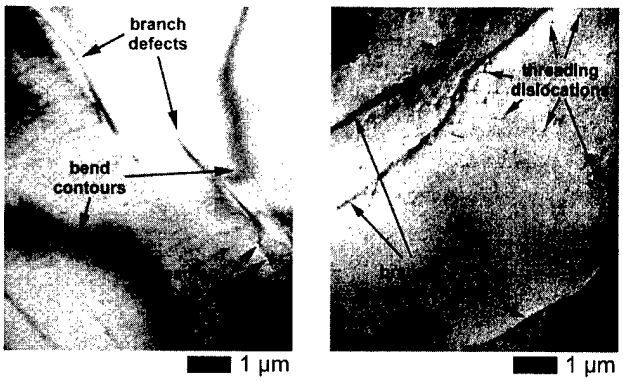


FIG. 2. PVTEM of $In_xGa_{1-x}P$/GaP graded to $x=0.26$ (left) and $x=0.31$ (right) at 760 °C. The branch defects present in both images have significant local strain fields and strongly pin dislocations. The sample graded to $x=0.31$ has a much higher defect density because the branch defects pin existing dislocations and force the nucleation of additional dislocations as lattice mismatch is increased by grading from $x=0.26$ up to $x=0.31$. The fine ridges are growth supersteps that do not appear to directly affect dislocation dynamics.

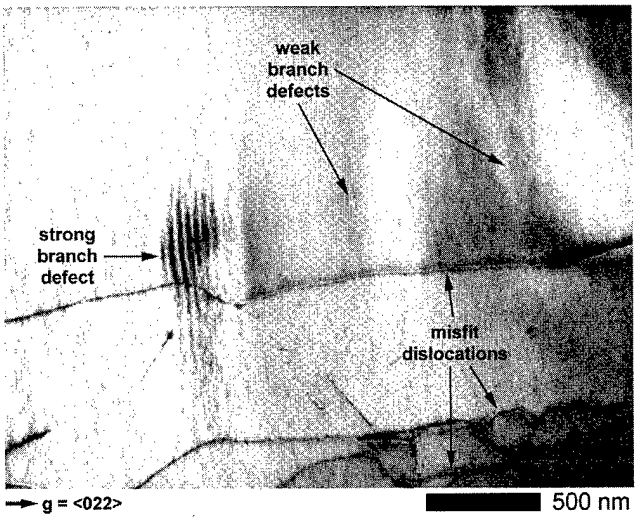


FIG. 3. XTEM image of $In_xGa_{1-x}P$/GaP graded to $x=0.26$ at 760 °C. The branch defects bend passing misfit dislocations and the degree of bending is proportional to their diffraction contrast strength, which is itself indicative of the strength of local strain fields associated with the branch defects.

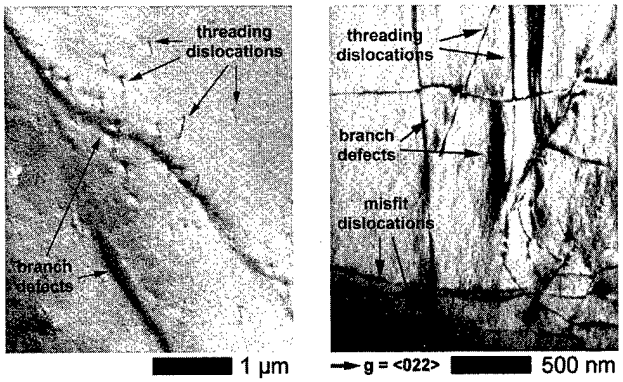


FIG. 4. PVTEM (left) and XTEM (right) images of pileups in $In_xGa_{1-x}P$/GaP grown at 760 °C. Branch defects in both images clearly pin dislocations and result in dislocation pileups.

tation are also seen in AFM. The ridges are growth supersteps that are roughly ⟨001⟩ oriented due to the {001} offcut of the substrate.[19] The supersteps are a surface phenomenon and do not appear to directly affect dislocation dynamics. The supersteps will be discussed in a later publication and they are simply noted here to prevent confusion in interpretation of PVTEM images.

Observed in XTEM, the branch defects are vertical features that extend from within the graded buffer up to the surface, as shown in Fig. 3. As in PVTEM, the diffraction contrast is strong for $\mathbf{g}=\langle 022\rangle$ and weaker for $\mathbf{g}=\langle 004\rangle$. The branches intersect the surface in gentle valleys that are less than 10 nm deep. Viewed down their ⟨011⟩ long axis, branch defects appear in close groups of one to four, consistent with the branching observed in PVTEM. The practical sampling length in XTEM is too small to obtain accurate statistics, but the $\rho_{branch}=3716\,cm^{-1}$ measured along a 62 μm collage of XTEM images is consistent with the transverse ρ_{branch} measured in PVTEM.

Figure 3 shows that the branches bend passing misfit dislocations, again indicating significant local strain. The degree of bending depends on the strength of the diffraction contrast in the branch defects, which is reasonable since diffraction contrast should be proportional to the magnitude of local strain variations. The strongest branch defects have a wavy interference contrast in XTEM and appear with a density of approximately $485\,cm^{-1}$.

The branch defects also strongly pin gliding threading dislocations in both samples, as seen in Fig. 4. Extensive PVTEM suggests that most pileups occur on branch defects. The dislocation density near a pileup exceeds the value in the field areas between pileups by approximately one order of magnitude. Both CL and PVTEM show that pileup formation is moderate at $x=0.26$, but very severe by $x=0.31$. At $x=0.26$, the overall defect density is low enough to see individual dislocations pinned to branch defects. There is also a high density of stacking faults near the branch defects at $x=0.31$, which suggests that the local strain fields are tensile.[20]

D. Series II: Effects of growth temperature in $In_xGa_{1-x}P$/GaP graded to $x\sim0.33$

To study the branch defects that dominate relaxation in series I, samples were graded to $x\sim0.33$ with a range of growth temperatures under otherwise identical conditions. All samples have a clear surface crosshatch pattern and, except for the sample grown at 810 °C, have specular surfaces.

TABLE III. Materials characterization results for series II: $In_xGa_{1-x}P$/GaP graded to $x\sim0.33$.

Temperature (°C)	700	760	810
x–indium composition	0.34	0.31	0.36
R_q–(10 μm)^2 scan (nm)	13±2	68±25	118±20
$\beta_{(004)}$–Bragg peak FWHM (arc sec)	275±10	420±30	···
ρ_{field}–CL (cm^{-2})	$4.9\times10^6\pm0.9\times10^6$	···	···
ρ_{field}–PVTEM (cm^{-2})	$4.4\times10^6\pm1.0\times10^6$	···	···
ρ_{pileup}–CL (cm^{-1})	92±23	>2000	···
ρ_{linear}–CL+PVTEM (cm^{-1})	2.1×10^5	···	···
$\rho_{overall}$–CL+PVTEM (cm^{-2})	$6.8\times10^6\pm2.0\times10^6$	$1.5\times10^9\pm0.3\times10^9$	···
ρ_{branch} (transverse)–PVTEM (cm^{-1})	7970±725	5433±801	···

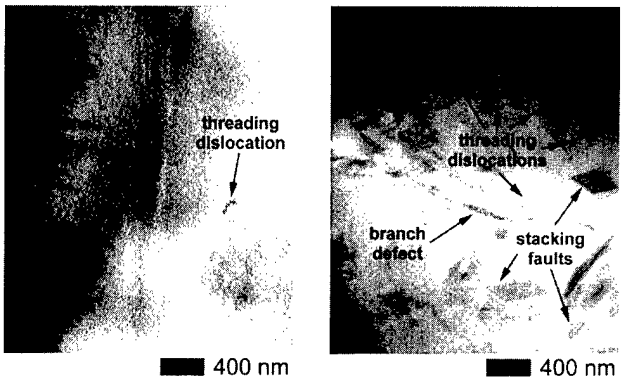


FIG. 5. PVTEM images of $In_xGa_{1-x}P$/GaP graded to $x\sim0.33$ at 700 °C (left) and 760 °C (right). The difference in defect density is very striking, with the sample grown at 760 °C possessing a high density of dislocations and stacking faults. The branch defect is much sharper and stronger in the sample grown at 760 °C compared to 700 °C.

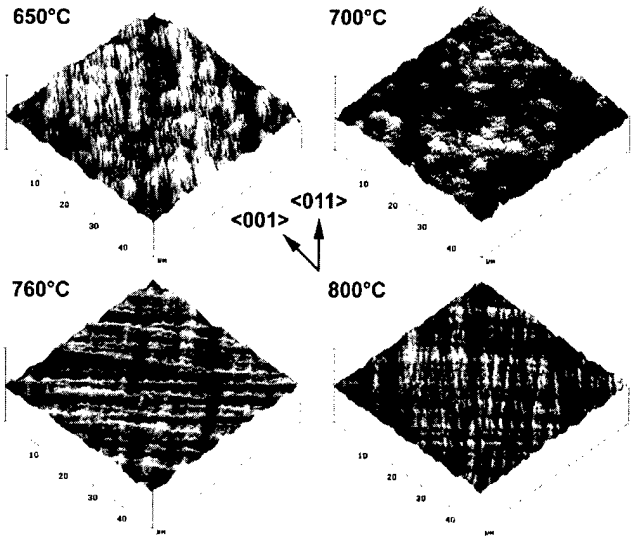


FIG. 6. AFM images of $In_xGa_{1-x}P$/GaP graded to $x\sim0.1$. All scans cover an area of $(50\,\mu m)^2$ with a vertical scale of 200 nm/division. The transition from a rough, cellular surface to a smooth, crosshatch surface suggests that relaxation becomes more efficient as growth temperature increases.

The quantitative indicators of material quality listed in Table III degrade rapidly with increasing growth temperature. The increase in $\rho_{overall}$ correlates strongly with increases in ρ_{pileup} and R_q. The sample at 810 °C is essentially polycrystalline and proved difficult to characterize.

Branch defects are present in all of the samples in this series. At 760 °C, the branch defects have the same density and general appearance as those in series I, but are stronger and sharper. The branch defects look very different at 700 °C. Figure 5 shows that the branches are weak, diffuse, and broad at 700 °C compared to 760 °C. The branching nature is also less pronounced and the branch defects appear as long ⟨011⟩ stripes with transverse $\rho_{branch}=7621\,cm^{-1}$ and an axial ρ_{branch} that is too low to measure. The branch defects are very faint for $\mathbf{g}=\langle004\rangle$ and do not strongly affect bend contours, which suggests that the branch defects possess less strain at 700 °C than at 760 °C.

Branch defects strongly pin dislocations in the sample grown at 760 °C, resulting in a very high density of pileups. While pileups are present at 700 °C, the weaker branch defects do not appear to pin dislocations as strongly as at 760 °C. The resulting differences in defect density are strikingly evident in Fig. 5.

CL of the sample grown at 700 °C shows a very low density of dislocation pileups compared to samples grown at higher temperatures. The samples grown above 700 °C are too rough and defective to obtain clear CL images, but the defect density is high enough to measure accurately with PVTEM alone.

E. Series III: Effects of growth temperature in $In_xGa_{1-x}P$/GaP graded to $x\sim0.1$

Speculating that optimal growth conditions would vary with composition, samples were graded to $x\sim0.1$ with a range of growth temperatures under otherwise identical conditions. All of the samples in this series have specular surfaces with very faint crosshatch. Table IV shows that material quality degrades steadily with decreasing growth temperature. The trend in R_q does not fully reflect the change in surface morphology, which Fig. 6 shows changes from a smooth, ordered crosshatch to a cellular, disordered structure with decreasing growth temperature.

The PVTEM images in Fig. 7 illustrate the change in microstructure with growth temperature. Branch defects oriented in one ⟨011⟩ direction are present in samples grown below 760 °C. At 650 °C, $\rho_{branch}=26\,911\,cm^{-1}$ and the

TABLE IV. Materials characterization results for series III: $In_xGa_{1-x}P$/GaP graded to $x\sim0.1$.

Temperature (°C)	650	700	760	800
x–indium composition	0.15	0.09	0.11	0.12
R_q–$(10\,\mu m)^2$ scan (nm)	8.5±1.5	7.7±1	6.0±0.5	4.9±0.5
R_q–$(50\,\mu m)^2$ scan	12.2±1.5	10.5±1	7.4±0.5	6.0±0.5
$\beta_{(004)}$–Bragg peak FWHM (arc sec)	93±5	58±3	54±3	53±3
ρ_{field}–PVTEM (cm^{-2})	1.2×10^7 $\pm0.3\times10^7$	3.9×10^6 $\pm1.4\times10^6$	1.1×10^6 $\pm0.6\times10^6$	4.0×10^5 $\pm1.9\times10^5$
ρ_{branch} (transverse)–PVTEM (cm^{-1})	26911±3611	9808±1144	...	...
ρ_{branch} (transverse)–AFM (cm^{-1})	24114±3246	...	...	...

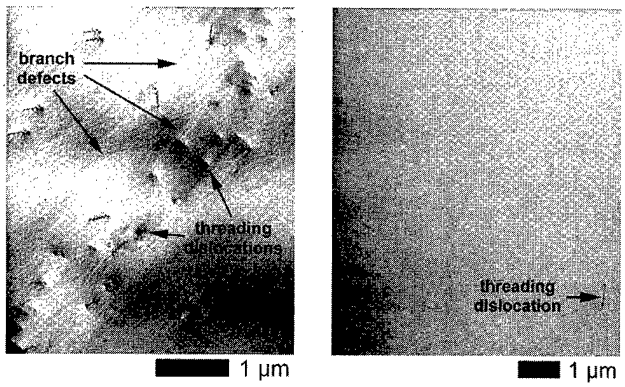


FIG. 7. PVTEM images of $In_xGa_{1-x}P/GaP$ graded to $x\sim0.1$ at 650 °C (left) and 760 °C (right). The weak and closely spaced branch defects in the sample grown at 650 °C appear to hinder dislocation glide, resulting in dislocation pileups parallel to the branches. There are no traces of branch defects or dislocation pileups in the sample grown at 760 °C.

branch defects are visible for both $\mathbf{g}=\langle022\rangle$ and $\mathbf{g}=\langle004\rangle$. At 700 °C, $\rho_{branch}=9808\,cm^{-1}$ and the branch defects are nearly invisible for $\mathbf{g}=\langle022\rangle$ and are not seen for $\mathbf{g}=\langle004\rangle$. The branch defects appear in AFM as 3 nm ridges at 650 °C, while no related topography can be resolved at 700 °C or above. Figure 7 shows that branch defects and dislocations interact at 650 °C, resulting in pileups parallel to the branch defects. The much weaker branch defects at 700 °C do not appear to interact with dislocations and PVTEM shows very few pileups. In general, the branch defects are much weaker in appearance and pin dislocations less effectively than in series I and II.

Figure 7 also shows that branch defects are not seen in the samples grown above 700 °C and consequently pileups are not observed. The microstructures of the samples grown at 760 and 800 °C are qualitatively similar, with the latter sample having better quantitative materials properties.

F. Series IV: Optimization of growth temperature as a function of composition

With the understanding of the temperature dependence of material quality at $x\sim0.33$ and $x\sim0.1$ gained in series II and III, respectively, growth temperature was optimized as a function of composition. The result is that the optimized 650 °C sample has the lowest $\rho_{overall}$ of any sample graded to $x>0.25$, despite being graded up to $x=0.39$. The relevant data are summarized in Table V.

Figure 8 shows that the microstructure is dominated by a high density of branch defects, but the dislocation density is low and relatively homogeneous. The branch defects are weak and only visible for $\mathbf{g}=\langle022\rangle$. As in series III at 650 °C, the branch defects appear as ridges in AFM. The branch defects appear to hinder dislocation glide, but do not strongly pin dislocations, so pileups are weak and infrequent in both PVTEM and CL images. While ρ_{pileup} is low, the biggest difference is that ρ_{linear} is more than an order of magnitude lower than any sample in series I or II.

TABLE V. Materials characterization results for series IV: $In_xGa_{1-x}P/GaP$ graded to $x\sim0.33$ with growth temperature optimization.

Temperature (°C)	760/700/650
x–indium composition	0.389
R_q–(10 μm)^2 scan (nm)	13±4
$\beta_{(004)}$–Bragg peak FWHM (arc sec)	277±10
ρ_{field}–CL (cm^{-2})	$4.1\times10^6\pm0.5\times10^6$
ρ_{field}–PVTEM (cm^{-2})	$3.7\times10^6\pm0.8\times10^6$
ρ_{pileup}–CL (cm^{-1})	71±18
ρ_{linear}–CL+PVTEM (cm^{-1})	8.5×10^3
$\rho_{overall}$–CL+PVTEM (cm^{-2})	$4.7\times10^6\pm1.1\times10^6$
ρ_{branch} (transverse)–PVTEM (cm^{-1})	24636±2053
ρ_{branch} (transverse)–AFM (cm^{-1})	28622±1795

IV. DISCUSSION

A. Series I: Mechanism of pileup formation on branch defects

The dramatic increase in $\rho_{overall}$ between $x=0.26$ and $x=0.31$ supports the speculation by Stinson *et al.*: degradation of $In_xGa_{1-x}P/GaP$ LED efficiency beyond 600 nm is due to increasing dislocation density with continued grading.[1] As noted earlier, graded buffer relaxation models predict a nearly steady-state $\rho_{overall}$.[10,11] Both microscopic and macroscopic evidence will show that branch defects are responsible for the escalation of defect density with continued grading.

The increase in $\rho_{overall}$ with continued grading correlates strongly with pileup formation. In both samples, approximately 60% of the total threading dislocations are trapped in pileups. This agrees with the observations of Samavedam and Fitzgerald in the Ge_xSi_{1-x}/Si system, where increasing $\rho_{overall}$ is due to increasing pileup formation.[13] While surface roughness is the primary cause of pileups in Ge_xSi_{1-x}/Si,

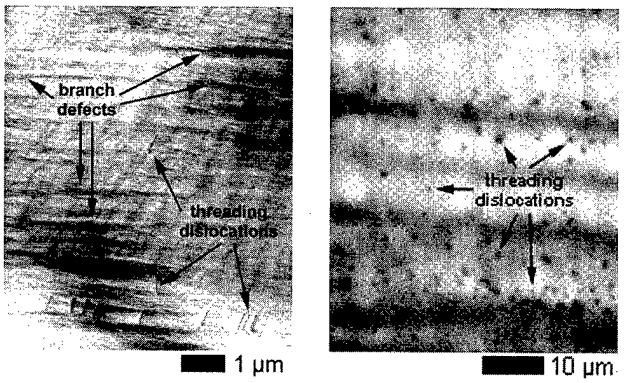


FIG. 8. PVTEM (left) and CL (right) images of $In_xGa_{1-x}P/GaP$ graded to $x\sim0.33$ with growth temperature optimized as a function of composition. The long horizontal features in the PVTEM image are branch defects and the fine vertical features are growth supersteps. The branch defects appear to slow dislocation glide, as evidenced by the light dislocation pileup that appears parallel to the branches. The black dots in the CL image are due to strong nonradiative trapping by individual threading dislocations. The larger scale afforded by CL imaging shows that the dislocation distribution is nearly homogeneous and pileup formation is very moderate.

most of the pileups observed in series I occur on or near branch defects, which have not been observed in Ge_xSi_{1-x}/Si. Branch defects pin individual dislocations even at $x=0.26$ and ρ_{pileup} increases with continued grading while ρ_{branch} remains constant, so branch defects clearly precede pileups. Branch defect topography is gentle and overall surface roughness is very low at $x=0.26$, so the sharp branch defect strain fields appear to be responsible for dislocation pinning.

Branch defects clearly dominate dislocation dynamics at the microscale, but the sampling area of TEM is too small to prove alone that branch defects are the primary cause of pileups at a macroscale. Indeed, the data show a correlation between surface roughness and defect density that qualitatively agrees with the pileup mechanism found in the Ge_xSi_{1-x}/Si system. Examined quantitatively, the increase in $\rho_{overall}$ with R_q in $In_xGa_{1-x}P/GaP$ is much more dramatic than expected, which suggests that something other than surface roughness dominates relaxation, namely the branch defects.

Relaxation modeling related to the approach of Abrahams *et al.*[11] confirms on a macroscale that branch defects are responsible for the escalation of defect density with continued grading in $In_xGa_{1-x}P/GaP$. The expected change in $\rho_{overall}$ between $x=0.26$ and $x=0.31$ can be calculated by assuming that dislocation pinning by the branch defects dominates relaxation. First, the misfit dislocation density, ρ_{misfit}, necessary in each $\langle 110\rangle$ direction is given by[8,11]

$$\rho_{misfit}=\frac{\Delta a_0}{a_0}\times\frac{2\sqrt{2}}{a_0}. \tag{2}$$

For the lattice mismatch between $x=0.26$ ($a_0=5.5593$ Å) and $x=0.31$ ($a_0=5.5802$ Å), the calculation yields $\rho_{misfit}=1.9\times10^5\,cm^{-1}$. Each misfit dislocation can be assumed to terminate with two threading dislocations, so $\rho_{overall}$ is twice ρ_{misfit} divided by the average misfit dislocation glide length, L. The branch defects strongly pin dislocations, so L can be set equal to the inverse of ρ_{branch} in each $\langle 110\rangle$. So, $\rho_{overall}$ is expressed as

$$\rho_{overall}=\frac{2\rho_{misfit}}{L_{\langle 110\rangle}}+\frac{2\rho_{misfit}}{L_{\langle \bar{1}10\rangle}}=2\rho_{misfit}(\rho_{branch}^{transverse}+\rho_{branch}^{axial}). \tag{3}$$

Using the average values of transverse and axial ρ_{branch} in series I, the result is $\rho_{overall}=2.4\times10^9\,cm^{-2}$. This calculated value is much larger than the observed $\rho_{overall}$ at $x=0.31$, because all branch defects are assumed to strongly pin dislocations, which is certainly inaccurate otherwise ρ_{pileup} would equal ρ_{branch}. Recalculation assuming that only branch defects with strong wavy interference contrast in XTEM effectively pin dislocations results in $\rho_{overall}=2.2\times10^8\,cm^{-2}$, which agrees very well with the measured value.

The model analysis suggests that dislocation pinning is proportional to branch defect strength and that pileups caused by strong branch defects ultimately limit how far $In_xGa_{1-x}P/GaP$ can be graded before serious material degradation occurs. Unfortunately, the limit to grading occurs near the indirect-direct band gap transition at $x\sim0.27$[18] where $In_xGa_{1-x}P$ becomes useful for efficient, direct band gap devices. Similar grading limits have also been observed at $x\sim0.35$ in HVPE[1] and $x\sim0.32$ in GSMBE.[12] Clearly, branch defects must be eliminated or suppressed to achieve high-quality device materials.

B. Series II: Control of branch defect morphology via growth temperature

Series I shows that branch defects dominate relaxation in samples grown at 760 °C. Varying growth temperature in series II dramatically changes both branch defect density and strength. Correspondingly, material properties degrade rapidly with increasing temperature, in particular between 700 and 760 °C. Analysis of the results will show that the material properties are dominated by the temperature dependence of the branch defects.

Since ρ_{branch} increases with decreasing growth temperature, L would decrease and both $\rho_{overall}$ and ρ_{pileup} would increase with decreasing temperature if the branch defects continued to strongly pin dislocations. The results show the opposite trend: $\rho_{overall}$ and ρ_{pileup} decrease dramatically with decreasing temperature, so the ability of branch defects to pin dislocations must decrease much more rapidly with decreasing temperature than ρ_{branch} increases. Branch defect diffraction contrast and strain fields seen in TEM are markedly weaker at lower temperatures, which suggests that dislocation pinning is proportional to the strength of branch defects.

Suppression of branch defects at 700 °C minimizes pileup formation and controls the escalation of $\rho_{overall}$ with continued grading, compared to higher growth temperatures. Note that the sample grown at 700 °C was graded farther than any sample grown at 760 °C, yet retains a much lower $\rho_{overall}$. Pileups still occur, but they are weaker and more widely spaced than at 760 °C, so L is relatively large and the escalation of $\rho_{overall}$ with grading is moderate.

The sample grown at 760 °C shows significant degradation even compared to series I. While ρ_{branch} is nominally identical in all of the samples grown at 760 °C, branch defects are much stronger in series II. Since the measured $\rho_{overall}$ matches the value calculated assuming all branch defects pin dislocations in the analysis of series I, it appears that most of the branch defects in the sample grown at 760 °C in series II pin dislocations. This further supports the observations that branch defects dominate relaxation in $In_xGa_{1-x}P/GaP$ at 760 °C and that dislocation pinning is proportional to branch defect strength.

The only difference between the samples graded to $x=0.31$ in series I and II is the grading rate profile. The variable grading rate profile used in series I produces higher strain rates at low indium compositions and lower strain rates at high indium compositions, compared to the constant grading rate in series II. Two general observations can be made: lower indium compositions appear to be less sensitive to strain rate, and the strength, but not density, of branch defects is sensitive to the strain rate.

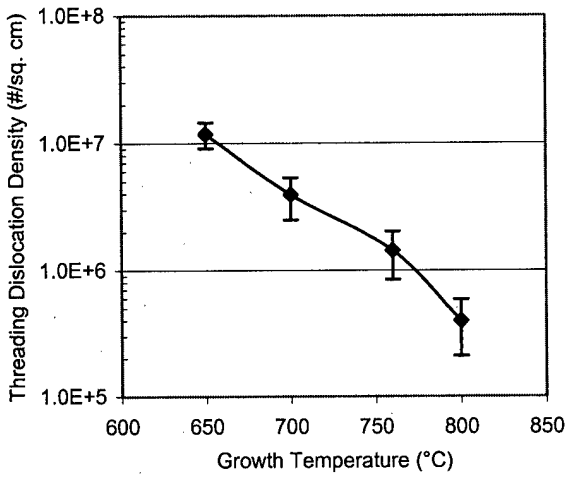


FIG. 9. Plot of ρ_{field} vs growth temperature for $In_xGa_{1-x}P/GaP$ graded to $x \sim 0.1$. The steady decrease in ρ_{field} with increasing temperature is opposite the trend seen in series II due to the diminished role of branch defects at $x \sim 0.1$.

The trend between 700 to 760 °C suggests that branch defects should be strong and widely spaced at 810 °C. Branch defect strength appears to remain the dominant factor at 810 °C, where the resulting material is essentially polycrystalline at $x \sim 0.33$. It is likely that this trend will continue to even higher temperatures.

The morphology of branch defects is affected in two primary ways by increasing growth temperature: branch defect strength increases and branch defect density decreases. The results show that branch defect strength dominates dislocation dynamics and that decreasing growth temperature to suppress branch defects is the effective strategy for producing high-quality material. It is also apparent that reducing grading rate at higher indium compositions can further suppress branch defect strength. As branch defects are suppressed, $\rho_{overall}$ approaches the nearly steady-state behavior predicted by ideal relaxation models.

C. Series III: Relaxation limited by dislocation glide kinetics in the absence of branch defects

Series I and II show that branch defects dominate relaxation in $In_xGa_{1-x}P/GaP$. Relaxation approaches the ideal steady-state behavior as branch defects are suppressed. Series III shows that branch defects can be avoided altogether at $x \sim 0.1$ and analysis will show that the resulting relaxation behavior is dominated by the kinetics of dislocation glide.

Branch defects are essentially absent in all of the samples, except for the sample grown at 650 °C where weak branch defects and pileups are observed. Without the influence of branch defects on dislocation dynamics, the resulting temperature dependence of ρ_{field} shown in Fig. 9 is the opposite of that seen in series II. Clearly, relaxation is very different when branch defects are not active.

Not only are branch defects absent, but the observed surface roughness is also far too low to affect dislocations in series III. The typical effective critical thickness, or "dislocation-free" thickness, in these experiments is on the order of 10^3 Å, while the worst surface undulations are an order of magnitude smaller. Also, the change in R_q with temperature is very moderate. Any interaction between dislocations and surface roughness should be weak.

With the absence of microstructural features that interact with dislocations, the kinetics of dislocation nucleation and glide should control relaxation. The observed temperature dependence of ρ_{field} is consistent with the behavior expected if relaxation is limited by dislocation glide kinetics: as temperature decreases, decreasing dislocation glide kinetics force the nucleation of a higher density of dislocations to maintain efficient strain relaxation, so ρ_{field} increases. If relaxation was limited by dislocation nucleation, then dislocation nucleation would be suppressed and ρ_{field} would decrease.

Relaxation modeling developed from the approach of Fitzgerald *et al.*[21] quantitatively demonstrates that dislocation glide kinetics dominate relaxation in series III. Assuming relaxation is isotropic in the two $\langle 011 \rangle$, the relaxation rate, $\dot{\delta}$, can be expressed as[8]

$$\dot{\delta} = \frac{\rho_{field} v b}{4}. \tag{4}$$

The dislocation glide velocity, v, is empirically expressed as[22]

$$v = A\left(\frac{\tau}{\tau_0}\right)^n \exp\left(-\frac{E_{glide}}{kT}\right). \tag{5}$$

A and n are materials constants, E_{glide} is the dislocation glide activation energy, τ is the resolved shear stress, and τ_0 is a constant equal to 1 MPa. It is reasonable to assume that the residual strain reflects the steady-state strain in the pseudomorphic film at the surface of the graded buffer during growth and TAXRD results show that residual strain does not vary significantly with temperature in series III. Therefore, τ can be considered constant in series III, especially since $n \sim 1$ and any moderate variations in τ are overwhelmed by changes in the exponential term.

The nearly constant residual strain in this series also suggests that $\dot{\delta}$ is the same in all of the samples. Therefore, ρ_{field} can be expressed as

$$\rho_{field} = \frac{4\dot{\delta}}{b} \times \frac{1}{v} = c \exp\left(\frac{E_{glide}}{kT}\right). \tag{6}$$

c is a constant. E_{glide} can be determined by measuring the slope in a plot of $\ln(\rho_{field})$ versus $1/T$:

$$\ln \rho_{field} = \ln c + \frac{E_{glide}}{kT}. \tag{7}$$

Figure 10 plots $\ln(\rho_{field})$ versus $1/T$ for series III, with the sample grown at 650 °C omitted due to the presence of significant branch defects. The measured slope corresponds to $E_{glide} = 2.0$ eV, which is in fair agreement with published values for E_{glide} in GaP and InP: 1.45 eV for α dislocations and 1.68 eV for β dislocations in GaP,[23] 1.6 eV for α dislocations and 2.3 eV for β dislocations in InP.[24]

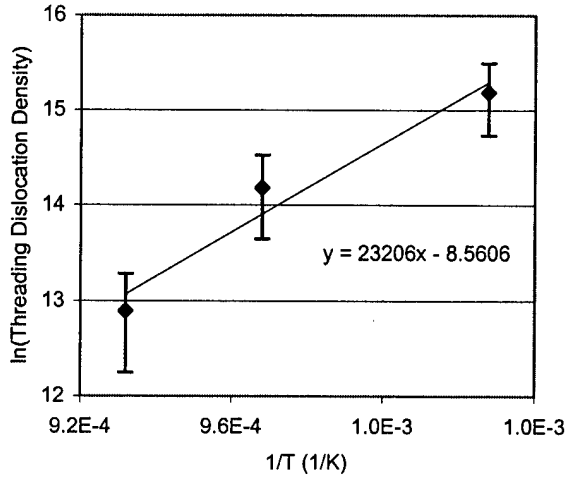


FIG. 10. Plot of $\ln(\rho_{field})$ vs $1/T$ for $In_xGa_{1-x}P/GaP$ graded to $x\sim0.1$ at 800, 760, and 700 °C. The activation energy calculated from the slope is 2.0 eV, which is in good agreement with the average value of 1.8 eV for the activation energy of dislocation glide for α and β dislocations in GaP and InP. This result suggests that strain relaxation at $x\sim0.1$ is dominated by the kinetics of dislocation glide.

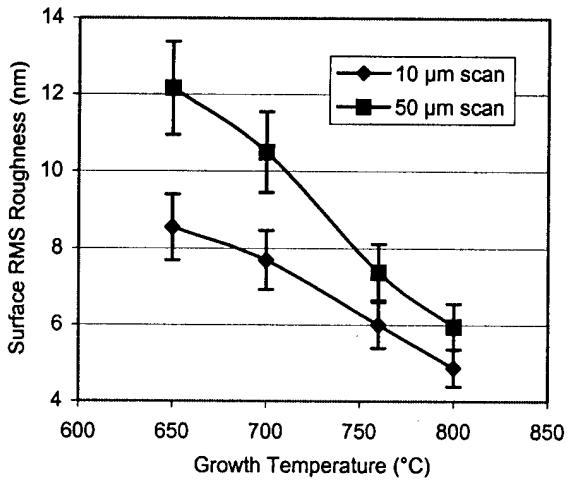


FIG. 11. Plot of surface rms roughness vs growth temperature for $In_xGa_{1-x}P/GaP$ graded to $x\sim0.1$. Scan areas of $(10\,\mu m)^2$ and $(50\,\mu m)^2$ were used to ensure that both small and large scale roughness were measured.

This two parameter model shows that the temperature dependence of ρ_{field} is consistent with the thermal activation energy of dislocation glide. The model simplifies relaxation by assuming that dislocation glide is entirely rate-limiting and that relaxation is isotropic. A more complex model can be derived, but such analysis does not provide additional insights given the size and resolution of the current data set.

The dominance of dislocation glide kinetics at $x\sim0.1$ is also apparent in the temperature dependencies of surface morphology. A smooth, regular crosshatch surface is indicative of efficient relaxation by dislocation glide with long misfit dislocation segments, so the transition from cellular to crosshatch surface morphology shown in Fig. 6 suggests that dislocation glide becomes more efficient with increasing growth temperature. It could be speculated that since crosshatch is the response of the growth surface to dislocation strain fields, the change in surface morphology is due to decreasing surface diffusivity. If diffusivity was primarily responsible for the change in surface morphology, the surface would become smoother with decreasing temperature, which Fig. 11 shows is not the case at $x\sim0.1$. So, while surface diffusivity certainly changes with temperature, the observed trends are not consistent with a surface morphology dominated by those changes in surface diffusivity. The results show that surface morphology and relaxation are controlled primarily by dislocation glide kinetics.

The analysis of series III demonstrates that dislocation glide kinetics dominate relaxation and microstructure in $In_xGa_{1-x}P/GaP$, in the absence of external or heterogeneous hindrances to dislocation dynamics such as branch defects. The branch defects at 650 °C are so weak and closely spaced that they can be considered an effective drag force on dislocation glide, hence, their effect on dislocation mobility is complimentary to that of decreasing temperature. As noted earlier in this section, the temperature dependence of defect density depends on whether relaxation is limited by dislocation nucleation or glide kinetics, therefore determining that dislocation glide kinetics dominate relaxation is critical for the rational design of high-quality graded buffers.

D. Series IV: Dislocation dynamics in graded buffers with optimized growth temperatures

The results of series II and III taken together show that a constant growth temperature is not optimal, since the temperature dependence of material quality at $x\sim0.1$ is the opposite of that seen at $x\sim0.33$. While 700 °C appears to be a reasonable compromise, growth temperature should be optimized as a function of composition to obtain higher quality $In_xGa_{1-x}P/GaP$. The superior material quality of series IV illustrates the effectiveness of even a simple growth temperature optimization.

Series II suggests that decreasing growth temperature below 700 °C should improve material quality by suppressing branch defect strength, but series III shows that ρ_{field} reaches $1.2\times10^7\,cm^{-2}$ by $x\sim0.1$ at 650 °C due to slower dislocation glide kinetics. An intuitive approach is to use the best growth temperature for each composition range: high growth temperature at low indium compositions and low growth temperature at high indium compositions. The results show that optimization suppresses branch defects and limits the degradation with continued grading seen in series I. Grading was continued $\Delta x=0.05$ beyond the sample grown at 700 °C in series II, yet the resulting defect densities are lower. With respect to the sample grown at 760 °C in series III, $\rho_{overall}$ only increases from 1.1×10^6 to $4.7\times10^6\,cm^{-2}$ for $\Delta x=0.28$. There is still pileup formation, so the $\rho_{overall}$ would have been even lower if grading had been stopped earlier.

The results also show that dislocation glide mobility must increase with indium composition, otherwise dropping the growth temperature to 650 °C should result in $\rho_{overall}\geqslant1.2\times10^7\,cm^{-2}$ according to the results of series III. Dislocation glide kinetics have not been previously studied

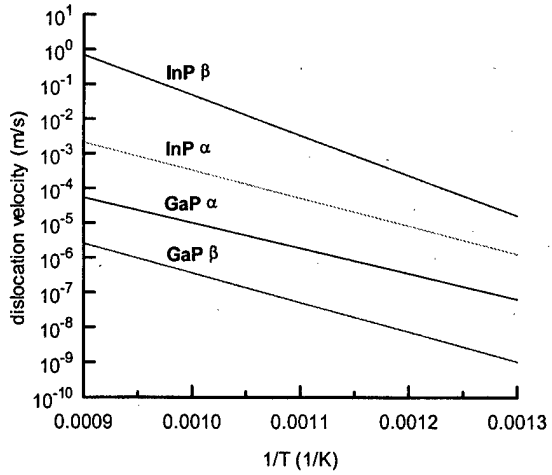


FIG. 12. Plot of dislocation glide velocity vs $1/T$ for α and β dislocations in GaP and InP (Refs. 23 and 24). A shear stress of 1 MPa was used for the calculations. Dislocations in InP glide several orders of magnitude faster than those in GaP.

in $In_xGa_{1-x}P$, but the general composition dependence can be inferred from available data. Figure 12 shows that dislocation glide velocities are several orders of magnitude greater in InP[24] than in GaP.[23] Furthermore, the $In_xGa_{1-x}P$ phase diagram[25] in Fig. 13 shows that most of the change in solidus temperature occurs between $x=0$ and $x=0.3$, so most of the change in materials properties between GaP and InP should occur in that range. Therefore, dislocation glide kinetics should increase by orders of magnitude between $x=0$ and $x=0.3$. The faster kinetics at $x=0.2$ allow growth temperature to be dropped to 650 °C without increasing $\rho_{overall}$ to the high value seen in series III at that temperature.

The optimized sample shows that changing growth temperature as a function of composition results in improved material quality. Dislocation glide kinetics set a lower limit for defect density, but glide kinetics increase with indium composition and allow growth temperature to be dropped to control branch defects. Further optimization requires a better understanding of the evolution of branch defects.

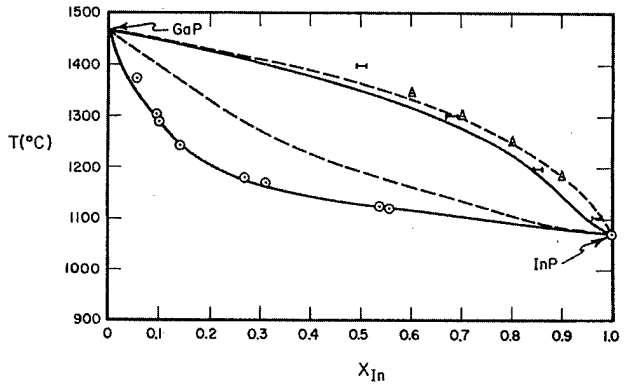


FIG. 13. Equilibrium phase diagram for $In_xGa_{1-x}P$. The solidus temperature drops mostly between $x=0$ and $x=0.3$, which suggests that materials properties should change dramatically in that composition range. The phase diagram was scanned from Ref. 25 and digitally edited.

E. Analysis of branch defects

The effect of branch defects on relaxation is clear: branch defects interact with dislocations to pin or slow glide, which decreases L and increases ρ_{pileup} and $\rho_{overall}$. The results of series II, III, and IV illustrate how branch defects can be avoided or suppressed to minimize their effect on dislocation dynamics. Additional progress depends on understanding the evolution of branch defects.

Several basic properties of branch defects can be observed directly. TEM shows that branch defects are crystallographic features with sharp tensile strain fields. Although AFM and XTEM reveal that branch defects form very shallow cusps at the surface, these cusps are much shallower than the effective critical thickness, so branch defects appear to hinder or pin dislocations primarily through their strain fields. Strong branch defects may also act as low activation energy sites for heterogeneous dislocation nucleation, but such an effect is difficult to identify and prove.

The evolution of branch defects can also be followed empirically. The formation and development of branch defects appears to be primarily controlled by growth temperature. Evaluation of all the available data reveals strong temperature dependencies for branch defect formation, strength, and density.

The onset of branch defect formation varies with growth temperature. At 760 °C, branch defects are completely absent at $x=0.11$ and are present at $x=0.26$, so they must form somewhere between those two compositions. The relatively low $\rho_{overall}$ at $x=0.26$ in series I indicates that branch defect formation must begin close to $x=0.26$. XTEM of the same sample shows that branch defects only penetrate the last several steps of the graded buffer. Overall, branch defects appear to form near $x\sim0.2$ at 760 °C. At 700 °C, the beginning of branch defect formation can be seen at $x=0.09$ in series III, so branch defects form near $x\sim0.1$ at 700 °C. At 650 °C, well developed branch defects are already present at $x=0.15$. Given the trend seen between 760 and 700 °C, branch defects probably appear for $x<0.1$ at 650 °C and for $x>0.2$ at 800 °C. This data can be used to develop a phase diagram for branch defects, as shown in Fig. 14 for our system. Although indium composition is used to describe the onset of branch defect formation, the current data cannot be used to differentiate whether branch defect formation depends on composition or total film thickness since the same grading rate was used in all of the samples.

Branch defect strength also varies with temperature. Series II and IV show that branch defects become stronger with increasing temperature. Series III appears to show the opposite trend, but this is due to the temperature dependence of the onset of branch defect formation: at $x\sim0.1$, branch defects in the samples grown at lower temperature have formed and evolved further than in samples grown at higher temperature. This also indicates that although the onset of branch defect formation occurs later with increasing temperature, branch defect strength develops much more rapidly at higher temperatures.

Comparison of the samples graded to $x=0.31$ at 760 °C in

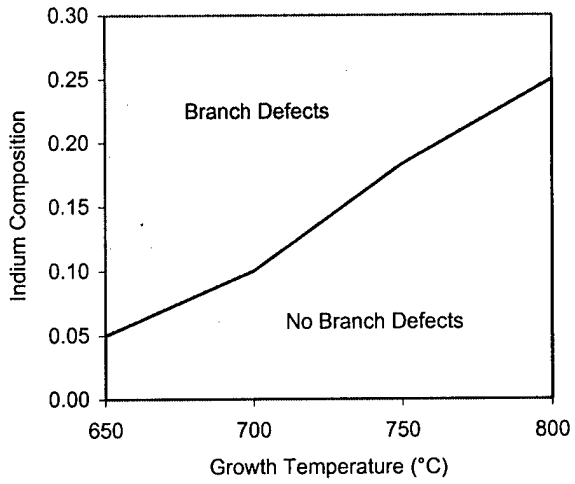


FIG. 14. Phase diagram for the presence of branch defects in $In_xGa_{1-x}P/GaP$ grown in this study. The onset of branch defect formation occurs later with increasing growth temperature.

series I and II shows that grading rate also affects the strength of branch defects. The constant grading rate sample has a higher grading rate at high indium compositions than the variable grading rate profile, which results in much stronger branch defects. Decreasing grading rate with increasing composition appears to suppress the branch defects, presumably because of the lower strain rate.

Branch defect density is also primarily a function of growth temperature. Figure 15 plots ρ_{branch} versus temperature for all samples with branch defects and shows that ρ_{branch} decreases exponentially with increasing temperature. There also appears to be a weaker dependence on composition where ρ_{branch} decreases with increasing composition. Increasing indium composition reduces the solidus temperature, which is roughly analogous to increasing growth

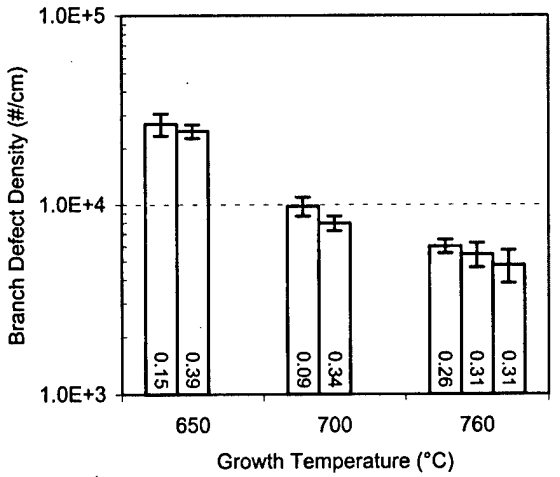


FIG. 15. Plot of ρ_{branch} vs growth temperature for all samples that exhibit branch defects. The final indium composition is inset in the bar for each sample to illustrate that branch defect density is primarily a function of growth temperature, but also appears to have a weaker composition dependence.

temperature at a fixed composition. Both trends therefore suggest that branch defect density is determined by a thermally activated process.

The precise nature and formation mechanism of branch defects are currently unclear. The branch defects themselves are difficult to isolate and quantitatively characterize. Nevertheless, several possibilities can be eliminated by analysis of the current data. Three known phenomena appear similar to branch defects: strain-relieving surface undulations, antiphase boundaries (APBs), and phase separation. The experimental evidence shows that none of these phenomena are likely to be responsible for branch defects in $In_xGa_{1-x}P/GaP$.

Branch defects are not merely the cusps of strain-induced surface undulations. PVTEM of polished samples show that branch defects are not simply topographic. Branch defect density does not match any period of the ubiquitous strain-induced crosshatch pattern. There is also no reason that surface undulations would occur in only one $\langle 011 \rangle$. Furthermore, the topography of the branch defects does not relieve strain. Strain-relieving surface undulations in a compressive film feature compressive strain in the valleys and tensile strain in the peaks, but branch defects form valleys with sharp tensile strain fields.

Branch defects are not APBs because they do not enclose and separate domains of different microstructure. Furthermore, no reported antiphase mechanism is consistent with the observed branch defects. The APBs due to sublattice displacement found in heteropolar epitaxy[26] are not expected for homopolar $In_xGa_{1-x}P$ growth on GaP. APBs between atomically ordered domains have been reported extensively in $In_{0.48}Ga_{0.52}P$ grown on GaAs,[27] but branch defects become stronger with increasing temperature, while ordering becomes weaker with increasing temperature and is almost completely absent at 760 °C.[28]

Branch defects also do not appear to be consistent with phase separation. Stringfellow has shown theoretically that there is no driving force for phase separation in $In_xGa_{1-x}P$ when coherency strain is considered.[29] Phase separation could be induced by the strain fields present in a graded buffer, but if the branch defects were due to strain-induced phase separation, then their periodicity should correspond to the strain fields present in the graded buffer. Since all of the graded buffers are nearly completely relaxed, the strain fields in all of the samples should be very similar and result in a constant ρ_{branch}. The strong temperature dependence of ρ_{branch} shows that branch defect formation is not dominated by dislocation strain fields.

The morphology of branch defects also does not appear likely to relieve strain by phase separation. The strain fields observed in branch defects are very strong and sharp, so any strain that they might relieve should also be similarly strong and sharp. The crosshatch morphology and recent strain field modeling[30] both show that dislocation strain fields are quite gentle at the growth surface, so branch defect morphology is not consistent with strain-induced phase separation at the surface. Strain fields are strong and sharp close to misfit

dislocations, but those dislocations are separated from the growth surface by the critical thickness. Strain-induced phase separation near buried misfit dislocations would rely on very slow solid diffusion and the resulting features would not extend to the surface, while XTEM shows that branch defects do indeed extend to the surface. Furthermore, branch defects appear in only one ⟨110⟩, while dislocation strain fields are arrayed in both ⟨110⟩. Therefore, the morphology of branch defects is not consistent with strain-induced phase separation at the surface or at the buried misfit dislocation network.

There are also no direct indications of phase separation in the data. Preliminary energy dispersive spectroscopy studies in both scanning transmission electron microscopy and scanning electron microscopy show no evidence of phase separation near branch defects, which suggests that any compositional variations are less than 1%. Room temperature CL spectra do not show any unusual features that could be attributed to phase separation. Overall, there is currently no empirical evidence that branch defects are due to strain-induced phase separation.

Overall, the behavior of branch defects depends primarily on growth temperature. The formation mechanism for branch defects is still unclear, but the temperature dependence of branch defect density suggests that it is a kinetic process. Further characterization and analysis of branch defects will be presented in a later publication.

F. Evolution of microstructure in epitaxial compound semiconductors

There is no reason to expect that branch defects are unique to $In_xGa_{1-x}P/GaP$. A survey of the literature suggests that branch defects could explain the material degradation commonly observed in indium-bearing arsenides and phosphides grown near 575 °C by both MBE and MOVPE. Combined with the results of this study, a picture of the evolution of microstructure in these materials between roughly 400 and 800 °C will be proposed.

Material degradation as growth temperature approaches 575 °C has been reported in $In_xGa_{1-x}As/GaAs$[30–32] and $In_xGa_yAl_{1-x-y}As/GaAs$,[33] as well as in lattice-matched $In_xGa_{1-x}As$[34] and $In_xGa_yAl_{1-x-y}As$[35] on InP. The properties observed to degrade in these reports include carrier mobility, PL intensity and peak width, Bragg peak width, and surface roughness. In particular, a common observation is that the surface degrades into ⟨110⟩ saw-tooth or ridge structures. A variety of mechanisms have been used to explain the roughening and degradation near 575 °C. Yoon *et al.* proposed that indium desorption somehow results in degradation with increasing growth temperature.[35] Valtueña *et al.* alternatively proposed that the evolution of surface features was primarily a strain relaxation mechanism in $In_xGa_{1-x}As/GaAs$ and that decreasing growth temperature kinetically limited surface roughening,[32] resulting in good quality materials.

Bulsara, Leitz, and Fitzgerald studied $In_xGa_{1-x}As/GaAs$ grown by MOVPE from 500 to 700 °C and found that degradation actually peaks near 550 °C, with material quality improving at both lower and higher growth temperatures.[31] In addition to the previously observed surface roughening, contrast modulation aligned primarily along one ⟨110⟩ was observed in PVTEM at 550 °C. Similar features were also observed in $In_xGa_{1-x}As/GaAs$ grown by MBE at 580 °C and were absent at 350 and 480 °C,[30] establishing a critical continuity between MBE and MOVPE work with $In_xGa_{1-x}As/GaAs$ and indicating the earlier explanations for degradation at 575 °C were incomplete.

In recent work, Bulsara also observed that the contrast modulation in $In_xGa_{1-x}As/GaAs$ evolved into discrete defects that grew increasingly far apart as growth temperature was increased.[15] These defects featured wavy interference contrast and appeared to pin dislocations. It was believed that the defects were boundaries between domains of different surface reconstructions, since $In_xGa_{1-x}As$ undergoes a transition in surface reconstruction between 550 and 580 °C in MBE, so the defects were called high-energy boundaries.

The behavior of high-energy boundaries parallels the behavior of branch defects in $In_xGa_{1-x}P/GaP$ and the two phenomena are probably equivalent. The degradation observed in $In_xGa_{1-x}As/GaAs$ has indeed been also seen in $In_xGa_{1-x}P$ near 575 °C. In $In_xGa_{1-x}P/GaP$ grown by MBE, Chin *et al.* observed an increase in Bragg peak width near 575 °C.[12] In lattice-matched $In_xGa_{1-x}P$ grown on GaAs by MOVPE, Ozasa *et al.* found surface roughening and faceting at 580 °C versus the smooth surfaces obtained at 610 °C,[36] while Yuan *et al.* observed strong degradation in PL intensity and peak width when growth temperature was decreased below 600 °C.[37] Speculating that $In_xGa_{1-x}P$ has the same continuity between MBE and MOVPE results as demonstrated in $In_xGa_{1-x}As/GaAs$ by Bulsara *et al.*,[30,31] these results suggest that strong degradation due to branch defects occurs in both materials systems near 575 °C.

In addition to analysis and unification of earlier reports, this study extends the understanding of the evolution of microstructure in $In_xGa_{1-x}P$ and $In_xGa_{1-x}As$. Surface ridges are only observed in the samples from series III and IV grown at the lowest growth temperature of 650 °C. The quantitative correlation of surface ridges seen in AFM and branch defects seen in PVTEM shows that branch defects are directly responsible for the reported surface roughening near 575 °C, although the reason that ridges appear only when the branch defects are very weak is unknown. Furthermore, series II shows that branch defects become stronger with increasing growth temperature, resulting in essentially polycrystalline materials above 760 °C. Therefore, these two degradation phenomena define a high temperature process window between roughly 575 and 760 °C for the growth of high-quality $In_xGa_{1-x}P$ and a similar process window is expected in $In_xGa_{1-x}As$.

Also, a low temperature process window has been observed in MBE. Material degradation in the $In_xGa_yAl_{1-x-y}As$ system has been reported below temperatures ranging from 300[33] to 450 °C,[35] while degradation was seen below 520 °C in $In_xGa_{1-x}P/GaP$.[12] A variety of kinetic

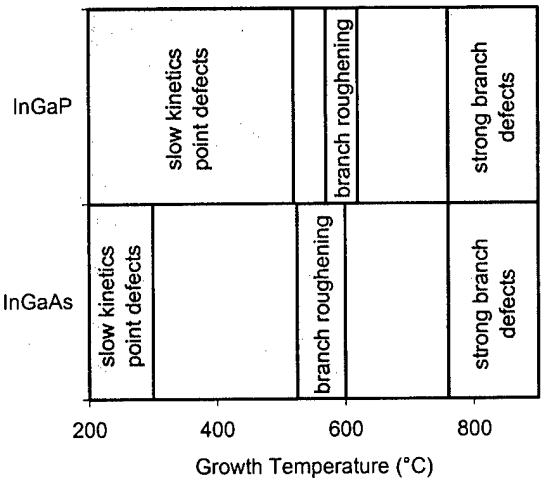


FIG. 16. Approximate growth temperature process windows for the growth of $In_xGa_{1-x}P$/GaP and $In_xGa_{1-x}As$/GaAs. The white regions represent process windows for the growth of high-quality material, while the shaded areas are regions of material degradation labeled with the mechanism responsible for the degradation.

factors could be responsible for poor material quality at extremely low growth temperatures, so the observed degradation at extremely low temperatures is not surprising. Along with the degradation seen near 575 °C, this very low temperature limit results in a low temperature process window for the growth of high-quality material.

Figure 16 illustrates the parallel process windows observed in $In_xGa_{1-x}P$ and $In_xGa_{1-x}As$, which we speculate are both dominated by the evolution of branch defects. At very high growth temperatures, branch defects are extremely strong and result in essentially polycrystalline material. Branch defect strength decreases with decreasing growth temperature, allowing high-quality film growth. Below 600 °C, increasing branch defect roughness and density produce severe material degradation. As growth temperature is decreased further, branch defects are suppressed and high-quality film growth is again possible. At very low growth temperatures, extremely slow growth kinetics result in material degradation.

This analysis shows that the overall qualitative evolution of microstructure in $In_xGa_{1-x}P$ and $In_xGa_{1-x}As$ is strikingly independent of whether the material is lattice-matched or mismatched and of growth technique. The very low temperature limit apparently varies significantly with the particular system being studied and the very high temperature limit may also vary as well, but the degradation observed at 575 °C is remarkably consistent. The mechanism for branch defect roughening is unclear at this time, but it must be something fundamental to the epitaxial growth of these materials.

If branch defects are indeed so universal, then it is surprising that they were not documented earlier. This study shows that branch defects evolve with continuing growth and are sensitive to strain, so it may be that branch defects are usually too weak to observe directly in most common lattice-matched heterostructures that are much thinner than the graded buffers in this study, although their effects are apparent in electrical and optical characterization. Also, much of the recent work in $In_xGa_{1-x}P$/GaP and $In_xGa_{1-x}As$/GaAs has been conducted with MBE, which is inherently limited to the low temperature process window where branch defects are probably suppressed by slow kinetics.

Optimization is possible in either process window, but the results of series III show that the high temperature process window is capable of much lower defect densities. A rational optimization strategy for achieving device-quality materials should therefore concentrate on the high temperature process window. The understanding of branch defect evolution developed in the preceding analysis is critical to improving material quality in $In_xGa_{1-x}P$/GaP.

G. Graded buffer design

With the new understanding of the evolution of microstructure and dislocation dynamics in $In_xGa_{1-x}P$/GaP, the challenge is to formulate processes to achieve device-quality materials. Control over material degradation with continued grading up to $x=0.39$ has already been demonstrated in series IV, so it is reasonable to expect that the degradation of LED efficiency beyond $x\sim0.35$ observed by Stinson *et al.*[1] can be avoided. Additional improvements in material quality depend on refining our understanding and control over dislocation kinetics and branch defects. Analysis of the current work and recent literature will be used to define an optimization strategy based on three primary design rules.

Fitzgerald *et al.* proposed a kinetic relaxation model for graded buffers to describe the Ge_xSi_{1-x}/Si system.[21] Expressions derived for the activation energy of homogeneous dislocation nucleation by Matthews, Blakeslee, and Mader[38] were used to qualitatively describe heterogeneous dislocation nucleation. The essential observation was that dislocation nucleation is activated by both temperature and strain, or put differently, that dislocation nucleation is thermally activated with an activation energy that is approximately inversely proportional to strain.[21] Dislocation glide is also thermally activated, but has only a power-law dependence on strain that is nearly linear,[22] so dislocation glide has a much weaker strain dependence than dislocation nucleation. So, the dislocation nucleation rate and glide velocity can be expressed as

$$\dot{\rho}\propto\exp\left(-\frac{E_{\text{nucleation}}}{\varepsilon\times kT}\right), \tag{8}$$

$$v\propto\varepsilon^n\exp\left(-\frac{E_{\text{glide}}}{kT}\right). \tag{9}$$

Fitzgerald *et al.* proposed that high growth temperatures result in very rapid dislocation glide, which leads to low strain and hopefully suppresses dislocation nucleation.[21] The implications of this model are clearer when a decrease in growth temperature is considered in Eqs. (8) and (9): both dislocation nucleation and glide slow exponentially, so strain increases. The increase in strain drastically increases dislocation nucleation and moderately increases dislocation glide

until a steady-state is once again achieved. At extremely low temperatures, dislocation kinetics may become so slow that strain simply accumulates in the material.

Series III is the first experimental demonstration of the kinetic graded buffer model. Comparing the exponential and nearly linear strain dependencies in Eqs. (8) and (9), dislocation nucleation varies significantly with strain, while dislocation glide is essentially fixed by the growth temperature. Dislocation nucleation should adjust to provide the necessary ρ_{field} to maintain a low strain state and efficient relaxation, as dictated by the dislocation glide velocity. Therefore, the observation of relaxation dominated by dislocation glide kinetics in series III is a verification of the model proposed by Fitzgerald *et al.*

A simple design rule can be derived from the kinetic graded buffer model: high growth temperatures should be used to maximize relaxation and minimize threading dislocation density in systems where dislocation glide kinetics dominate relaxation. Growth rate and grading rate could also be used to affect dislocation glide kinetics, but both have only linear effects and are constrained by the limitations of the growth process: order of magnitude changes in either parameter are not practical.[21] Dislocation glide mobility is thermally activated, therefore relatively small changes in growth temperature result in dramatic changes in dislocation dynamics. These arguments presume that the growth temperatures used are within the regime where dislocation glide kinetics dominate relaxation, which series III shows ranges at least from 650 to 800 °C for $In_xGa_{1-x}P/GaP$.

The new design rule that evolves from the results of this article is that branch defects must be controlled to achieve high-quality materials, because they clearly dominate relaxation when they appear with any reasonable strength and density. The temperature dependence of branch defect evolution frustrates efforts to find a single ideal growth temperature. At high temperatures, the onset of branch defect formation is delayed and the branch defect density is low, but the branch defects quickly become very strong once they begin to form. At low temperatures, branch defects are very weak, but they form almost immediately and with a high density.

The choice is clear for low indium compositions: a high growth temperature results in optimum material quality. The current results show that very high-quality materials can be grown without branch defects up to $x\sim0.2$ and higher growth temperatures may extend this range. For useful direct band gap compositions of $x>0.27$, the formation of branch defects appears inevitable and additional process design is necessary.

If a single growth temperature is desired, Fig. 17 shows that 700 °C appears to be the optimal growth temperature. While series II suggests that $\rho_{overall}$ would continue to decrease with decreasing growth temperature due to the suppression of branch defects, series III shows that dislocation glide kinetics set a basic lower limit for $\rho_{overall}$. Threading dislocations present early in the growth process cannot simply disappear and will persist with continued grading, even though faster dislocation glide kinetics could support efficient relaxation with increasingly lower $\rho_{overall}$ according to Eqs. (4) and (5). For example, ρ_{field} is greater than 10^7 cm^{-2} by $x=0.15$ at 650 °C due to slow dislocation glide kinetics, so ρ_{field} must be greater than 10^7 cm^{-2} at $x\sim0.3$ if a single growth temperature of 650 °C is used. Some reduction of $\rho_{overall}$ is possible with continued grading if dislocations glide to the edge of the material, terminate on heterogeneous defects, or annihilate each other and are not replaced,[8] but none of these mechanisms are expected to be very strong in typical $In_xGa_{1-x}P/GaP$ samples. Therefore, 700 °C appears to be the process window for the growth of high-quality $In_xGa_{1-x}P/GaP$ at a single growth temperature in our system.

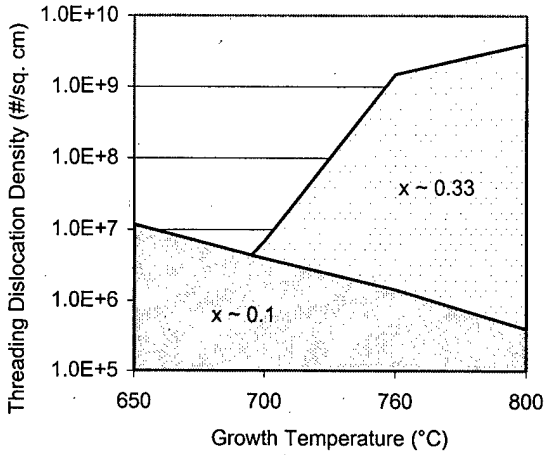


FIG. 17. Plot of $\rho_{overall}$ vs growth temperature for $In_xGa_{1-x}P/GaP$ graded to $x\sim0.1$ and $x\sim0.33$ at a single growth temperature, i.e., series II and III. A process windows exists near 700 °C for the growth of high-quality material. Note that $\rho_{overall}$ and ρ_{field} are used interchangeably at $x\sim0.1$, since pileup formation is negligible in those samples.

The success of series IV suggests a strategy to control branch defects: avoid branch defect formation for as long as possible and then suppress their strength once they become inevitable. A high growth temperature is used initially to avoid the formation of branch defects for as long as possible, then the growth temperature is abruptly dropped before the onset of branch defect formation to make the branch defects as weak as possible. Also, the lowest practical grading rate should be used to further suppress branch defect strength, as observed in the analysis of series II.

An optimization strategy for $In_xGa_{1-x}P/GaP$ results from combining the high growth temperature and branch defect suppression design rules. The highest practical growth temperature should be used initially to satisfy both design rules: maximize dislocation glide kinetics and delay the onset of branch defect formation. Shortly before branch defect formation is inevitable, the growth temperature should be dropped to suppress branch defect strength. Figures 12 and 13 can be used to estimate how large a temperature drop is possible without increasing ρ_{field} due to slow dislocation glide kinetics, as demonstrated in Fig. 18. If growth temperature must be dropped further, then the benefits of suppressing branch defect strength must be weighed against the limitation of slow dislocation glide kinetics and the material degradation

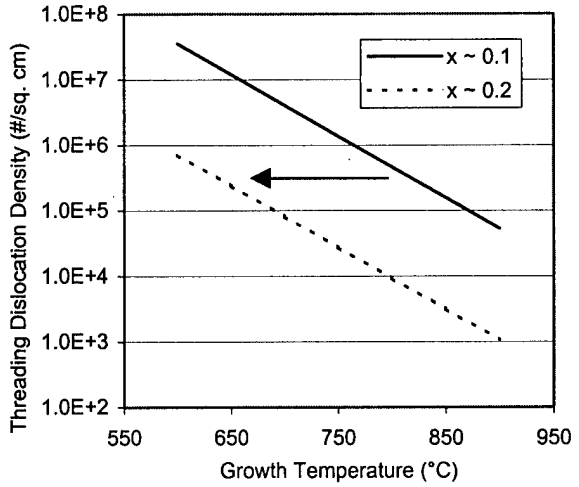


FIG. 18. Plot of ρ_{field} vs growth temperature for $In_xGa_{1-x}P/GaP$ graded to $x \sim 0.1$ and $x \sim 0.2$, assuming relaxation is dominated by dislocation glide kinetics. The line for $x \sim 0.1$ is taken from the data in series III, while the line for $x \sim 0.2$ is estimated using the information in Figs. 12 and 13. The arrow shows the largest temperature drop possible during grading without increasing ρ_{field} due to slow dislocation glide kinetics. The degradation expected at 575 °C places a hard limit on the maximum practical temperature drop within the high temperature process window.

that results near 575 °C due to branch defect roughening.

Optimization can be further complicated by other design rules that must be observed simultaneously. The other important design rule is to minimize surface roughness, otherwise interaction between dislocations and surface roughness will produce pileups through the mechanism described by Samavedam and Fitzgerald.[13] Surface roughness is in general minimized in the process of satisfying the first two design rules, especially if the grading rate is kept very low to suppress branch defects: Fitzgerald *et al.* observed that while $\rho_{overall}$ remained nearly constant over a wide range of grading rates, surface morphology degraded with faster grading.[21] Surface roughness and pileup formation have been controlled by planarization during periodic growth interruptions in Ge_xSi_{1-x}/Si[14] and this technique may also be useful in $In_xGa_{1-x}P/GaP$ if branch defects can be sufficiently suppressed during growth. Branch defects do not rely on surface topography to pin dislocations, so planarization cannot free dislocations pinned by strong branch defects. If branch defects are suppressed, $In_xGa_{1-x}P/GaP$ should behave similarly to Ge_xSi_{1-x}/Si and the planarization technique may yield significant improvements in defect density.

There are also basic limitations to the range of process parameters available. The growth technique limits growth temperature: GSMBE is limited to temperatures below 650 °C, while practical MOVPE temperatures range roughly from 500 to 900 °C. The ranges of practical growth rates and grading rates are also constrained by the capabilities of the growth technique. In general, the growth process must be optimized within the phase space allowed by the growth technique and materials system.

The most rational strategy for process optimization appears to have three main design rules. First and foremost, the formation of defect microstructures that pin dislocations must be controlled. Second, the growth temperature should be as high as possible to maximize dislocation glide kinetics and minimize defect density. Finally, the growth process must be designed to minimize surface roughness. These design rules operate simultaneously and any one of the phenomena can become dominant, but they are ordered here by their apparent relative importance in $In_xGa_{1-x}P/GaP$.

V. CONCLUSIONS

We have grown and carefully characterized $In_xGa_{1-x}P/GaP$ to extend our understanding of the closely coupled evolution of microstructure and dislocation dynamics in graded buffers. The growth experiments cover a temperature range of 650 to 810 °C and compositions ranging from $x=0.09$ to $x=0.39$. Within this phase space, three orders of magnitude of control over dislocation density has been demonstrated.

The mechanism for escalating defect density with continued grading in earlier $In_xGa_{1-x}P/GaP$ work has been identified as pileup formation on branch defects. Experiments and modeling show that branch defects are controlled primarily by growth temperature. A simple process optimization has recovered a nearly steady-state dislocation density up to $x=0.39$, which appears to eliminate the barriers to device-quality materials encountered in earlier work.

At low indium compositions, where the effect of branch defects can be nearly completely avoided, experiments and modeling show that dislocation glide kinetics dominate dislocation dynamics. This is the first experimental verification and extension of the kinetic graded buffer model proposed by Fitzgerald *et al.* The fact that the model was formulated for Ge_xSi_{1-x}/Si, yet finds application and verification in $In_xGa_{1-x}P/GaP$, suggests that it will be both powerful and widely applicable.

Through analysis of the literature in light of the new understanding of branch defects, we speculate that the evolution of branch defects dominates the microstructure of indium-bearing arsenides and phosphides over a very wide temperature range, roughly at least 500 to 800 °C. Degradation due to branch defects appears to affect both lattice-matched and mismatched materials grown by MBE or MOVPE. The unified picture that emerges is that there are low and high temperature process windows for the high-quality growth of these materials.

By mapping the evolution of microstructure and dislocation dynamics throughout the experiments in this study, we are able to summarize our findings in three design rules: branch defects must be avoided or suppressed, growth temperature must be maximized, and surface roughness must be minimized. If the branch defect design rule is generalized to say that any defect microstructure that hinders dislocation dynamics must be controlled, then these design rules are universally applicable to graded buffers.

Finally, an optimization strategy was developed and applied to $In_xGa_{1-x}P/GaP$. The first simple optimization resulted in the lowest accurate defect density in

$In_xGa_{1-x}P$/GaP graded to $x>0.27$ reported to date, $\rho_{overall}=4.7\times10^6\,cm^{-2}$ at $x=0.39$, thus proving the utility of the optimization strategy. With more sophisticated optimization, we hope to achieve $\rho_{overall}<10^6\,cm^{-2}$ for a full range of useful compositions, thus realizing the potential of $In_xGa_{1-x}P$/GaP as a high performance substrate for optoelectronic device applications.

ACKNOWLEDGMENTS

The authors would like to thank Glen Carey and Steve Stockman at the Optoelectronics Division of Hewlett–Packard for the donation of GaP, GaAsP, and GaAs wafers. Funding for this work was provided by a Department of Defense NDSEG Fellowship and by ARO Contract No. DAAG55-97-1-0111. This work made use of the MRSEC Shared Facilities supported by the National Science Foundation under Award No. DMR-940034.

[1]L. J. Stinson, J. G. Yu, S. D. Lester, M. J. Peanasky, and K. Park, Appl. Phys. Lett. **58**, 2012 (1991).
[2]W. Nakwaski, J. Appl. Phys. **64**, 159 (1988).
[3]F. A. Kish *et al.*, Appl. Phys. Lett. **64**, 2839 (1994).
[4]A. Y. Kim and E. A. Fitzgerald, Proc. SPIE **3621**, 187 (1999).
[5]R. E. Enstrom and D. G. Fischer, J. Appl. Phys. **46**, 1976 (1975).
[6]D. G. Fischer, R. E. Enstrom, J. S. Escher, H. F. Gossenberger, and J. R. Appert, IEEE Trans. Electron Devices **ED–21**, 641 (1974).
[7]Y. Komatsu, K. Hosotani, T. Fuyuki, and H. Matsunami, Jpn. J. Appl. Phys., Part 1 **36**, 5425 (1997).
[8]E. A. Fitzgerald, Mater. Sci. Rep. **7**, 87 (1991).
[9]A. Y. Kim and E. A. Fitzgerald, Mater. Res. Soc. Symp. Proc. **510**, 131 (1998).
[10]M. S. Abrahams, C. J. Buiocchi, and G. H. Olsen, J. Appl. Phys. **46**, 4259 (1975).
[11]M. S. Abrahams, L. R. Weisberg, C. J. Buiocchi, and J. Blanc, J. Mater. Sci. **4**, 223 (1969).
[12]T. P. Chin, J. C. P. Chang, K. L. Kavanagh, C. W. Tu, P. D. Kirchner, and J. M. Woodall, Appl. Phys. Lett. **62**, 2369 (1993).
[13]S. B. Samavedam and E. A. Fitzgerald, J. Appl. Phys. **81**, 3108 (1997).
[14]M. T. Currie, S. B Samavedem, T. A. Langdo, C. W. Leitz, and E. A. Fitzgerald, Appl. Phys. Lett. **72**, 1718 (1998).
[15]M. T. Bulsara, Ph.D. thesis, MIT (1998).
[16]E. A. Fitzgerald, S. B. Samavedam, Y. H. Xie, and L. M. Giovane, J. Vac. Sci. Technol. A **15**, 1048 (1997).
[17]S. Mahajan, Mater. Sci. Eng., B **30**, 187 (1995).
[18]H. Casey and M. Panish, *Heterostructure Lasers Part B* (Academic, San Diego, 1978).
[19]G. B. Stringfellow, L. C. Su, Y. E. Strausser, and J. T. Thornton, Appl. Phys. Lett. **66**, 3155 (1995).
[20]W. Wegscheider and H. Cerva, J. Vac. Sci. Technol. B **11**, 1056 (1993).
[21]E. A. Fitzgerald, Y. H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, A. R. Kortan, F. A. Thiel, and B. E. Weir, J. Vac. Sci. Technol. B **10**, 1807 (1992).
[22]K. Sumino, Mater. Sci. Rep. **13**, 269 (1974).
[23]I. Yonenaga and K. Sumino, J. Appl. Phys. **73**, 1681 (1993).
[24]I. Yonenaga and K. Sumino, Appl. Phys. Lett. **58**, 48 (1991).
[25]G. Stringfellow, *Organometallic Vapor-Phase Epitaxy* (Academic, Boston, 1989).
[26]H. Kroemer, J. Cryst. Growth **81**, 193 (1987).
[27]O. Ueda, M. Takikawa, M. Takechi, J. Komeno, and I. Umebu, J. Cryst. Growth **93**, 418 (1988).
[28]C. Nozaki, Y. Ohba, H. Sugawara, S. Yasuami, and T. Nakanisi, J. Cryst. Growth **93**, 406 (1988).
[29]G. B. Stringfellow, J. Electron. Mater. **11**, 903 (1982).
[30]M. T. Bulsara, V. Yang, A. Thilderkvist, E. A. Fitzgerald, K. Haüsler, and K. Eberl, J. Appl. Phys. **83**, 592 (1998).
[31]M. T. Bulsara, C. Leitz, and E. A. Fitzgerald, Appl. Phys. Lett. **72**, 1608 (1998).
[32]J. F. Valtueña *et al.*, J. Cryst. Growth **182**, 281 (1997).
[33]M. Haupt, K. Köhler, P. Ganser, S. Emminger, S. Müller, and W. Rothemund, Appl. Phys. Lett. **69**, 412 (1996).
[34]F. Peiró, A. Cornet, J. R. Morante, M. Beck, and M. A. Py, J. Appl. Phys. **83**, 7537 (1998).
[35]S. F. Yoon, P. H. Zhang, H. Q. Zheng, K. Radhakrishnan, and S. Swaminathan, J. Cryst. Growth **186**, 315 (1998).
[36]K. Ozasa, M. Yuri, S. Nishino, and H. Matsunami, J. Cryst. Growth **93**, 177 (1988).
[37]J. S. Yuan, M. T. Tsai, C. H. Chen, R. M. Cohen, and G. B. Stringfellow, J. Appl. Phys. **60**, 1346 (1986).
[38]J. W. Matthews, A. E. Blakeslee, and S. Mader, Thin Solid Films **33**, 253 (1976).

Strain in AlGaN layer studied by Rutherford backscattering/channeling and x-ray diffraction

M. F. Wu[a] and Shude Yao
Department of Technical Physics, Peking University, Beijing 100871, People's Republic of China and Instituut voor Kern-en Stralingsfysika, University of Leuven, B-3001 Leuven, Belgium

A. Vantomme, S. M. Hogg, and G. Langouche
Instituut voor Kern-en Stralingsfysika, University of Leuven, B-3001 Leuven, Belgium

J. Li and G. Y. Zhang
Department of Physics, Mesoscopic Physics Laboratory, Peking University, Beijing 100871, People's Republic of China

(Received 26 March 1999; accepted 30 April 1999)

An AlGaN layer with good crystalline quality (χ_{min}=2.1%) was grown by metalorganic vapor phase epitaxy on a sapphire (0001) substrate using a GaN intermediate layer. The Al composition, which cannot be unambiguously determined by x-ray diffraction (XRD) or by photoluminescence, was determined by Rutherford backscattering (RBS). The perpendicular and parallel elastic strain of the $Al_{0.28}Ga_{0.72}N$ layer, $e^{\perp}=-0.16\%$ and $e^{\parallel}=+0.39\%$, respectively, were derived using XRD and RBS/channeling. The small ratio $|e^{\perp}/e^{\parallel}|=0.41$ indicates that the $Al_{0.28}Ga_{0.72}N$ lattice is much stiffer in the *c*-axis direction than in the *a*-axis direction. A comparison of the strain data for GaN, InGaN, and AlGaN layers is presented. © *1999 American Vacuum Society.* [S0734-211X(99)02404-X]

I. INTRODUCTION

GaN based materials, consisting of AlN, GaN, InN, and their alloys AlGaN and InGaN, are very attractive materials for applications in visible light emitters, high-temperature/high-power electronics, and high-power optoelectronic devices due to their large direct band gaps, good thermal and chemical stability, and physical hardness, and because band gap engineering is possible from 1.89 to 6.2 eV by adjusting the composition of the ternary alloys.[1] These materials, which crystallize preferentially in the hexagonal wurtzite structure, are difficult to grow as large bulk crystals. Therefore they are prepared as heteroepitaxial films on various substrates, mostly on an Al_2O_3(0001) substrate, despite the very large lattice mismatch of 16% between GaN and Al_2O_3(0001). It is well known that heteroepitaxial layers are normally strained due to the lattice mismatch and thermal mismatch between the epilayer and the substrate. The elastic strain is an important issue and it affects the band gap of GaN based materials[2–5] and linewidth of electroluminescence (EL) or photoluminescence (PL) spectra.[6] Moreover, strain relaxation results in the formation of misfit dislocations in the film which correlates to the observation of a yellow luminescence band in the cathodoluminescence (CL) or PL spectrum.[7] Even though GaN based materials and devices have developed very rapidly in recent years, there have been very few reports concerning the elastic strain in GaN based materials, especially in AlGaN and InGaN films.

Xiao *et al.*[8] have measured the perpendicular and parallel elastic strain, $e^{\perp}=+0.2\%$ and $e^{\parallel}=-2.0\%$, using transmission electron microscopy (TEM) for an α-GaN film (~1 μm) grown by gas-source molecular beam epitaxy on an Al_2O_3(0001) substrate. By using x-ray diffraction (XRD) combined with TEM, Powell *et al.*[9] reported that an epitaxial GaN layer on an Al_2O_3(0001) substrate has lattice constants of $a_{epi}=0.3185$ nm and $c_{epi}=0.5189$ nm. Hence, $e^{\perp}=(c_{epi}-c_b)/c_b=+0.077\%$ and $e^{\parallel}=(a_{epi}-a_b)/a_b=-0.125\%$ can be derived using $a_b=0.3189$ nm and $c_b=0.5185$ nm for bulk GaN materials. Detchprohm *et al.*[10] measured the strain of GaN films with thicknesses ranging from 1 to 1200 μm using XRD. They found that the strain of the GaN films is thickness dependent for thicknesses less than 100 μm (the largest $e^{\perp}$ is ~+0.12%), while the ratio of the perpendicular strain and parallel strain $|e^{\perp}/e^{\parallel}|=0.38$ is thickness independent. The strain is almost completely relaxed in films thicker than 100 μm, and the intrinsic lattice constants of fully relaxed α-GaN, i.e., a_b and c_b, are determined to be 0.318 92 ±0.000 09 and 0.518 50±0.000 05 nm, respectively, values which are now widely accepted.[11,12] Kim *et al.*[13] measured the lattice constant of very thin GaN films on AlN using synchrotron x-ray diffraction. For 10 and 5 nm GaN films, $e^{\parallel}$ of −0.7% and −1.7%, respectively, can be derived. However, these very thin GaN films are still not fully strained. A critical thickness of 2.9±0.4 nm (below which the film is pseudomorphic, i.e., fully strained) is thus determined by extrapolation from data points for films with four different thicknesses. It should be pointed out that the XRD results have a smaller error compared to TEM strain measurements, and therefore are more reliable.

For InGaN and AlGaN, the data related to the strain are very scarce. El-Masry *et al.*[14] reported that, using TEM, an InGaN film grown at 750 °C with 10% InN is highly strained, while an InGaN film containing 49% InN has an extremely high strain, however no values were given. Bremser *et al.*[3] reported that the AlGaN layers grown on 6H-SiC(0001) are highly strained and that an $Al_{0.12}Ga_{0.86}N$ layer with different strain (grown with or without an AlN buffer

[a]Electronic mail: mingfang.wu@fys.kuleuven.ac.be

TABLE I. Strain data of GaN, InGaN, and AlGaN layers presented in or deduced from the literature.

Layer	Thickness	$e^{\perp}$ (%)	$e^{\parallel}$ (%)	$e^{\perp}/e^{\parallel}$	Techniques	Reference
GaN	~1 μm	+0.2	−2.0	−0.1	TEM	8
GaN		+0.077[a]	−0.125[a]	−0.62	XRD+TEM	9
GaN	~1 μm[b]	+0.12[b]	−0.32[b]	−0.38	XRD	10
GaN	5 nm		−1.7%		XRD	13
$In_{0.18}Ga_{0.82}N$[c]	40 nm	+2.0	−1.97[c]	−1.01	XRD	2
$In_{0.18}Ga_{0.82}N$	270 nm	+0.21	−0.53	−0.40	XRD+RBS	15
$Al_{0.25}Ga_{0.75}N$[c]	350–650 nm	−0.56	+0.57[c]	−0.98	XRD	2
$Al_{0.28}Ga_{0.72}N$	765 nm	−0.16	+0.39	−0.41	XRD+RBS	This work

[a]Values derived from the a_{epi} and c_{epi} data.
[b]The thickness varied in a range of 1–1200 μm; here only the $e^{\perp}$ value for thickness ~1 μm is listed. The $e^{\parallel}$ value is derived from $e^{\perp}/e^{\parallel}=-0.38$ which is thickness independent.
[c]The compositions of In and Al ranged from 0 to 0.2 and from 0 to 0.25, respectively; here we only list one of them. The $e^{\parallel}$ values are deduced from the reported a_{epi} value of 0.3182 nm and the calculated a_b value using Vegard's law.

layer) results in a 75 meV shift in the CL spectra. However, no strain data were presented. Takeuchi *et al.*[2] reported on the perpendicular elastic strain $e^{\perp}$ of 40 nm $In_xGa_{1-x}N$ layers (with x values ranging from 0 to 0.2) and of 350–650 nm $Al_xGa_{1-x}N$ layers (with x values ranging from 0 to 0.25) grown by metalorganic vapor phase epitaxy (MOVPE) on a 2 μm GaN layer on a sapphire (0001) substrate. For $In_{0.18}Ga_{0.82}N$ and $Al_{0.25}Ga_{0.75}N$, rather high $e^{\perp}$ values of +2.0% and −0.56%, respectively, were given. In the meantime, they concluded that all the samples in their experiment, both InGaN and AlGaN, are fully strained and have a constant lattice parameter of $a_{epi}=0.3182\pm0.0001$ nm. Therefore, the parallel elastic strain can be deduced using Vegard's law as shown in Table I. In comparison with the results in Refs. 2 and 13, it is unclear why the critical thickness h_c of the $In_{0.18}Ga_{0.82}N$ layer on GaN is so thick (>42 nm)[2] while the h_c of the GaN layer on AlN is only 2.9 nm,[13] although the lattice mismatch only differs slightly between $In_{0.18}Ga_{0.82}N$/GaN (+2.0%) and GaN/AlN (+2.5%). A quantitative result of $e^{\perp}=+0.21\%$ and $e^{\parallel}=-0.53\%$ for 270 nm $In_{0.18}Ga_{0.82}N$ grown by metalorganic chemical vapor deposition was presented recently[15] and shows that the strain in the 270 nm $In_{0.18}Ga_{0.82}N$ layer is partially relaxed. Table I lists the typical strain data for GaN, InGaN, and AlGaN layers presented in or deduced from the literature.

It should be emphasized that, in order to study the strain of the ternary alloys, $In_xGa_{1-x}N$ or $Al_xGa_{1-x}N$, the determination of the composition, i.e., the x value of the alloy, is a prerequisite. In principle. the x value can be determined by XRD using Vegard's law assuming the layers are fully relaxed.[14,16–18] However, if the strain in the InGaN and AlGaN layers is not negligible as was the case in Refs. 2, 3, 14, and 15, then the x value deduced by XRD will be inaccurate. Alternatively, the x value can also be determined by PL measurement using the formula $E_g(In_xGa_{1-x}N)=(1-x)E_g(GaN)+xE_g(InN)-bx(1-x)$, where b is the bowing parameter and E_g is the band gap, respectively. However, due to the very different b values reported in the literature[2,5,19] and because the existence of elastic strain in the epilayer will alter the value of the band gap,[2–5] the x value deduced using this approach is also questionable. On the other hand, because nuclear elastic-scattering cross sections are well understood. Rutherford backscattering (RBS) is a reliable and accurate method for directly and nondestructively determining the composition of an alloy as a function of depth.[20] In this work, we first determine the x value, the thickness, and the crystalline quality of an $Al_xGa_{1-x}N$ layer by RBS/channeling and subsequently present a quantitative analysis of the perpendicular and parallel elastic strain of the $Al_{0.28}Ga_{0.72}N$ layer, measured by XRD and RBS/channeling.

II. EXPERIMENT

AlGaN layers were grown by MOVPE over an $Al_2O_3(0001)$ substrate in a horizontal reactor. Trimethygallium (TMGa), Trimethyaluminium (TMAl), ammonia (NH_3), and hydrogen (H_2) were used as the Ga, Al, and N source and carrier gas, respectively. The growth was conducted at a pressure of ~76 Torr. First, the substrate was heated to 1150 °C in a stream of hydrogen. Then, the temperature was lowered to 550 °C to grow a thin (~25 nm) GaN buffer layer. Subsequently, the temperature was elevated to 1030°C to anneal the buffer layer in the mixed ammonia and hydrogen, and followed by growth of an intermediate GaN layer of ~250 nm. Finally, an $Al_xGa_{1-x}N$ layer was grown at 1100 °C with a x value ~0.3. The thickness of the AlGaN layer was controlled to be ~700 nm. The sample is transparent and the surface is mirror like.

For RBS/channeling measurements, a collimated 1.57 MeV He^+ beam produced by an accelerator (5SDH-2 Pelletron) was used with a backscattering angle of 172°. The sample was mounted on a high precision (±0.01°) three-axis goniometer in a vacuum chamber so that the orientation of the sample relative to the He^+ beam can be controlled precisely. Off-normal axis angular scans were used to measure the angle Φ_{epi} between the ⟨0001⟩ and ⟨$1\bar{2}13$⟩ axes of AlGaN, which is related to the strain of the AlGaN epilayer.

A computer controlled 7.5 kW rotating anode Rigaku double-crystal diffractometer with monochromator Cu $K\alpha_1$ x rays ($\lambda=0.154\,05$ nm) was used to perform symmetric

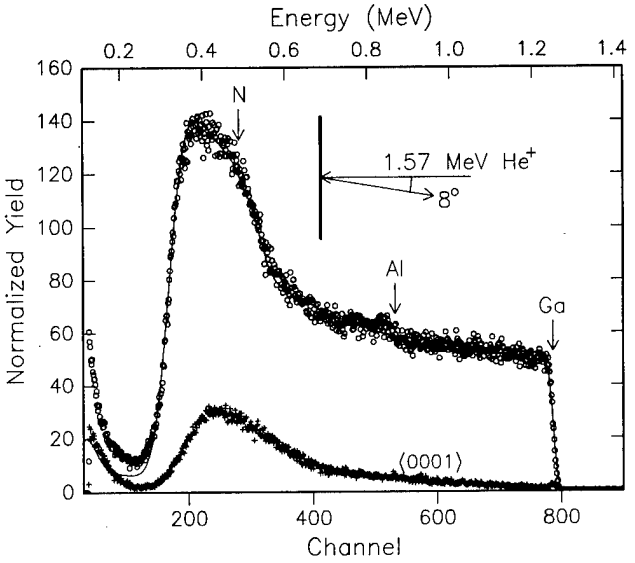


FIG. 1. Random (○), aligned (+), and simulated (solid line) RBS spectra of an $Al_{0.28}Ga_{0.72}N/GaN/Al_2O_3(0001)$ sample prepared by MOVPE.

x-ray θ–2θ scans and ω scans (rocking curve). To obtain a higher angular resolution, some measurements were carried out by adding a third crystal, Si(220).

III. RESULTS AND DISCUSSION

A. RBS/channeling: Composition and crystallinity of the AlGaN layer

Figure 1 shows the random and aligned RBS spectra of the sample. The inset shows the geometry used in the backscattering measurements. The arrows (labeled Ga, Al, and N) indicate the energy for backscattering from these elements at the surface. A simulation of the random spectrum reveals that the thicknesses of the AlGaN and GaN layers are 765 and 250 nm, respectively. The simulation further indicates that the composition of the $Al_xGa_{1-x}N$ layer is not uniform, with x gradually increasing from 0.24 at the surface to 0.31 near the AlGaN/GaN interface. On average, the AlGaN layer can be considered to be a layer of $Al_{0.28}Ga_{0.72}N$. The aligned spectrum indicates that the χ_{min} (i.e., the ratio of the backscattering yields of aligned and random spectra in the near-surface region) of the $Al_{0.28}Ga_{0.72}N$ layer is 2.1%, which means that the crystalline quality of this layer is very good. The aligned spectrum also shows that the crystalline quality of the 250 nm GaN intermediate layer (~ channels 170–320) is rather poor due to the large lattice mismatch (+16%) between the GaN layer and the Al_2O_3 substrate. In contrast, the crystalline quality of the $Al_{0.28}Ga_{0.72}N$ layer is greatly improved due to the buffering function of the GaN layer and to the small lattice mismatch (−0.68%) between the $Al_{0.28}Ga_{0.72}N$ layer and the GaN intermediate layer. However, since the 250 nm GaN layer is not thick enough, the high density of defects partially spreads into the $Al_{0.28}Ga_{0.72}N$ layer near the AlGaN/GaN interface (~channel 320).

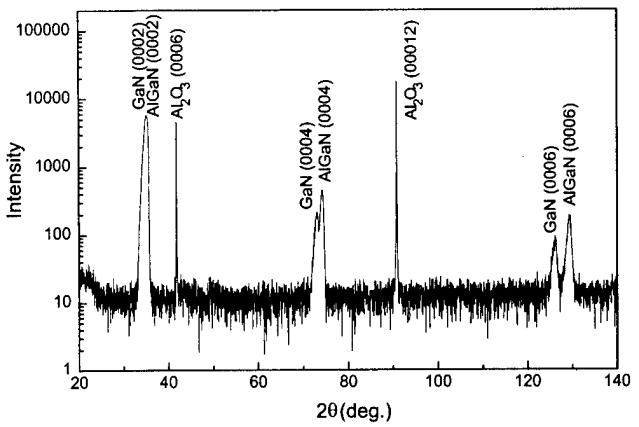


FIG. 2. X-ray diffraction of the sample in Fig. 1 with 2θ ranging from 20° to 140°.

B. XRD θ–2θ scan, ω scan, and perpendicular strain

Figure 2 shows a symmetric x-ray diffraction obtained from this sample with 2θ ranging from 20° to 140° showing the unresolved GaN(0002)+AlGaN(0002) peak, the partially resolved GaN(0004)+AlGaN(0004) peak, and the resolved GaN(0006)+AlGaN(0006) peak. To resolve the GaN(0002) and AlGaN(0002) peaks, a symmetric θ–2θ scan ranging from 34.5° to 36.2° (Fig. 3) was carried out using a third crystal. While the intensity decreased significantly, the angular resolution greatly improved. The intensity of the AlGaN(0002) peak is about three times that of the GaN(0002) peak implying that the thickness of the AlGaN layer is approximately three times that of the GaN layer, in agreement with the RBS measurement. Moreover, the width of the AlGaN peak (0.096°) is larger than that of the GaN peak (0.082°) which is also consistent with the conclusion obtained from the RBS measurement, i.e., the nonuniform composition of the AlGaN layer results in a spread of d spacing and thus a wider AlGaN diffraction peak. Figure 4 shows the ω scans of the GaN(0002) and AlGaN(0002) diffractions. We note that the width of the GaN(0002) peak (0.379°) is larger than that of the AlGaN(0002) peak (0.346°) which is opposite to the situation in Fig. 3. This is not sur-

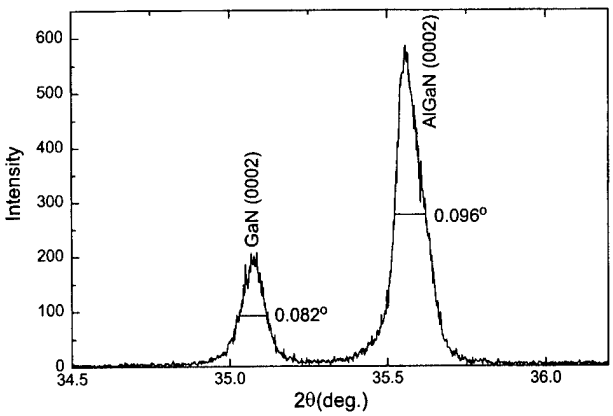


FIG. 3. High-resolution θ–2θ scan of the sample in Fig. 1 with 2θ ranging from 34.5° to 36.2°.

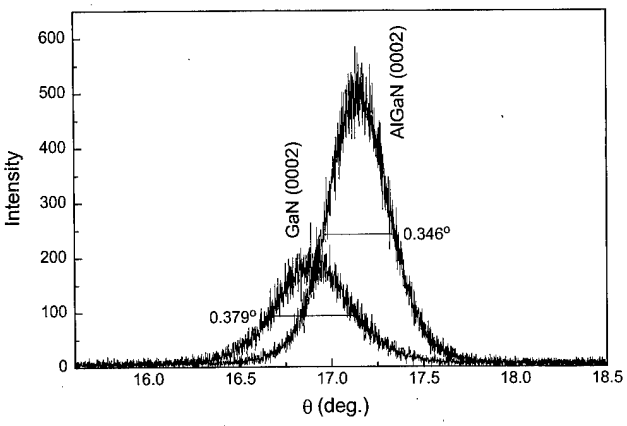


FIG. 4. ω scans of the sample in Fig. 1 showing the GaN(0002) and AlGaN(0002) diffractions.

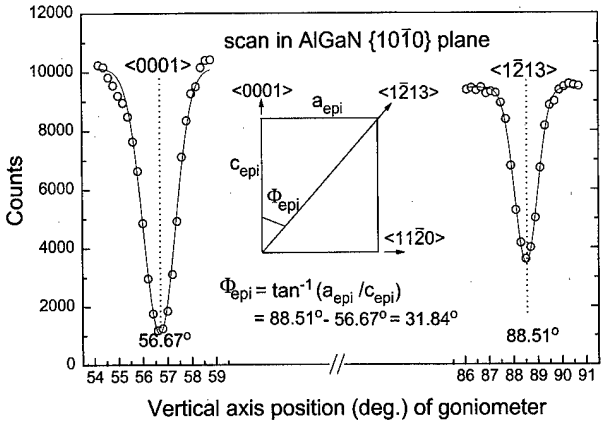


FIG. 5. Angular scan along the AlGaN{10$\bar{1}$0} plane showing the AlGaN⟨1$\bar{2}$13⟩ and AlGaN⟨0001⟩ dips. A Φ_{epi} angle of 31.84° is determined from this scan.

prising because the width of the diffraction peaks in the θ–2θ scan is related to the spread of the d spacing while the width of the peaks in the ω scan is sensitive to the spread of the orientation (mosaicity) which is related to the crystalline quality of the epilayer. The result in Fig. 4 shows that the crystalline quality of the GaN layer is poorer than that of the AlGaN layer, also in agreement with the RBS result in Fig. 1.

Taking the position of the Al_2O_3(00012) peak in Fig. 2 as a reference, the c_{epi} of the AlGaN epilayer is calculated to be 0.5119 nm. On the other hand, using Vegard's law and the lattice constants of AlN (a_b=0.3111 nm, c_b=0.4978 nm) and GaN (a_b=0.3189 nm, c_b=0.5185 nm),[11] values of a_b =0.3167 nm and c_b=0.5127 nm for bulk (=fully relaxed) $Al_{0.28}Ga_{0.72}N$ are obtained. Therefore the deduced perpendicular elastic strain, $e^{\perp}=(c_{epi}-c_b)/c_b$, of the $Al_{0.28}Ga_{0.72}N$ layer is $-0.16\%\pm0.05\%$ indicating that the layer is under compressive strain in the perpendicular direction. This perpendicular strain is smaller than the value of -0.56% reported for a fully strained 350–650 nm $Al_{0.25}Ga_{0.75}N$ layer,[2] which can be explained by the relaxation of the thicker (765 nm) layer used in this study.

C. RBS angular scan: Parallel strain

To measure the parallel elastic strain, a RBS angular scan (Fig. 5), similar to the study in Ref. 21, was carried out along the AlGaN{10$\bar{1}$0} plane by choosing a region of integration (channels 370–800 in Fig. 1) corresponding to the Ga and Al atoms in the AlGaN layer. The dip of the angular scan shown in Fig. 5 indicates that along the direction of the AlGaN⟨1$\bar{2}$13⟩ axis, the backscattering yield has a lower value compared to the other off-axis directions nearby due to the channeling effect.[20] The position of the AlGaN⟨1$\bar{2}$13⟩ dip reveals that the angle Φ_{epi} between the ⟨0001⟩ and the ⟨1$\bar{2}$13⟩ axes is 31.84°±0.05°. The value of Φ is equal to $\tan^{-1}(a/c)$ (as shown in the inset of Fig. 5), which is Φ_b $=\tan^{-1}(a_b/c_b)=31.70°$ for bulk $Al_{0.28}Ga_{0.72}N$. Therefore, $\Delta\Phi=\Phi_{epi}-\Phi_b=+0.14°\pm0.05°$. The positive $\Delta\Phi$ value means that the AlGaN layer is under compressive strain in the perpendicular direction, which is consistent with the XRD result. The tetragonal distortion deduced from the RBS/channeling study[22] is $e_T=e^{\parallel}-e^{\perp}=\Delta\Phi/\sin\Phi\cos\Phi$ $=+0.55\%\pm0.20\%$. Combining this result with the value for $e^{\perp}$ derived by XRD, a parallel elastic strain $e^{\parallel}=+0.39\%$ $\pm0.25\%$ is obtained, indicating that the layer is under tensile strain in the parallel direction. The resulting ratio $e^{\perp}/e^{\parallel}$ $=-0.41$, is rather close to the value of -0.38 reported for the GaN films[10] and the value of -0.40 for the $In_{0.18}Ga_{0.82}N$ film,[15] however it is rather different from the values of ~-1.0 and -0.1 deduced from Refs. 2 and 8, respectively. For a strained layer, $|e^{\perp}/e^{\parallel}|$ should be equal to 2 if the volume is conserved. The value of 0.41, which is much smaller than 2, indicates that the volume of the strained layer is not conserved and the $Al_{0.28}Ga_{0.72}N$ lattice is much stiffer along the c axis than along the a axis.

In Table I, we note that the signs of $e^{\parallel}$ and $e^{\perp}$ of AlGaN are opposite to those of InGaN. This is reasonable, because the lattice mismatch between AlGaN and GaN is negative (for $Al_{0.28}Ga_{0.72}N$, it is -0.68%), whereas the lattice mismatch between InGaN and GaN is positive (for $In_{0.18}Ga_{0.82}N$, it is $+2.0\%$).

IV. CONCLUSION

A 765 nm $Al_{0.28}Ga_{0.72}N$ layer with good crystalline quality (χ_{min}=2.1%) was prepared on GaN/Al_2O_3(0001) using MOVPE at 1100 °C. The Al composition was determined directly by RBS and a variation of composition with depth is revealed. The tetragonal distortion measured from the position of the ⟨1$\bar{2}$13⟩ dip in an angular scan is useful to determine the strain of the epilayer. Combined with XRD, the perpendicular and parallel elastic strains of the $Al_{0.28}Ga_{0.72}N$ layer were determined as $e^{\perp}=-0.16\%$ and $e^{\parallel}=+0.39\%$, respectively, implying that the epilayer is compressively strained in the perpendicular direction and the strain is partially relaxed. The ratio, $|e^{\perp}/e^{\parallel}|=0.41$, indicates that the $In_{0.18}Ga_{0.82}N$ lattice is much stiffer along the c axis than along the a axis.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos. 19775001 and 69789601), the Bilateral Cooperation between China and Flanders (96/32), the Belgian Fund for Scientific Research, Flanders (FWO), Concerted Action (GOA) programs, and the Inter-University Attraction Poles (IUAP P4/10). One of the authors (A.V.) is a postdoctoral researcher for FWO.

[1]H. Morkoç, S. Strite, G. B. Gao, M. E. Lin, B. Sverdlov, and M. Burns, J. Appl. Phys. **76**, 1363 (1994).
[2]T. Takeuchi, H. Takeuchi, S. Sota, H. Sakai, H. Amano, and I. Akasaki, Jpn. J. Appl. Phys., Part 2 **36**, L177 (1997).
[3]M. D. Bremser, W. G. Perry, T. Zheleva, N. V. Edwards, O. H. Nam, N. Parikh, D. E. Aspnes, and R. F. Davis, MRS Internet J. Nitride Semicond. Res. **1** (1996).
[4]X. Zhang, P. Kung, A. Saxier, D. Walker, T. C. Wang, and M. Razeghi, Appl. Phys. Lett. **67**, 1745 (1995).
[5]C. Wetzel, T. Takeuchi, S. Yamaguchi, H. Katoh, H. Amano, and I. Akasaki, Appl. Phys. Lett. **73**, 1994 (1998).
[6]K. P. O'Donnell, T. Breitkopf, H. Kalt, W. Van der Stricht, I. Moerman, P. Demeester, and P. G. Middleton, Appl. Phys. Lett. **70**, 1843 (1997).
[7]F. A. Ponce, MRS Bull. **22**, 51 (1997).
[8]H. Z. Xiao, N.-E. Lee, R. C. Powell, Z. Ma, L. J. Chou, L. H. Allen, J. E. Greene, and A. Rockett, J. Appl. Phys. **76**, 8195 (1994).
[9]R. C. Powell, N.-E. Lee, Y.-W. Kim, and J. E. Greene, J. Appl. Phys. **73**, 189 (1993).
[10]T. Detchprohm, K. Hiramatsu, K. Itoh, and I. Akasaki, Jpn. J. Appl. Phys., Part 2 **31**, L1454 (1992).
[11]T. Kachi, K. Tomita, K. Itoh, and H. Tadano, Appl. Phys. Lett. **72**, 704 (1998).
[12]M. S. Shur and M. A. Khan, MRS Bull. **22**, 44 (1997).
[13]C. Kim, I. K. Robinson, J. Myoung, K. Shim, K. Kim, and M.-C. Yoo, Mater. Res. Soc. Symp. Proc. **423**, 557 (1996).
[14]N. A. El-Masry, E. L. Piner, S. X. Liu, and S. M. Bedair, Appl. Phys. Lett. **72**, 40 (1998).
[15]M. F. Wu, A. Vantomme, S. M. Hogg, G. Langouche, W. Van der Stricht, K. Jacobs, and I. Moerman, Appl. Phys. Lett. **74**, 365 (1999).
[16]W. Van der Stricht, K. Jacobs, I. Moerman, P. Demeester, L. Considine, E. J. Thrush, J. A. Crawley, and P. Ruterana, Mater. Res. Soc. Symp. Proc. **482**, 107 (1998).
[17]R. Singh, D. Doppalapudi, T. D. Moustakas, and L. T. Romano, Appl. Phys. Lett. **70**, 1089 (1997).
[18]D. Korakakis, K. F. Ludwig, Jr., and T. D. Moustakas, Appl. Phys. Lett. **71**, 72 (1997).
[19]D. Korakakis, H. M. Ng, M. Misra, W. Grieshaber, and T. D. Moustakas, MRS Internet J. Nitride Semicond. Res. **1** (1996).
[20]W. K. Chu, J. W. Mayer, and M. A. Nicolet, *Backscattering Spectrometry* (Academic, New York, 1978).
[21]M. F. Wu, A. Vantomme, J. De Wachter, S. Degroote, H. Pattyn, G. Langouche, and H. Bender, J. Appl. Phys. **79**, 6920 (1996).
[22]M. F. Wu, A. Vantomme, H. Pattyn, G. Langouche, Q. Yang, and Q. Wang, J. Appl. Phys. **80**, 5713 (1996).

Blueshift of $In_{0.2}Ga_{0.8}N$/GaN single quantum well band gap energy by rapid thermal annealing

G. Li[a] and S. J. Chua[b]
Institute of Materials Research and Engineering, Singapore 119260, Singapore

J. H. Teng
Center for Opto-electronics, Department of Electrical Engineering, National University of Singapore, Singapore 119260, Singapore

W. Wang, Z. C. Feng, and Y. H. Huang
Institute of Materials Research and Engineering, Singapore 119260, Singapore

T. Osipowicz
Research Centre for Nuclear Microscopy, National University of Singapore, Singapore 119260, Singapore

(Received 5 March 1999; accepted 14 May 1999)

Up to 15 nm blueshift in the low-temperature photoluminescence (LT-PL) peak wavelength of Si-doped $In_{0.2}Ga_{0.8}N$/GaN single quantum well (SQW) is reported after rapid thermal annealing in atmospheric N_2 ambient at temperatures below 1000 °C. We found that the thermal annealing not only blueshifts the LT-PL peak wavelength but it also improves the optical properties in terms of an increase in LT-PL peak intensity and a reduction in the spectrum width. Under the same thermal annealing conditions, however, a dielectric encapsulant layer (a spin-on silica layer) suppresses the blueshift of the LT-PL peak wavelength of the $In_{0.2}Ga_{0.8}N$/GaN QW. The effect of the thermal annealing temperature and time on the blueshift of the LT-PL peak wavelength is also reported. The possible mechanism responsible for the thermal annealing effect on the band gap energy of the $In_{0.2}Ga_{0.8}N$/GaN SQW is discussed. © *1999 American Vacuum Society.*
[S0734-211X(99)06604-4]

Post-tuning of the optical band gap energy of quantum well (QW) structures has attracted enormous interest due to its potential application in integration of opto-electronic and photonic devices. Due to the thermal stability of heterojunctions, such as AlGaAs/GaAs, a number of approaches have been used to enhance interdiffusion of the group III or V or both atoms across a heterojunction at an elevated temperature.[1] The most promising technique is impurity-free vacancy enhanced (QW) intermixing (IFVEI). In IFVEI, a dielectric encapsulant layer (such as SiO_2) is deposited on the top surface. At an elevated thermal annealing temperature, the group III atomic vacancies are generated at the interface between the dielectric layer and the top surface. The diffusion of those vacancies into the structure can significantly enhance the QW intermixing, leading to a large blueshift of the QW band gap energy with minimum effect on their electrical and optical properties. Previous work in post-growth tuning of the QW band gap energy has mainly been focused on non-nitride based III-V QW structures.[1]

McCluskey *et al.* recently reported interdiffusion of In and Ga in InGaN multiple QW (MQW) structures after thermal annealing at temperatures ranging from 1300 to 1400 °C.[2] At such a high temperature, hydropressure of up to 15 kbar in N_2 is required to prevent surface decomposition. There are no results reported of its thermal annealing effect on the optical properties of those MQW structures. In this work, the thermal annealing was conducted using Si-doped $In_{0.2}Ga_{0.8}N$/GaN single QW (SQW) structures at 900 or 990 °C. The effect of a dielectric encapsulant layer on post-growth tuning of the $In_{0.2}Ga_{0.8}N$/GaN QW band gap energy is reported for the first time.

An $In_{0.2}Ga_{0.8}N$/GaN SQW was grown on (0001) orientated sapphire by metalorganic vapor phase epitaxy (MOVPE). Trimethylgallium (TMGa), trimethylindium (TMIn), and NH_3 were used as the precursors. The composition of bulk $In_{0.2}Ga_{0.8}N$ was determined using room temperature photoluminescence (PL) and the unity bowing parameter. An 8 nm thick $In_{0.2}Ga_{0.8}N$ well was grown on a 1.5 μm thick undoped GaN buffer and capped by 0.2 μm thick undoped GaN. Si doping of $In_{0.2}Ga_{0.8}N$ led to an electron density of about 2.5×10^{18} cm^{-3}. Thermal annealing was conducted in a rapid thermal annealing furnace with *in situ* pyrometer temperature monitoring. The ambient was pure N_2 at atmospheric pressure. The samples were face to face covered by clean undoped GaN to avoid potential contamination. Low-temperature PL (LT-PL) was carried out at 7 K using a He–Cd laser at 325 nm. The luminescence signal was dispersed into a 0.75 m monochromator and detected by a GaAs photomultiplier.

Figure 1 illustrates an up to 15 nm blueshift of the LT-PL peak wavelength of an $In_{0.2}Ga_{0.8}N$/GaN SQW after 20 s of thermal annealing at 990 °C. It is further noticed that the PL peak intensity was even slightly increased, and the peak width was reduced by about 2 nm from 12 to 10 nm after the thermal annealing. These results demonstrate that blueshift of the LT-PL peak wavelength of an $In_{0.2}Ga_{0.8}N$/GaN SQW

[a]Present address: Blue Laboratory Pty. Ltd., P.O. Box 413, Burwood NSW 1805, Australia; electronic mail: likuang@ficnet.net
[b]Electronic mail: elecsj@nus.edu.sg

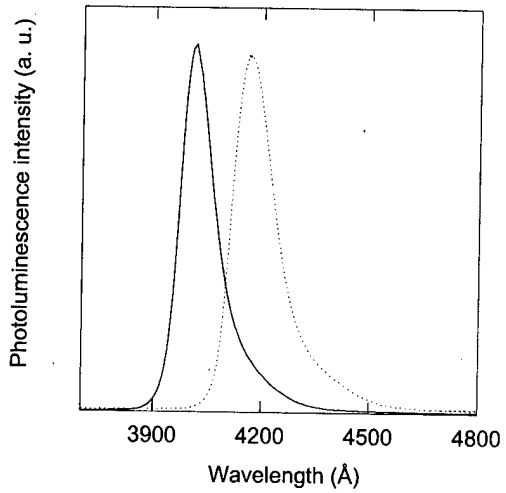


FIG. 1. Low-temperature photoluminescence of the as-grown (dotted line) and thermally annealed (solid line) $In_{0.2}Ga_{0.8}N$/GaN SQW. The thermal annealing conditions are 990 °C for 20 s in N_2.

can be achieved after thermal annealing in atmospheric N_2 ambient at temperatures below 1000 °C. More important, the thermal annealing seems to not degrade the optical properties of $In_{0.2}Ga_{0.8}N$/GaN SQWs in terms of the peak intensity and the width of low temperature photoluminescence spectra at 7 K. We also note an asymmetric LT-PL peak with a tail on the long wavelength side. Thermal annealing does not alter this spectral shape. A similar asymmetrical broadening of PL peaks was observed in Si-doped GaN.[3] We believe that the asymmetric PL peak may also arise from impurity band broadening since our $In_{0.2}Ga_{0.8}N$ was Si doped.

Figure 2 illustrates effect of the thermal annealing time on the blueshift of the LT-PL peak wavelength of an $In_{0.2}Ga_{0.8}N$/GaN SQW. As can be seen, the blueshift rapidly increases when the thermal annealing time was increased to 20–30 s at 990 °C. Afterwards, further increase of the thermal annealing time does not further increase the blueshift. At 900 °C, the same thermal annealing time gave rise to about 3–4 nm less blueshift of the LT-PL peak wavelength than that obtained at 990 °C. The dependence of the blueshift on the thermal annealing time observed at 900 °C was similar to that obtained at 990 °C (see Fig. 2).

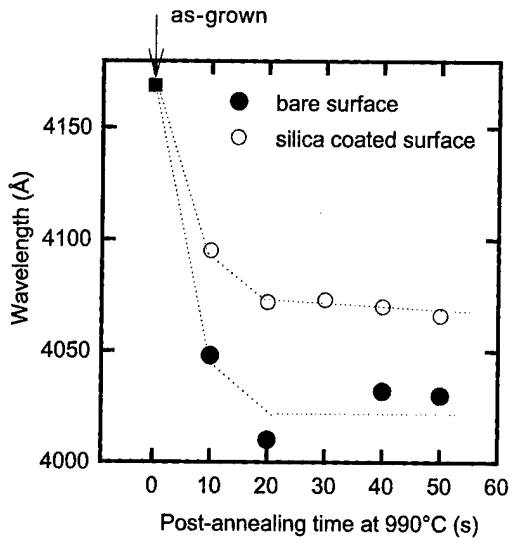


FIG. 2. Effect of the thermal annealing time on the blueshift of the LT-PL peak wavelength of an $In_{0.2}Ga_{0.8}N$/GaN SQW. The thermal annealing temperature was 990 °C. The dashed lines are a guide to the eyes.

In non-nitride based III-V QW structures, QW intermixing relies on the use of a dielectric encapsulant layer (such as SiO_2). Recently, we reported that a spin-on silica layer was used as a more reliable and more effective dielectric encapsulant layer in impurity-free vacancy enhanced QW intermixing.[4] So far, there have been no reports on the thermal annealing of nitride based QW structures with capping of the dielectric encapsulant layer. In this work, the same spin-on silica layer as that used in Ref. 4 was applied to an $In_{0.2}Ga_{0.8}N$/GaN SQW. This thin silica layer was formed by spinning commercial liquid silica onto the surface at a rotation speed of 3000 rpm for 30 s. The spin-coating process was repeated after an interval of 10 s. The process resulted in a silica capping layer with a total thickness of 2000±60 Å. The coated wafers were then baked at 300 °C for 2 h in air. Silica coated samples were placed side by side to noncoated samples and subjected to the same thermal annealing. Contrary to what we found in other QW structures, the spin-on silica actually plays the role of suppressing rather than enhancing the blueshift of the LT-PL peak wavelength of $In_{0.2}Ga_{0.8}N$/GaN SQWs at 990 and 900 °C. The effect of the thermal annealing time on the silica coated samples is also presented in Fig. 2. Obviously, apart from a reduced blueshift of the LT-PL peak wavelength of the samples coated with silica, the dependence of the blueshift on the thermal annealing time is similar to that obtained using the samples without a silica coating.

With the use of a spin-on silica layer, enhanced QW intermixing can take place at about 850 °C for InGaAs/GaAs and AsGaAs/GaAs QW structures,[4] which is the result of re-distribution of point defects generated at the interface between the spin-on silica layer and the top surface (usually GaAs). The inertness of GaN may prevent any effective interfacial reactions from taking place at the interface between silica and GaN. The blueshift of the LT-PL peak wavelength observed in $In_{0.2}Ga_{0.8}N$/GaN QWs may not rely on the same mechanism as that commonly known and has been proven very effective in non-nitride based QW structures.[1,4] We found that the variation of $In_{0.2}Ga_{0.8}N$ well thickness by up to 20% did not lead to a significant change of the PL peak energy. So, the slight interdiffusion of In and Ga across an InGaN/GaN interface might not necessarily lead to a significant blueshift of the LT-PL peak wavelength of $In_{0.2}Ga_{0.8}N$/GaN QWs. We propose that the blueshift of the LT-PL peak wavelength of $In_{0.2}Ga_{0.8}N$/GaN QWs is the result of spatial re-distribution of atomic In and Ga within the $In_{0.2}Ga_{0.8}N$ well at an elevated thermal annealing temperature.

Several recent studies suggested that solid phase immiscibility[5] might induce phase separation in InGaN, which could eventually lead to formation of quantum dots

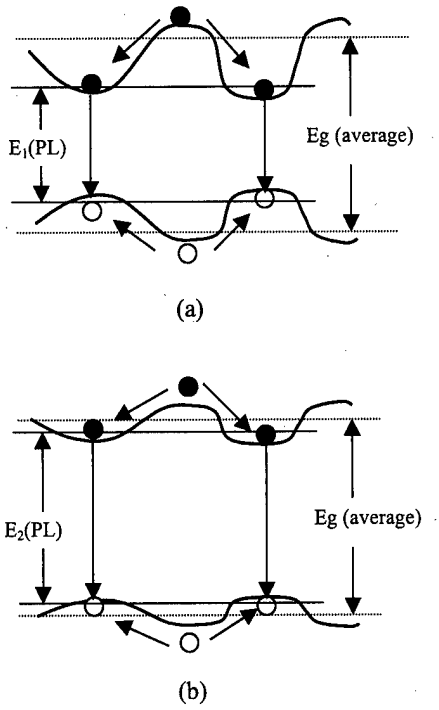


FIG. 3. Schematic diagrams of potential fluctuation induced by compositional inhomogeneity in InGaN. (a) Before thermal annealing and (b) after thermal annealing. Eg (average) is the band gap corresponding to the average In composition in the well. E_1 (PL) and E_2 (PL) are the PL peak energies before and after thermal annealing.

and/or spatially compositional inhomogeneity in InGaN.[6–8] Previous study also suggested that a broad PL profile such as that seen in Fig. 1 arises from compositional inhomogeneity rather than from piezoelectric fields or the strain effect.[9] The potential fluctuation induced by the spatially compositional inhomogeneity is schematically shown in Fig. 2(a). For compositional inhomogeneity induced potential fluctuation, the conduction band minima always correspond to the valence band maxima. So, photoexcited electrons will first relax by diffusion into the lowest available energy states (or local conduction band minima) and then recombine with holes at the corresponding valence band maxima. Those recombination processes will dominate photoluminescence emission.

At an elevated thermal annealing temperature, compositional inhomogeneity drives spatial migration of atomic In and Ga, respectively, from their rich regions to their poor regions. This spatial re-distribution flattens the conduction band mixima and valence band maxima. The energy gap between the conduction band mixima and valence band maxima widens, leading to the blueshift in the LT-PL peak wavelength of $In_{0.2}Ga_{0.8}N$/GaN QWs [see Fig. 3(b) $E_2(\mathrm{PL}) > E_1(\mathrm{PL})$]. These atomic diffusions in the QW structure are usually enhanced by the presence of point defects.[10] For InGaN/GaN QWs, high densities of point defects like Ga, In and N vacancies exist in both InGaN and GaN layers most likely at nonequilibrium concentrations and can substantially enhance atomic In and Ga migration. These nonequilibrium point defects will be annihilated through reactions with other large defects like dislocations in the course of thermal annealing. So, the large blueshift in the LT-PL peak wavelength occurs at the beginning of thermal annealing. As long as these point defects no longer exist or their concentrations are substantially reduced, the spatial migration of atomic In and Ga cannot occur at moderate thermal annealing temperatures. A simple increase of the thermal annealing time does not lead to further blueshifting of the QW band gap energy (see Fig. 2).

The suppression of the blueshift with a spin-on silica layer means that interfacial reactions do not take place at the interface between silica and GaN. Otherwise, the generation of point defects would enhance the spatial migration of atomic In and Ga in the well. At the thermal annealing temperatures used in this work, a bare GaN surface starts to decompose and release atomic N.[11] This decomposition will produce a large number of atomic N vacancies on the surface. The gradient in the atomic N vacancy concentration adjacent to the surface will subsequently drive the diffusion of N vacancies into the QW. A larger blueshift in the LT-PL peak wavelength obtained in the samples without a silica coating than those with a silica coating means that the spin-on silica actually prevent the decomposition of GaN and therefore suppress the blueshift in the LT-PL peak wavelength.

In conclusion, blueshifting of the $In_{0.2}Ga_{0.8}N$/GaN SQW band gap energy was reported after the samples were subjected to thermal annealing in atmospheric N_2 ambient at temperatures below 1000 °C. The blueshift of the LT-PL peak wavelength can be enhanced by increasing the thermal annealing temperature and suppressed by coating a spin-on silica layer on the surface. Increased thermal annealing time only increases the blueshift at an early stage. The blueshift of the LT-PL peak wavelength of $In_{0.2}Ga_{0.8}N$/GaN SQWs is explained by flattening of the potential fluctuation as a result of spatial re-distribution of atomic In and Ga in the InGaN quantum well.

[1]See, for example, a review article, H. Marsh, Semicond. Sci. Technol. **8**, 1136 (1993).
[2]M. D. McCluskey, L. T. Romano, B. S. Krusor, N. M. Johnson, T. Suski, and J. Jun, Appl. Phys. Lett. **73**, 1281 (1998).
[3]E. Iliopoulos, D. Doppalapudi, H. M. Ng, and T. D. Moustakas, Appl. Phys. Lett. **73**, 375 (1998).
[4]G. Li, S. J. Chua, S. J. Xu, and X. C. Wang, Appl. Phys. Lett. **73**, 3393 (1998).
[5]T.-h. Ho and G. B. Stringfellow, Appl. Phys. Lett. **69**, 2701 (1996).
[6]K. C. Zeng, M. Smith, J. Y. Lin, and H. X. Jiang, Appl. Phys. Lett. **73**, 1724 (1998).
[7]S. Chichibu, K. Wada, and S. Nakamura, Appl. Phys. Lett. **71**, 2346 (1997).
[8]S. Chichibu, T. Sota, K. Wada, and S. Nakamura, J. Vac. Sci. Technol. B **16**, 2204 (1998).
[9]W. Shan *et al.*, J. Appl. Phys. **84**, 4452 (1998).
[10]R. M. Cohen, Mater. Sci. Eng., R. **20**, 167 (1997).
[11]C. B. Vartuli, S. J. Pearton, C. R. Abernthy, J. D. MacKenzie, E. S. Lambers, and J. C. Zolper, J. Vac. Sci. Technol. B **14**, 3523 (1996).

New self-aligned processes for III–V electronic high speed devices

J. Etrillard[a)]
OPTO+, Groupement d'Intérêt Economique, Route de Nozay, 91460 Marcoussis, France

C. Besombes, L. Bricard, and A. M. Duchenois
France Telecom-CNET, Lab de Bagneux, P. 249A, 92225 Bagneux, Cedex, France

(Received 11 September 1998; accepted 7 May 1999)

We report on two new self-aligned processes intended for microelectronic devices realization exhibiting a significant reduction of the number of elementary technological levels. These processes allow to obtain a very short self-aligned contact separation without using spacers or wet etched mesa overhangs. As an example, only three lithographic masks are used to realize a triple mesa self-aligned heterojunction bipolar transistor (HBT). A bilayer resist process is used to define the specific shape of the upper contact that is used for self-alignment. A combination of selective, isotropic, or anisotropic processes and very simple selective lift-off processes are used to define the mesa and the contacts which are also used as masks during etching. The alloying of contacts may be performed just after deposition and lift-off. These processes can reduce the production cost and increase the reliability for integration in comparison with conventional self-alignment using the selective wet etched emitter overhang in the HBT application. Furthermore, the parasitic access resistance can be reduced both by using thin mesa active layers and decreasing the contacts separation. This separation length can be determined by the aspect ratio of the bilayer resist, the characteristics, and parameters of the contact deposition equipment. At last, low induced damage inductively coupled plasma dry etch processes are partly used to reduce the dry etch damages.
© *1999 American Vacuum Society.* [S0734-211X(99)06104-1]

I. INTRODUCTION

InP-based electronic devices have been extensively investigated for high speed optical fiber transmission. As an example, heterojunction bipolar transistors (HBTs) and high electron mobility transistors are promising for high-speed microwave and optoelectronic applications. InP devices have many advantages over other III–V devices: high speed characteristics especially with the use of InGaAs for the base and collector materials in the case of HBT, low $1/f$ noise, low turn-on voltage allowing reduction of power consumption in circuits. Impressive high-frequency results were reported for metalorganic chemical beam epitaxy or molecular beam epitaxy InP/InGaAs HBTs.[1] High-frequency performances are obtained through a significant reduction of parasitic resistances or capacitance which are controlled from the technological processes, mainly by reducing extrinsic device dimensions. This reduction implies the use of a self-alignment process at least for the emitter-base contacts to minimize the extrinsic base resistance It is also necessary to decrease the base-collector capacitance and then the collector contact dimensions.[2,3]

We report on new self-alignment processes that allow one to obtain very short contact separation without using spacers or wet chemical mesa overhangs. For instance, a self-aligned bipolar device is realized with only three lithographic masks which can increase the reliability of integration processes without reducing the alignment precision. It could also be suitable for an application to devices which use low contact separation such as waveguide photodiodes.

a)Electronic mail: jacky.etrillar@cnet.francetelecom.fr

II. EXPERIMENT

The processes were performed on InP wafers for simulation of a complete realization of devices.

Both processes use a bilayer resist to obtain a good lift-off profile for the first contact as described in Fig. 1(a). The bilayer consists of a 400-nm-thick layer of poly(methyl/methacrylate) (PMMA) which is hard baked and a 2-μm-thick layer of 4533 from Hoestch.

Other thicknesses could be used as explained later.

This bilayer is spinned over a 600-nm-thick layer of plasma enhanced chemical vapor deposition (PECVD) Si_3N_4. The resist of layer is ultraviolet (UV) exposed and developed in the conventional way to open the area of the first contact.

The PMMA layer is then etched by a conventional downstream reactor (Plasmatechnology μP 80) in a NO_2/O_2 plasma mixture at a 1 Torr pressure to obtain the overhang profile as illustrated in Fig. 1(b). The characteristic lengths D, H, and L of the structure are defined in that figure.

For that process, the resist layer is used as a selective mask after being reinforced by UV and hard baked.

At that point, the processes differ: in the first process, the contact will be made with a "T" shape, in the second process with a "V" shape.

We describe in the following the first of these processes.

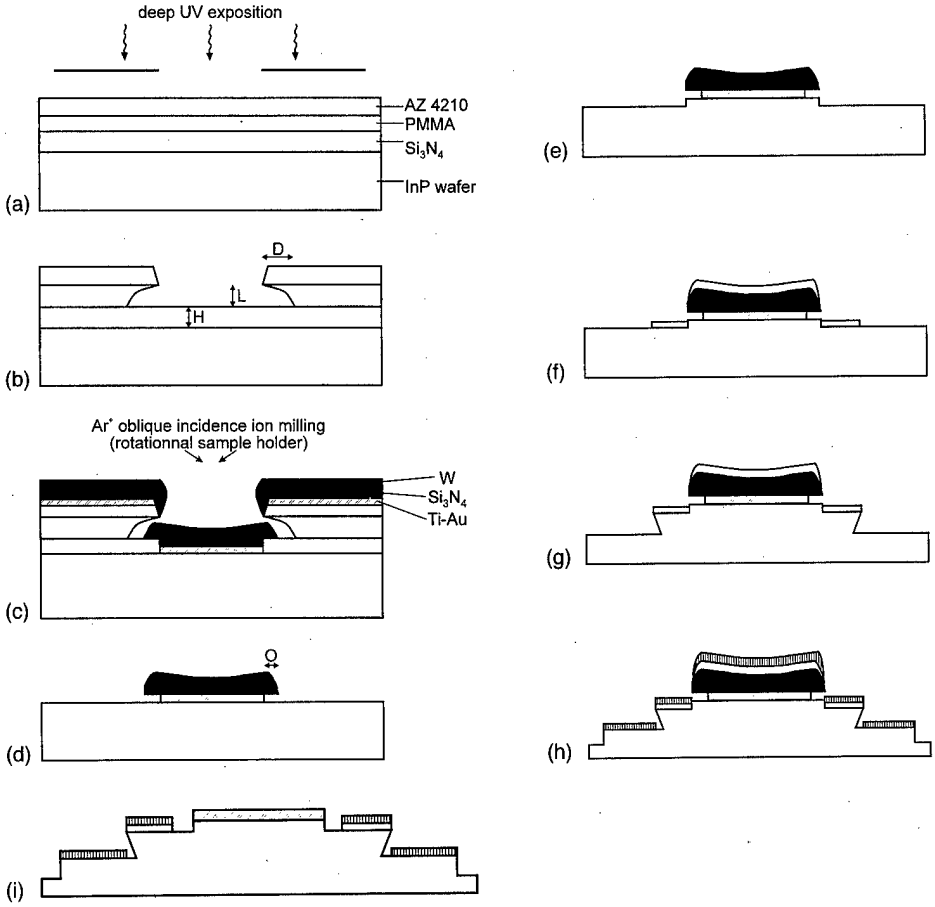


FIG. 1. Description of the T-shape process.

A. T-shape contact process

A conventional reactive ion etching (RIE) plasma process is used to obtain an anisotropic etch of the Si_3N_4 layer.

The first contact is then deposited by e-beam evaporation as shown in Fig. 1(c). In our experiment we use a Ti–Au contact.

After this first deposition a 100 nm layer of Si_3N_4 is deposited by PECVD and covered by a thin layer of tungsten obtained by low pressure sputtering with Alcatel 600 Equipment. A short (1 min long) oblique incidence Ar^+ ion milling etch may be performed using Veeco Microetch equipment before the deposition of silicon nitride to eliminate sidewall Ti–Au thin deposits as shown in Fig. 1(c).

A first lift off is then performed in a hot trichlorethylene bath to eliminate the resist bilayer and the deposition that are laying over, and the wafer is cleaned by solvent rinse. The Si_3N_4 sublayer is then wet etched by buffered hydrofluoric acid with an etch rate of 0.2 μm/min. During this wet process the 100-nm-thick Si_3N_4 layer is almost not etched since covered by the unetched tungsten layer as illustrated in Fig. 1(d). The T side width O is defined in that figure.

The wafer is then etched using a conventional RIE in methane hydrogen plasma or by inductively coupled plasma (ICP) to reduced the induced damages. The etch process is anisotropic to obtain a first mesa which can simulate the emitter mesa in the case of the HBT specific application. This dry-etch process is followed by a selective thin wet etch process for simulation of an etch stop on the base layer as seen in Fig. 1(e). A parasitic conduction area between emitter and base is seen in this figure. This area is etched away in the following steps, as explained farther.

The second inverted resist lift-off mask is used to obtain the second contact that is aligned with a relative coarse precision on the first contact as illustrated in Fig. 1(f). This process is followed by a second dry etching by ICP or RIE to obtain the third mesa using the first two contacts as masks during etching. The dry process which is used exhibits an isotropic profile to avoid the parasitic contact between the second and third contacts on the sidewall of the second mesa as seen in Fig. 1(g). This second mesa is high enough to obtain an overhang of about 250 nm with an isotropic dry process. That is not the case for the first mesa whose height is too low (about 250 nm) to allow a reproducible overhang.

The third lithographic level is then used to obtain the last contact (which simulates the collector contact in the HBT application) as illustrated in Fig. 1(h), by inverted resist as used for the second contact. A solvent lift off is then performed.

The three device contacts are used to etch the isolation mesa as illustrated in Fig. 1(h). This dry etch also uses an anisotropic process.

At last, a second lift off is used to eliminate the tungsten deposition which has been made on the first contact by dissolving the silicon nitride layer in diluted hydrofluoric acid.

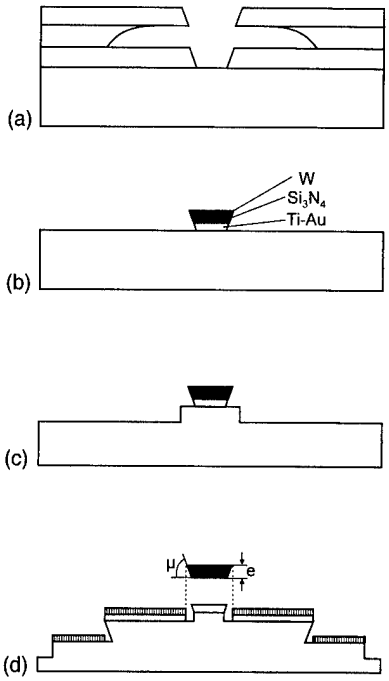


FIG. 2. Description of the V-shape process.

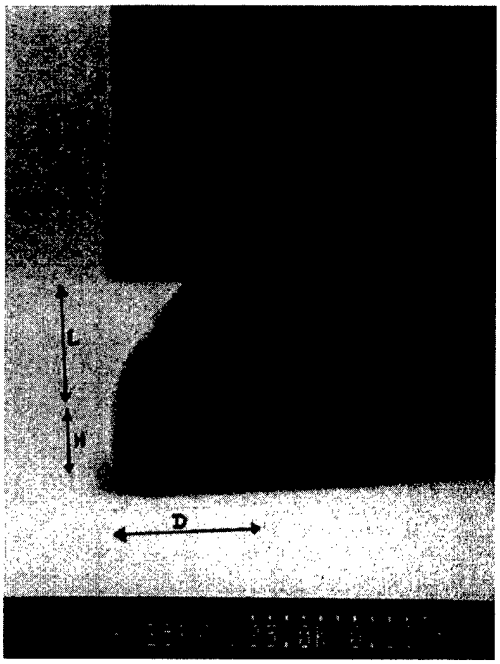


FIG. 3. Bilayer over the anisotropically etched silicon nitride layer. L is the thickness of the PMMA, H the thickness of the Si_3N_4, and D the overhang of the AZ resist.

This process allows one to dry etch the first mesa layer area that is between the first and second contacts. Then, no more parasitic conduction is observed between the two first contacts whose separation is given by the width O of the T side. This width is obtained by the depth of the PMMA lateral etching and by the deposition process equipment, as explained in Sec. III. The final result is shown in Fig. 1(i).

B. V-shape contact process

The V-shape process uses the same bilayer resist as illustrated in Fig. 1(a). The main difference with the T-shape process appears in Fig. 2(a). Instead of an anisotropic etching of the silicon nitride which is used in the T shape, we use a slope-transfer etching process illustrated in Fig. 2(a). This kind of process is sufficiently reproducible and uniform to guarantee a constant angle of the V slope.

Once the V shape is obtained in the bilayer and in the silicon nitride layer, the first contact deposition is made on the wafer followed by a deposition of silicon nitride and tungsten as already described in the T-shape realization. The lift off is then performed and a V-shape multilayer is obtained as illustrated in Fig. 2(b).

After lift off in hot trichlorethylene the wafer is dry etched as illustrated in Fig. 2(c).

All the following technological processes are then similar to those exposed in the T-shape process. The selective lift off which is performed at the end of the process gives a separation between the first and the second contact. The value of this separation distance is given by the formula

$$d = e.\cot g\,\mu,$$

where μ is the angle transferred in the silicon nitride layer and e is the total thickness of the selective lift-off layer (second silicon nitride layer and tungsten layer) as seen in Fig. 2(d).

It is easy to obtain a good precision of the deposited layers thickness and then a good precision of the value of the separation length of the two contacts. In fact e can be obtained with a precision of 2% in recent deposition equipments and can also be defined with a great precision in conventional RIE reactors.

In our experiments the results were reproducible and correctly modeled by the earlier formula. It is also clear that with such a process, separation distance of less than 100 nm can be obtained as exposed in the next section.

III. RESULTS

As described earlier, the self-alignment processes are based on a bilayer resist technique which has to be well controlled and reproducible.

Moreover, for both processes, the aspect ratio of the bilayer structure is defining the size and shape of the first contact. An example of a resist bilayer is seen in Fig. 3. We define again in that figure the characteristic lengths D, L, and H of this structure. The width of the T shape, O [see Fig. 1(d)] is determined by the trajectories of the atoms coming from the deposition source. Then, the width O is a function of the ratio D/L.[4]

This width O is also the separation between the first two contacts as explained earlier. The geometric characteristics of the deposition equipment are involved. For the same pressure and e-beam or radio frequency (rf) power, the deposi-

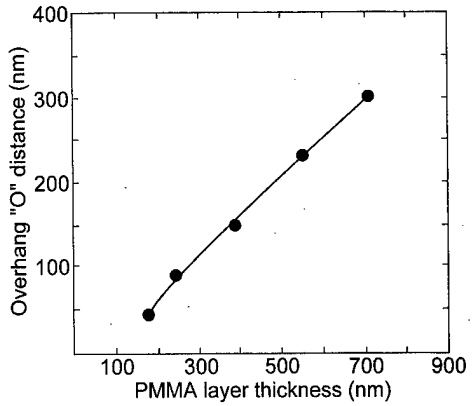


FIG. 4. Overhang O distance as a function of the PMMA layer thickness. The depth of the overetch is about twice the PMMA layer thickness in each case.

tion rate, and angular distribution are constant. O was obtained with a good reproducibility (less than 5%) in our study.

The thickness [H is defined in Fig. 1(b)] of the silicon nitride is chosen as a constant since it also determines the T-shape contact height. For the tungsten layer we used a rf sputtering conventional equipment with a high deposition rate to reduce sidewall deposition.[4]

The variation of overhang length O versus the PMMA thickness L is illustrated in Fig. 4 in the case of e-beam evaporation. The plot is almost linear which is explained by the low variation (with the PMMA thickness) of the solid angle formed by the deposition source position and the wafer which is the case in our e-beam equipment.

The overhang D of the resist bilayer has been chosen larger than the value of the T-shape overhang O to avoid the parasitic deposition on the PMMA sidewalls.

For a given thickness L of the PMMA layer, the value of D is a function of the plasma etch time as seen in Fig. 5. As observed in this figure, the lateral etch rate of PMMA is slower for low PMMA thickness than for thick PMMA layer.

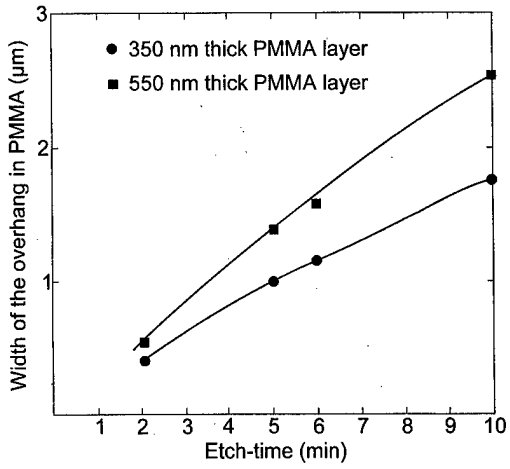


FIG. 5. Overhang width D under the AZ resist as a function of downstream plasma etch time.

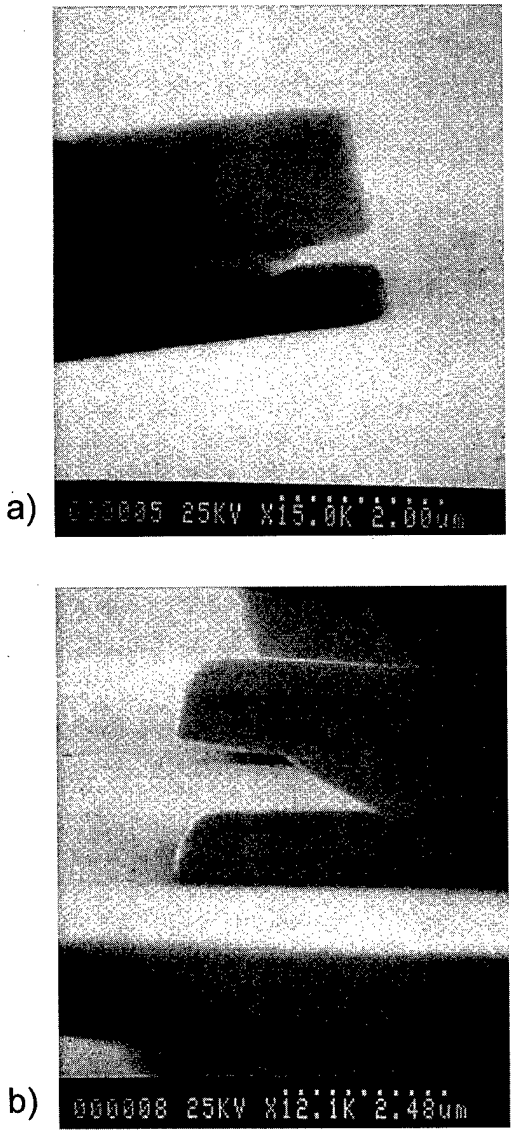


FIG. 6. Shape of the overhang of the bilayer is dependent of the overetch time. For (a), the etch time is 3 and 5 mn for (b).

The shape of the overhang is also noticeable when a long overetch is performed on the bilayer as seen in Fig. 6. The lateral etch rate and the profile pattern is also dependent of the baking temperature of the PMMA.

The pattern dimension is given by the opened AZ resist top layer of the bilayer structure. These bilayer materials are both etched in an oxygen plasma. Then the selectivity between the vertical direction etching for AZ resist and the lateral direction for the PMMA layer has to be high. This selectivity is about 15. The dimension loss is then very low.

The T-shape contact is observed in Fig. 7, the V shape in Fig. 8. In each case, the complete simulation of realization are illustrated in Figs. 9 and 10. Anisotropic etching of the mesa have been processed by conventional RIE using CH_4/H_2 plasma with a ratio of 10% of CH4 in flux at a pressure of 50 mT and power density of 0.8 W/cm^2. The bias voltage was 450 V. For induced damage sensible mesa etch-



FIG. 7. T-shape contact. The *T* sides are constituted by Si_3N_4 and tungsten.

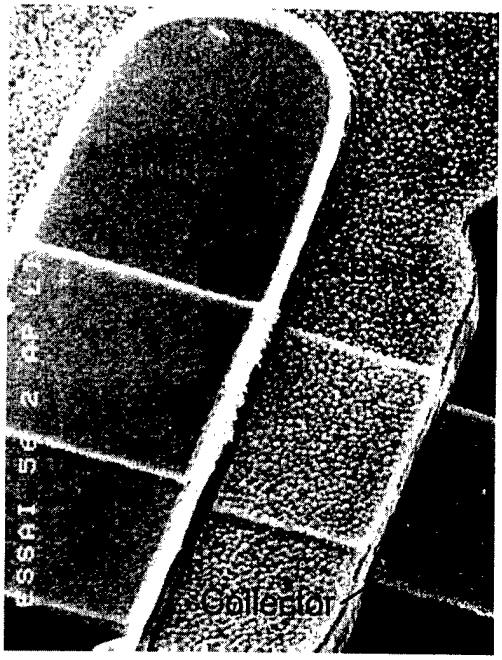


FIG. 9. Simulation of a full self-aligned HBT with a T-shape emitter contact. Separation between emitter and base is about 100 nm.

ing, ICP[5] was used in a $CH_4/O_2/H_2$ mixture with a bias voltage of 20 V.

The second lift off to eliminate the cap layer of Si_3N_4/W at the end of both processes was performed using diluted hydrofluoric acid. The lateral etch rate of Si_3N_4 under the tungsten layer is about 2 μm/min. No chemical etching of the III–V is observed. The lateral etch rate is an increasing function of the Si_3N_4 layer thickness. That is also the case for

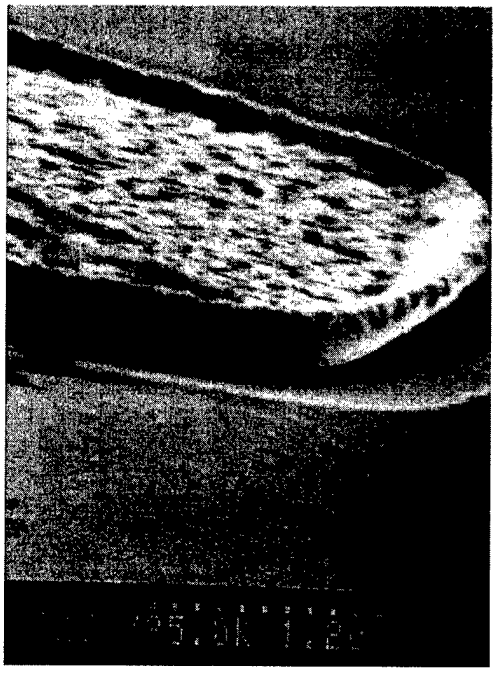


FIG. 8. V-shape contact.

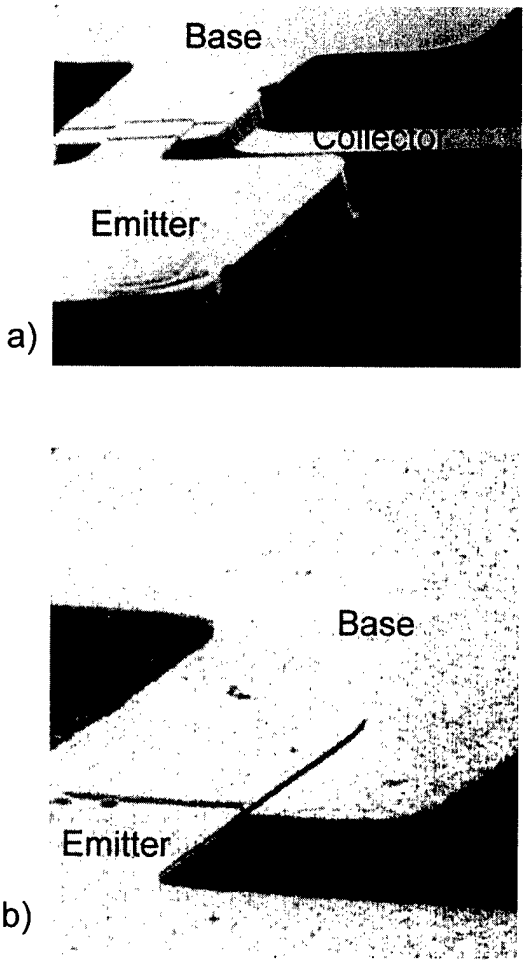


FIG. 10. (a) Simulation with V-shape emitter contact. (b) Magnification of the separation between emitter and base contact.

titanium. Then the layer of titanium is kept at 15 nm to improve the selectivity.

The separation between the self-aligned contacts is about 100 nm for the T-shape process and about 65 nm for the V-shape contact as observed in Figs. 9 and 10.

IV. CONCLUSION

Simple processes have been studied to simplify the realization of self-aligned devices and applied to the simulation of the realization of bipolar transistors on InP wafers. Selective wet lift off and isotropic and anisotropic dry etch processes were used. It is shown that very short separation distance can be obtained for the self-aligned contacts reducing the number of lithographic masks. These processes may allow one to increase the devices performance, the reliability of integration process, and to reduce the fabrication cost.

ACKNOWLEDGMENTS

The authors appreciated the assistance of A. Scavennec and J. Godin for reviewing the manuscript and for technical discussions.

[1]M. Hafizi, W. E. Stanchina, F. Williams, Jr., and J. F. Jensen, IEEE Microwave and Millimeter-wave Monolithic Circuits Symposium, 1995.
[2]M. K. Mishra, J. F. Jensen, D. B. Rensch, A. S. Brown, W. E. Stanchina, R. J. Trew, M. W. Pierce, and T. S. Kargodorian, IEEE Electron Device Lett. **10**, 467 (1989).
[3]H. F. Chau and E. A. Beam, IEEE Electron Device Lett. **14**, 388 (1993).
[4]J. L. Vossen and W. Kern, *Thin Film Process II* (Academic, New York, 1991), p. 79; J. L.Vossen, J. Vac. Sci. Technol. **8**, 13 (1971).
[5]J. J. Etrillard *et al.*, J. Vac. Sci. Technol. A **14**, 1056 (1996).

Effect of short-time helicon-wave excited N_2–Ar plasma treatment on the interface characteristic of GaAs

Satoshi Wada, Keisuke Kanazawa, Nariaki Okamoto, and Hideaki Ikoma[a)]
Faculty of Science and Technology, Science University of Tokyo, Noda, Chiba 278, Japan

(Received 7 August 1998; accepted 23 April 1999)

The short-time helicon-wave excited N_2–Ar plasma treatments of *n*-type GaAs (100) substrates were performed in order to investigate the effect of these treatments on the interface characteristics of GaAs and to explore a possibility of a robust surface passivation of GaAs. X-ray photoelectron spectroscopic measurements indicated formation of GaN at the insulative-layer GaAs interface. The current-density–voltage ($J-V$) characteristics of the Schottky or tunnel metal–insulator–semiconductor diodes were measured before and after the plasma treatment. The analysis of these data suggested that the "intrinsic" surface Fermi level pinning near the midgap of the GaAs forbidden band was partially removed and the Fermi level was newly pinned at $E_C-0.2-0.3$ eV (E_C: the conduction band edge) after the plasma treatment. This is probably due to generation of high-density plasma-induced donor-like interface states having the energy level (or levels) in this energy region. The reverse leakage current increased with this plasma treatment. However, it decreased after the post-thermal annealing at moderate temperatures (200–500 °C) in N_2 ambient. The thermal and air-exposure stability were substantially improved with the plasma treatment. The $J-V$ characteristics did not show any deterioration after air exposure for three months for the plasma-treated samples, whereas these greatly degraded with air exposure for three weeks for the untreated ones. © *1999 American Vacuum Society.* [S0734-211X(99)01704-7]

I. INTRODUCTION

Gallium arsenide is a promising material for the high-speed and low-power device applications. However, the GaAs metal–insulator–semiconductor (MIS) devices are not yet in practical use because the very high density of interface states brings about the surface Fermi level pinning near or at the midgap of the forbidden band of bulk GaAs,[1–7] which inhibits the formation of the inversion and accumulation regions under the gate electrode. As well, the prominent passivation techniques have not been established for the GaAs-based devices such as the high-electron mobility transistor, heterobipolar transistors, and semiconductor lasers. This is again due to the unstable surface or interface of GaAs. The sulfur passivation is one of the techniques for stabilizing the interface or surface of compound semiconductors. However, the sulfur-passivated GaAs surface shows the gradual degradation when it is exposed to air.

Recently, we investigated the helicon-wave excited N_2–Ar plasma treatments for lasting about 4–8 h in order to obtain a robust surface passivation and explore the possibility of GaAs MIS devices.[8] X-ray photoelectron spectroscopic (XPS) data showed the formation of gallium nitride (GaN) or Ga–N bonds at the interface between the grown insulative layer and the GaAs (100) substrate. Fairly good capacitance–voltage ($C-V$) characteristics were observed for the MIS diodes fabricated with these insulative-layer/GaAs structures, which suggested a partial removal of the Fermi level pinning at the midgap. However, some anomalous behaviors of $C-V$ characteristics were often observed, such as the decrease in C values and frequency dispersion effect in the "accumulation" region. This probably suggests that the Fermi level pinning near the midgap is once removed and the Fermi level is newly pinned at the energy position near the conduction band edge.[9]

More recently,[9] we investigated the effect of the short-time (5–15 min) helicon wave excited N_2 plasma treatments on the Fermi level pinning in GaAs (100) substrate by measuring the electrical characteristics of Schottky or tunnel MIS diodes in order to check the earlier consideration. The observed results indicated that the Fermi level pinning at the midgap was removed and the Fermi level was newly pinned at the energy position of about 0.3–0.4 eV below the conduction band edge ($E_C-0.3-0.4$ eV). This is probably due to the formation of a large amount of the plasma-induced defects whose energy level lies in these energy regions. However, in the case of N_2 plasma (without Ar mixing) treatment, only the oxide was formed on the GaAs surface and GaN did not exist at the interface with GaAs.

In the present study, the short-time helicon-wave excited N_2–Ar plasma treatment of GaAs was performed to investigate the effect of the coexistence of the GaN adjacent to the interface on the Fermi level pinning effect. GaN was formed at the interface with Ar mixing in N_2 gas. The Fermi level pinning at the midgap was also once removed, however, it occurred at the energy position of $E_C-0.2-0.3$ eV (E_C: the conduction band edge of the bulk GaAs) similar to the case of the N_2-plasma treatment.[9] The current-density–voltage ($J-V$) characteristics of the Schottky (or tunnel MIS diodes) somewhat degraded after the N_2–Ar plasma treatment compared with those for nontreated Schottky diodes. However, those were recovered after the post-thermal annealing at moderate temperature in N_2 ambient. The thermal stability

[a)]Electronic mail: ikoma@ee.noda.sut.ac.jp

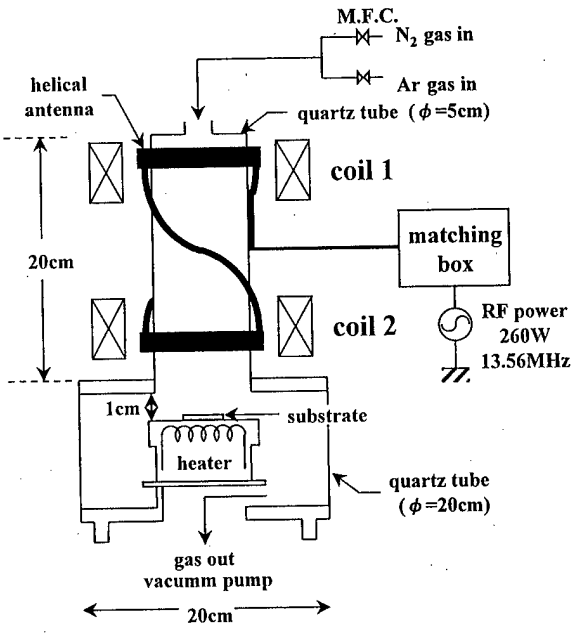


FIG. 1. Schematic diagram of the plasma processing system.

was greatly improved by the N_2–Ar plasma treatment. The $J-V$ characteristics did not show any degradation after the air exposure for three months, while these largely deteriorated for the untreated samples only after the air exposure for three weeks. The experimental procedures are presented in Sec. II. The experimental results are described and discussed in Sec. III and the summary and conclusions are given in Sec. IV.

II. EXPERIMENTAL PROCEDURES

A. Sample preparations

The *n*-type GaAs (100) substrates with the free carrier concentration of about 7×10^{16} cm^{-3} were employed. The substrates were ultrasonically cleaned in dichlorethane, acetone, ethanole, and deionized water each for 5 min and then soaked in diluted HCl solution (18%) to remove the native oxide. The Au–Ge alloy was deposited on the back surface of the substrate and annealed at 260 °C for 10 min in N_2 ambient to form the ohmic contact. Then, the same organic and HCl treatments were performed to clean the front surface of the substrate. After rinsing in deionized water and blow drying in N_2 gas, the substrates were inserted into the processing chamber of the remote plasma system schematically shown in Fig. 1. In Fig. 1, the upper portion of the systems is a plasma generation chamber in which N_2 and Ar gases are imported from an inlet in the top. The mass flow controlled system finely controlled the flow rate of each gas. Plasma was created by applying the radio frequency (rf) (frequency of 13.56 MHz) power to a helical antenna wound around the chamber. Vertical magnetic fields were generated by currents in two coils (coil 1 and 2). The lower portion of this plasma system is a processing chamber in which the GaAs substrates are placed on a sample holder located about 1 cm lower from an outlet of the upper plasma generation chamber. The sample holder was earthed. The substrates were not intentionally heated and the sample temperature was estimated to be about 30 °C by a thermocouple attached to the substrate. In this plasma system, the helicon wave is excited as the input rf power exceeds about 250 W. All the plasma treatments here were carried out in the helicon-wave excited plasma (HWP) state. The helicon-wave plasma was used here since the plasma density (10^{12}–10^{13} cm^{-3}) was about one order of magnitude higher than other plasma sources such as the capacitively coupled and electron-cyclotron-resonance plasmas. The reactivity of N_2 was then enhanced when using HWP. The processing condition is shown in Table I. The gas flow rates of N_2 and Ar were 6.0 and 3.0 sccm, respectively. The pressure during processing was kept to be about 2.5×10^{-3} Torr by evacuating the whole system with combined rotary and turbomolecular pumps. The plasma exposure time was 5–60 min to grow very thin insulator films to make the "Schottky" or tunnel MIS diodes. Al or Au was vacuum-evaporated on the plasma-treated and untreated GaAs surfaces through a metal mask in a vacuum of about $1-2\times10^{-6}$ Torr. Some of the Schottky and tunnel MIS diode samples were post-thermally annealed at 200–500 °C for 30 min in N_2 ambient to investigate the thermal stabilities of these samples. Some of them were also exposed to laboratory air for three weeks–three months to study the air-exposure stability.

TABLE I. Processing condition.

Gas flow rate	N_2:Ar=6.0:3.0 sccm
Pressure	2.5×10^{-3}
Plasma exposure time	5–60 min
Coil 1 current	0.3 A
Coil 2 current	0.0 A
Substrate temperature	30 °C
Post-anneal temperature	200–500 °C

B. Measurements

$I-V$ characteristics of the Schottky or tunnel MIS diodes were measured with the conventional digital electrometer controlled by a personal computer. $C-V$ characteristics were also measured at a frequency of 1 MHz by an HP-4280A $C-V$ meter to obtain the carrier density profile along the thickness of the GaAs substrate before and after the plasma treatment. X-ray photoelectron spectroscopic measurements were carried out in a combination with the *in situ* Ar ion etching to study the chemical composition and its depth profile in the grown insulator film due to the plasma treatment. A Mg $K\alpha$ line was used as an x-ray source. The Ga $3d$, As $3d$, N $1s$, and O $1s$ core levels were investigated.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. XPS data

Figures 2(b) and 2(b) show the XPS spectra of the Ga $3d$ and As $3d$, respectively, for the helicon-wave excited N_2–Ar

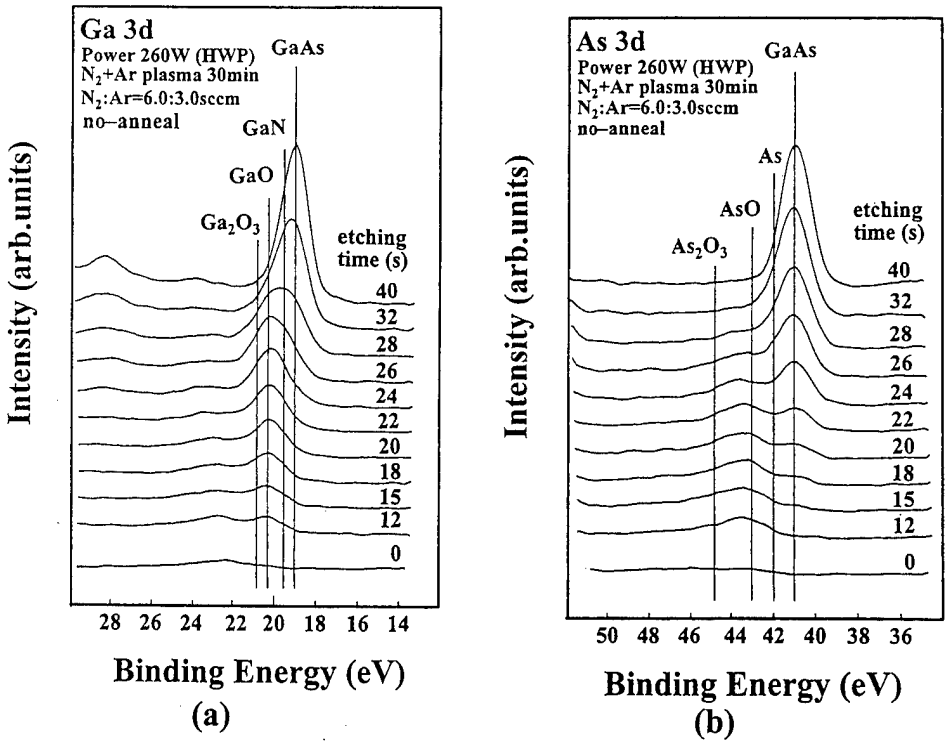


FIG. 2. XPS spectra of (a) the Ga 3*d* and (b) As 3*d* core levels for the GaAs sample treated in HWP N_2–Ar plasma for 30 min.

plasma treated GaAs sample for 30 min without the post-thermal annealing. The binding energies of Ga bonded in bulk GaAs, GaN, GaO, and Ga_2O_3 are indicated by vertical lines and the *in situ* Ar ion etching time is noted as a parameter. From the top surface to the middle portion of the grown film, the XPS signal peak coincides with the binding energy of Ga bonded in GaO (suboxide of Ga), indicating that GaO is the main component of the grown film. However, the XPS peak is at the binding energy of Ga in GaN in the spectrum for the *in situ* Ar ion etching time of 28 s, which shows that GaN is formed at the insulator/GaAs interface in contrast with the case of N_2-plasma treatment [Fig. 5(d) in Ref. 9]. The amount of the Ga oxide (Ga_2O_3 and GaO) is much smaller than those in the native oxide of GaAs.[11]) In Fig. 2(b), the binding energies of As bonded in GaAs,[10] As_2O_3,[10,11] AsO,[11] and elemental arsenic (As)[12] are indicated by vertical lines and the *in situ* Ar ion etching time is also noted. It is clearly shown that the amount of the Ar oxides such as As_2O_3 and AsO, and elemental arsenic are far smaller than those in the native oxide of GaAs.[11] These indicate that the N_2–Ar plasma treatment substantially suppressed oxidation of GaAs as previously shown.[8] Figures 3(a) and 3(b) show the XPS spectra of N 1*s* and O 1*s* core levels, respectively, for the same sample shown in Fig. 2. In Fig. 3(a), the XPS signal from N is observed throughout the whole depth of the grown film though the sensitivity is very small for this signal, showing that nitrogen is distributed along the whole thickness of the film. In Fig. 3(b), it is shown that oxygen is also distributed throughout the entire depth of the film. These are inconsistent to the XPS data shown in Figs. 2(a) and 2(b).

The XPS peak signal assigned here to GaN is possively the superposition of the XPS signals from Ga-bonded in GaAs and GaO. Then, we attempted again to measure the XPS spectra of the Ga 3*d* and N 1*s* core levels for the GaAs samples plasma-treated in the same manner as the sensitivity of the XPS apparatus was raised higher by increasing the intensity of the incident x ray than in the former case (Figs. 2 and 3). Figures 4(a) and 4(b) show the XPS spectra of, respectively, the Ga 3*d* and N 1*s* core levels for the GaAs sample treated in the HWP N_2–Ar plasma for 30 min, measured with the higher sensitivity of the XPS apparatus. The *in situ* Ar ion etching time is denoted as a parameter in both figures. In Fig. 4(a), the XPS spectra for the *in situ* Ar ion etching time of 8 s shows a peak at the binding energy between those corresponding to GaN and GaO. This indicates that both GaN and GaO are present at the interface with GaAs. In Fig. 4(b), the XPS peak is clearly observed at the binding energy (about 397.2 eV) of N bonded in GaN in the spectra for the *in situ* Ar ion etching time of 8 s. (The XPS spectra of the N 1*s* core level shows a peak at the binding energy of about 397 eV for the high-quality GaN single crystal.[13]) From earlier findings, GaN is probably formed (together with a small amount of GaO) at the interface with GaAs. This conclusion is supported by the following discussions. The electrical characteristics of the GaO/GaAs interface is very poor because GaO is incomplete oxide (suboxide) of Ga having a fair amount of the dangling bonds. The leakage current is very high since the electron tunneling through the insulative layer is greatly enhanced by the presence of the high density of the interface states at the GaO/GaAs interface. The reliability is also very inferior to those reported later in the present article for the plasma-treated GaAs samples. The electrical characteristics of the untreated GaAs interface showed substantial deterioration after exposure to the laboratory air as reported later. Very large

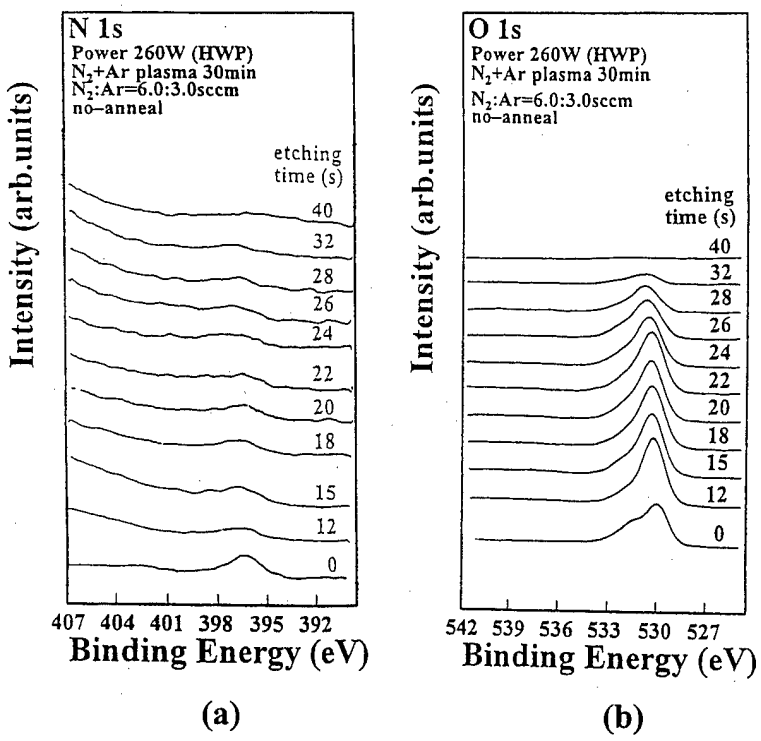


FIG. 3. XPS spectra of (a) the N 1*s* and (b) O 1*s* core levels for the GaAs sample treated in HWP N_2–Ar plasma for 30 min.

amounts of the suboxides such as GaO are present on the surface of the untreated GaAs.[11]

The conclusion here is that GaN (or Ga–N bonds) are generated at the insulator/GaAs interface as is opposed to the case of N_2-only-plasma treatment, where only the oxide is formed.[9]

B. Electrical characteristics of Schottky or tunnel MIS diodes

Figure 5(a) shows the J–V characteristics for the Schottky or tunnel MIS samples fabricated using the N_2–Ar plasma-treated (for 30 min) and nontreated GaAs substrates

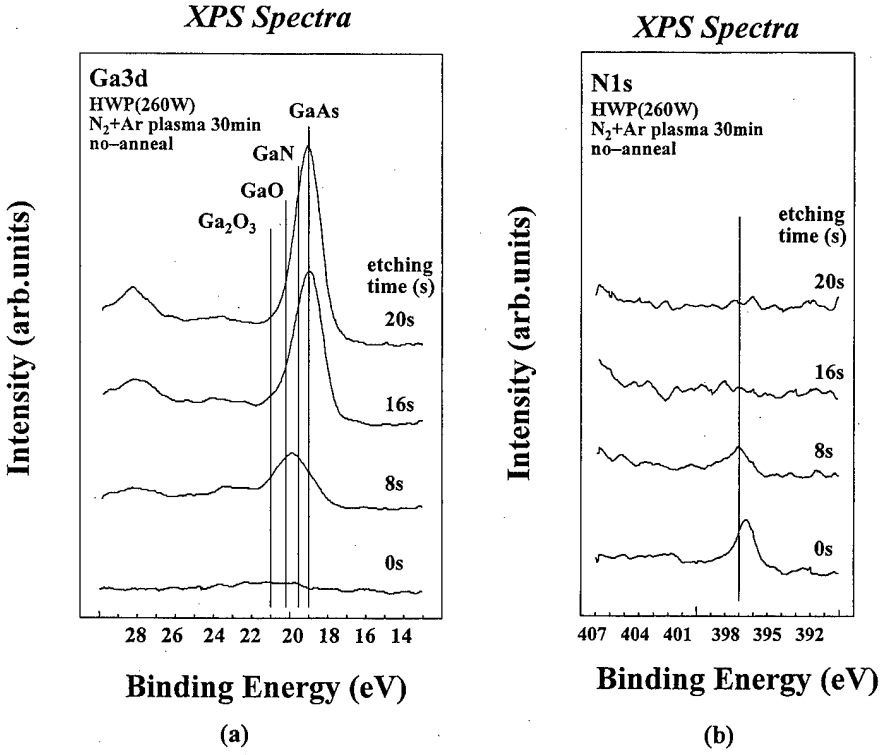


FIG. 4. XPS spectra of (a) the Ga 3*d* and (b) N 1*s* core levels for the GaAs sample treated in HWP N_2–Ar plasma for 30 min with higher sensitivity of the XPS apparatus.

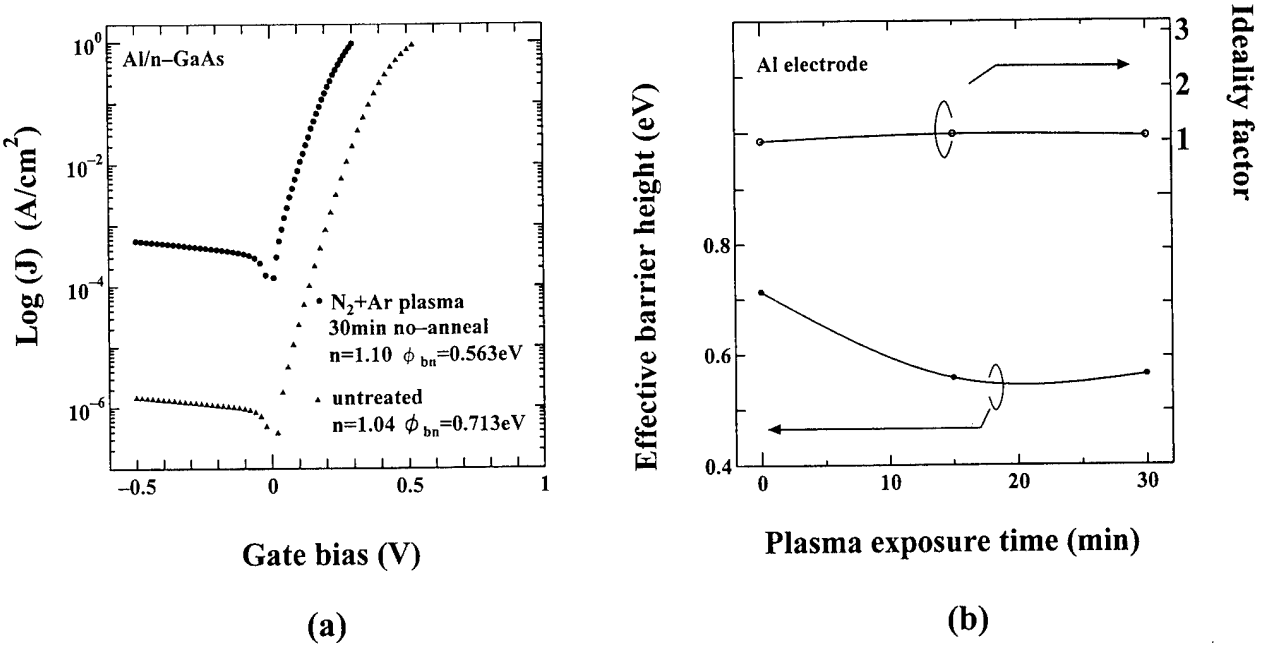


FIG. 5. (a) $J-V$ characteristics and (b) the ideality factor n and effective Schottky barrier height ϕ_{bn} as a function of the plasma-exposure time for the GaAs sample treated in N_2–Ar plasma for 30 min and untreated one. An Al electrode was used.

with using Al as a gate metal. The ideality factor n was calculated from the slope of the forward $J-V$ characteristics. The effective Schottky barrier height ϕ_{bn} are indicated in this figure. The ideality factor n is somewhat higher than that for the untreated sample and the effective Schottky barrier height ϕ_{bn} is lower than that for the untreated one. The reverse leakage current is about 2–3 orders of magnitude larger for the plasma-treated sample than for the untreated one. These findings indicate a high density of the plasma-induced defects are produced at or near the interface with GaAs. These new interface states are probably donor-like ones, resulting in lowering of the effective Schottky barrier height. Figure 5(b) shows the ideality factor n and effective Schottky barrier height ϕ_{bn} as a function of the plasma exposure time. The ideality factor n does not show any remarkable increase as the plasma exposure time is increased. On the other hand, the effective Schottky barrier height ϕ_{bn} decreases from about 0.7 eV to about 0.56 eV as the GaAs sample is plasma-treated and nearly constant with the plasma exposure time. Figures 6(a) and 6(b) show, respectively, the $J-V$ characteristics and the dependences of the ideality factor n and the effective Schottky barrier height ϕ_{bn} on the plasma-exposure time for the N_2–Ar plasma-treated (for 30 min) and untreated GaAs samples fabricated with Au as a gate metal. The quite same trends as shown in Figs. 5(a) and 5(b) are observed, though the ideality factors are nearly the same for both samples. The effective Schottky barrier height ϕ_{bn} also decreases by the plasma treatment from about 0.8 eV for the untreated sample to about 0.65 eV for the plasma-treated one.

The Richardson plot was performed to obtain the true Schottky barrier height ϕ_{bno}. At first, the backward saturation current density J_S was calculated as shown in Fig. 7. Figure 7 shows the backward $J-V$ characteristics for various temperatures between 260 and 300 K. The backward saturation current J_S is defined as a current-density value at an intersecting point of the tangential line of the $J-V$ curve at $V=-0.1$ V and the vertical axis. In Fig. 7, J_s is plotted as a function of the reciprocal of temperature ($1/T$) for the plasma-treated and untreated samples. The true Schottky barrier height ϕ_{bno} is obtained from the slope of the straight line shown in Fig. 8 to be 0.30 and 0.70 eV for, respectively, the N_2–Ar plasma-treated and untreated samples. The true Schottky barrier height thus obtained is nearly equal to the difference between the Fermi level E_{Fs} and the conduction band edge E_C at the interface, i.e., $\phi_{bno}=E_C-E_{Fs}$. In Fig. 9, $E_C-E_{Fs}(=\phi_{bno})$ is shown as a function of the plasma-exposure time. This clearly indicates that the surface Fermi level is pinned at about $E_C-0.7$ eV for the untreated GaAs which is consistent with the previous works. The Fermi level moves towards the conduction band edge as the plasma exposure time increases. E_C-E_{Fs} nearly saturates at 0.2–0.3 eV as the plasma exposure time increases further. This tendency is very similar to the reported results in the previous paper,[9] but the final energy position is somewhat different (in Ref. 9, the Fermi level saturated at $E_C-0.4$ eV). This finding indicates that the "intrinsic" Fermi level pinning near the midgap is once removed and the Fermi level is newly pinned at the energy position of $E_C-0.2$–0.3 eV with the helicon-wave excited N_2–Ar plasma treatment. This strongly suggests that the rather high density of the interface states are created in this energy region probably due to the plasma-induced damages as discussed in the earlier paper for the N_2-only-plasma treatment.[9]

The energy distribution of the interface state density N_{ss} was calculated for the N_2–Ar plasma-treated and untreated

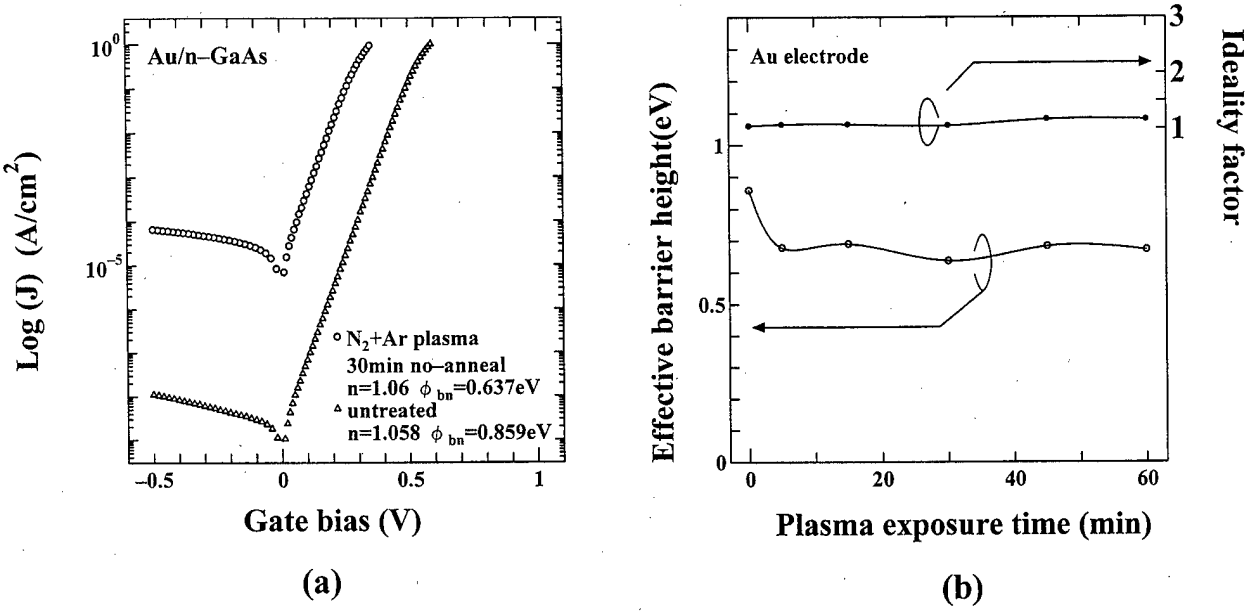


FIG. 6. (a) $J-V$ characteristics and (b) the ideality factor n and effective Schottky barrier height ϕ_{bn} as a function of the plasma-exposure time for the GaAs sample treated in N_2–Ar plasma for 30 min and untreated one. Au was used as a gate electrode.

samples from $J-V$ characteristics using the analysis proposed by Hovárth.[14] The interface state density N_{ss} is written as[14]

$$N_{ss}=\frac{\epsilon_o}{q}\left[\frac{\epsilon_i}{\delta}(n_i-1)-\frac{\epsilon_s}{\omega_F}\right], \tag{1}$$

where

$$\frac{\epsilon_i}{\delta}=\left[\frac{\epsilon_s}{\omega_R}\left(\frac{q}{s_i kT}-1\right)\right]^{-1}. \tag{2}$$

Here, δ is the thickness of the interfacial layer, ϵ_i and ϵ_s are the relative dielectric constant of the interfacial layer and semiconductor, respectively, ω_F and ω_R are the depletion depth for, respectively, the forward and reverse applied voltage, s_i is the slope of the $J-V$ curve in the reverse bias region, and finally, n_i is given by the following equation:[14]

$$n_i=1+\frac{(\epsilon_s/\omega)+)(qN_s/\epsilon_o)}{(\epsilon_i/\delta)+(qN_m/\epsilon_o)}, \tag{3}$$

where N_s and N_m are the densities of the interface states which are in the equilibrium with the semiconductor and the metal, respectively. In the forward bias region of the $J-V$ characteristics, $N_s=N_{ss}$ and $N_m=0$,[14] while $N_s=N_m=0$[14] in the reverse bias voltage region. From Eqs. (1)–(3), the interface state density N_{ss} is obtained as a function of the energy. Figures 10 and 11 show the N_{ss} as a function of the energy relative to the conduction band edge E_C of the bulk GaAs

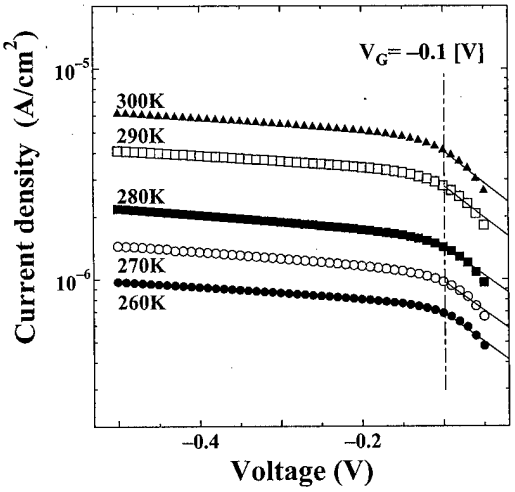


FIG. 7. Current density as a function of the reverse applied voltage for various different temperatures (typical example).

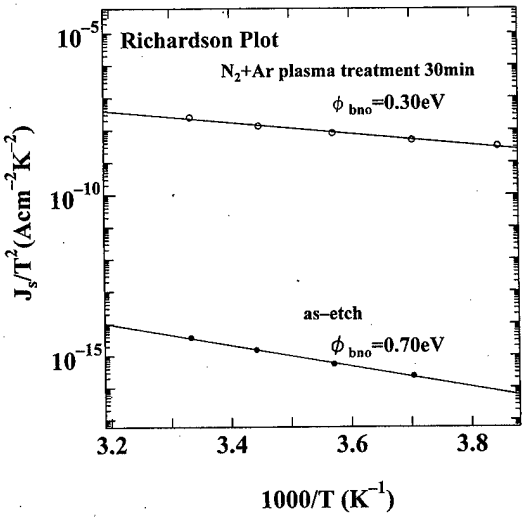


FIG. 8. Richardson plots for the GaAs sample treated in HWP N_2–Ar plasma for 30 min and the untreated one.

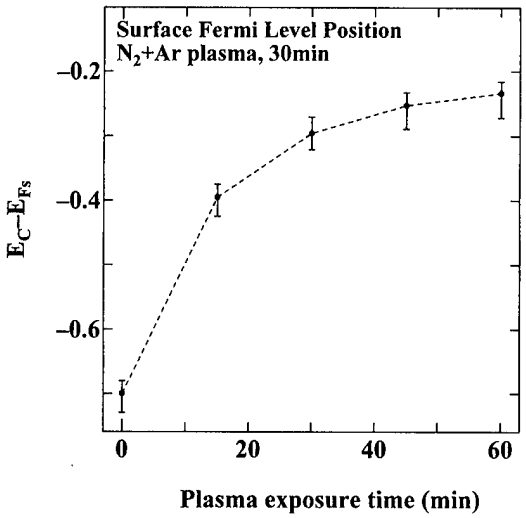


FIG. 9. "True" Schottky barrier height deduced from Richardson plot as a function of the plasma-exposure time.

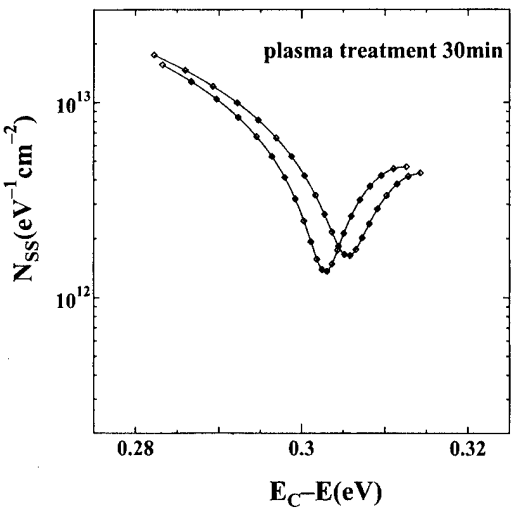


FIG. 11. Interface-state density N_{ss} as a function of E_C-E, for the GaAs samples treated in HWP N_2–Ar plasma for 30 min.

for, respectively, the two untreated samples and the samples treated in the N_2–Ar plasma for 30 min. In Fig. 10, the interface state density N_{ss} has its minimum value $N_{ss_{\min}}$ at the energy around $E_C-0.7$ eV, which is consistent with the reported ones.[1–7] $N_{ss_{\min}}$ is about $2-6\times10^{11}$ $eV^{-1}cm^{-2}$. The intrinsic Fermi level pinning usually occurs near the midgap. In Fig. 11, on the other hand, the interface state density N_{ss} shows a minimum ($N_{ss_{\min}}$) at the energy position of about $E_C-0.3$ eV for the N_2–Ar plasma-treated GaAs samples. The value of $N_{ss_{\min}}$ is about $1-2\times10^{12}$ $eV^{-1}cm^{-2}$ which is higher than those observed for the untreated samples. The Fermi level is then considered to be strongly pinned at about $E_C-0.2$–0.3 eV, consistent with the data for the true Schottky barrier height shown in Fig. 9.

Figure 12 shows the carrier profiles obtained by $C-V$ characteristics for the N_2–Ar plasma-treated and untreated samples. The Au electrode was used in this case. It is shown that the carrier density decreases near the surface of the GaAs substrate for the plasma-treated sample while it is nearly constant for the untreated one. This probably indicates that the acceptor-like defects are introduced in the bulk of GaAs near the surface by the plasma treatment and these acceptor-like defects compensate the bulk donor leading to the decrease in the carrier density near the interface. In the $J-V$ curves shown in Figs. 5(a) and 5(b), the effect of the series resistance is observed at the lower applied forward voltage for the plasma-treated sample than that for the untreated one, indicating that the series resistance is higher for the former than the latter. This increase in the series resistance is due to the presence of the low-carrier density region in the plasma-treated sample.

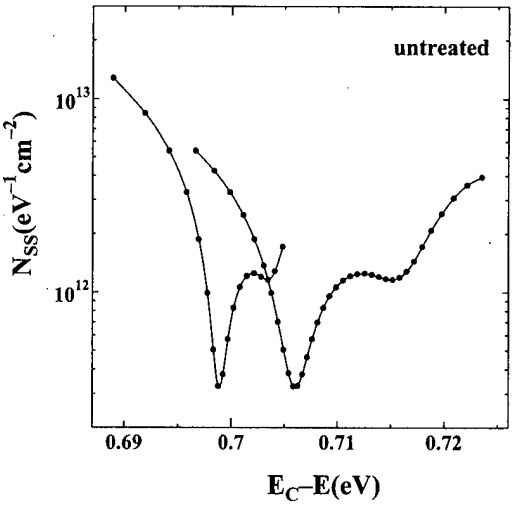


FIG. 10. Interface-state density N_{ss} as a function of E_C-E, where E_c is the conduction band minimum, for the two untreated GaAs samples.

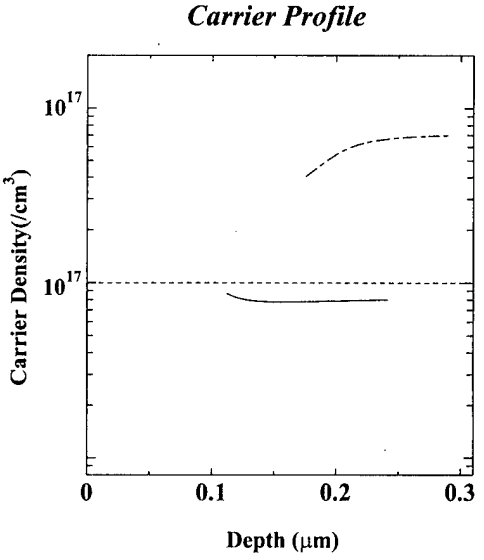


FIG. 12. Carrier-density profiles along the depth of the GaAs substrates treated in HWP N_2–Ar plasma for 30 min and untreated one.

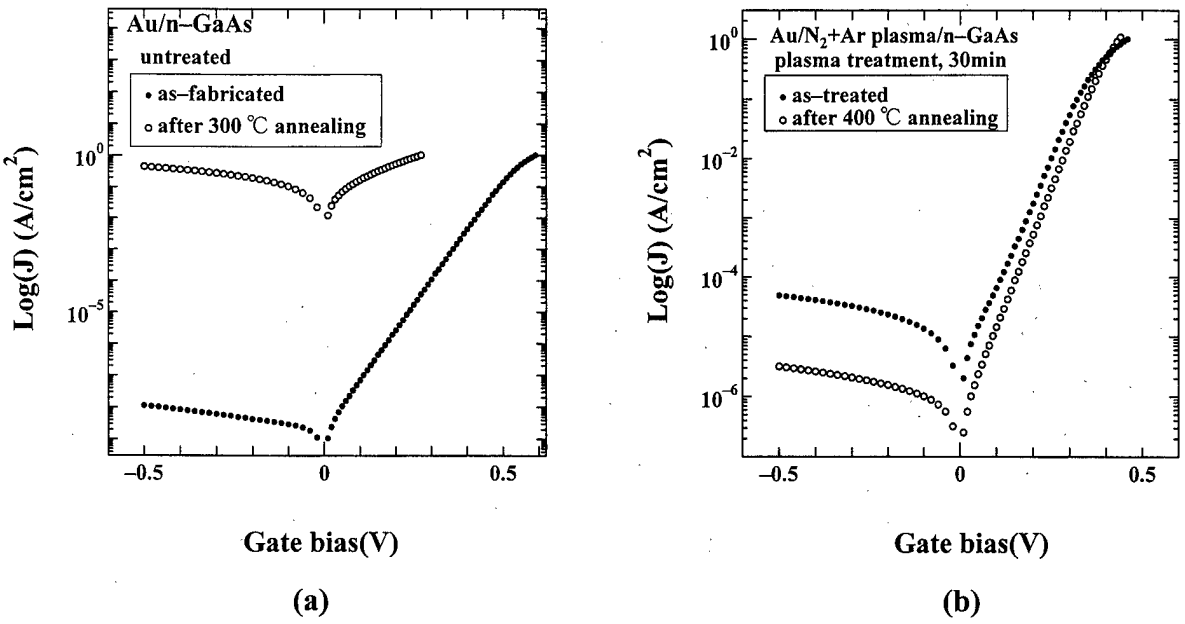


FIG. 13. $J-V$ characteristics for (a) the untreated GaAs samples without and with the post-thermal annealing at 300 °C for 10 min in N_2 ambient and for (b) the GaAs samples treated in HWP N_2–Ar plasma for 30 min without and the post-thermal annealing at 400 °C for 10 min in N_2 ambient.

C. Thermal and air-exposure stability

Thermal stability of the N_2–Ar plasma-treated GaAs surface was investigated by measuring of the $J-V$ characteristics before and after the post-thermal annealing of the Schottky or tunnel MIS diodes at 200–500 °C for 30 min in nitrogen ambient. Figure 13(a) shows the $J-V$ characteristics for the untreated Schottky samples before and after the post-thermal annealing at 300 °C for 30 min in N_2 gas. The $J-V$ characteristics considerably degraded by the post-thermal annealing. The rectification effect is no more observed for the post-thermal annealed sample. Figure 13(b) shows the $J-V$ characteristics for the N_2–Ar plasma-treated samples before and after the post-thermal annealing at 400 °C for 30 min in N_2. The leakage current in the reverse bias region for the annealed sample is about one order of magnitude lower than that for the nonannealed one. Then the electrical characteristics are sufficiently improved by the post-thermal annealing in contrast with the case of the non-plasma treated samples. Figure 14 shows the ideality factor n and the effective Schottky barrier height ϕ_{bn} as a function of the post-thermal annealing temperature for the plasma-treated and untreated GaAs samples. The ideality factor n increases at about 250 °C for the untreated sample while it is constant until the annealing temperature is 400 °C for the plasma-treated done. The effective Schottky barrier height ϕ_{bn} shows an abrupt decrease at 200 °C for the untreated sample. For the plasma-treated sample, it is nearly constant even the annealing temperature is 500 °C. These findings can be explained by the suppression of the chemical reaction or interdiffusion between Au and GaAs by the intermediate GaN layer. The post-thermal annealing is then expected to recover the plasma-induced damages and to improve the electrical characteristics of the metal/GaAs surfaces.

Figure 15 shows the $J-V$ characteristics before and after the air exposure for the N_2–Ar plasma-treated sample and untreated sample. The plasma-treated sample was exposed to laboratory air for one to three months and the untreated one was also left in air for three weeks. The ideality factor n and the effective Schottky barrier height are noted in the figure. For the untreated sample, the $J-V$ characteristics shows a distinct deterioration after exposure to air only for three weeks. The reverse leakage current increased about one order of magnitude compared with that for the as-fabricated sample. The effective barrier height decreased somewhat by

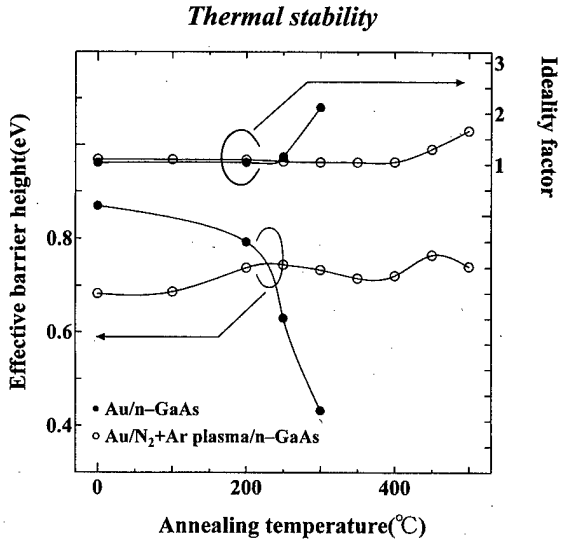


FIG. 14. Effective Schottky barrier height ϕ_{bn} and the ideality factor n as a function of the annealing temperature for both the GaAs samples treated in HWP N_2–Ar plasma for 30 min and untreated one.

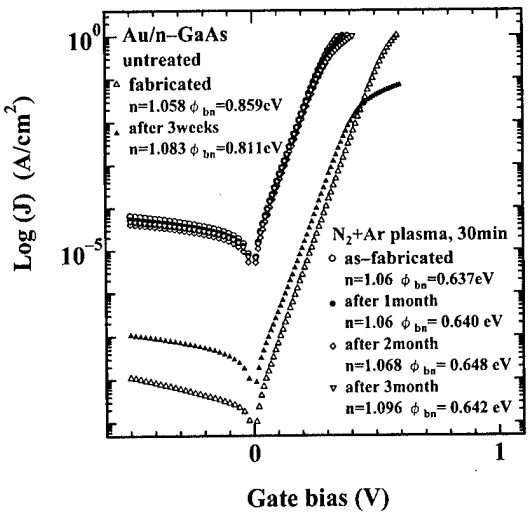


FIG. 15. $J-V$ characteristics before and after air exposure in the laboratory for both the GaAs sample treated in HWP N_2–Ar plasma for 30 min and untreated one. The air-exposure times were three weeks for the untreated sample and one to three months for the plasma-treated one.

air exposure. On the other hand, the $J-V$ curves did not change after air exposure for three months for the plasma-treated sample. The ideality factor and the effective barrier height is almost constant after air exposure even for three months. These indicate a possibility of the robust surface passivation of GaAs with the N_2–Ar plasma treatment though the plasma-induced damages need be removed.

IV. SUMMARY AND CONCLUSIONS

A short-time helicon-wave excited N_2–Ar plasma treatment of the GaAs (100) substrate was performed in order to investigate the effect of the helicon-wave excited N_2–Ar plasma treatment on the Fermi level pinning effect in GaAs. XPS measurements showed that GaN was formed at the insulator/GaAs interface (with a small amount of GaO). The analysis of the $J-V$ characteristics indicated that the intrinsic Fermi level pinning near the midgap of the forbidden band of bulk GaAs was once removed and the Fermi level was newly pinned at the energy of about $E_C-0.2-0.3$ eV, where E_C was the conduction band edge by the plasma treatment. This is due to formation of the high-density donor-like interface states correlated with the plasma-induced damages. The acceptor-like defects were also introduced in the GaAs substrate near the interface by the plasma treatment resulting in the decrease in the carrier density near the GaAs surface. The thermal stability of the Schottky or tunnel MIS diode was greatly improved by the N_2–Ar plasma treatment. The GaN layer formed at the interface probably prevents the chemical reaction between Au and GaAs. The $J-V$ characteristics did not change after air exposure even for three months, indicating a possibility of the robust surface passivation of GaAs. The problem is that the effective Schottky barrier height decreases and the reverse leakage current increases with the plasma treatment. However, these deficiencies are considered to be remedied by the post-thermal annealing. The effects of the post-thermal annealing in nitrogen and hydrogen ambient are now under investigation. Hydrogen is known to terminate the dangling or incomplete bonds and stabilize the surface characteristics, especially in the case of Si. The results will be reported in a forthcoming article.

[1]W. E. Spicer, P. W. Chye, P. R. Skeath, C. V. Su, and I. Lindau, J. Vac. Sci. Technol. **16**, 1422 (1979).
[2]W. E. Spicer, I. Lindau, P. R. Skeath, and D. V. Su, J. Vac. Sci. Technol. **17**, 1019 (1980).
[3]J. Tersoff, Surf. Sci. **168**, 275 (1986).
[4]H. Hasegawa and H. Ohno, J. Vac. Sci. Technol. B **4**, 1130 (1986).
[5]H. Hasegawa, LiHe-H, H. Ohno, T. Sawada, T. Haga, T. Abe, and H. Takahashi, J. Vac. Sci. Technol. B **5**, 1097 (1987).
[6]E. Ikeda, H. Hasegawa, S. Otsuka, and H. Ohno, Jpn. J. Appl. Phys., Part 1 **27**, 180 (1988).
[7]W. E. Spicer, N. Newman, C. J. Spindt, A. Liliental-Weber, and E. R. Weber, J. Vac. Sci. Technol. B **8**, 2084 (1990).
[8]A. Hara, R. Nakamura, and H. Ikoma, J. Vac. Sci. Technol. B **16**, 183 (1998).
[9]A. Hara, F. Kasahara, S. Wada, and H. Ikoma, J. Appl. Phys. **85**, 3234 (1999).
[10]J. A. Talor, J. Vac. Sci. Technol. **19**, 751 (1982).
[11]T. Ishikawa and H. Ikoma, Jpn. J. Appl. Phys., Part 1 **31**, 3981 (1992).
[12]M. K. Bahl, R. O. Woodall, R. L. Watson, and K. J. Irgolic, J. Chem. Phys. **20**, 751 (1982).
[13]S. Wada (unpublished data).
[14]Zs. J. Horváth, J. Appl. Phys. **63**, 976 (1988).

Nitridation of GaAs(001) surface: Auger electron spectroscopy and reflection high-energy electron diffraction

Igor Aksenov[a] and Hidekazu Iwai
Joint Research Center for Atom Technology (JRCAT), Angstrom Technology Partnership (ATP), Higashi 1-1-4, Tsukuba, Ibaraki 305, Japan

Yoshinobu Nakada
Central Research Institute, Mitsubishi Materials Corporation, 1-297 Kitabukuro-cho, Omiya, Saitama 330, Japan

Hajime Okumura
Electrotechnical Laboratory, Umezono 1-1-4, Tsukuba, Ibaraki 305, Japan

(Received 16 October 1998; accepted 16 April 1999)

Auger electron spectroscopy (AES) and reflection high-energy electron diffraction (RHEED) have been used to investigate the processes taking place during the initial stages of nitridation of GaAs(001)–2×4 surface by active nitrogen species. The results of analysis of the spectral shapes and the spectral positions of the Auger electron signals from Ga, As, and N, as well as their dependencies on the nitrogen exposure combined with RHEED results show that the processes taking place during nitridation greatly differ depending on the nitridation temperature. On the one hand, at low temperatures ($T_s \leqslant 200\,^{\circ}C$) nitridation is hindered by kinetic restrictions on the atomic migration, leading to island growth with formation of the disordered GaAsN phase in the subsurface region, which cannot be completely re-crystallized even after prolonged annealing. On the other hand, at high temperatures ($T_s \geqslant 500\,^{\circ}C$) the process of nitridation takes place simultaneously with etching of the surface due to decomposition of the substrate, leading to growth of facets and resulting in a rough interface between the GaN and GaAs phases. However, for intermediate nitridation temperatures ($300\,^{\circ}C \leqslant T_s \leqslant 400\,^{\circ}C$) it was found that the disordered GaAsN phase, formed at the initial stage of nitridation, partly crystallizes even without annealing. Moreover, AES results indicate that a monolayer of N atoms may well be formed on the substrate during the initial stage of nitridation. The postnitridation annealing of the samples nitrided at the intermediate temperatures results in the formation of a crystalline GaN layer. The line shape of the AES signals from this layer is identical to that for a thick GaN reference sample. © *1999 American Vacuum Society.* [S0734-211X(99)01004-5]

I. INTRODUCTION

GaN compound, which is a prospective material for high power electronics and blue light emitting devices,[1–5] is very difficult to grow on technologically well-established substrates, i.e., Si and GaAs (001), which is mostly due to the significant lattice mismatch of GaN with these materials. One of the methods of overcoming the above difficulty is to grow, prior to GaN growth, a thin amorphous "low-temperature" GaN buffer layer on the top of GaAs. The amorphous layer is then crystallized upon subsequent thermal annealing at 600–650 °C, leading to confinement of defects in the interfacial region and to relaxation of the lattice mismatch. The subsequent molecular beam epitaxial (MBE) growth of the GaN compound on the top of the buffer layer results in a very good quality cubic phase GaN.[6]

While the above results verify great promise for the application of MBE-grown GaN, a deeper understanding of the mechanisms governing initial stages of nitridation and subsequent nitride growth is vital for further improvement of device performance. However, direct comparison between reported studies of the GaN growth is difficult even for the same substrate material due to differences in the methods used to obtain active nitrogen. While most of the previous results indicate that nitridation of GaAs (001) leads to a disordered surface exhibiting a rough and relaxed morphology,[6] recently published results of Hauenstein *et al.*[7] and Gwo *et al.*[8] show the formation of the atomically smooth and coherently strained (3×3)-ordered GaN monolayer on GaAs (001) after a limited exposure nitridation process, the obtained smooth morphology presumably being due to the unique characteristics of the nitrogen sources.

The formation of Ga–N and As–N bonds on the GaAs (001) surface during nitridation has been verified by infrared reflectance spectroscopy, where a hot tungsten filament was used to supply active nitrogen species,[9] as well as by x-ray photoelectron spectroscopy (XPS), where nitrogen[10] or ammonia[11] plasmas, and photolyzing of preadsorbed ammonia,[12,13] were used as methods to obtain the active nitrogen. The published XPS results show the chemical shifts of the core levels of gallium, arsenic, and nitrogen caused by the nitridation process, with the formation of what is generally believed to be a disordered Ga–As–N ternary compound right at the beginning of the nitridation process. The last belief, however, is questionable in view of the recent results

[a]Electronic mail: aksenov@jrcat.or.jp

of Gwo *et al.*[8] and also because no detailed study of the growth mechanism (by quantitative surface techniques such as XPS or Auger electron spectroscopy) at the very first stages of nitridation of GaAs (001) surface has been reported so far.

Moreover, all XPS experiments, to the best of our knowledge, have been carried out *after* nitridation by either transferring the nitrided sample from the "nitridation" chamber to the "XPS" chamber or by switching between nitridation and measurement cycles in the same chamber. It is important to note the inherent limitations of this experimental protocol, where nitridation is interrupted for several minutes or more prior to recording the spectra. Indeed, it is not clear, especially at elevated temperatures, that the surface conditions during measurements would be the same as those directly after the nitridation, which may be caused by a variety of processes such as atom migration, desorption, adsorption of impurity atoms, etc.

In the present study we used Auger electron spectroscopy (AES) and reflection high-energy electron diffraction (RHEED) to investigate the processes taking place during the initial stages of interaction of active nitrogen species with the As-stabilized GaAs (001)–2×4 surface at various surface temperatures in the range from room temperature (RT) to 600 °C. The measurements at all temperatures studied have been carried out *simultaneously* with the nitridation process, providing, thereby, an insight into the process of nitridation not hindered by the "postnitridation" relaxation effects.

II. EXPERIMENT

An experimental machine combining a differentially pumped electron beam column for primary electron beam generation and RHEED imaging, a hemispherical electron energy analyzer for AES, as well as a radio frequency (rf) plasma source for generation of active nitrogen species, has been constructed. Differential pumping of the column allows us to make AES and RHEED measurements while supplying active nitrogen to the surface, when the pressure in the chamber goes up to 10^{-5} Torr (base pressure in the system is 2×10^{-10} Torr). After degassing at 300 °C and surface oxide layer desorption at 550 °C, GaAs homoepitaxial layers with a thickness of about 500 Å were deposited on a Si-doped GaAs (001) substrate held at a temperature of 580 °C. The resultant surface exhibited As-terminated (2×4) surface reconstruction, this reconstruction remaining for at least 10 h after cooling the sample down to room temperature. During the nitridation process, as well as during subsequent thermal annealing, the Ga and As cell shutters were closed, and the nitridation rate was controlled by using a mass flow controller, through which a steady flow of 1 ccm/s of the high-purity (99.9999%) N_2 gas was introduced into the rf plasma source (SVT Associates, Inc.), the distance from the tip to the sample being 20 cm. The chamber pressure during nitridation was around 10^{-5} Torr and the input rf power of the plasma source was 300 W. The substrate temperature during nitridation (T_s) was kept constant at various levels, i.e., at T_s=RT, 200, 300, 400, 500, and 600 °C.

To characterize the bonding and composition of the buffer layer growing during the nitridation process, the Auger signals of Ga-$L_3M_{45}M_{45}$ transition at about 1070 eV, As-$L_3M_{45}M_{45}$ transition at about 1230 eV, and N-$KL_{23}L_{23}$ transition at about 380 eV have been used. No traces of impurities, such as oxygen and carbon, were detected on the original surface or at the beginning of the nitridation processes, although some oxygen (in concentrations lower than 2%) tends to appear after prolonged exposures to nitrogen flow (after about 2 h of exposure). The acceleration voltage used for the excitation of the Auger process was 10 keV, the primary beam current was 40 nA with the beam diameter of about 20 nm, giving a current density of an order of 10^4 A/cm^2. Although the power density of that order may be expected to lead to the electron beam damage of the surface combined with electron irradiation stimulated desorption (EDS), we believe that the above effects do not play a significant role in the nitridation process under investigation since lowering of the electron beam current density by rastering down to about 5×10^2 A/cm^2 or taking the spectra in a scanning mode over the area of 1×1 mm^2 (with resulting mean current density of about 5 μA/cm^2) does not result in any noticeable changes in the spectra. We, therefore, used the focused electron beam with high power density in order to achieve a high signal-to-noise ratio during a reasonable time frame.

The energy resolution of the hemispherical analyzer was kept constant at $\Delta E/E=0.2\%$ (const relative resolution mode). Since the spectral shape of the AES signals changes in the process of nitridation (discussed below), we could not use the peak-to-peak intensities to characterize the compositional changes on the surface and in the subsurface region, but used, instead, the integrated intensities of the AES signals.

III. RESULTS AND DISCUSSION

A. General observations

Figure 1 (curve a) shows a wide kinetic energy (KE) range differentiated Auger spectrum of the initial GaAs (001)–2×4 surface for the substrate temperature T_s=300 °C. It can be seen from this figure that no signals originated from either oxygen (~510 eV) or carbon (~275 eV) impurity have been detected, which means that those impurities amount to no more than 1% of a monolayer (ML). Nitridation of the initial surface at any temperature in the range studied results in a decrease in intensity of the As signal, as well as in an appearance of a N signal at around 380 eV, which is shown in Fig. 1 for T_s=300 °C (curve b). This should have been expected in view of an anion exchange nature of the nitridation process. Indeed, considering the large difference in Ga–N and Ga–As bond energies, it is not surprising that the N atoms readily replace the As atoms on the surface with the formation of the Ga–N bonds.[7] Figure 1 (curve c) shows the spectrum of the sample nitrided at 300 °C for 100 min and subsequently annealed at 630 °C for 3 min in As-free environment, whereas Fig. 1 (curve d) ex-

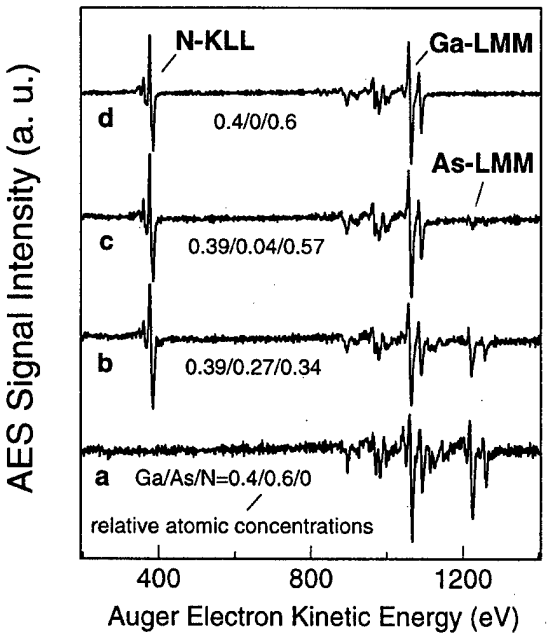


FIG. 1. Wide energy range differentiated Auger spectra from the (a) initial GaAs (001)–2×4 surface, (b) from the same surface after nitridation at 300 °C for 100 min, and (c) subsequent annealing at 630 °C for 3 min, as well as (d) the spectrum of a thick (about 1.5 μm) MBE-grown GaN epitaxial layer used as a reference.

hibits the spectrum of a thick (about 1.5 μm) cubic phase GaN epitaxial layer grown by MBE and used as a reference.

The values of the relative atomic concentrations of the three elements under consideration (Ga, As, and N) have been calculated from the nondifferentiated spectra using the sensitivity factors of 0.6, 0.84, and 0.41 for N, Ga, and As, respectively,[14] and are shown below each spectrum in Fig. 1. It can be seen that, whereas the initial surface is As rich (which should be expected for the As-terminated surface), nitridation leads to the substitution of As atoms by N atoms.

It can be seen from Fig. 1 (curve c) that thermal annealing of the nitrided sample even for a short time (3 min for the spectrum shown in Fig. 1) results in a further decrease in the amount of arsenic in the subsurface region, which is due to the desorption of arsenic from the surface at elevated temperatures. Moreover, prolongation of the duration of thermal annealing up to 20 min (at 630 °C) leads to the complete disappearance of the As-originated signal (not shown), which means that the relative concentration of As drops below the AES detection limit of about 1%. The apparent composition of the subsurface region becomes Ga/N=0.39/0.61, which is almost identical to that of the GaN reference (Fig. 1). It is important to note here that annealing results in a complete disappearance of the As-originated signal only for the samples nitrided at 300 °C$\leq T_s \leq$400 °C, whereas the samples annealed after nitridation at higher or lower temperatures still exhibit some As in the subsurface region (~2%–5%) even after prolonged annealing (up to 30 min).

B. Spectral shape of the Auger signals

The detailed nondifferentiated spectra of nitrogen, gallium, and arsenic Auger transitions for various nitrogen exposures are shown in Figs. 2 and 3 for T_s=RT and 300 °C, respectively, together with the spectra for N and Ga taken after subsequent annealing at 630 °C for 3 min in the As-free environment (dotted curves) and those for the thick (~1.5 μm) MBE-grown GaN reference sample (dashed curves). It can be seen that both Ga and As signals exhibit, in addition

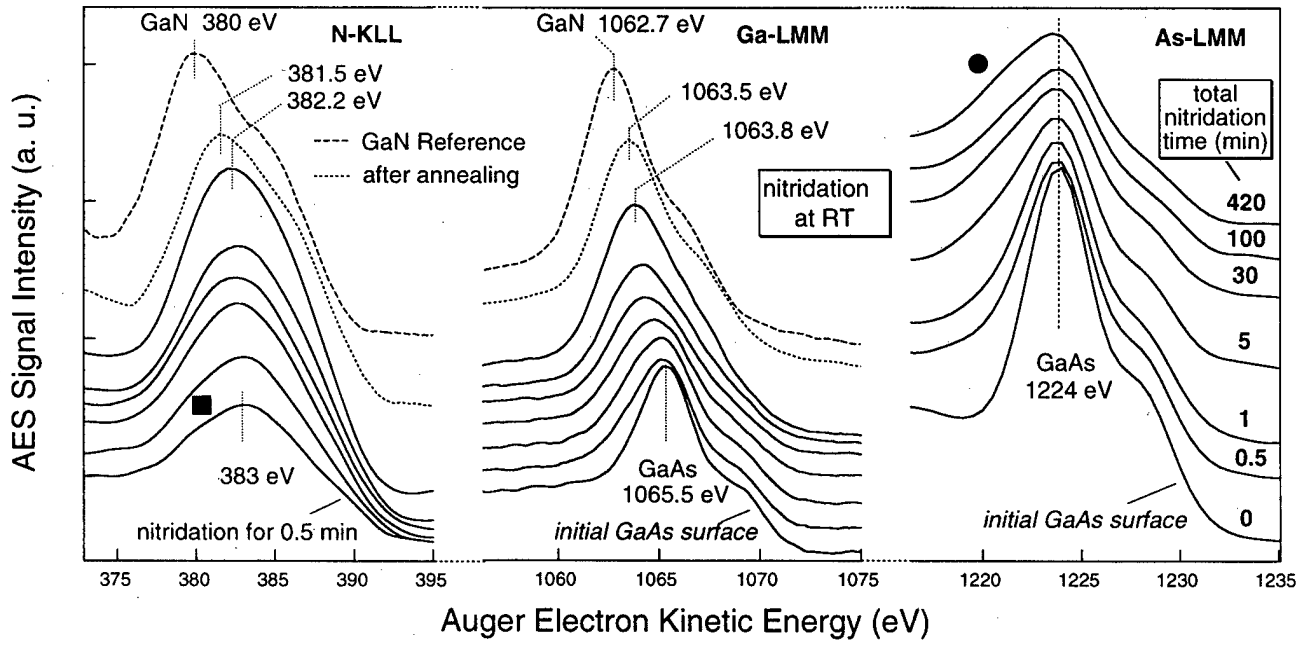


FIG. 2. Detailed nondifferentiated Auger spectra of N-*KLL*, Ga-*LMM* and As-*LMM* transitions for various nitrogen exposures for the nitridation temperature T_s=RT, together with that of N and Ga taken after subsequent annealing at 630 °C for 5 min (dotted curves) and those for the GaN reference sample (dashed curves).

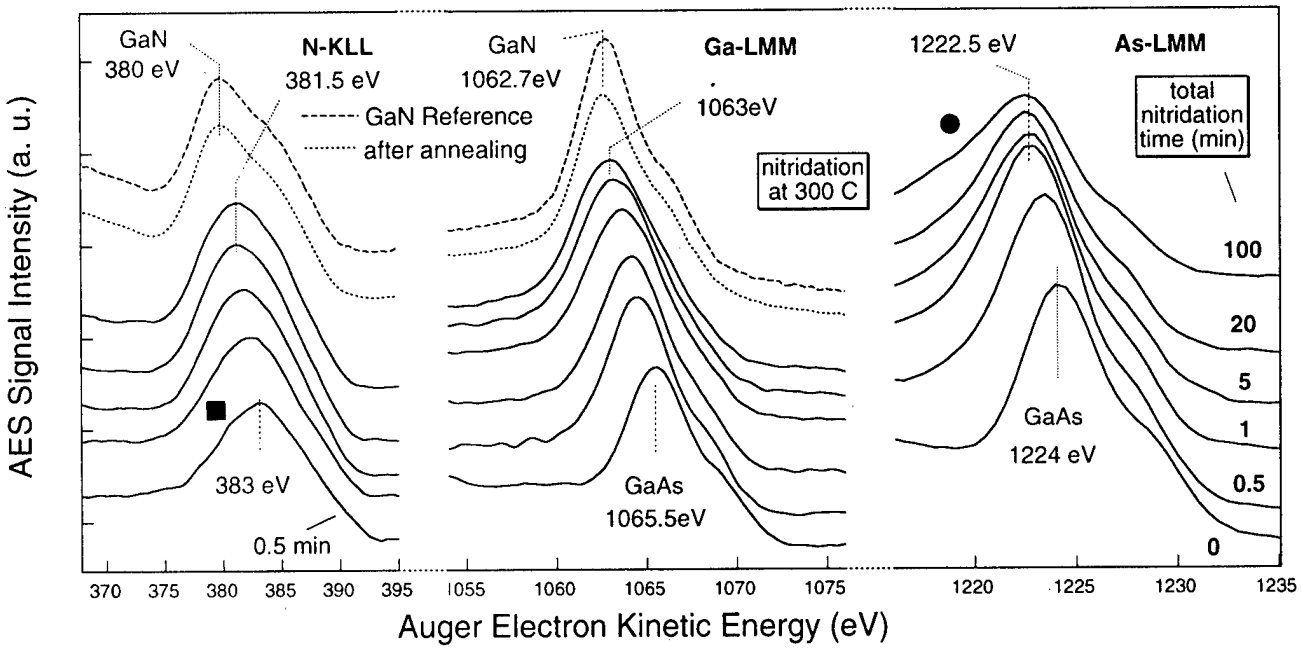


FIG. 3. Detailed nondifferentiated Auger spectra of N-*KLL*, Ga-*LMM* and As-*LMM* transitions for various nitrogen exposures for the nitridation temperature $T_s = 300\,^{\circ}\mathrm{C}$, together with that of N and Ga taken after subsequent annealing at 630 °C for 5 min (dotted curves) and those for the GaN reference sample (dashed curves). Notice that both the energy positions and the spectral shapes of the N- and Ga-originated signals after annealing are identical to those of the GaN reference sample.

to the dominant *LMM* peak, a shoulder on the high-KE side of the spectra, the similar shoulders also having been observed for the Ga and N signals of the GaN reference sample, the spectra of which are shown by the topmost dashed curves in Figs. 2 and 3.

The experimental spectra of the Ga and As signals from the initial GaAs surface, as well as that of the N signal from the GaN reference sample, can be deconvoluted (after the removal of a cosine background) into Voigt (Lorentzian–Gaussian) doublets as shown in Fig. 4. The energy separation between the Voigt peaks constituting each doublet is estimated to be 3.6, 4.3, and 5.0 eV for Ga, As, and N-originated signals, respectively. These values are in good agreement with the energy separations between final state multiplet terms of the Auger electronic transitions for Ga-$L_3M_{45}M_{45}$, [15] As-$L_3M_{45}M_{45}$,[16] and N-$KL_{23}L_{23}$[17] Auger multiplets. The splitting of the above multiplets occurs due to the difference in the total (intra-atomic plus extra-atomic) relaxation energy of two core holes created during the Auger process; and the total relaxation energy, in its turn, can be viewed as an amount by which the KE of the Auger electron is increased due to rearrangement of the intra-atomic and ligand charge distributions in order to shield the two core holes created by the Auger process.

However, the above shoulders on the Ga- and As-originated spectra become less pronounced in the spectral shapes as the nitridation proceeds (see Figs. 2 and 3), and they are not discernible at all in the N-originated spectra during nitridation at any temperature studied. Moreover, both Ga- and As-originated spectra also exhibit about 30% broadening after 100 min of nitrogen exposure as compared to that for the initial surface, and the gradual increase in this additional broadening is especially apparent after the first several minutes of nitridation. The broadening of the spectra is believed to be responsible for the disappearance (in the case of Ga and N signals) or less pronounced character (in the case of As signal) of the shoulders on the high-energy side of the spectra during nitridation. It can also be seen that the broadening of the N-originated spectra greatly exceeds that for the Ga and As signals, which is because the line shape of the N signal is determined largely by a self-fold of the valence band density of states and is therefore strongly affected by the chemical surrounding (which will be discussed later).

In addition to the above discussed shoulder on the high-KE side of the spectra, the As-originated signal during nitridation also exhibits a shoulder at the low-KE side of the spectra (shown by filled dots in Figs. 2 and 3), this shoulder becomes more pronounced with increase in the nitridation time or decrease in the nitridation temperature. The results of a deconvolution of the As signal into three Voigt peaks (the Voigt doublet identical to that for the As signal from the initial GaAs surface shown in Fig. 4 plus an additional Voigt peak at the low-KE side of the spectra) are shown in Fig. 5 for two nitridation temperatures, i.e., $T_s = \mathrm{RT}$ and 500 °C. It can be seen from this figure that the energy separation of the peak at the low-KE side of the spectra from the dominant As-originated peak is about 4 eV. The intensity of that peak diminishes drastically with an increase in the nitridation temperature, the peak becoming indiscernible in the spectra taken at $T_s = 600\,^{\circ}\mathrm{C}$. On the basis of the previous XPS results the peak at the low-KE side of the As spectra has been attributed to the surface AsN species, and the disappearance

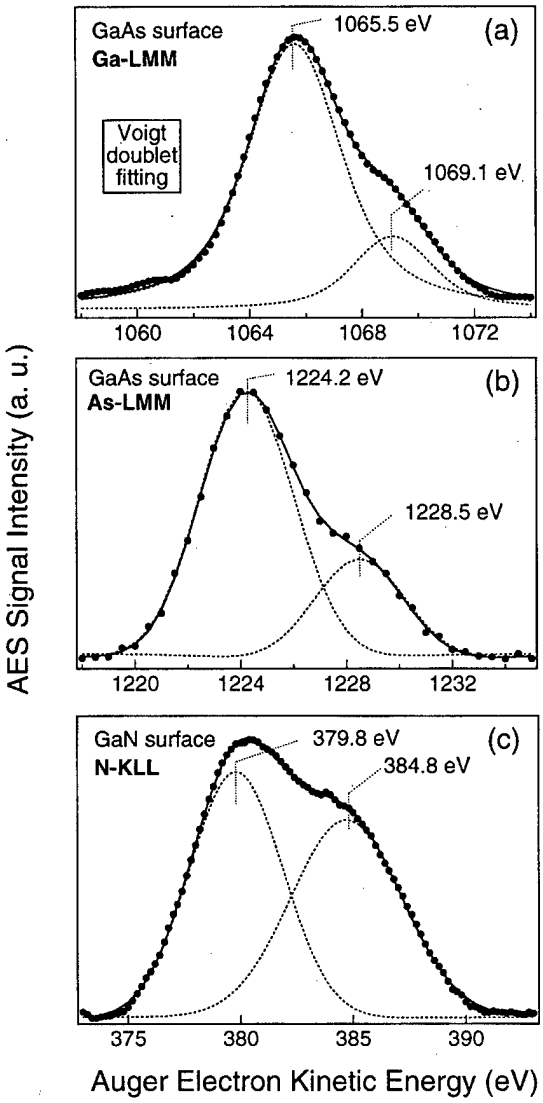


FIG. 4. Detailed nondifferentiated Ga-*LMM* (a) and As-*LMM* (b) Auger spectra taken at RT from the initial GaAs surface together with N-*KLL* (c) spectra from the GaN reference sample. Each spectrum was deconvoluted into a Voigt doublet after removal of a cosine background. Dots represent the experimental data points, the solid curves are overall fits to the experimental data and dashed curves are the deconvoluted peaks of the Voigt doublets.

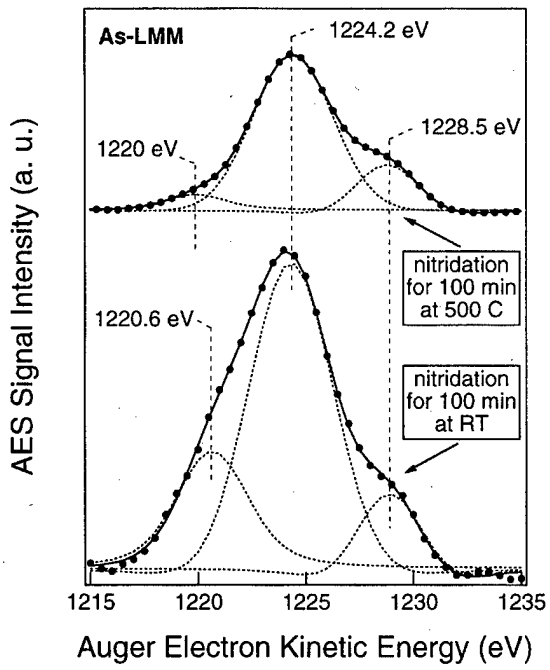


FIG. 5. Detailed nondifferentiated As-*LMM* Auger spectra taken from the GaAs surface after nitridation for 100 min at RT (lower curve) and 500 °C (upper curve) deconvoluted into three Voigt peaks, i.e., the Voigt doublet identical to that for the As-originated signal from the initial GaAs surface shown in Fig. 4 plus an additional Voigt peak at the low-KE side of the spectra, the last peak being attributed to the surface AsN species. The cosine background has been numerically removed.

of this peak at $T_s > 500$ °C is caused by desorption of the AsN species from the surface.[11–13] It should be noted that our results show that the As signal originated from AsN is discernible in the spectra up to the nitridation temperature of 500 °C, which is in contradiction with the previous XPS results showing a complete disappearance of that peak for $T_s \geqslant 300$ °C. This discrepancy, however, can be explained by a difference in the experimental protocols used. Indeed, the XPS measurements have been carried out after the nitridation was interrupted for some finite period of time, the period of time long enough for the desorption of AsN from the nitrided surface for $T_s \geqslant 300$ °C. On the other hand, our measurements have been carried out simultaneously with the nitridation process and, therefore, the magnitude of the AsN-originated As peak in our case shows the amount of AsN on the surface resulting from a dynamic equilibrium between the two processes, i.e., the process of desorption of AsN species from the surface and the process of formation of new AsN molecules from As and N atoms in the subsurface region. Since the AsN species are formed on the surface during the process of nitridation we were able to observe the AsN-originated As signal at higher nitridation temperatures than those at which the signal was detectable in the previous XPS studies.

The N spectra during nitridation also exhibit a shoulder at the low-energy side of the spectra separated by ~2.0 eV from the main N-*KLL* peak (filled squares in Figs. 2 and 3), which disappears at higher temperatures $T_s > 500$ °C just like the low-KE side shoulder in the As spectra. This shoulder has been assigned to the N signal from As–N bonds in the AsN species, which is in accordance with XPS results of Zhu *et al.*[12]

C. Core levels shifts

It can also be seen from Figs. 2 and 3 that the spectral positions of the peaks of Auger spectra under consideration exhibit shifts to lower kinetic energies as the nitridation proceeds. The magnitude of these shifts, however, depends not only on the element, the electronic transitions from which these shifts are observed, but also on the surface (nitridation) temperature T_s. Dependencies of the apparent energy shifts of the peaks of Ga-*LMM* and As-*KLL* transitions on nitrogen exposure for various nitridation temperatures are shown in

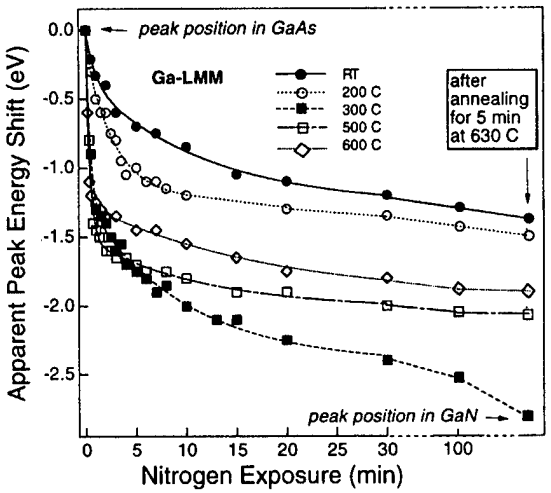


FIG. 6. Dependencies of the apparent energy shift of the peak of Ga-*LMM* transition on nitrogen exposure time for various substrate temperatures together with that after subsequent annealing at 630 °C for 5 min. The data for T_s = 400 °C are similar to that for T_s = 300 °C and are not shown.

Figs. 6 and 7, respectively. It can be seen from Fig. 6 that the magnitude of the low-KE shift for Ga-*LMM* transition increases as the nitridation proceeds. The magnitude of the shift also increases with increase in the nitridation temperature in the range RT$\leqslant T_s \leqslant$300 °C, levels off for 300 °C $\leqslant T_s \leqslant$400 °C, and starts to decrease for $T_s \geqslant$500 °C. Therefore, the dependence of the magnitude of the shift on the nitridation temperature is not monotonic, but exhibits a maximum for the nitridation temperature in the range 300–400 °C.

The situation is similar for the shift observed for As-*LMM* transition (Fig. 7). Indeed, while the shift is negligible for low surface temperatures (RT and 200 °C), it shows a drastic increase up to about −1.5 eV for 300 °C$\leqslant T_s \leqslant$400 °C. A further increase in the temperature, however, results in a small (but discernible) shift to *high* energies, the magnitude of this high-KE shift increasing with an increase in the nitridation temperature. For the N-*KLL* signal, the most pronounced low-KE shift (−1.5 eV) has also been observed for T_s in the range 300–400 °C, whereas nitridation at the lower or higher temperatures results in the low-KE shifts in the range of 0.4–0.8 eV.

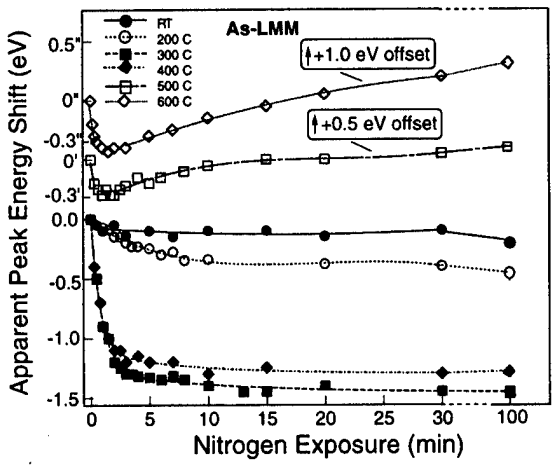


FIG. 7. Dependencies of the apparent energy shift of the peak of As-*LMM* transition on nitrogen exposure time for various substrate temperatures. Notice the positive offset of the curves for T_s=500 and 600 °C.

Annealing at 600–650 °C of the samples, nitrided at T_s in the range 300–400 °C, results in the additional low-KE shifts for all signals, i.e., −0.3 eV for Ga (Fig. 6) and As (not shown) and −1.5 eV for N (not shown). It is important to note that after annealing both the spectral position and the spectral shape of the Ga- and N signals for the samples nitrided at a surface temperature 300 °C$\leqslant T_s \leqslant$400 °C (and *only* for T_s in that range) becomes *identical* to that of the GaN reference sample. Let us compare, for example, the dotted and dashed curves for the N- and Ga-originated signals shown in Figs. 2 and 3. It can be seen by examining Fig. 2 that, while both N- and Ga-originated spectra exhibit the low-KE shift discussed earlier, the last spectra taken after the longest nitridation time of 100 min and subsequent annealing are not yet completely shifted to their spectral positions in the GaN reference sample. This is due to the fact that even after the annealing the subsurface GaAsN layer does not transform completely into GaN. Moreover, as has been mentioned previously, annealing results in a complete disappearance of the As-originated signal only for the samples nitrided at the surface temperature in the range 300 °C$\leqslant T_s \leqslant$400 °C, whereas the samples annealed after nitridation at higher or lower temperatures still exhibit some As in the subsurface region (~2%–5%) even after prolonged annealing.

In general there may be two factors determining the shift of the Auger spectra taken from the surface of a semiconductor. The first one is a chemical shift reflecting change in a geometrical or chemical environment and the second is a shift of the Fermi level due to band bending. The first factor depends on the element under consideration and is directly related to the charge transfer (or difference in ionicities) between the atoms forming chemical bonds. The second factor, i.e., the shift of the spectra caused by the band bending, may be expected to be nearly independent on the element experiencing Auger transitions in the absence of effects caused by strain. However, due to the large lattice mismatch (~20%) between GaAs and GaN we can expect a great deal of strain at the GaAs/GaN interface. Since the core levels taking part in the Auger transitions do not track the top of the valence band in the presence of strain effects,[18] the second factor is also expected to depend on the element emitting Auger electrons. Now let us consider the above two factors.

It has been verified by XPS studies[10–13,19] that nitridation of the GaAs (001) surface results in the formation of the disordered Ga–As–N compound as incoming N atoms substitute for As atoms. The formation of the Ga–As–N phase leads to a broadening of the Auger signals and is partly responsible for the observed shifts in the spectra. Since the ionicity of the Ga–N bonds is higher than that of the Ga–As

bonds due to an increased charge transfer from Ga to anion as arsenic is substituted by nitrogen,[20] we can expect a shift of the core levels of Ga to higher binding energies (BE) and, therefore, a low-KE shift of the Ga-*LMM* Auger transition, which is in accordance with the experimental results. Since the low-KE shift of the spectra is mostly pronounced during the first 2–5 min of nitridation, we believe that the nitridation of the surface takes place mostly during this short time interval, after which it is much slowed down due to kinetic restrictions as N atoms begin to substitute As atoms in the backbonds and deeper in the bulk.

Comparing N–As bonds with the Ga–As ones we can see that the direction of charge transfer is reversed, i.e., while the electron charge was partly transferred in the direction to the As atom in the case of GaAs, it is transferred from the As atom in the N–As bonds. Therefore, the formation of the N–As bonds should result in a low-KE shift of the As-originated signal, which is observed experimentally, but only for the nitridation temperatures $T_s \leq 400$ °C. At higher temperatures the As-originated signal, while slightly shifting towards low-KE during the first 2 min of nitridation, exhibits a small (but discernible) shift towards high-KE as the nitridation proceeds. Therefore, we believe that the shift of the As signal is determined by an interplay of two (or several) effects. Since the binding energy of $3p$ core levels of As is fairly deep (~140 eV), one of the effects may be a positive (causing upward shift of the core levels) Madelung potential caused by the surrounding charges on the sites of As atoms,[21] which increases as the ionicity of the surrounding bonds increases due to the substitution of more and more arsenic atoms by nitrogen.

The charge transfer to N atoms is greater in the Ga–N bonds as compared to the As–N bonds. Therefore, the N signal from the AsN compound, residing on the surface, should be observed at lower KE than that from GaN, which is in accordance with our results (see filled squares in Figs. 2 and 3). However, the small (but discernible) low-KE shift of the N signal during nitridation cannot be interpreted in terms of the chemical shift. Indeed, as we proceed with nitridation, the amount of As in the subsurface region decreases and, therefore, more and more N atoms form chemical bonds with Ga (not As). Since the charge transfer to N atoms in the Ga–N bonds is greater than that in the As–N bonds, the N-originated spectra should exhibit a high-KE shift if that shift is caused by changing chemical surrounding. This is contrary to our experimental results, and, therefore, some other effect, in addition to the change in the chemical surrounding, may be responsible for the observed low-KE shift of the N signal.

As was previously discussed, the most pronounced low-KE shift for all elements under consideration, i.e., −2.5 eV for Ga and −1.5 eV for As and N, has been observed for the nitridation temperature in the range 300–400 °C, whereas nitridation at the lower or higher temperatures results in smaller shifts. We can estimate the magnitude of this additional shift by subtracting the apparent shift for the nitridation temperature T_s=RT (when the smallest shifts were observed) from the values of that for every other nitridation temperature studied. The resulting additional low-KE shift was found to be nearly independent of the element and amounts, after 100 min of nitridation, to about −0.2 eV for T_s=200 °C and about −1.2 eV for T_s in the range 300–400 °C. On the other hand, the additional energy shift observed for T_s=500 and 600 °C depends on the element and amounts to about −0.6 eV, −0.4, and 0.3 for the Ga, N, and As signals, respectively.

On the basis of the above results we believe that, whereas nitridation at low temperatures $T_s < 300$ °C results only in the chemical shift of the Auger signals under consideration, that for the temperatures $T_s \geq 300$ °C leads to both the chemical shift and a shift caused by the band bending effect. It remains unclear why the band bending effect is the most pronounced at the temperatures in the range 300–400 °C. One possible explanation may be that a uniform layer of GaN (possibly creating a heterojunction with GaAs) is formed on the surface at these nitridation temperatures, which, however, is not formed upon nitridation at lower or higher temperatures due to kinetic restrictions on atomic migration (with the resulting island-like structures on the surface) or roughening of the surface due to etching effects[6] for lower and higher temperatures, respectively (to be discussed later).

Thermal annealing results in a further low-KE shift of the Auger signals, the magnitude of the shift being −0.3 eV for the Ga and As signals and −1.5 eV for the N signal. This annealing-induced shift is believed to be caused by both the change in the chemical surrounding (due to the desorption of arsenic from the surface) and the band bending effect, although it is impossible at the present stage to separate the contributions of those two effects into the observed apparent shift. It is not surprising that the band bending causes the largest low-KE shift for the N-*KLL* transitions. Indeed, whereas Ga- and As-*LMM* Auger transitions take place "between" $3p$ and $3d$ core levels, which are deep enough to be truly "core" levels, the N-*KLL* transitions involve $2p$ levels of nitrogen, which actually form the valence band of GaN. Therefore, the line shape of the N signal is determined largely by a self-fold of the valence band density of states and is strongly affected by the band bending effects.

Masuda *et al.*,[11] while proposing the formation of the Ga–As–N compound from their XPS results, stressed that this compound could not be a $GaAs_xN_{1-x}$ mixed crystal due to the very low solubility of nitrogen in the bulk GaAs ($\sim 10^{14}$ cm^{-3}). However, recent experimental[22] and theoretical[23] results suggest that the solubility of nitrogen may be greatly (up to $\sim 10^{21}$ cm^{-3}) enhanced in the vicinity of the GaAs (001) surface due to two reasons, i.e., (1) because of the fact that strain can be more easily relieved near the surface than in the bulk due to an increased movability of the near-surface atoms, and (2) due to the fact that surface reconstructions can effectively relieve subsurface mismatch strain. Therefore, we believe that the GaAsN compound, formed in the process of nitridation, may be the $GaAs_xN_{1-x}$ surface phase.

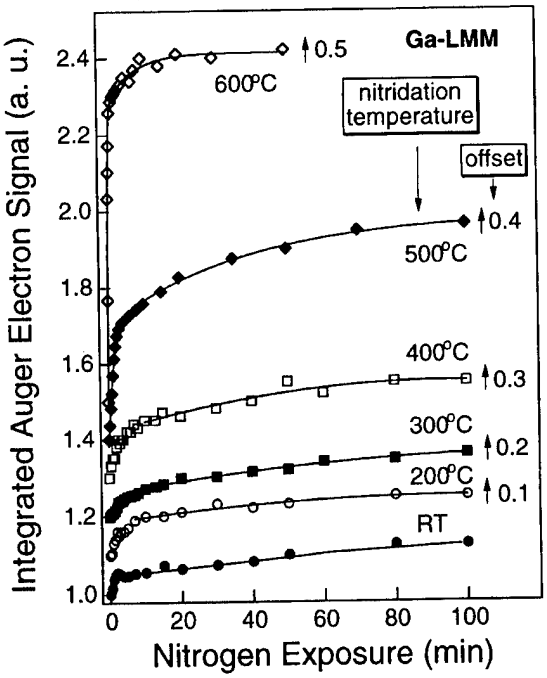


FIG. 8. Dependencies of the integrated Ga-*LMM* Auger signal intensities on the nitrogen exposure for various nitridation temperatures. Notice the positive offset of the curves, increasing with an increase in the nitridation temperature.

D. Dependencies of the Auger intensities on nitrogen exposure

Figures 8–10 show the dependencies of the integrated intensities $I(t)$ for the Ga, As, and N-originated Auger signals on the nitrogen exposure time t during nitridation. Since, as was discussed above, the spectra of all three elements exhibit

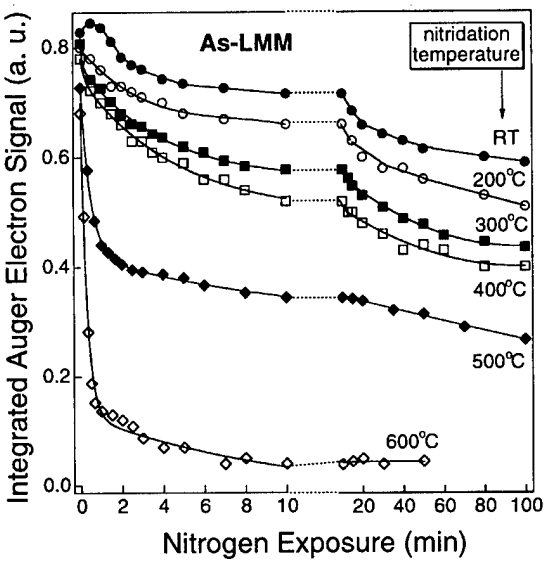


FIG. 9. Dependencies of the integrated As-*LMM* Auger signal intensities on the nitrogen exposure for various nitridation temperatures. Notice an initial (prior to nitridation, data for $x=0$) decrease in the intensity of the As-originated signal with an increase in the nitridation temperature, which is due to the desorption of As from the initial surface before nitridation. Notice also the change in the temporal scale of the x axis after 10 min of nitridation.

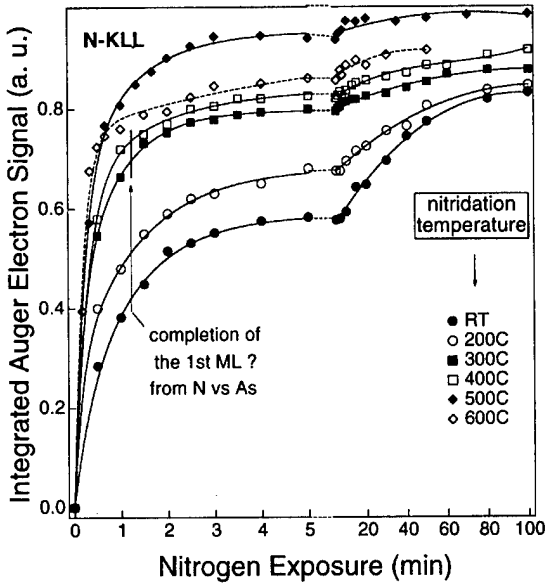


FIG. 10. Dependencies of the integrated N-*KLL* Auger signal intensities on the nitrogen exposure for various nitridation temperatures. Notice the change in the temporal scale of the x axis after 5 min of nitridation.

a change in their spectral shape as the nitridation proceeds, we could not use the peak-to-peak values of the differentiated Auger signals to determine intensities of the signals. Instead, the areas of nondifferentiated signals have been used for quantitative evaluation of the nitridation process, the areas having been determined by a trapezoidal integration with the cosine background correction. The data for the intensities of the As and N signals has been normalized to that of the Ga signal prior to the nitrogen exposure, the intensity of which at any nitridation temperature was taken as 1.0.

It can be seen from Fig. 8 that, as the nitridation proceeds, the intensity $I_{Ga}(t)$ of the Ga signal increases slightly for the nitridation temperatures in the range RT$\leqslant T_s \leqslant$400 °C, the increase being the most pronounced during the first several minutes of nitridation. The rate of increase in the intensity of the Ga signal, however, rises sharply with an increase in the nitridation temperature up to 500 °C and then up to 600 °C, when the intensity of the Ga signal after several minutes of nitridation amounts to two times its value for the initial GaAs surface. Therefore, we can conclude that with an increase in the nitrogen exposure time and in the nitridation temperature the nitrided surface becomes enriched with gallium.

Figures 9 and 10 show that, as we proceed with nitridation, the intensity $I_{As}(t)$ of the As signal decreases, whereas that of the N signal $[I_N(t)]$ increases. The solid curves between the experimental points have been obtained as a result of double-exponential fitting to the experimental results. It can also be seen that the intensities of the As and N signals exhibit the most drastic changes during the first several minutes (or even seconds, depending on the temperature) of nitridation, after which the signals level off at levels strongly dependent upon the nitridation temperature. The magnitude of the changes in the intensities of the N and As signals

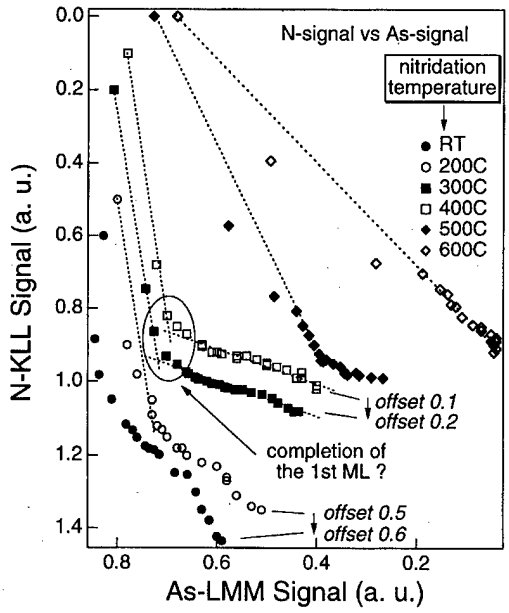


FIG. 11. Integrated intensities of the N-*KLL* (adsorbate) Auger signal as a function of that for the As-*LMM* (substrate) signal for various nitridation temperatures. Ellipse shows the position of the breaks on the curves which probably arise due to the completion of the first monolayer. Notice the offsets of the curves.

greatly exceeds that for the Ga signal, which is consistent with the substitutional (anion exchange) nature of the nitridation process, when N atoms substitute for those of As, whereas Ga atoms mostly retain their positions in the subsurface region.

An increase in the nitridation temperature results (for the same duration of nitrogen exposure) in a drastic decrease in the intensity of the As signal due to the increase in the desorption rate of arsenic and in a drastic increase in the intensity of the N signal. However, it should be noted (Fig. 10) that, although the rate of increase in the intensity of the N signal (related to the sticking coefficient) for the highest nitridation temperature studied (T_s=600 °C) during the initial 40 s of nitridation exceeds that for T_s=500 °C, further nitridation results in the leveling off the signal for T_s=600 °C at a level *lower* than that for T_s=500 °C. The last phenomenon can be explained by taking into account an increased desorption rate of the nitrogen species from the nitrided surface at T_s=600 °C, which is in accordance with the results of Makimoto *et al.*[24]

It can also be seen from Figs. 9 and 10 that the N signal saturates after shorter duration of nitridation as compared to the As-originated signal. This difference in the signals behavior can be explained taking into account the difference in the mean escape depth λ of Auger electrons for N (λ_N =13 Å)[25] and As (λ_{As}=20.5 Å).[26] Since λ_N is much smaller than λ_{As}, the changes in the intensity of the N signal will become less detectable than those for the As signal.

The dependencies of intensities of both As and N signals shown in Figs. 9 and 10 are exponential-like smooth curves exhibiting no apparent breaks. However, plots of intensity I_N of the N signal (adsorbate) against that I_{As} of the As signal (substrate) shown in Fig. 11, exhibit pronounced breaks, but *only* for the data taken for the nitridation temperatures T_s =300 and 400 °C (see ellipse in Fig. 11). The apparent breaks on the I_N vs I_{As} dependencies occur at the points where the normalized signals from N and As equal 0.70 (both N and As) for T_s=300 °C, and 0.74 (N) and 0.69 (As) for T_s= 400 °C. The break in an adsorbate signal *vs* substrate signal dependence usually occurs at the completion of a monolayer (ML) provided that the escape depth of Auger electron for the adsorbate atoms is different from that for the substrate atoms.[27,28] The values of the escape depth for N and As differ and, therefore, the breaks on the curves shown in Fig. 11 are believed to correspond to the completion of the first ML. The position of the breakpoints in Fig. 11 corresponds to the nitridation time of about 1.5 min and is shown as a vertical line on the $I_N(t)$ dependencies in Fig. 10. It can be seen from examination of this figure that no apparent breaks are discernible in the $I_N(t)$ dependencies after 1.5 min of nitridation, which seems to contradict the results shown in Fig. 11. The above discrepancy, however, can be explained by taking into account possible (and very likely to occur in the submonolayer stages)[27] changes of sticking probability S_N of the active nitrogen species. Indeed, whereas gradual changes in the value of S_N result in smoothing of the $I_{As}(t)$ and $I_N(t)$ dependencies, it will not affect the plot of I_N vs I_{As} since the time variable is eliminated in this plot.

Next, we can estimate the nitrogen coverage Θ_N at the breakpoint in the I_N vs I_{As} plot from the relative attenuation of the As (substrate) signal intensity after 1.5 min of nitridation as[29] $I_{As}(N)/I_{As}(\text{clean}) = 1-\Theta_N[1-\exp(-h_{As-N}/\lambda_{As}\cos\phi)]$, where $I_{As}(N)$ is the intensity of the As signal after nitridation, I_{As}(clean) is that for the initial GaAs surface, h_{As-N} is the distance between the layer of N atoms being adsorbed on the surface and that of As atoms in the second layer (bonded to backbonds of the Ga atoms in the topmost Ga layer), and ϕ is the angle between the sample surface normal and the direction of electron detection. Taking the value of h_{As-N} as $h_{As-N}=(h_{As-As}+h_{N-N})/2=2.54$ Å (where $h_{As-As}=2.83$ Å and $h_{N-N}=2.25$ Å are the interlayer distances of the As–As and N–N layers in GaAs and GaN, respectively), and ϕ=45 deg, we obtain the value of nitrogen coverage Θ_N=0.81 and 0.73 ML for the nitridation temperature T_s=300 and 400 °C, respectively. The estimated nitrogen coverage at the breakpoint in the I_N vs I_{As} plots is, therefore, less than 1 ML.

The last discrepancy, however, may be expected considering the diffusion of the As atoms from the substrate into the growing nitride layer. Indeed, as the nitridation proceeds, some of the N atoms do not accumulate on the surface, but form bonds with Ga atoms as they substitute for As atoms in the second (or deeper) layers. Although we believe that during the first 1.5 min of nitridation nitrogen mostly substitutes for the arsenic on the surface, some of it goes deeper into the subsurface region. In other words, the formation of the disordered GaAsN compound begins from the first moments of nitrogen exposure. Moreover, as the substituted As atoms move up to the surface, they do not desorb immediately at

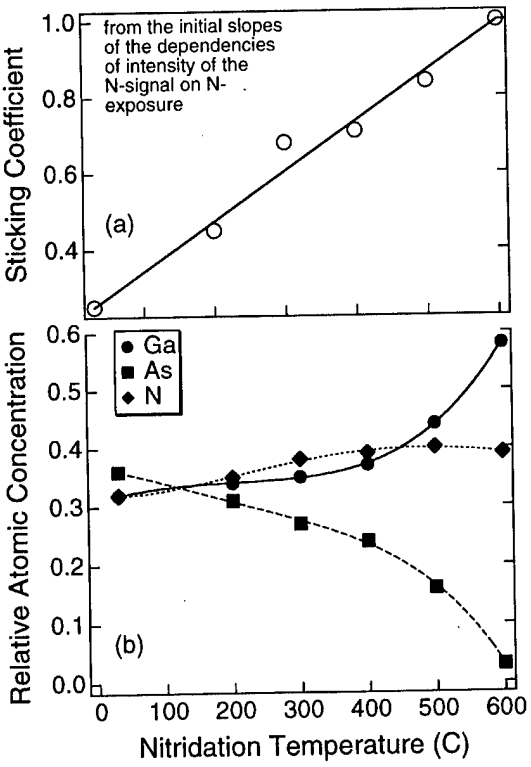


FIG. 12. Dependencies of the (a) initial sticking coefficient S_{no} (determined under an assumption that $S_{no}(600\,°C) = 1.0$), as well as (b) relative atomic concentrations of gallium, arsenic, and nitrogen in the subsurface region after 100 min of nitridation, on the nitridation (surface) temperature.

the nitridation temperature used in this study, but, instead, may form bonds with other As atoms. In addition, the formation of the As–N species is also expected. Therefore, we may expect that, on the one hand, from the first moments of nitrogen exposure some nitrogen atoms penetrate into the second or deeper layers of the GaAs substrate and, on the other hand, some arsenic atoms originated from the subsurface region diffuse to the surface and add up to those whose initial position was on the surface. Due to the above admixture of the substrate atoms into the growing nitride layer the breaks in the $I_{As}(t)$ and $I_N(t)$ dependencies will occur earlier than in the case of pure nitride growth,[30] i.e., the contamination of the nitride layer with arsenic will reduce the apparent value of the intensity for the N-originated signal and increase that for the As-originated one. The apparent break in the I_N vs I_{As} dependence, therefore, will correspond to the nitrogen coverage less than 1 ML, which is in accordance with our results.

It can also be seen from Fig. 10 that an increase in the nitridation temperature results in a noticeable rise of the initial slopes of the uptake $I_N(t)$ curves. The slope of those curves is the ratio of the adsorption rate to the impingement rate, i.e., the initial (or zero-coverage) sticking coefficient S_{no}. Assuming that the value of S_{no} for $T_s = 600\,°C$ equals 1.0 we can build the dependence of the initial sticking coefficient on the nitridation temperature shown in Fig. 12(a). It can be seen that the sticking coefficient increases with an increase in the surface temperature, and, therefore, adsorption of nitrogen is a temperature-activated process, where the impinging nitrogen species must overcome the potential barrier to chemisorption in order to be adsorbed on the surface.

The values of the relative atomic concentrations of Ga, As, and N in the subsurface region after 100 min of nitridation have been calculated for various nitridation temperatures using the sensitivity factors of 0.6, 0.84, and 0.41 for N, Ga, and As, respectively,[14] and are shown in Fig. 12(b). It can be seen that as the nitridation temperature increases, the amount of As in the subsurface region decreases (due to an increase in the desorption rate of As), the amount of Ga increases (with the surface becoming Ga rich), and that of N saturates at the level of about 0.4 for higher temperatures $T_s \geqslant 400\,°C$, the nitrogen concentration reaching its value for the surface of the GaN reference sample.

It is important to repeat here that (1) the plots of intensity I_N of the N-originated signal against the I_{As} of the As-originated signal, shown in Fig. 11, exhibit pronounced breaks *only* for the data taken for the nitridation temperatures $T_s = 300$ and 400 °C (see ellipse in Fig. 11), whereas no meaningful breaks are discernible in the I_N vs I_{As} dependencies for higher or lower nitridation temperature. At the same time, as was discussed above, (2) the most pronounced low-KE shifts of the Auger spectra for all elements under consideration, i.e., −2.5 eV for Ga and −1.5 eV for As and N, have been observed for the nitridation temperature in the range 300–400 °C, whereas nitridation at the lower or higher temperatures results in smaller shifts (Figs. 6 and 7). In addition, it is also important to consider that (3) after annealing of the samples at 600–650 °C both the spectral position and the spectral shape of the Ga and N signals, for the samples nitrided at the surface temperature $300\,°C \leqslant T_s \leqslant 400\,°C$ (and *only* for T_s in that range), becomes identical to that of the GaN reference sample (Fig. 3). Moreover, (4) the thermal annealing results in a complete disappearance of the As signal from the surface only for the samples nitrided at $300\,°C \leqslant T_s \leqslant 400\,°C$, whereas the samples annealed after nitridation at higher or lower temperatures still exhibit some As in the subsurface region even after prolonged annealing.

The above four important points (1)–(4) could not arise from a pure coincidence. Instead, they are closely related to the temperature-dependent mechanism of nitridation of the surface, the last conclusion being supported by the RHEED results discussed in the following section.

E. RHEED results

RHEED pattern for the As-terminated (2×4) surface reconstruction of the initial GaAs (001) surface prior to nitridation, observed along the incidence azimuth of [110], is shown in Fig. 13. Nitridation results in drastic changes of the initial RHEED pattern, these changes depending on the nitridation temperature T_s, which is shown in Figs. 14(a)–14(c) for T_s = RT, 300, and 600 °C, respectively. It should be noted that the changes in the RHEED patterns observed for the nitridation temperatures $T_s = 200$, 400, and 500 °C (not shown) are essentially the same as for the temperatures T_s = RT, 300, and 600 °C, respectively (shown in Fig. 14).

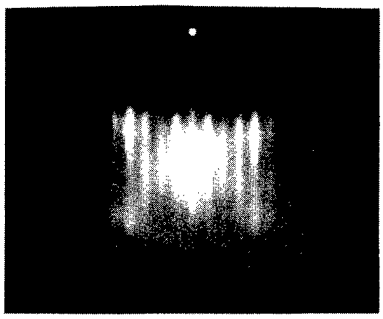


FIG. 13. RHEED pattern of the initial GaAs (001)–2×4 surface observed along [110] azimuth.

It can be seen from Fig. 14(a)(1)–(3) that at low temperature $T_s \leq 200$ °C, once the nitrogen plasma is struck the GaAs reconstruction features fade out, the intensity of the diffracted beam being transferred to the background. The original (2×4) reconstruction of the initial GaAs surface is completely destroyed after 1.5 min of nitridation with an appearance of (1×1) pattern from the unreconstructed surface, which, in its turn, disappears after 3 min of nitridation, giving rise to no pattern except for a smooth background [Fig. 14(a)(1)], indicating that a highly disordered amorphous and/or polycrystalline GaAsN layer is formed in the subsurface region.

Thermal annealing of the samples, nitrided at $T_s \leq 200$ °C, at 630 °C results in an appearance of an arc pattern, the shape of the arcs becoming more spot-like with an increase in the annealing duration [see Fig. 14(a)(4),(5)]. The formation of the arc pattern indicates an appearance of preferred orientations in the disordered layer formed during nitridation, and the partial transformation of the arc-like features into the spot-like ones indicates the partial crystallization of that layer. The crystallization, however, is not complete for this nitridation temperature since the diffraction features, while exhibiting some spot-like character, mostly retain their arc-like appearance even after prolonged annealing [Fig. 14(a)(5)].

The situation is different for the nitridation temperature in the range 300–400 °C. Just as in the above case of lower T_s, nitridation during the first several minutes results in the destruction of the (2×4) reconstruction with the formation of the highly disordered layer in the subsurface region [Fig. 14(b)(1),(2)]. However, after approximately 10 min of nitridation the disordered layer begins to recrystallize, which is evidenced by an appearance of diffuse (but discernible) spots in the RHEED pattern corresponding to the lattice spacing of GaN (4.5 Å), which is shown in Fig. 14(b)(3).

The thermal annealing, even at the fairly low temperature of 500 °C [Fig. 14(b)(4)], then results in an appearance of the RHEED pattern characteristic of a mixture of cubic and hexagonal GaN phases, the cubic phase being dominant.[6,31] An increase in the annealing temperature up to 630 °C leads to the sharpening of the diffraction spots, the intensity of the pattern being gradually transferred from the background to the GaN-originated spots, which is shown in Fig. 14(b)(5). The above results indicate the crystallization (simultaneously with a loss of arsenic from the surface at these annealing temperatures) of the GaN phase from the GaAsN phase on the surface. The spotty character of the final RHEED pattern is due to some roughening of the surface, whereas diagonal crossing lines between the diffraction spots, which can be seen in the patterns during the annealing, indicate the formation of (111) facets due to the desorption of arsenic, bringing about the growth of stacking faults.[6,32,33]

An increase in the nitridation temperature up to $T_s = 500$ and 600 °C results in a different progression as shown in Fig. 14(c). Upon striking the nitrogen plasma the (2×4) reconstruction (after about 20 s of nitridation) changes to a sharp streaked (3×3) pattern, which is shown in Fig. 14(c)(1) and is consistent with previous reports on nitridation at this temperature.[7,19,34] However, the surface retains the (3×3) reconstruction for only 10 s, after which the (3×3) features fade giving rise to the (1×1) pattern from the unreconstructed surface shown in Fig. 14(c)(2). The (1×1) pattern, in its turn, also fades as the nitridation proceeds, but after a further 2 min of nitridation the GaN-originated spots, superposed on the blurred (1×1) streaked pattern from GaAs, become discernible in the pattern [Fig. 14(c)(2)]. As we proceed with nitridation, the intensity of the GaN-originated spots increases, whereas that of the GaAs-originated streaks decreases, the streaks becoming indiscernible after about 5 min from the beginning of nitridation [Fig. 14 (c)(4)].

At the same time, however, a blurred spot appears above the (004)-diffraction spot of the GaN phase, the intensity of this spot increasing as we proceed with nitridation, which is shown in Fig. 14(c)(5). The spot is marked as "GaAs" in the figure since its position in the pattern corresponds to that for the (004)-diffraction beam from the GaAs phase. It can also be seen that the (115)-diffraction spots on the left and right of the GaAs spot are elongated in the vertical direction on the RHEED patterns shown in Figs. 14(c)(4),(5), indicating that those (115) spots consist of unresolved overlapping spots corresponding to the GaN and GaAs phases. Therefore, whereas the positions of the GaAs-originated diffraction spots in the vertical direction of the RHEED patterns correspond, in real space, to the lattice spacing of GaAs, in the horizontal direction they correspond to that of GaN. Since a vertical direction of the RHEED pattern corresponds to the real space direction perpendicular to the surface, and a horizontal direction reflects the lateral spacing, the positions of the GaAs-originated spots indicate that the GaAs phase is strained (with the lattice constant of GaN) in the lateral direction and relaxed (with the lattice constant of GaAs) in the direction perpendicular to the surface.

Annealing the samples, nitrided at $T_s = 500$ and 600 °C, at 650–750 °C results in a disappearance of the GaAs spot in the pattern, as well as in a more circular shape of the (115) spots, the intensity of these spots being transferred to their lower parts corresponding to diffraction from GaN (not shown). Therefore, we believe that at high nitridation temperatures the GaAs islands are formed on the top of the nitrided GaN layer due to out-diffusion of arsenic, substituted

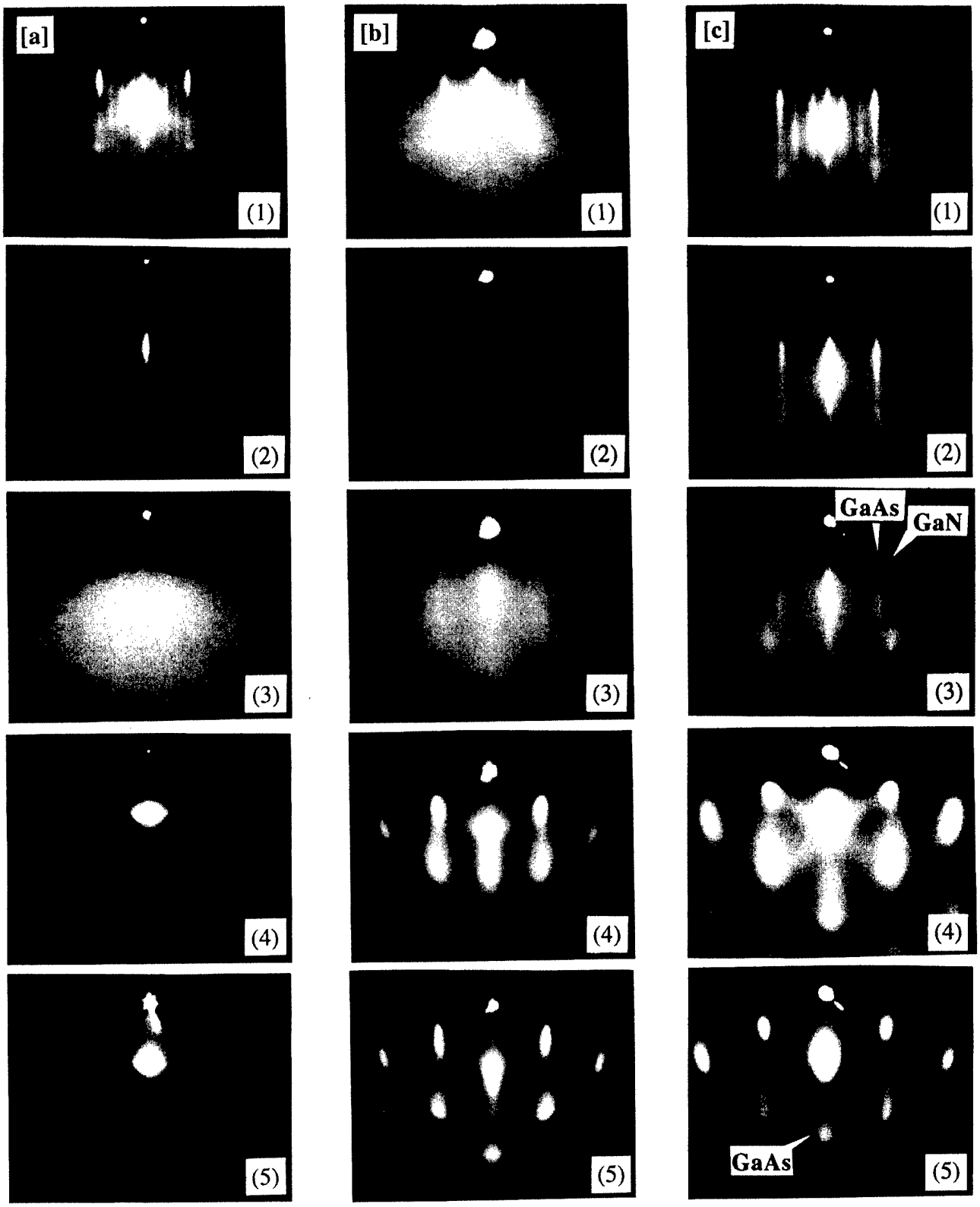


FIG. 14. RHEED patterns observed along [110] azimuth during nitridation of the GaAs (001)–2×4 surface at (a) RT, (b) 300 °C, and (c) 600 °C. (a) (From the top to the bottom) the patterns taken after (1) 20 s, (2) 3 min, and (3) 5 min of nitridation at RT, together with those taken after 100 min of nitridation and subsequent annealing at 630 °C for (4) 3 min and (5) 20 min. (b) Patterns taken after (1) 20, (2) 5 min, and (3) 10 min of nitridation at 300 °C, together with those taken after 100 min of nitridation and subsequent annealing at (4) 500 °C for 1 min and (5) 630 °C for 3 min. (c) Patterns taken after (1) 20 s, (2) 40 s, (3) 3 min, (4) 5 min, and (5) 30 min of nitridation at 600 °C. The results for the nitridation temperatures T_s = 200, 400, and 500 °C (not shown) are similar to those for T_s = RT, 300 and 600 °C, respectively.

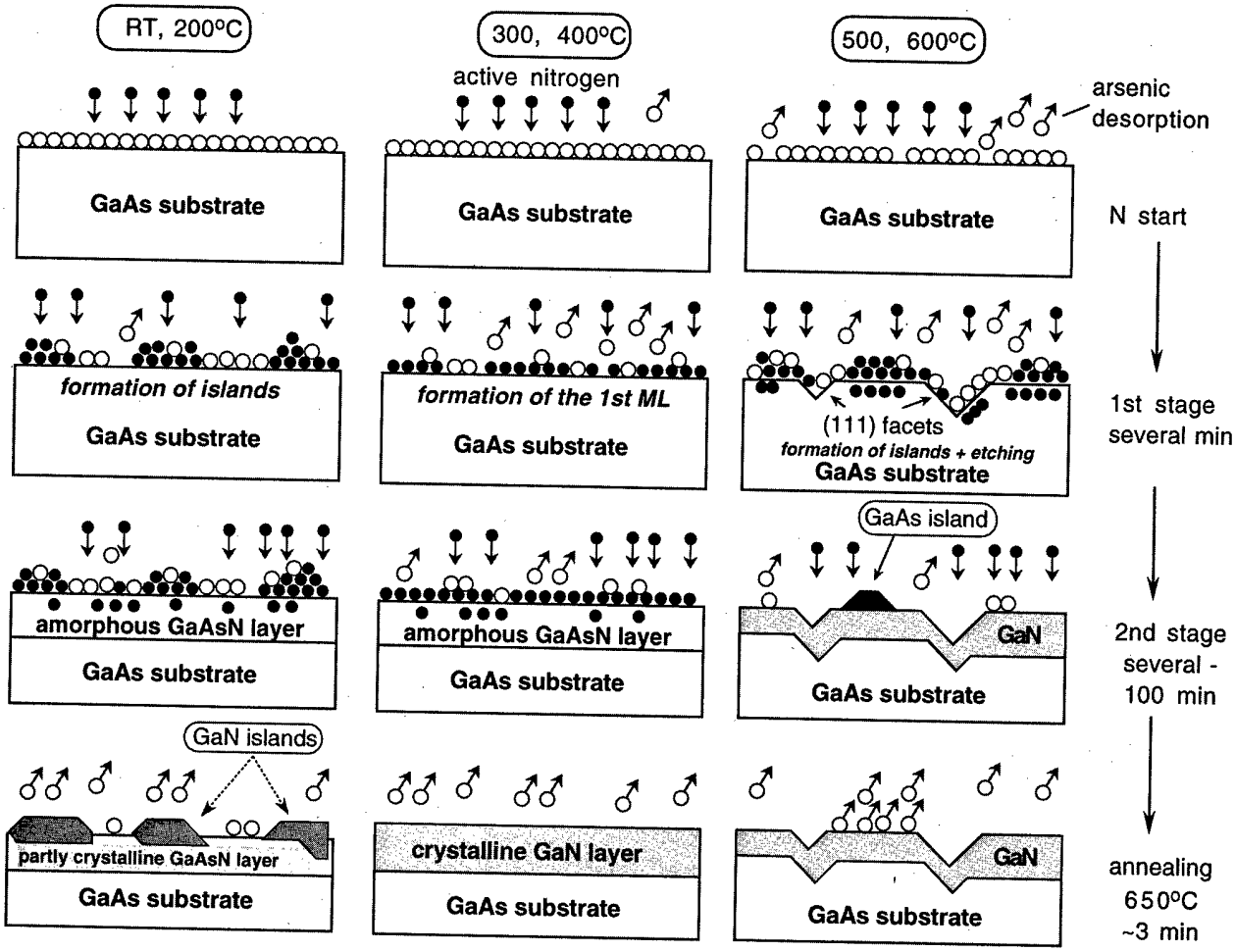


FIG. 15. Tentative model proposed for the nitridation process of GaAs (001) surface at various temperatures derived from our AES and RHEED data.

by nitrogen, from the GaAs substrate. The islands are strained in the lateral direction due to the influence of an underlying GaN layer and relaxed in the direction perpendicular to the surface, and they desorb from the surface at the temperatures higher than 650 °C due to the increased rate of desorption of arsenic. It can also be seen from Figs. 14(c)(4),(5) that the diagonal lines between the diffraction spots are more pronounced in this case as compared to that for the lower nitridation temperature [Figs. 14(b)(4),(5)], and, therefore, the formation of the (111) facets takes place more intensely for higher nitridation temperatures.

It should be noted that the above RHEED results, i.e., the growth of the mixture of cubic and hexagonal GaN phases, do not agree with the latest results of Okumura *et al.*[35] showing that the presence of As on the surface prevents the formation of wurtzite stacking on (111) facets and promotes the growth of the cubic phase. This discrepancy may be explained by the use of different methods to obtain active nitrogen since a flux of the nitrogen species from the rf plasma source used by us is expected to be more intense compared to that generated by an electron cyclotron resonance plasma source used by Okumura *et al.* In addition, the amount of As on the surface in our case is limited by the diffusion of As atoms, released in the subsurface region, to the surface, whereas the results of Okumura *et al.* were obtained with the incoming As ambient flux.[35]

F. Model of the nitridation process

The tentative model of the nitridation process of the GaAs (001) surface derived from our AES and RHEED data is schematically shown in Fig. 15, where the nitridation process is presented as occurring through two main stages with the subsequent annealing stage, the nitridation process being different for three distinct temperature ranges, i.e., RT–200 °C, 300–400 °C, and 500–600 °C.

The process of nitridation at the temperature in the range RT$\leqslant T_s \leqslant$200 °C is believed to take place as follows:

(1) At the first stage the impinging active nitrogen species destroy As dimers [forming the (2×4) surface reconstruction] on the surface and substitute for the As atoms in the surface layer. Due to the low surface diffusivity of atoms at this low temperature,[36] the lateral migration of atoms is kinetically hindered; and we believe that nitridation occurs through the formation of islands or "simultaneous multilayers" which do not cover the underlying surface entirely.

(2) At the second stage, as the nitridation proceeds, more and more impinging N atoms penetrate deeper into the sub-

surface region, substituting the As atoms deeper in the bulk and disrupting the crystal lattice in the subsurface region with the formation of a disordered GaAsN phase. Since the desorption rate of the As atoms from the surface is low at this temperature, most of the substituted As atoms remain on the surface forming the AsN species. The accumulation of the AsN species on the surface (Fig. 5, lower curve) leads to much sharper slopes for the uptake $I_N(t)$ curves comparing to that for the downward $I_{As}(t)$ curves (Figs. 8 and 9). The nitridation rate, however, saturates after several minutes of nitridation, presumably due to the requirement of transporting the incoming nitrogen through the surface layer.

(3) Thermal annealing of the samples, nitrided at $T_s \leqslant 200$ °C, at 630 °C leads to the desorption of AsN species from the surface and to partial recrystallization of the disordered GaAsN phase. However, the crystallization is not complete, which is evidenced by the RHEED results shown in Fig. 14(a). The GaN islands presumably are formed on the partly crystalline GaAsN layer, whereas the portions of the surface which were not nitrided initially remain exposed. Moreover, the GaN islands may coexist with those of GaAs on the surface due to phase separation in the GaAsN phase. The last considerations are supported by the detection (by AES) of arsenic on the surface even after the prolonged annealing.

The progression changes for the nitridation temperature in the range 300 °C$\leqslant T_s \leqslant$400 °C, i.e.:

(1) At the first stage, just like in the above case for a lower nitridation temperature, nitrogen substitutes for arsenic on the surface with the destruction of the (2×4) reconstruction, the substituted arsenic atoms desorbing from the surface. Whereas most of the N atoms accumulate on the surface, some of them penetrate deeper into the subsurface region, substituting As atoms deeper in the bulk. The first stage continues until a complete ML of N atoms (with some admixture of arsenic still remaining on the surface) is formed, which is evidenced by the breaks on the I_N vs I_{As} plots shown in Fig. 11. The formation of the complete ML of GaN results in the band bending, which leads to the most pronounced low kinetic energy shifts of the Auger signals for the nitridation temperatures in this range (Figs. 6 and 7). The shifts of the Auger signals to low kinetic energies may well be caused by the formation of a GaN/GaAs heterojunction since the band gap offset in this heterojunction has been found to be concentrated predominantly in the valence band, the valence band of GaN being about 1.8 eV lower than that of GaAs.[37]

(2) At the second stage the sticking coefficient of N atoms significantly decreases because of a complete coverage of the surface by N atoms forming bonds with Ga atoms, and incoming nitrogen penetrate into the subsurface region. An increased diffusivity of atoms at this temperature results in a deeper penetration depth of N atoms, as well as in a lower arsenic content in the disordered GaAsN phase, which, in its turn, leads to a partial crystallization of the disordered layer as we proceed with nitridation [Fig. 14(b)(3)].

(3) Annealing of the partially crystallized GaAsN phase results in desorption of arsenic from the surface and a complete crystallization of the phase with the formation of cubic/hexagonal GaN [Fig. 14(b)(4),(5)]. Although some faceting takes place during annealing, it is much less pronounced in comparison with that for the higher nitridation temperature discussed below. The complete transformation of the GaAsN phase into the GaN phase is also supported by the fact that the line shapes of the Ga- and N-originated AES signals for the samples annealed after nitridation in this temperature range, are identical to that for the GaN reference sample (Fig. 3).

Since (1) no arsenic has been detected from the samples, nitrided at these temperatures, after annealing, and (2) the sensitivity of AES is about $0.01 = \exp(-4.5)$,[38] the thickness h of the GaN layer can be estimated from the attenuation factor $[\exp(-h/\lambda_{As})]$ of the As-originated AES signal by using the mean escape depth $\lambda_{As} = 20.5$ Å of *LMM*-Auger electrons for arsenic as being $h \sim 4.5 \times \lambda_{As} \approx 100$ Å.

An increase in the nitridation temperature up to 500 °C and then to 600 °C results in the following processes:

(1) Soon after the beginning of nitridation the (2×4) surface reconstruction transforms into the (3×3) reconstruction structure, which may be due to the formation of nitrogen dimers on the surface.[8] This reconstruction, however, does not live long and is quickly replaced by the bulk GaAs (1×1) structure, on the top of which the growth of the GaN phase begins [Fig. 14(c)(3)]. The growth, however, does not proceed smoothly, which is believed to be caused by two factors, i.e., (i) an exposure to the nitrogen plasma causes degradation of the surface with the loss of As, probably due to an impact of energetic species from the plasma source; and (ii) the As atoms released in the subsurface region diffuse to the surface. Since it is unlikely that As, a relatively large atom, can migrate to the surface at the same rate as N atoms migrate into the bulk, it is highly possible that the release of arsenic at the GaN/GaAs interface results in the disruption of the surface region.

In other words, the processes of nitridation and etching of the surface takes place simultaneously at different portions of the surface. It seems that the N atoms agglomerate to the nitrided portions of the surface to form GaN islands, whereas the surface portions which are not nitrided initially are etched by exposing the (111) facets.

(2) As the nitridation proceeds, the GaN islands extend laterally on the surface and into the bulk of the substrate, with the resulting GaN layer covering all the initially etched and rough GaN/GaAs interface. Therefore, the (111) facets of GaAs, which were exposed during the first stage of nitridation, transform into the (111) facets of GaN, the surface morphology following that of the interface region. The desorption rate for both Ga and N at these temperatures drastically increases, which results (1) in a decrease or a complete elimination of the AsN species on the surface (Fig. 5, upper curve) and (2) in a drastic increase in the nitridation rate together with the rate of arsenic loss (Figs. 9, 10, and 12), as well as (3) in an increase of the relative amount of Ga on the surface (Figs. 8 and 12). Moreover, some of the As

atoms diffused to the surface do not desorb immediately, but form the GaAs islands on top of the GaN layer, which is evidenced by the RHEED results shown in Fig. 14(c)(5). These GaAs islands, however, desorb from the surface at even higher temperatures.

IV. SUMMARY

In conclusion, we have used AES and RHEED to study the processes taking place during nitridation of the GaAs (001)–2×4 surface by active nitrogen species generated by a rf plasma source. The results of analysis of the spectral shapes and the spectral positions of the Auger electron signals from Ga, As, and N, as well as their dependencies on the nitridation duration combined with RHEED results show that the processes taking place during nitridation greatly differ depending on the nitridation (surface) temperature. On the one hand, at low temperatures ($T_s \leqslant 200\,^{\circ}\mathrm{C}$) nitridation is hindered by kinetic restrictions on the atomic migration, leading to island growth with the formation of disordered GaAsN phase in the subsurface region, which cannot be completely recrystallized even after prolonged annealing. On the other hand, at high temperatures ($T_s \geqslant 500\,^{\circ}\mathrm{C}$) the process of nitridation takes place simultaneously with the etching of the surface due to decomposition of the substrate, leading to the growth of facets and resulting in the rough interface between the GaN and GaAs phases. However, for intermediate nitridation temperatures ($300\,^{\circ}\mathrm{C} \leqslant T_s \leqslant 400\,^{\circ}\mathrm{C}$) it was found that the disordered GaAsN phase, formed at the initial stage of nitridation, partly crystallizes even without annealing. Moreover, AES results indicate that one monolayer of N atoms (partly mixed with arsenic) may well be formed on the substrate during the initial stage of nitridation. The postnitridation annealing of the samples nitrided at the intermediate temperatures results in the formation of a crystalline GaN layer, the line shape of the AES signals from which is identical to that for a thick GaN reference sample.

ACKNOWLEDGMENTS

Fruitful discussions with Dr. Masashi Ozeki are gratefully acknowledged. This work was partly supported by New Energy and Industrial Technology Development Organization (NEDO).

[1]H. Sakai, T. Kiode, H. Suzuki, M. Yamaguchi, S. Yamasaki, M. Kioke, H. Amano, and I. Akasaki Jpn. J. Appl. Phys., Part 2 **34**, L1429 (1995).
[2]J. M. Van Hove, G. Carpenter, E. Nelson, A. Wowchak, and P. Chow, J. Cryst. Growth **164**, 154 (1996).
[3]R. P. Vaudo, I. D. Goepfert, T. D. Moustakas, D. M. Beyea, T. J. Frey, and K. Meehan, J. Appl. Phys. **79**, 2779 (1996).
[4]H. Tews, R. Averbeck, A. Graber, and H. Reichert, Electron. Lett. **32**, 2004 (1996).
[5]S. Mohammad, W. Kim, A. Salvador, and H. Morkoç, MRS Bull. **22**, 22 (1997).
[6]H. Okumura, K. Ohta, G. Feuillet, K. Balarkrishnan, S. Chichibu, H. Hamaguchi, P. Hacke, and S. Yoshida, J. Cryst. Growth **178**, 113 (1997).
[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]S. Gwo, H. Tokumoto, and S. Miwa, Appl. Phys. Lett. **71**, 362 (1997).
[9]K. Uwai, Y. Yamauchi, and N. Kobayashi, Appl. Surf. Sci. **100/101**, 412 (1996).
[10]S. Gourrier, L. Smit, P. Friedel, and P. K. Larsen, J. Appl. Phys. **54**, 3993 (1983).
[11]A. Masuda, Y. Yonezawa, A. Morimoto, and T. Shimizu, Jpn. J. Appl. Phys., Part 1 **34**, 1075 (1995).
[12]X.-Y. Zhu, M. Wolf, T. Huett, and J. M. White, J. Chem. Phys. **97**, 5856 (1992).
[13]X.-Y. Zhu, T. Huett, M. Wolf, and J. M. White, Appl. Phys. Lett. **61**, 3175 (1992).
[14]Φ *Handbook of Auger Electron Spectroscopy*, 3rd ed. (Physical Electronics, Eden Prairie, MN, 1995).
[15]J. Hedman and N. Martensson, Phys. Scr. **22**, 176 (1980).
[16]M. K. Bahl, R. O. Woodall, R. L. Watson, and K. J. Irgolic, J. Chem. Phys. **64**, 1210 (1976).
[17]R. V. Berghe and R. Vlaeminck, Surf. Interface Anal. **10**, 316 (1987).
[18]G. P. Schwartz, M. S. Hybertsen, J. Bevk, R. G. Nuzzo, J. P. Mannaerts, and G. J. Gualtieri, Phys. Rev. B **39**, 1235 (1989).
[19]P. Hill, D. I. Westwood, L. Haworth, J. Lu, and J. E. Macdonald, J. Vac. Sci. Technol. B **15**, 1133 (1997).
[20]C. Falter, W. Ludwig, and M. Selmke, Solid State Commun. **54**, 321 (1985).
[21]D. Briggs and M. P. Seah, *Practical Surface Analysis* (Wiley, New York, 1996), p. 120.
[22]M. Weyers, M. Sato, and H. Ando, Jpn. J. Appl. Phys., Part 2 **31**, L853 (1992).
[23]S. B. Zhang and A. Zunger, Appl. Phys. Lett. **71**, 677 (1997).
[24]T. Makimoto and N. Kobayashi, Appl. Surf. Sci. **100/101**, 403 (1996).
[25]H. Gant and W. Monch, Surf. Sci. **105**, 217 (1981).
[26]W. Ranke and K. Jakobi, Surf. Sci. **47**, 525 (1975).
[27]C. Argile and G. E. Rhead, Surf. Sci. Rep. **10**, 277 (1989).
[28]J. P. Biberian and G. A. Somorjai, Appl. Surf. Sci. **2**, 352 (1979).
[29]G. Ertl and J. Kuppers, *Low Energy Electrons and Surface Chemistry* (VCH, Weinheim, 1985).
[30]M. Ece and R. W. Vook, J. Vac. Sci. Technol. A **6**, 1036 (1988).
[31]H. Okumura, S. Misawa, and S. Yoshida, Appl. Phys. Lett. **59**, 1058 (1991).
[32]S. Yoshida and M. Sasaki, Appl. Surf. Sci. **82/83**, 28 (1994).
[33]H. Okumura, K. Ohta, T. Nagatomo, and S. Yoshida, J. Cryst. Growth **164**, 149 (1996).
[34]O. Brandt, H. Yang, B. Jenichen, Y. Suzuki, L. Daweritz, and K. H. Ploog, Phys. Rev. B **52**, R2253 (1995).
[35]H. Okumura, H. Hamaguchi, G. Feuillet, Y. Ishida, and S. Yoshida, Appl. Phys. Lett. **72**, 3056 (1998).
[36]O. Brandt, H. Yang, and K. H. Ploog, Phys. Rev. B **54**, 4432 (1996).
[37]S. A. Ding, S. R. Barman, K. Horn, H. Yang, O. Brandt, and K. Ploog, Appl. Phys. Lett. **70**, 2407 (1997).
[38]H. Luth, *Surfaces and Interfaces of Solid Materials* (Springer, Berlin, 1995).

Inductively coupled plasma damage in GaN Schottky diodes

X. A. Cao
Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611

A. P. Zhang and G. T. Dang
Department of Chemical Engineering, University of Florida, Gainesville, Florida 32611

H. Cho
Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611

F. Ren
Department of Chemical Engineering, University of Florida, Gainesville, Florida 32611

S. J. Pearton[a)]
Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611

R. J. Shul and L. Zhang
Sandia National Laboratories, Albuquerque, New Mexico 87185

R. Hickman and J. M. Van Hove
SVT Associates, Eden Prairie, Minnesota 55344

(Received 2 April 1999; accepted 14 May 1999)

The effects of H_2 or N_2 plasma exposure on the current–voltage characteristics of GaN Schottky diodes were examined as a function of source power and rf chuck power. Under all conditions there was a strong reduction in diode reverse breakdown voltage and an increase in forward and reverse currents. The results are consistent with creation of a thin ($\leq$600 Å) *n*-type conducting surface region after ion bombardment of the GaN surface. Much of the degradation in diode quality can be recovered by annealing in N_2 at 750 °C. © *1999 American Vacuum Society.*
[S0734-211X(99)06804-3]

I. INTRODUCTION

With the commercialization of GaN-based laser diodes,[1] most of the development efforts in this material system have turned to fabrication of the electronic devices for high power/high temperature applications. These include power microwave for radar and high voltage switches and rectifiers for utility power management.[2–7] Dry etching has proven an effective technique for the formation of mesas in GaN laser structures.[8–12] However while the etch depth is relatively large (2–3 μm) in these structures, the demands on electrical quality of the etched surface are not stringent. Indeed, in the conventional laser structure, the etching terminates on an n^+ contact layer, to which ohmic metallization is applied. Preferential loss of N_2 from the near-surface region during the etch step is actually beneficial in this case because it leads to increased *n*-type doping levels and hence lower contact resistances.[13–15]

The etching requirements for electronic devices are more demanding that those for photonic devices, at least from the electrical quality viewpoint. One of the most sensitive tests of near-surface electrical properties is the quality of rectifying contacts deposited on the etched surface. There has been relatively little work in this area to date. Ren *et al.*[16,17] found that rectifying contacts on electron cyclotron resonance plasma etched GaN and InAlN surfaces were very leaky, though some improvement could be obtained by postetch annealing at 400 °C. Ping *et al.*[9] found that reactively ion etched *n*-GaN surfaces had poor Schottky contact properties, but that plasma chemistries with a chemical component (e.g., Cl_2-based mixtures) produced less degradation than purely physical etching.

Inductively Coupled Plasma (ICP) etching has many attractive attributes for pattern transfer in GaN, including a wide range of control of both ion energy and ion flux, as well as reactive neutral flux.[18,19] To this point there have been no reports of electrical damage effects in ICP etched GaN. In this work we have examined current–voltage (*I–V*) characteristics of GaN Schottky diodes formed on surface exposed to ICP H_2 or N_2 discharges. The effects of both ion energy and ion flux were examined, as well as subsequent annealing or wet etching on the diode properties.

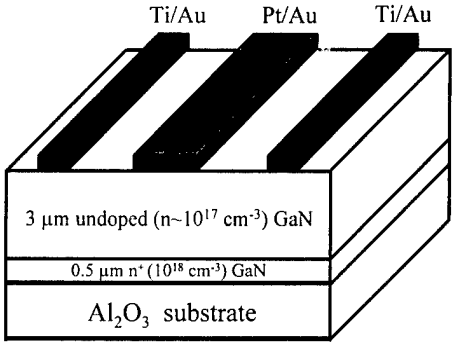


FIG. 1. Schematic of GaN Schottky diode structure.

[a)]Electronic mail: spear@mse.ufl.edu

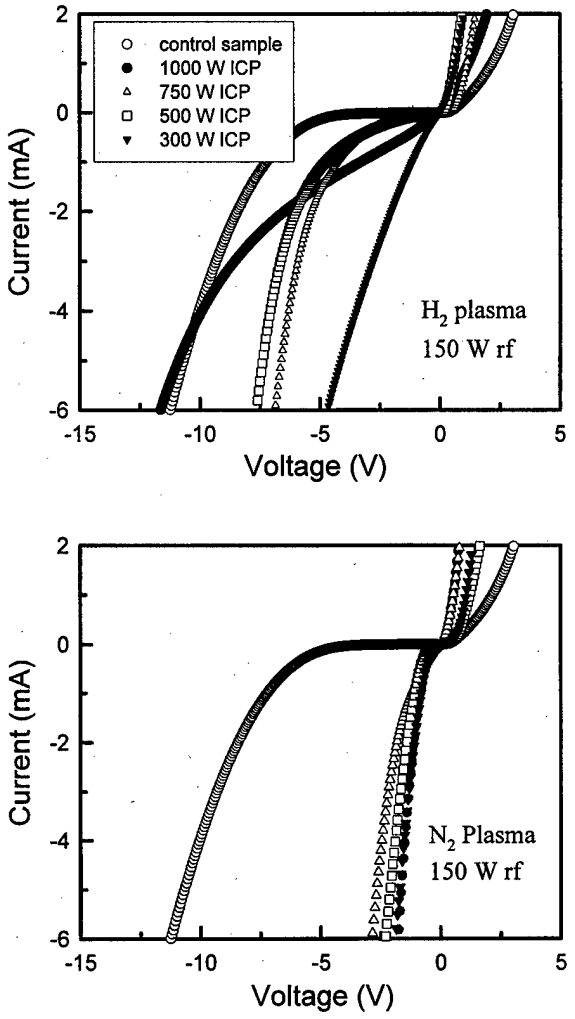


FIG. 2. I–V characteristics from GaN diodes before and after H_2 (top) or N_2 (bottom) plasma exposure (150 W rf chuck power, 5 mTorr) at different ICP source powers.

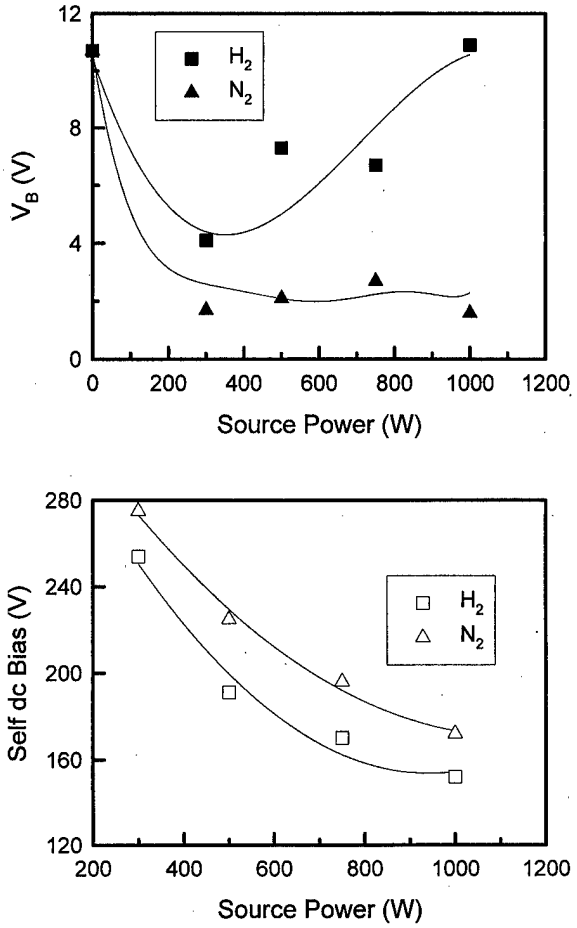


FIG. 3. Variation of V_B in GaN diodes (top) and dc chuck self-bias (bottom) as a function of ICP source power in H_2 or N_2 plasmas (150 W rf chuck power, 5 mTorr).

II. EXPERIMENT

The layer structure and contact metals are shown schematically in Fig. 1. The GaN was grown by rf plasma-assisted molecular beam epitaxy on c-plane Al_2O_3 substrates.[20] The Ti/Au ohmic contacts were patterned by lift-off and annealed at 750 °C, producing contact resistances in the 10^{-5} Ω cm^{-2} range. Samples were exposed to either pure N_2 or H_2 discharges in a Plasma Therm 790 ICP system at a fixed pressure of 5 mTorr. The gases were injected into the ICP source at a flow rate of 15 standard cubic centimeters per minute (sccm). The experimentally varied parameters were source power (300–1000 W) and rf chuck power (40–250 W), which control ion flux and ion energy, respectively. In some cases the samples were either annealed in N_2 for 30 s at 300–850 °C, or photoelectrochemically etched in 0.2 M KOH solutions at 25 °C after plasma exposure. The Pt/Au Schottky metallization was then deposited through a stencil mass by e-beam evaporation. Current–voltage characteristics were recorded on a HP4145A parameter analyzer, and we defined the reverse breakdown voltage (V_B) as the voltage at which the leakage current was 10^{-3} A. We found in all cases that plasma exposure caused significant increases in forward and reverse currents, with ideality factors increasing from typical values of 1.4–1.7 on control samples to >2. For this reason we were unable to extract meaningful values of either ideality factor or barrier height.

III. RESULTS AND DISCUSSION

Figure 2 shows a series of I–V characteristics from the GaN diodes fabricated on samples exposed to either H_2 or N_2 discharges at different source powers. It is clear that N_2 plasma exposure creates more degradation of the diode characteristics than does H_2 exposure. This implicates the ion mass ($^{28}N_2^+$, $^2H_2^+$ for the main positive ion species) as being more important in influencing the electrical properties of the GaN surface than a chemical effect, since H_2 would be likely to preferentially remove nitrogen from the GaN as NH_3.

The variations of V_B of the diodes with the source power during plasma exposure are shown in Fig. 3. For any exposure to the N_2 discharges, V_B is severely reduced. By contrast there is less degradation with the H_2 plasma exposures at higher source powers. This is likely related to the lower average ion energy at those conditions, as shown at the bottom of Fig. 3. The average ion energy is approximately equal

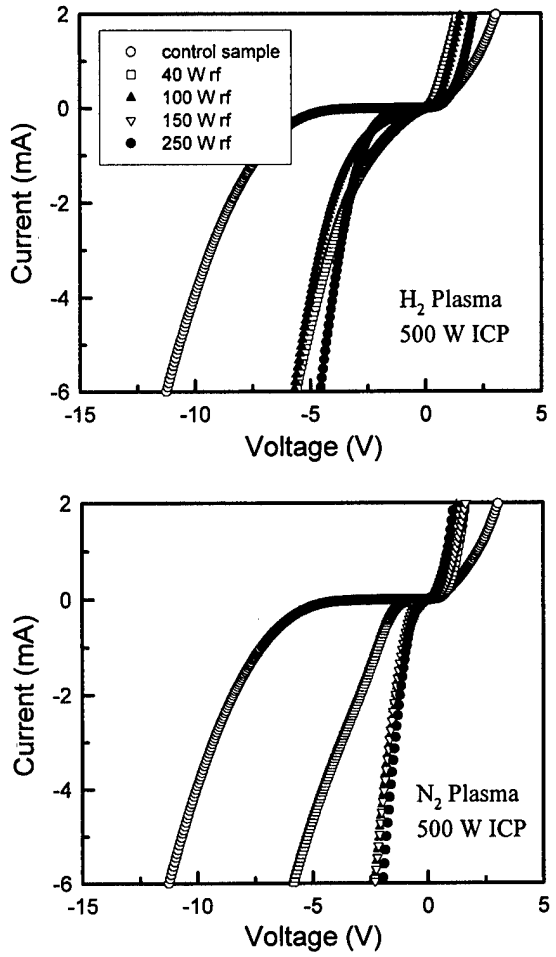


FIG. 4. *I–V* characteristics from GaN diodes before and after H_2 (top) or N_2 (bottom) plasma exposure (500 W source power, 5 mTorr) at different rf chuck powers.

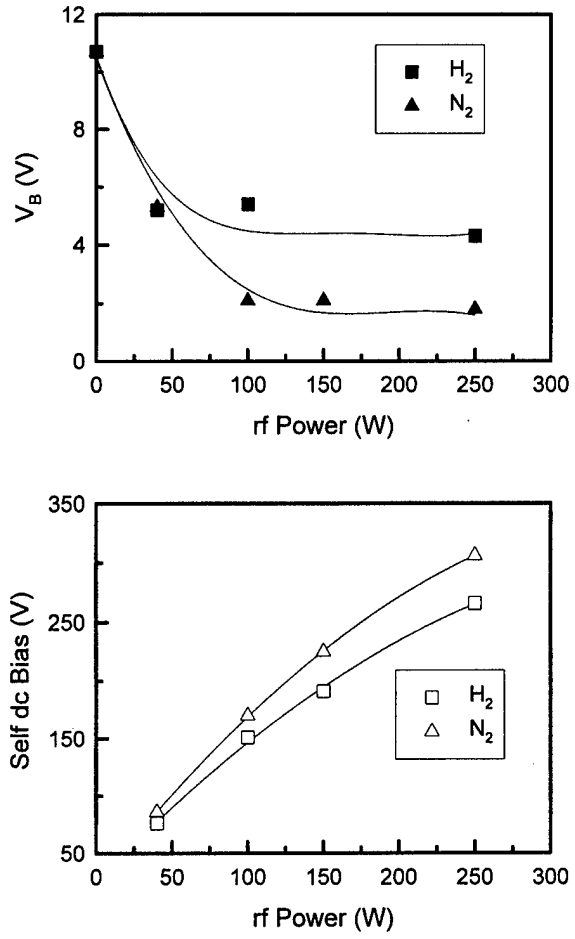


FIG. 5. Variation of V_B in GaN diodes (top) and dc chuck self-bias (bottom) as a function of rf chuck power in H_2 or N_2 plasmas (500 W source power, 5 mTorr).

to the sum of dc self-bias and plasma potential, with the latter being in the range -22 to -28 V as determined by Langmuir probe measurements. Ion induced damage in GaN displays *n*-type conductivity,[21,22] and in addition the heavy N_2^+ ions are also more effective in preferential sputtering of the N relative to Ga, compared to the H_2^+ ions. The net result is that N_2^+ ions will lead to more degradation of the surface electrical properties of GaN than do H_2^+ ions of similar energy.

Similar conclusions can be drawn from the data on the effect of increasing rf chuck power. Figure 4 shows the diode *I–V* characteristics from H_2 or N_2 plasma exposed samples at fixed source power (500 W) but varying rf chuck power. There are once again very severe decreases in breakdown voltage and increase in leakage current. The dependence of V_B on rf chuck power during the plasma exposures is shown in Fig. 5, along with the dc self-basis. The V_B values fall by more than a factor of two even for very low self-biases, and emphasize how sensitive the GaN surface is to degradation by energetic ion bombardment. The degradation saturates beyond ~100 W chuck power, corresponding to ion energies of ~175 eV. We assume that once the immediate surface becomes sufficiently damaged, the contact properties basically cannot be made any worse and the issue is then whether the damage depth increases with the different plasma parameters. Since ion energy appears to be a critical factor in creating the near-surface damage, we would expect damage depth to increase with ion energy in a nonetching process. In the case of simultaneous etching and damage creation (e.g., in Cl_2/Ar etch processing), higher etch rates would lead to lower amounts of residual damage because the disordered region would be partially removed.

The damage depth was established by photoelectrochemically wet etching different amounts of the plasma-exposed GaN surfaces, and then depositing the Pt/Au metal. Figure 6 (top) show the effect on the *I–V* characteristics of this removal of different depth of GaN. There is a gradual restoration of the reverse breakdown voltage, as shown at the bottom of the figure. Note that the forward part of the characteristics worsens for removal of 260 Å of GaN, and shows signs of high series resistance. This would be consistent with the presence of a highly resistive region underneath the conducting near-surface layer, created by point defect diffusion from the surface. A similar model applies to ion-damaged InP, i.e., a nonstoichiometric near-surface region (deficient in P in that case),[23] followed by a transition to a

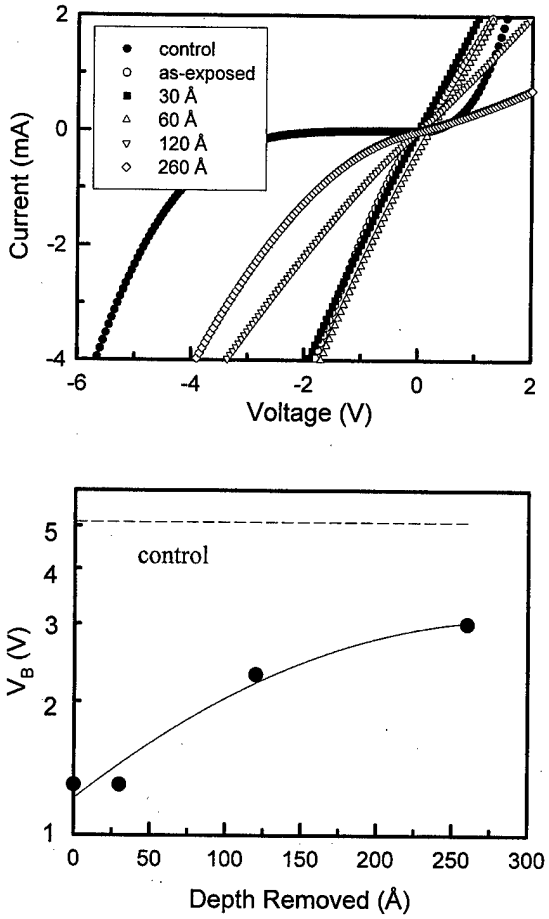


FIG. 6. I–V characteristics from N_2 plasma exposed GaN diodes before and after wet etch removal of different amounts of GaN prior to deposition of the Schottky contact (top), and variation of V_B as a function of the amount of material removed (bottom).

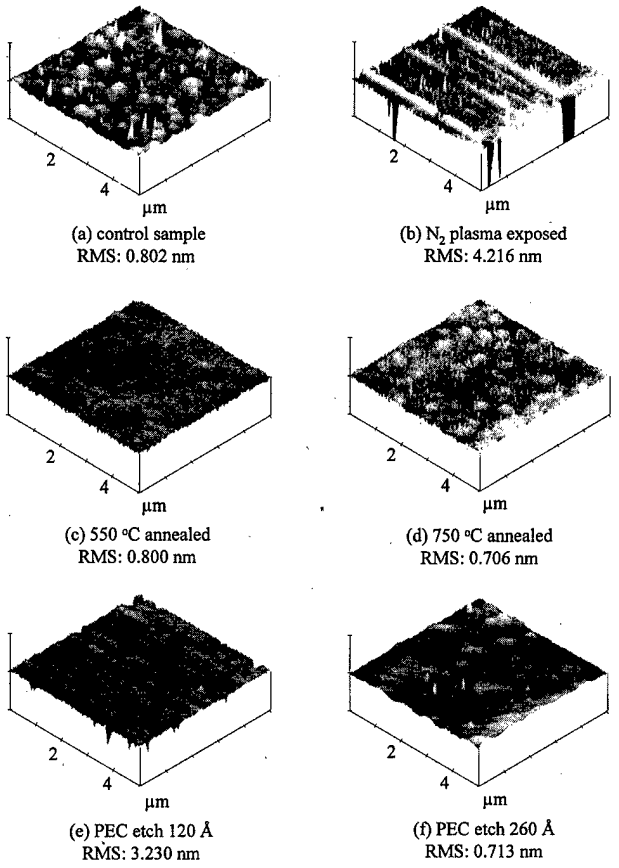


FIG. 7. AFM scans before and after N_2 plasma exposure (500 W source power, 150 W rf chuck power, 5 mTorr) and subsequent annealing or photochemical etching.

stoichiometric but point defect-compensated region, and finally to unperturbed InP.

The fact that plasma exposure severely degraded the surface is clear from the atomic force microscopy (AFM) data of Fig. 7. Exposure to the 500 W source power, 150 W rf chuck power (dc self-bias −221 V), 5 mTorr N_2 discharge increased the root-mean-square (rms) surface roughness from 0.8 to 4.2 nm. Subsequent photoelectrochemical etching restored the initial morphology. However, we observed the onset of increasingly rough surface for deeper etch depths,[24] producing a relatively inaccurate measure of how much of the surface had to be removed in order to restore the diode breakdown voltage to its original value. We were able to estimate this depth as ~600±150 Å for the N_2 plasma conditions mentioned above.

Another method for trying to restore the electrical properties of the plasma-exposed surface is annealing. Figure 7 also show AFM scans from samples after annealing at 550 or 750 °C, with no significant change in rms values. I–V data from annealed samples are shown in Fig. 8. At the top are characteristics from samples which were plasma exposed (N_2, 500 W source power, 150 W rf chuck power, 5 mTorr), then annealed and the contact deposited. These samples show that increasing annealing temperature up to 750 °C brings a substantial improvement in V_B (Fig. 8, bottom). However for annealing at 850 °C the diode began to degrade and this is consistent with the temperature at which N_2 begins to be lost from the surface.[22] In the case where the samples were exposed to the N_2 plasma, and then the Pt/Au contact was deposited prior to annealing, the I–V characteristics show continued worsening upon annealing (Fig. 8, center). In this case, the Pt/Au contact is stable to 700 °C on unetched samples. The poorer stability in etched samples could be related to the surface damage enhancing interfacial reaction between the Pt and GaN.

IV. SUMMARY AND CONCLUSIONS

The main findings of this study can be summarized as follows:

(1) There is a severe degradation in the electrical quality of GaN surfaces after ICP H_2 or N_2 discharge exposure. Under all conditions there is a strong reduction of V_B in diode structures to the point at which the Schottky contacts show almost ohmic-like behavior. These observations are consistent with the creation of a conducting n-type surface layer resulting from energetic ion bombardment. Heavier ions (N_2^+) create more damage than

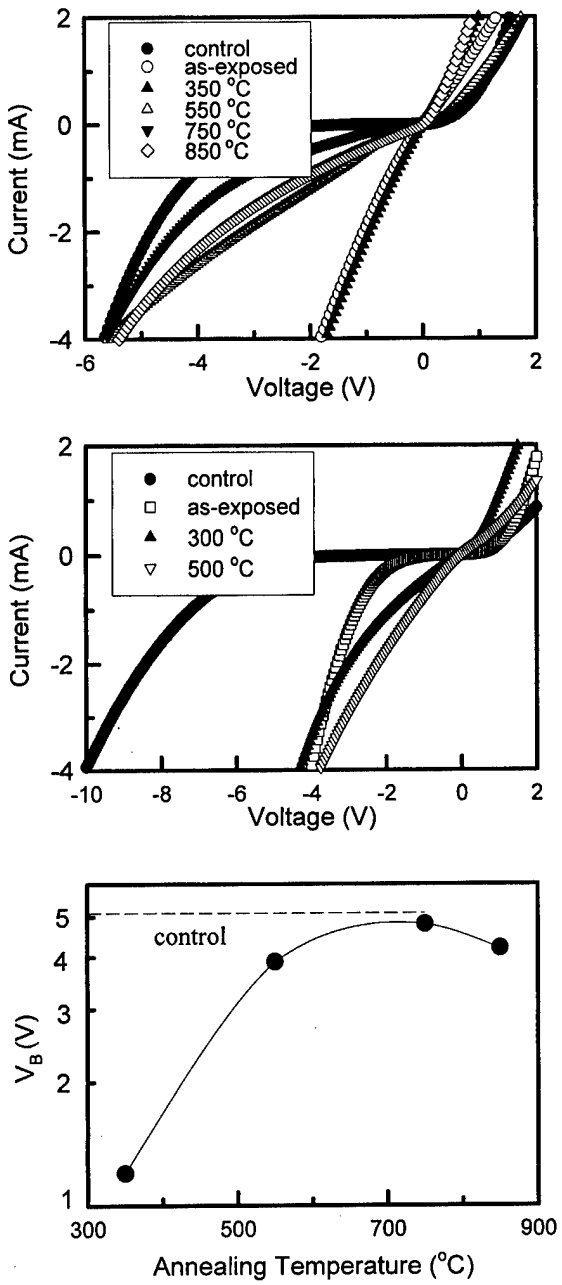


FIG. 8. *I–V* characteristics from GaN diodes before and after N_2 plasma exposure (500 W source power, 150 W rf chuck power, 5 mTorr) and subsequent annealing either prior (top) or subsequent (center) to the deposition of the Schottky metallization. The variation of V_B in the samples annealed prior to metal deposition is shown at the bottom of the figure.

lighter ions (H_2^+) in this situation, where damage is accumulating without any concurrent etching of the surface.

(2) The depth of the damage is approximately 600 Å, as judged by the return of the diode characteristics to their control values.

(3) Annealing at 750 °C is also effective in helping remove the effects of plasma exposure. Higher temperatures lead to degradation in GaN diode properties for uncapped anneals.

ACKNOWLEDGMENTS

The work at UF is partially supported by a DARPA/EPRI Grant No. MDA 972-98-1-0006 (D. Radack and J. Melcher) monitored by ONR (J. C. Zolper), and a NSF Grant No. DMR 97-32865 (L. D. Hess). Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin company, for the US Department of Energy under Contract No. DEAC04-94-AL-85000.

[1]S. Nakamura, IEEE J. Sel. Top. Quantum Electron. **4**, 483 (1998).
[2]M. Shur, Solid-State Electron. **42**, 2131 (1998).
[3]J. C. Zolper, Solid-State Electron. **42**, 2153 (1998).
[4]S. C. Binari, J. M. Redwing, G. Kelner, and W. Kruppa, Electron. Lett. **33**, 242 (1997).
[5]Y.-F. Wu, B. P. Keller, P. Fini, S. Keller, T. J. Jenkins, L. T. Kehias, S. P. DenBaars, and U. Mishra, IEEE Electron Dev. Lett. **19**, 50 (1998).
[6]B. E. Foutz, L. F. Eastman, U. V. Bhapkar, and M. S. Shur, Appl. Phys. Lett. **70**, 2849 (1997).
[7]Q. Chen, R. Gaska, M. A. Khan, M. S. Shur, A. T. Ping, I. Adesida, J. Burn, W. J. Schaff, and L. F. Eastman, Electron. Lett. **33**, 637 (1997).
[8]See, for example, R. J. Shul, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), Chap. 12.
[9]A. T. Ping, A. C. Schmitz, I. Adesida, M. A. Khan, Q. Chen, and J. W. Yang, J. Electron. Mater. **26**, 266 (1997).
[10]H. P. Gillis, D. A. Choutov, K. P. Martin, M. D. Bremser, and R. F. Davis, J. Electron. Mater. **26**, 301 (1997).
[11]A. T. Ping, C. Youtsey, I. Adesida, M. A. Khan, and J. N. Kuznia, J. Electron. Mater. **24**, 229 (1995).
[12]S. J. Pearton, C. R. Abernathy, and F. Ren, Appl. Phys. Lett. **64**, 2294 (1994).
[13]Z. F. Fan, S. N. Mohammad, W. Kim, Ö. Aktas, A. E. Botchkarev, and H. Morkoç, Appl. Phys. Lett. **69**, 1672 (1996).
[14]J. Y. Chen, C. J. Pan, and G. C. Chi, Solid-State Electron. **43**, 649 (1999).
[15]A. T. Ping, Q. Chen, J. W. Yang, M. A. Khan, and I. Adesida, J. Electron. Mater. **27**, 261 (1998).
[16]F. Ren, J. Lothian, Y. K. Chen, J. D. MacKenzie, S. M. Donovan, C. B. Vartuli, C. R. Abernathy, J. W. Lee, and S. J. Pearton, J. Electrochem. Soc. **143**, 1217 (1996).
[17]F. Ren, J. Lothian, Y. K. Chen, R. Karlicek, L. Tran, M. Schurmann, R. Stall, J. W. Lee, and S. J. Pearton, Solid-State Electron. **41**, 1819 (1997).
[18]R. J. Shul, G. B. McClellan, S. A. Casalnuovo, D. J. Rieger, S. J. Pearton, C. Constantine, C. Barratt, R. F. Karlicek, Jr., C. Tran, and M. Schurman, Appl. Phys. Lett. **69**, 1119 (1996).
[19]S. A. Smith, C. A. Wolden, M. D. Bremser, A. D. Hanser, R. F. Davis, and W. V. Lampert, Appl. Phys. Lett. **71**, 3631 (1997).
[20]J. M. Van Hove, R. Hickman, J. J. Klaassen, P. P. Chow, and P. P. Ruden, Appl. Phys. Lett. **70**, 282 (1997).
[21]S. J. Pearton, J. W. Lee, J. D. MacKenzie, C. R. Abernathy, and R. J. Shul, Appl. Phys. Lett. **67**, 2329 (1995).
[22]J. C. Zolper, D. J. Rieger, A. G. Baca, S. J. Pearton, J. W. Lee, and R. A. Stall, Appl. Phys. Lett. **69**, 538 (1996).
[23]S. J. Pearton, U. K. Chakrabarti, and F. A. Baiocchi, Appl. Phys. Lett. **55**, 1633 (1989).
[24]H. Cho, K. H. Auh, J. Han, R. J. Shul, S. M. Donovan, C. R. Abernathy, E. S. Lambers, F. Ren, and S. J. Pearton, J. Electron. Mater. **28**, 288 (1999).

Effect of rapid thermal annealing on radio-frequency magnetron-sputtered GaN thin films and Au/GaN Schottky diodes

Ching-Wu Wang,[a] Jin-Yuan Liao, Chih-Liang Chen, and Wen-Kuan Lin
Electronic Engineering Department, I-Shou University, Ta-Hsu Hsiang, Kaohsiung County, Taiwan, Republic of China

Yan-Kuin Su and Meiso Yokoyama
Electrical Engineering Department, National Cheng-Kung University, Tainan, Taiwan, Republic of China

(Received 30 December 1998; accepted 30 April 1999)

The structural and optical properties of rf magnetron-sputtered GaN thin films on p^+-Si substrates have been accessed as a function of rapid thermal annealing (RTA) temperatures from 800 to 1000 °C. The evidence has revealed that higher RTA temperatures not only assist the GaN films in recrystallizing into stable hexagonal form but also enhance the near-band-edge emission of GaN films in the photoluminescence spectrum. Moreover, a deep electron trap (E_t) with activation energy $E_c - E_t \cong 0.39$ eV detected at the surface of higher-RTA-temperature-treated GaN films was asserted to be a nitrogen-vacancy-related defect that takes a defect-assisted-tunneling role in the forward conduction process of Au/GaN Schottky diode. The greater reverse leakage current and lower breakdown voltage are suggested to be due to the effects of a lower barrier height and higher ideality factor that occurred in the higher-RTA-temperature-treated samples. © *1999 American Vacuum Society.* [S0734-211X(99)02504-4]

I. INTRODUCTION

As suggested by a growing number of authors, gallium nitride (GaN) is becoming an important material with regard to the development of optoelectronic devices operating in either the blue region of the visible spectrum or the near-ultraviolet (UV) regime due to its appropriate band-gap magnitude. In most cases, GaN films are fabricated by metalorganic chemical vapor deposition (MOCVD),[1,2] and molecular beam epitaxy (MBE)[3–5] methods and have with a GaN or AlN buffer on various substrates. Recently, Nonomura *et al.*[6] have demonstrated for the first time that by the radio-frequency (rf) reactive sputtering method amorphous GaN thin films can be deposited onto quartz substrates. However, such an amorphous GaN film cannot be employed in the utilization of high-quality optoelectronic devices. Therefore, if GaN is to be used in these devices, an effective thermal treatment to improve the film quality of GaN needs to be developed. Unfortunately, to our knowledge, a detailed systematic investigation of GaN film improvement by thermal treatment has not been presented until recently.[7]

On the other hand, it is widely accepted that defects in the band gap of GaN can affect both the electrical and optical properties of the material and they are known to be detrimental to its device applications. For instance, the multiphonon-hopping-conduction phenomenon, which is due to the deep defect centers, has been observed in MBE-grown GaN.[8] Additionally, a combination of deep interface states and deep bulk levels within a certain tunneling distance of the Schottky barrier has been suspected as being responsible for the degradation mechanism in GaN ultraviolet photodetectors.[9] Thus, the investigation of the origins of deep levels and their influence on GaN-made devices is an important issue.

The purpose of this work, then, is to report the rapid thermal annealing (RTA) effects on the structural and optical characteristics of GaN thin films and Au/GaN Schottky diodes grown on p^+-Si(111) substrates using a rf magnetron sputtering technique. A nitrogen-vacancy-related deep defect level, E_t, near the surface of RTA-treated GaN thin films, measured by deep level transient spectroscopy (DLTS)[10] and secondary ion mass spectrometry (SIMS), was demonstrated to play a defect-assisted-tunneling[11–14] role in the conduction process of Au/GaN Schottky diodes. This result is shown in Sec. III. In Sec. II, the experimental details are described while the conclusions are presented in Sec. IV.

II. EXPERIMENT

GaN thin films were prepared by a two-stage-growth[15] rf magnetron sputtering method. The source material, made by Superconductive Components, Inc. (USA), was a GaN target (3 in. diam) with 99.999% purity. Prior to loading the p^+-Si(111) substrates into the sputter chamber, the wafers were rinsed in 10:1 HF acid to remove the native oxide layer and blown dry. During the growth process, the rf power was controlled at 50 W and a mixed sputtering gas of N_2 and Ar was utilized to keep the sputtered pressure at 5×10^{-3} Torr. For the first stage, a thin GaN buffer layer (~500 Å) was deposited onto the Si substrate at 400 °C. Then the substrate temperature was elevated to 500 °C to anneal the GaN buffer under a purified N_2 atmosphere for 30 min. Next, a 3000 Å thick GaN epilayer was grown on top of the GaN buffer at 700 °C. After growth, various RTA treatment temperatures (800–1000 °C) on the sputtered GaN thin films were carried out in 0.5 Torr purified N_2 ambient for 1 min. For the DLTS measurement, Schottky diodes were

[a]Electronic mail: cwwang@csa500.isu.edu.tw

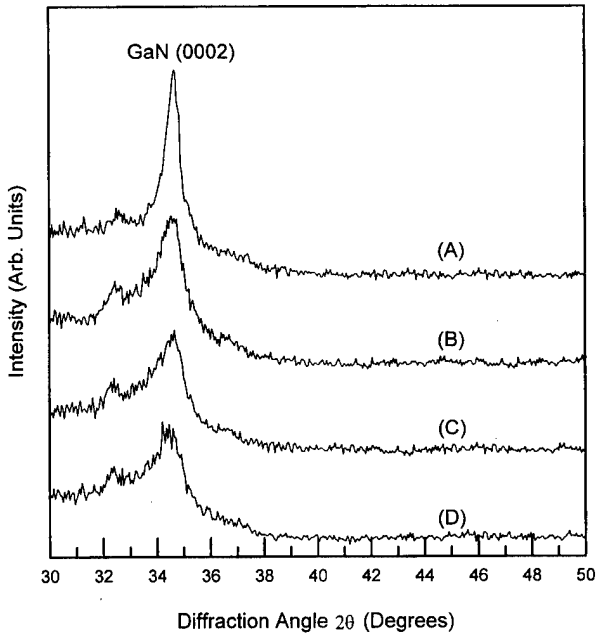


FIG. 1. Typical x-ray diffraction patterns of sputtered GaN thin films treated by different RTA temperatures, where curves (A), (B), (C), and (D) represent RTA temperatures of, respectively, 1000, 950, 900, and 800 °C for 1 min.

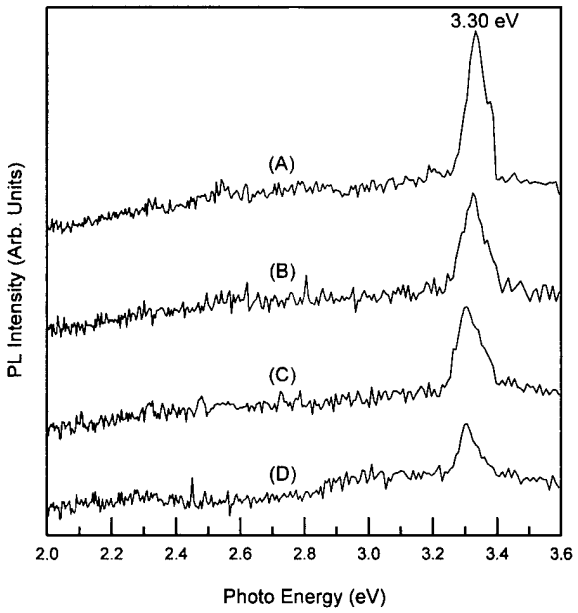


FIG. 2. Room temperature PL spectra of sputtered GaN thin films. The RTA temperatures for curves (A), (B), (C), and (D) are, respectively, 1000, 950, 900, and 800 °C for 1 min.

made by evaporating gold (Au) and aluminum (Al) on top of GaN thin film which provide, respectively, a Schottky contact at Au/GaN and an ohmic contact at Al/GaN junctions.

III. RESULTS AND DISCUSSION

The as-grown samples exhibited *n*-type conductivity and amorphous film quality. Furthermore, typical x-ray diffraction (XRD) patterns showed that a RTA treatment can effectively improve the crystallinity of the GaN film property, which is illustrated in Fig. 1. As is evident from Fig. 1, the peak intensity associated with the (0002) hexagonal GaN reflection at $2\theta=34.6°$ clearly increases with increasing RTA temperatures. No peak attributable to cubic GaN (commonly corresponding to the $2\theta=39.6°$) is found. This implies that high temperature RTA treatment promotes the sputtered GaN to form a stable wurtzite (α) structure other than the metastable zinc-blende (β) phase.[16]

Figure 2 shows the photoluminescence (PL) spectra of the RTA-treated samples excited by a 325 nm He–Cd laser at room temperature (RT). The blue emission of 3.30 eV corresponding to the near-band-edge transition was observed as becoming stronger as the RTA temperature increased. This band emission has been assigned to excitons bound to neutral donors[17] and is considered a PL fingerprint of *N*-type GaN.

To examine the deep defects existing within these RTA-treated GaN materials, DLTS measurements were performed on the different RTA-temperature-treated Au/GaN Schottky diodes. The DLTS results are shown in Fig. 3. As can be seen, a deep electron trap (designated as E_t) was consistently detected in all the samples. Moreover, the Arrhenius plot yielding the activation energy of E_t below the conduction band-edge minimum is $E_c-E_t\cong 0.39$ eV (± 0.02 eV). Interestingly, the DLTS signal intensity of E_t was clearly observed to increase by raising the RTA temperature. We suspect that E_t is a nitrogen-vacancy-related deep trap and most likely exists at the surface of higher-temperature-RTA-treated GaN film. This inference is based on the SIMS observation that the higher the temperature of RTA-treated samples the more nitrogen deficiency detected at the film

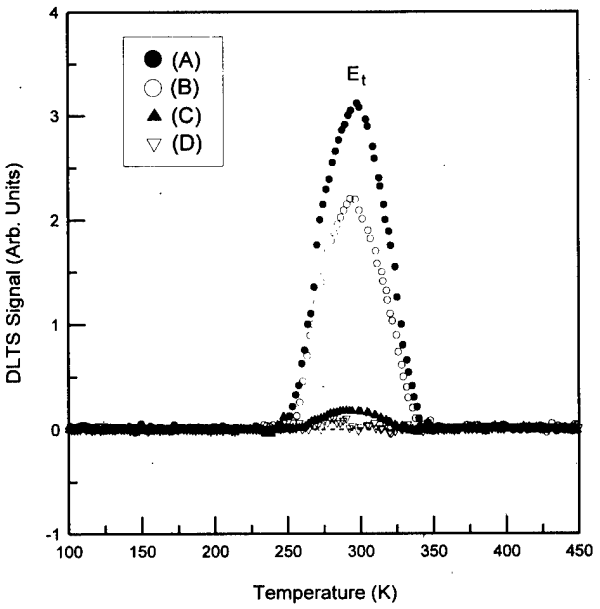


FIG. 3. DLTS signals of different RTA-treated Au/GaN Schottky diodes. Curves (A), (B), (C), and (D) correspond to RTA temperatures of, respectively, 1000, 950, 900, and 800 °C for 1 min.

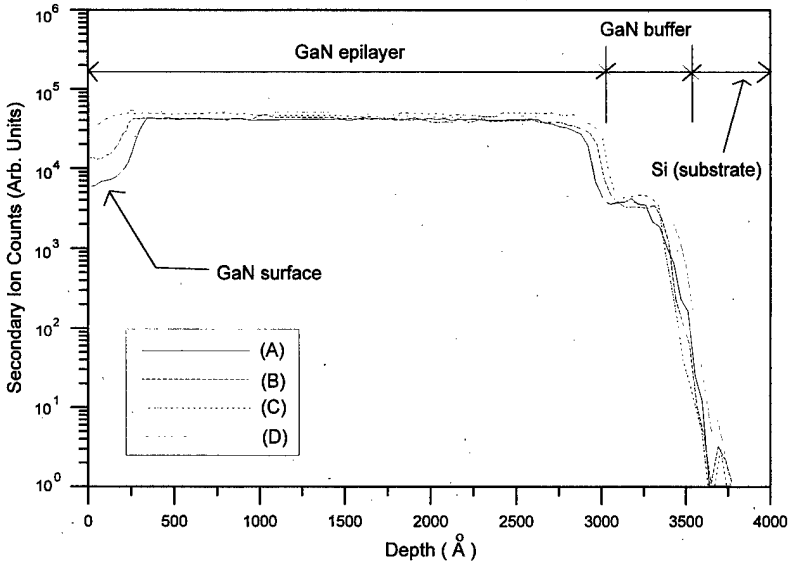


FIG. 4. SIMS depth profiles of N on GaN (epilayer)/GaN (buffer)/p^+-Si devices with different RTA treatment temperatures of, respectively, (A) 1000, (B) 950, (C) 900, and (D) 800 °C for 1 min.

surface, as illustrated in Fig. 4. How this observed E_t trap affects the electrical behavior of GaN Schottky diodes will be discussed below.

In recent years, a great amount of effort has focused on the investigation of electrical characteristics [current-voltage (I-V)] of Schottky contacts of Au on MOCVD and MBE-made n-GaN.[18–21] Here, we report for the first time that the RTA temperature has a significant effect on Au/sputtered GaN Schottky diodes. In Fig. 5, it is clearly seen that for samples treated with higher RTA temperatures superior forward conducting characteristics resulted, i.e., a smaller turn-on voltage and a bigger forward conduction current despite the fact that the reverse leakage current is a little bit larger. It is our assertion that the near-surface deep electron trap, E_t, caused by the nitrogen vacancy plays a major role in affecting the conduction process of Au/sputtered GaN Schottky diodes. Accordingly, the I-V characteristics of most Schottky contacts may be expressed by interfacial layer-thermonic emission theory[22,23] as

$$J = A^* \times T^2 \exp\left(-\frac{q\phi_b}{KT}\right)\left[\exp\left(\frac{qV}{nKT}\right) - 1\right], \tag{1}$$

where J is the current density, A^* is the effective Richardson constant, including the transmission probability across the interfacial layer, T is the temperature, q is the electron charge, ϕ_b is the barrier height, n is the ideality factor, K is the Boltzmann constant, and V is the applied voltage. In the forward-biased situation, if there were many defects near the surface region, electrons would go through the barrier from n-type GaN to metal by defect-assisted tunneling,[11–14] thus greatly enhancing the tunneling probability. In this work, higher-RTA-temperature treated samples possessing a higher density of deep electron traps (E_t) near the surface of n-type sputtered GaN fit with the prediction of defect-assisted-tunneling, producing smaller forward turn-on voltage and larger conduction current. Moreover, such a defect-assisted-tunneling phenomenon has also been identified as occurring at Ni–GaN Schottky diodes,[11] at PtSi/GaN Schottky contacts,[12] and at ITO/P-InP structures,[13] respectively. Nevertheless, in the reverse-biased condition, the large leakage current and the smaller breakdown voltage were deduced by the presence of Ga–Au compounds at the metal-semiconductor interface. This formation of Ga–Au compounds has been identified by Sheu *et al.*[24] in their x-ray diffraction measurements and has been demonstrated to reduce the Schottky barrier height, yielding the larger reverse

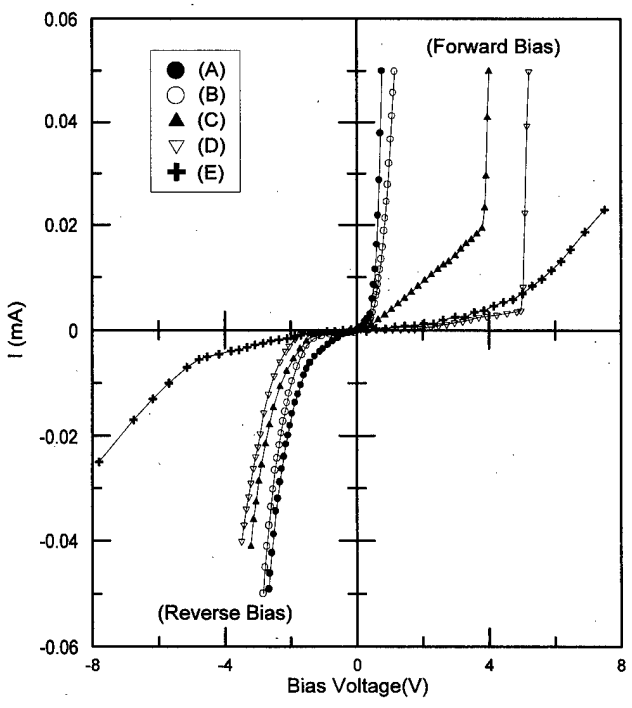


FIG. 5. Current-voltage characteristics of Au/GaN Schottky diodes. Curves (A), (B), (C), and (D) represent the samples treated by RTA temperatures of, respectively, 1000, 950, 900, and 800 °C for 1 min. Curve (E) represents the as-deposited sample.

TABLE I. Summarizing the results for I-V characteristics of 1000 °C RTA-treated Au/sputtered GaN Schottky diodes obtained in this work and comparisons to other work.

Schottky diode	Fabrication technique	Substrate	Ideality factor, n	Barrier height, ϕ_b (eV)	Source
Au/GaN	Sputtering	Si(111)	1.56	0.77	This work
Au/GaN	MOCVD	Sapphire (0001)	1.04	0.87	Ref. 20
Au/GaN	MBE	Sapphire	1.4	0.98	Ref. 21

leakage and smaller breakdown voltage. The ideality factor (n) and barrier height (ϕ_b) obtained from I-V measurements for a 1000 °C RTA-treated Au/sputtered GaN Schottky diode and a comparison with results of other authors are presented in Table I. In reviewing Table I, our results show a 1000 °C RTA-treated Au/GaN Schottky diode possessing a lower barrier height and higher ideality factor in comparism with MOCVD and MBE-made samples. Strictly speaking, it is a well-known fact that the quality of easily made sputtered GaN thin films can hardly compete with that of MBE or MOCVD-grown thin films. The lower barrier height measured in our sputtered sample is attributed to the formation of Ga–Au compounds at the Au–GaN interface. However, the higher ideality factor is suggested to be due to the recombination and trapping of carriers in the depletion region. These two factors all contribute to a lower breakdown voltage and a higher leakage current in the devices tested.

IV. CONCLUSIONS

We have investigated the RTA effects on the structural and optical characteristics of rf magnetron-sputtered GaN thin films. The evidence shows that the higher temperature RTA treatment not only promotes recrystalization of the GaN films into a stable hexagonal structure but also significantly enhances the near-band-edge emission of sputtered GaN films in the PL spectrum. The deep electron trap, E_t, attributed to the nitrogen-vacancy-related defect occurring at the GaN film surface through the higher temperature of the RTA treatment, was demonstrated as playing the role of defect-assisted tunneling in the forward conduction process of Au/GaN Schottky diodes. The greater reverse leakage current and lower breakdown voltage are believed to be due to the effects of the lower barrier height and higher ideality factor that occurred in the higher-RTA-temperature treated samples. Further work in this direction is necessary.

ACKNOWLEDGMENT

One of the authors (C.W.W.) gratefully acknowledges the financial support from the National Science Council (NSC) of Taiwan under Contract No. NSC88-2215-E-214-002.

[1]S. Fujieda and Y. Matsumoto, Jpn. J. Appl. Phys., Part 2 **30**, L1665 (1991).
[2]S. Miyoshi, K. Onabe, N. Ohkouchi, H. Yaguchi, R. Ito, S. Fukatsu, and Y. Shiraki, J. Cryst. Growth **124**, 439 (1992).
[3]Z. Q. He, X. M. Ding, X. Y. Hou, and X. Wang, Appl. Phys. Lett. **64**, 315 (1994).
[4]R. C. Powell, N. E. Lee, Y. W. Kim, and J. E. Greene, J. Appl. Phys. **73**, 189 (1993).
[5]A. Kikuchi, H. Hoshi, and K. Kishino, Jpn. J. Appl. Phys., Part 1 **33**, 688 (1994).
[6]S. Nonomura, S. Kobayashi, T. Gotoh, S. Hirata, T. Ohmori, T. Itoh, S. Nitta, and K. Morigaki, J. Non-Cryst. Solids **198–200**, 174 (1996).
[7]M. W. Cole, F. Ren, and S. J. Pearton, Appl. Phys. Lett. **71**, 3004 (1997).
[8]D. C. Look, D. C. Reynolds, W. Kim, O. Aktas, A. Botchkarer, A. Salvador, and H. Morkoç, J. Appl. Phys. **80**, 2960 (1996).
[9]J. C. Carrano, T. Li, P. A. Grudowski, C. J. Eiting, R. D. Dupuis, and J. C. Campbell, J. Appl. Phys. **83**, 6148 (1998).
[10]D. V. Lang, J. Appl. Phys. **45**, 3023 (1974).
[11]L. S. Yu, Q. Z. Liu, Q. J. Xing, D. J. Qiao, S. S. Lau, and J. Redwing, J. Appl. Phys. **84**, 2099 (1998).
[12]Q. Z. Liu, L. S. Yu, S. S. Lau, J. M. Redwing, N. R. Perkins, and T. F. Kuech, Appl. Phys. Lett. **70**, 1275 (1997).
[13]J. K. Luo and H. Thomas, Appl. Phys. Lett. **62**, 705 (1993).
[14]M. H. Park, C. L. Yeh, L. C. Wang, F. Deng, Q. Z. Liu, A. R. Clawson, and S. S. Lau, J. Appl. Phys. **77**, 2056 (1995).
[15]T. Sasaki and T. Matsuoka, J. Appl. Phys. **77**, 192 (1995).
[16]T. S. Cheng, L. C. Jenkins, S. E. Hooper, C. T. Foxon, J. W. Orton, and D. E. Lacklison, Appl. Phys. Lett. **66**, 1509 (1995).
[17]R. Dingle, D. D. Sell, S. E. Stokowski, and M. Ilegems, Phys. Rev. B **4**, 1211 (1971).
[18]L. D. Bell, R. P. Smith, B. T. McDermott, E. R. Gertner, R. Pittman, R. L. Pierson, and G. J. Sullivan, J. Vac. Sci. Technol. B **16**, 2286 (1998).
[19]C. L. Wu and A. Kahn, J. Vac. Sci. Technol. B **16**, 2218 (1998).
[20]A. C. Schmitz, A. T. Ping, M. A. Khan, Q. Chen, J. W. Yang, and I. Adesida, Semicond. Sci. Technol. **11**, 1464 (1996).
[21]Y. Kribes, I. Harrison, B. Tuck, T. S. Cheng, and C. T. Foxon, Semicond. Sci. Technol. **12**, 913 (1997).
[22]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), p. 166.
[23]C. Y. Wu, J. Appl. Phys. **53**, 5949 (1982).
[24]J. K. Sheu *et al.*, J. Appl. Phys. **83**, 3172 (1998).

Plasma assisted two stage selenization process for the preparation of low resistivity ZnSe films

R. Jeyakumar
National Physical Laboratory, Dr. K. S. Krishnan Road, New Delhi 110012, India

G. K. Chadda
Department of Physics and Astrophysics, University of Delhi

S. T. Lakshmikumar[a] and A. C. Rastogi
National Physical Laboratory, Dr. K. S. Krishnan Road, New Delhi 110012, India

(Received 29 May 1998; accepted 23 April 1999)

Preparation of low resistivity, ZnSe films using a plasma assisted two stage selenization process is reported. The presence of N_2 plasma enhances the crystallinity of the films. The films are optically transparent and have a band gap of 2.65 eV which remains unchanged by the plasma enhancement process. Plasma grown ZnSe films exhibit lower resistivities possibly due to doping by the incorporation of N_2 ions. ZnSe films formed at low reactor pressures (<1.5 mbar) in the presence of N_2 plasma showed a significant lowering of the resistivity to 10 Ω cm. XPS analysis shows a chemical shift of 1.6 eV for the Se 3*d* line of the plasma selenized sample instead of 0.9 eV for the nonplasma sample. © *1999 American Vacuum Society.* [S0734-211X(99)01604-2]

I. INTRODUCTION

In polycrystalline thin film solar cells, CdS is conventionally used as the window layer even though there have been some recent research efforts to identify alternatives.[1] ZnSe, a wide gap II–VI semiconductor, widely studied for application in photonic devices, is a promising candidate. ZnSe has a higher band gap (2.7 eV) than CdS (2.4 eV). Its use would enhance the blue response and open circuit voltage of the solar cell. In addition, it would eliminate the environmental concerns regarding large scale use of Cd. Polycrystalline ZnSe films have been grown by various vacuum evaporation procedures,[2,3] sputtering in hydrogen selenide,[4] metalorganic chemical vapor deposition (MOCVD),[5] and the two stage process.[6] Difficulties in the control of film stoichiometry and defects lead to carrier compensation and the consequent high resistivity of the ZnSe films has restricted their application. Numerous investigations are directed towards forming films with higher conductivities. Incorporation of nitrogen dopant using a plasma source has been extensively employed for epitaxial ZnSe layers grown by molecular beam epitaxy (MBE).[7] N_2 in ionic form has been incorporated during growth by microwave, electron cyclotron resonance (ECR), and rf plasma excitation. Compensating capabilities of the dopant are not critically dependent on the plasma characteristics.

In this article, we report preparation of polycrystalline ZnSe films by a plasma enhanced two stage selenization of metallic Zn precursor films. Selenization at reduced pressures permits easy ionization of the reactant species and the N_2 atoms. This modifies the reaction kinetics and facilitates incorporation of nitrogen dopant. X-ray diffraction (XRD) investigation of the structure, scanning electron microscopy (SEM) analysis of the microstructure, optical band gap measurements, and electrical properties of ZnSe films grown by plasma enhanced selenization are reported here. X-ray photoelectron spectroscopy (XPS) measurements have been employed to indirectly support nitrogen incorporation in the films.

II. EXPERIMENT

Zn precursor films used in this study were deposited on cleaned glass substrates by the modified thermal evaporation process in which initially a thin ≈80 Å layer of In is deposited which enhances the adhesion of Zn atoms and provides a compact microstructure for the precursor.[8] The thickness of the precursor films is monitored by a quartz crystal thickness monitor (Maxtek TM200). These measurements are correlated with the thicknesses of the selenized films determined by Tally step (Rank Taylor Hobson Model 2223). Excellent correlation between the two measurements was observed. Substrates were maintained at room temperature during the deposition. The substrates are then transferred into the selenization reactor details of which were published earlier.[9] The Se flux is generated by heating Se to 240–260 °C and carried by a flow of N_2 to the Zn precursor film held at the preselected reaction temperature of 400 °C in a separate zone for selenization. In the present work, Se vapor and N_2 gas are ionized by creating a plasma in the region of the Se source and substrate. For this, a metal electrode was provided in the reactor assembly to couple the output of a 10 MHz spark coil oscillator. The pressure in the reactor was varied between 0.5 and 5 mbar for these studies. At these pressures, confinement of the plasma to the reaction zone and the maintenance of reasonably high Se flux even with modest temperatures of the Se source can be ensured.

III. RESULTS AND DISCUSSION

The results of x-ray diffraction studies (performed using a Siemens D-500 x-ray powder diffractometer) of the yellow

[a]Electronic mail: stlkumar@csnpl.ren.nic.in

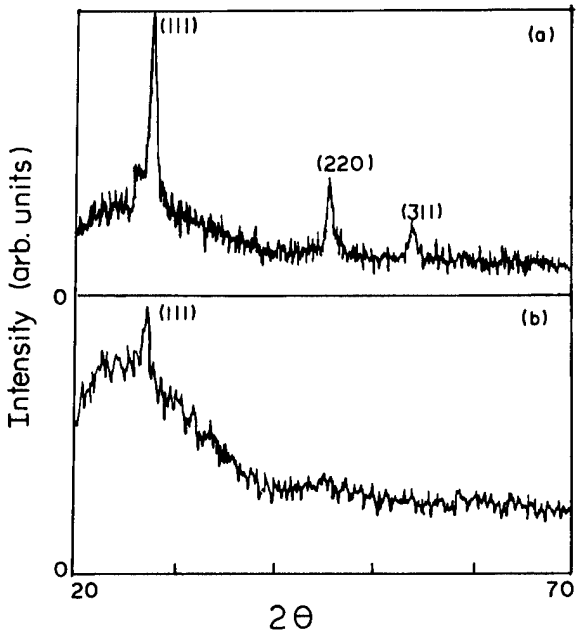


FIG. 1. XRD of ZnSe samples formed by selenization (a) at 400 °C for 15 min with plasma and (b) at 400 °C for 15 min without plasma assistance.

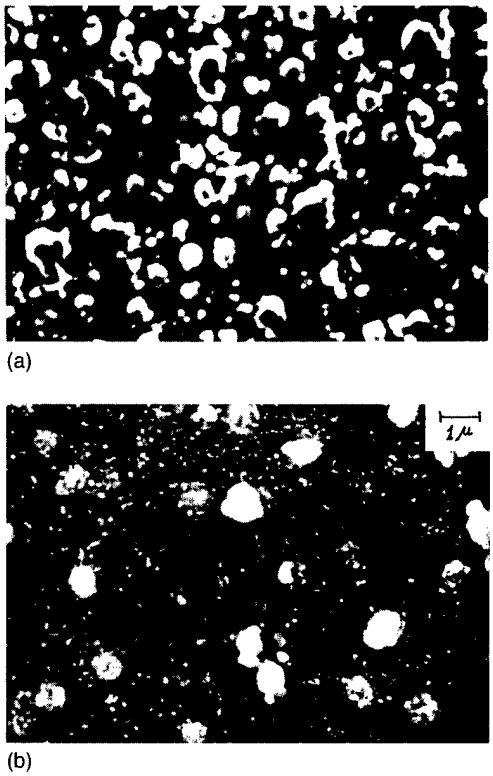


FIG. 2. SEM micrographs of ZnSe films formed by selenization at 400 °C for 10 min (a) with and (b) without plasma assistance.

transparent polycrystalline ZnSe films formed by selenization with and without plasma, under identical substrate temperature of 400 °C and reaction time of 10 min are shown in Figs. 1(a) and 1(b), respectively. For films of this overall thickness, typically less than 1000 Å, only the major reflection from (111) planes of the cubic phase have been identified for the samples selenized without plasma. Plasma selenized films have a highly crystalline structure as seen from the increase in the intensity of the reflections. The integrated intensity of the (111) reflection, determined by integrating the spectrum increases by a factor of 20 as compared to the film selenized without plasma. The three main diffraction peaks corresponding to the (111), (220), and (311) planes have been identified in Fig. 1(a). The relative intensities of these reflections do not indicate any texture in the films. The absence of any of the strong hexagonal reflection, from among (100), (101), (102), and (103) planes confirms the formation of single phase ZnSe films. When much thicker ZnSe films (>5000 Å) are formed, selenization reaction proceeds to completion in more than 30 min. In the thick films, intensity of the XRD peaks increases and the important diffraction peaks are identified.[6] The influence of the plasma enhancement is perceptible for the thinner films (<1000 Å) and short reaction times of 10 min or lower. Traces of the reflections from the unreacted precursor are observed when the reaction time is not sufficient. This provides an indirect means of assessing the reaction kinetics. In the case of plasma enhanced selenization smaller reaction times are sufficient, indicating an enhancement of the kinetics. The increase in the intensity of the XRD reflections, which indicates enhanced crystallinity of the films is also more easily detected for thinner films (<1000 Å).

SEM micrographs of ZnSe films formed by selenization with and without plasma in Fig. 2 show significant differences in the microstructure. Plasma assists the growth of increased concentration of large ZnSe crystallites [Fig. 2(a)] whereas those formed without plasma [Fig. 2(b)] show a few nonuniform crystallites in a microstructure dominated by crystallites of average size <0.1 μm. A similar microstructure prevailed in thicker films. The basic microstructural features also do not change when the reactor pressure is changed from 0.5 to 5 mbar which alters the reactant availability.

Optical absorption of the ZnSe films was measured in the spectral range 200–2500 nm. The optical band gap was determined by plotting the relevant portion of the data as $(\alpha h\nu)^2$ vs $(h\nu)$, where α is the absorption coefficient. The band gap values for ZnSe films formed with and without plasma assistance are the same. The direct band gap of 2.65 $\pm$0.08 eV (Fig. 3) determined from this study is in agreement with the published value for cubic ZnSe. The optical studies were also performed on ZnSe films formed at different reactor pressures between 0.5 and 5 mbar but identical results were obtained.

Electrical resistivities of as-formed samples were measured by four probe technique using gold contacts and film thickness measured by optical and stylus techniques. The resistivity of ZnSe films is considerably effected when formed by plasma assisted selenization. Figure 4 shows the changes in the ZnSe resistivity as a function of the reactor pressure used during selenization with and without plasma assistance. Resistivities of the as-formed ZnSe films, grown without plasma are typically 10^5 Ω cm which is comparable

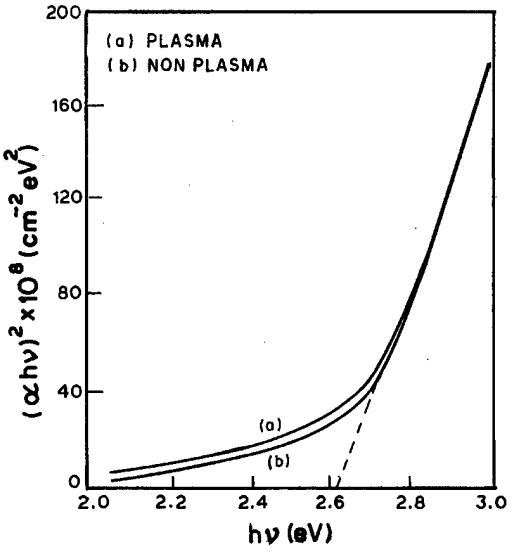


FIG. 3. Variation of $(\alpha h)^2$ with the photon energy for ZnSe film grown by selenization with and without plasma assistance. The extrapolation of the data gives an optical band gap of 2.6±0.05 eV in both cases.

to polycrystalline ZnSe films deposited by MOCVD and thermal deposition techniques.[2,3,5] An important result is that the resistivities of all samples selenized with plasma assistance are lower than the corresponding samples made without plasma. The reduction in the resistivity follows the lowering of reactor pressure used in the selenizing in the presence of N_2 plasma. At high reactor pressure, ≈5 mbar, decrease in resistivity is 5% but it decreases by a factor of 2.6 at a lower pressure of 2.5 mbar (Fig. 4). A remarkable and abrupt lowering of resistivity is observed at ≈1 mbar, the lowest reactor pressure studied. Here, a decrease of more

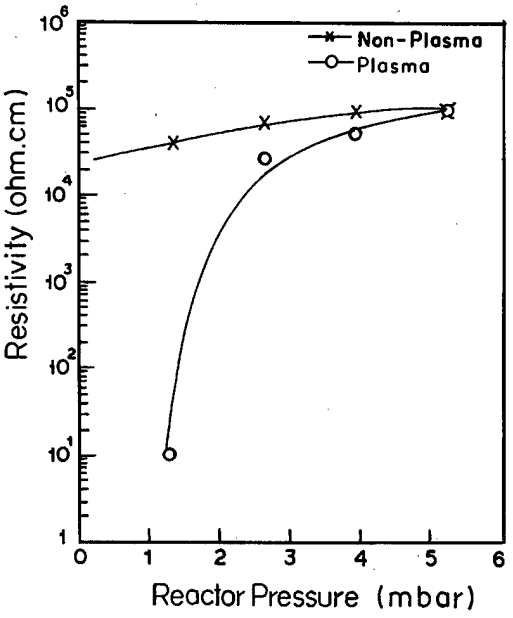


FIG. 4. Variation of the ZnSe film resistivity with the reactor pressure formed with and without plasma assistance.

than four orders of magnitude is recorded. The ZnSe film formed under these conditions have a resistivity of 10 Ω cm.

XPS analysis of the low resistivity film was performed using a 1257 PHI Perkin Elmer multitechnique system and the spectrum compared with ZnSe film formed similarly but without plasma.[6] Mg $K\alpha$ radiation with a photon energy of 1253.6 eV was used. The energy resolution was ~0.2 eV during the high resolution analysis. Surface cleaning was performed using a 1 keV Ar^+ ion gun for 15–30 s. In several samples, it was confirmed by sputter profiling that the surface and bulk compositions are identical. The peak intensities indicate a stoichiometry of $Zn_{(0.48-0.52)}Se_{(0.52-0.55)}$ indicating a marginal Se excess. General features of both the spectra are similar. The selenium $3d$ XPS line is located at 53.7 eV for the sample selenized with plasma assistance and 54.4 eV for the sample selenized without plasma in comparison to 55.3 eV for pure Se showing a chemical of 1.6 and 0.9 eV, respectively.[10] The Zn $2p_{3/2}$ line is located at 1021.8 eV for the nonplasma sample and at 1021.5 eV for the plasma sample in comparison to 1021.5 eV for pure Zn. Since the XPS line of Zn is very close to the photon energy, the Auger parameter has been calculated. The Auger parameter for Zn, calculated from the experimental binding energies of the Zn $(2p_{3/2})$ XPS line and the *LMM* Auger line is 2009.65 in comparison to the 2011.1 eV for the nonplasma sample and 2014 for pure Zn. Standard value reported for ZnSe are 2010.2–2011.5.[11] These results indicate that electronic charge transfer of the plasma selenized and nonplasma selenized sample are different. However, no physical mechanism that could justify such changes due to doping or long range order are known. The line widths (full width at half maximum) of the Zn and Se XPS lines are 3.3 and 2.5 eV for plasma grown ZnSe films, while these values are 2.0 and 2.3 eV for the nonplasma sample which are comparable with the values of 1.5 and 1.8 eV for pure elements.[10] Peak shape analysis of the high resolution scans indicated single peaks suggesting that neither Zn nor Se atoms exist in multiple valance states. This further supports the XRD evidence regarding absence of minor phases.

At the selenization temperatures employed in this study, the longer molecular fractions, Se_5,Se_6,Se_7 are abundant in the gas phase. It has been experimentally observed that interaction of the Se_{5-7} molecular segments with the rf radiation induces ionization and dissociation of these longer molecular fragments into smaller molecular fragments which are more reactive.[12] Thus, it is expected that the consequent modification of the selenization kinetics is responsible for the increased crystallinity of the films which has been observed in the XRD and SEM analysis. It is known that excited nitrogen species act as dopants during epitaxial growth of ZnSe.[7] Presence of ionized nitrogen during plasma selenization would provide similar opportunities for incorporation of nitrogen ions in the growing ZnSe film. Incorporation of ionized nitrogen as a dopant in the ZnSe film could account for the observed lowering of resistivity. It seems that at higher reactor pressures, a larger availability of Se flux enables the reaction to proceed at a much faster rate and reduce

the probability of nitrogen incorporation. Detailed gas dynamics calculations of the reactor[13] shows that the reactor pressure determines the availability of Se reacting species at the precursor. The net availability of 10^{13-14}/cm/s at 1 mbar, which translates to a deposition rate of $\approx$0.1 Å/s in a corresponding deposition system. This would be the minimum reactant requirement for a reaction front to advance into the precursor material at a reasonable rate. The lowering of availability of the Se to impractical levels at lower reactor pressures suggests that there is a narrow window of opportunity for the lowering of resistivities of ZnSe films by this method. Further optimization of the polycrystalline film deposition process would be required to lower the ultimate resistivities for actual device application.

IV. CONCLUSIONS

A novel plasma enhanced selenization process for the preparation of polycrystalline ZnSe films of low resistivity and high crystallinity has been demonstrated. Plasma assisted growth at low reactor pressure of $\approx$1 mbar causes significant (>3 orders of magnitude) reduction of the resistivity to 10 Ω cm. XPS analysis shows a chemical shift of 1.6 eV for the Se 3*d* line for the plasma selenized samples is 1.6 eV rather than 0.9 eV for the nonplasma sample. ZnSe films exhibit an optical band gap of 2.65±0.08 eV which is invariant with selenization pressure and plasma condition. The two stage selenization process for ZnSe films described here is quite compatible with the growth process for $CuInSe_2$ (CIS) absorber layer. This represents considerable process simplification as the growth of low resistivity ZnSe by the present technique lends itself to a single step processing of window (ZnSe) and absorber (CIS) layers for the fabrication of polycrystalline thin film CIS solar cells in a superstrate configuration.

ACKNOWLEDGMENTS

The authors thank the Director, National Physical Laboratory, for permission to publish this work. Thanks are due to the x-ray group for use of the XRD facilities. One of the authors (R.J.) acknowledges the Council of Scientific and Industrial Research, New Delhi for a research fellowship.

[1]M. Konagai, Y. Ohtake, and T. Okamoto, Mater. Res. Soc. Symp. Proc. **426**, 153 (1996).
[2]J. B. Yoo, A. L. Fahrenbruch, and R. H. Bube, Sol. Cells **31**, 171 (1991).
[3]J. M. Powlikowski, Thin Solid Films **127**, 9 (1985).
[4]A. Nouhi and R. J. Stirn, Sol. Cells **21**, 225 (1987).
[5]T. L. Chu, S. S. Chu, G. Chen, J. Britt, C. Ferekides, and C. Q. Wu, J. Appl. Phys. **71**, 3865 (1992).
[6]S. T. Lakshmikumar and A. C. Rastogi, Thin Solid Films **259**, 150 (1995).
[7]T. Yao, Z. Zhu, Y. H. Yu, C. D. Song, F. Nishiyama, K. Kimura, H. Kajiyama, S. Miwa, and T. Yasuda, J. Cryst. Growth **159**, 214 (1996).
[8]R. Jeyakumar, S. T. Lakshmikumar, A. C. Rastogi, and G. K. Chadda, *Physics of Semiconducting Devices*, edited by K. Lal (Narosa, New Delhi, 1996), p. 569.
[9]S. T. Lakshmikumar and A. C. Rastogi, J. Appl. Phys. **79**, 3585 (1996).
[10]*Handbook of X-Ray Photoelectron Spectroscopy*, edited by G. E. Mullenberg (Perkin Elmer, Eden Prairie, MN, 1979).
[11]D. Briggs and M. P. Seah, *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (Wiley, New York, 1983), pp. 125 and 497.
[12]J. Berkowitz and W. A. Chupka, J. Chem. Phys. **45**, 4289 (1966).
[13]S. T. Lakshmikumar, J. Vac. Sci. Technol. A **14**, 2968 (1996).

Arrays of ungated GaAs field emitters fabricated by wet or dry etching

F. Ducroquet,[a] P. Kropfeld, O. Yaradou, and A. Vanoverschelde
Institut d'Electronique et de Microélectronique du Nord (IEMN-DHS—UMR CNRS 9929), University of Lille-1, Cité Scientifique, B.P. 69, 59652 Villeneuve d'Ascq cedex, France

(Received 3 March 1999; accepted 28 May 1999)

Dry and wet etching methods are used to fabricate arrays of ungated GaAs field emitters. Comparisons were made among tips, wedges, pillars and walls. Effects of the etching technique and subsequent deoxidation procedure on the GaAs surface are evaluated in terms of emitter geometry, uniformity, heating and integration. Moreover, the current emission capability of field emitter arrays is investigated. In particular, the influence of a thin surface oxide layer on the emission stability is observed. An extraction of geometrical parameters is attempted on the basis of a statistical examination of the emission characteristics and the validity of this approach is discussed. © *1999 American Vacuum Society.* [S0734-211X(99)09604-3]

I. INTRODUCTION

GaAs field emitter arrays (FEAs) are among the emerging candidates for the next generation of high power field emission sources for microwave applications.[1] Indeed, emission capabilities of GaAs are expected to be slightly better than Si (Ref. 2) with the additional advantage of favorable thermal dissipation.[3] In order to satisfy typical rf amplifier requirements (high total array current, large current density per tip, high transconductance and low array capacitance), various shapes of emitters have been proposed (cone, pyramidal,[4] tip-on-post,[5]...). Wedge structures have also been considered for high power emitters, since Joule heating is expected to be an order of magnitude lower than for tips, making higher packing densities possible.[3]

This article describes some aspects of the fabrication of GaAs FEAs. Both chemical etching and reactive ionic etching (RIE) processes have been developed. Numerous studies have been reported in the literature describing the etching of GaAs;[6,7] a wide variety of techniques have been used, resulting in various profiles, etching rates, morphologies and uniformities. The techniques we have studied were selected based on criteria such as high isotropy, moderate etching rate, sharp profile and high aspect ratio. Their evaluation is based not only on the resulting geometry of the structures but also on the nature of the surface after etching and its evolution after a further treatment. Indeed, the surface properties before emission are found to greatly influence the emission capabilities. The potentialities in terms of microemitter integration are discussed for both wet and dry etching processes. Finally, emission characteristics of various arrays of GaAs emitters are presented.

II. FABRICATION

A. Wet etching process

In a first approach, the emitter arrays were fabricated using a simple technology of wet etching. This method is expected to provide sharp structures. It is based on a one-step process of chemical underetching beneath circular (4 μm in diameter), square ($2\times2\ \mu m^2$) or rectangular ($2\times100\ \mu m^2$) masks patterned on a highly doped (1 0 0) GaAs substrate. Silicon nitride masks are commonly used except for HF based mixtures for which resist is employed. The edges of the masks are in most cases aligned along the (1 1 0) and ($\bar{1}$ 1 0) cleavage planes. The etching is gradually controlled by optical microscopy and precisely stopped when the contact limit between the mask and GaAs is reached, in order to achieve the sharpest shape. Various shapes of emitters have thus been obtained (Fig. 1). The features of the different investigated etching solutions are summarized in Table I. The geometry of the structures are evaluated from several parameters: the aspect ratio, the curvature radius and the verticality angle at the apex. The ratio of mask size to emitter height is given as an estimation of the integration potentialities.

At first, considering the system HF:HNO_3:H_2O, roundish profiles for both natural cleavage planes on GaAs could be expected from a 1HF:1HNO_3:1H_2O mixture.[7] However, the etching speed is too fast to allow an accurate control in the fabrication of 1 μm height structures, and was slowed down by increasing the H_2O concentration. In this way, a 1HF:1HNO_3:2H_2O solution was found to provide pyramidal tips with good geometry features[8] [Fig. 1(a)], showing also good uniformity over the array and interesting emission characteristics. For further dilutions, the etching rate decreases down to rapidly become zero. The oxidation of GaAs by HNO_3 appears then as a limitation for this system because the resulting formation of As_2O_3 produces a very rough surface after etching, and can even block the reaction when the HF concentration is too low to immediately remove the formed oxide.[9]

The next investigated system is the previous HF based system where the oxidizing agent HNO_3 is replaced by the hydrogen peroxide for which the oxidation of GaAs is known to form a thin and stable oxide of GaAs.[10] A composition of (1:10:21.2) is chosen for the HF:H_2O_2:H_2O solution by reference to profile studies already reported for this system on GaAs.[11] This etching appears highly nonisotropic for

[a]Electronic mail: ducroquet@insa-lyon.fr

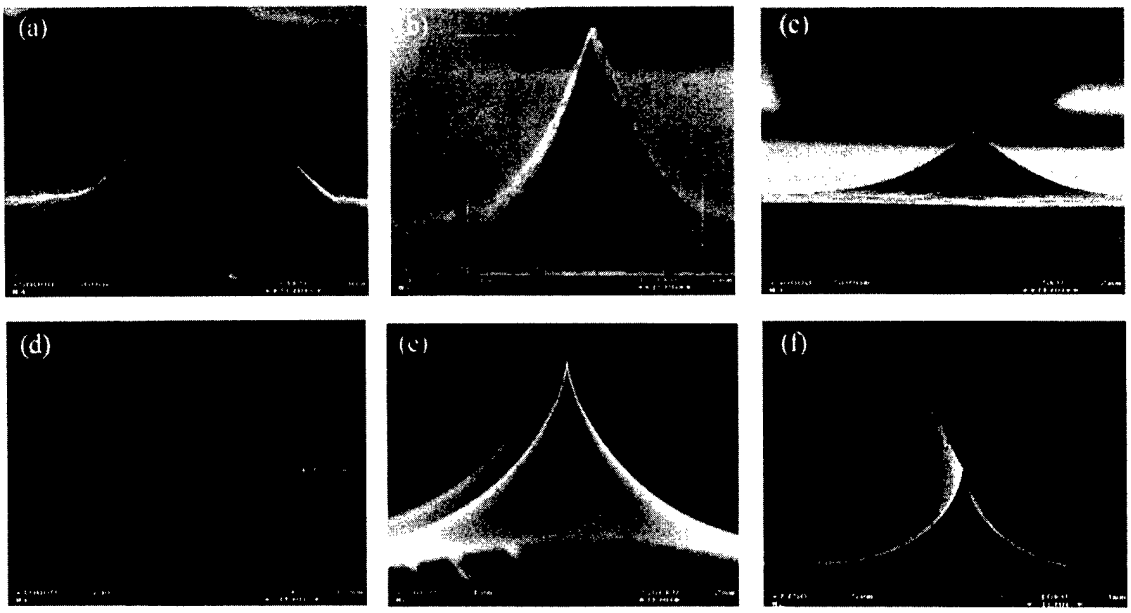


FIG. 1. Various profiles of emitters shaped by wet chemically etch: (a) $1HF:1HNO_3:2H_2O$; (b) $1HF:10H_2O_2:21.2H_2O$; (c) $1HF:20H_2O_2:100H_2O$; (d) $1NH_4OH:1H_2O_2:8H_2O$; (e) and (f) $40HCl:4H_2O_2:1H_2O$.

both directions of cleavage. Microwedges with a sharp "pyramidal" profile were obtained from an underetch of square masks [Fig. 1(b)]. The solution was cooled down to reduce the etching rate and to improve the aspect ratio and sharpness of the structures. The relatively high aspect ratio of the structures, the good verticality at the apex and especially the very small radius of curvature (5–6 nm) are the main advantages in using this mixture. On the other hand, the moderate emitter height compared to the initial mask size limits the possibility of integration. No further improvements were achieved, neither by increasing the water concentration which leads to a degradation of the aspect ratio, nor by increasing the hydrogen peroxide concentration which leads to the formation of a bottleneck at low etch temperatures (although a gain on the aspect ratio was noted). Taking into account the etch anisotropy, tip shaped structures were formed by orienting the mask edges at 45° to the natural cleavage planes [Fig. 1(c)]. In this case, the best results were obtained with the solution: $1HF:20H_2O_2:100H_2O$. The tip height and aspect ratio remain nevertheless modest despite a very good radius of curvature.

A similar pattern orientation at 45° was used to demonstrate the capacity to chemically etch GaAs to nearly vertical sidewalls with a $1NH_4OH:1H_2O_2:8H_2O$ solution.[12] Using this highly anisotropic etching through rectangular masks, wall shaped structures combining an excellent verticality with a very high aspect ratio have been obtained [Fig. 1(d)]. Due to a relatively high etch rate, the emitter uniformity over

TABLE I. Summary of the geometrical emitter parameters for different etchings (h: emitter height, r: radius of curvature, γ: aspect ratio, α: verticality angle at the apex, a: size of mask edge).

Etching solution	Mixture composition	Mask shape	Emitter shape	Etch time [h (μm)]	r (nm)	γ^a	α^b	a/h	Fig.
$HF:HNO_3:H_2O$	1:1:2	square	pyramidal tip	1′ [0.93]	30–40	1.07	43°	2.15	1(a)
$HF:H_2O_2:H_2O$	1:10:21.2[c]	square	micro-wedge	30″ [0.77]	6–8	1.8	29°	2.6	1(b)
$HF:H_2O_2:H_2O$	1:20:100	45° square[d]	pyramidal tip	35″ [0.47]	5–7	1	45°	4.25	1(c)
$NH_4OH:H_2O_2:H_2O$	1:1:2	45° rectangular[d]	wall	2′ [4.6]	–	12–15	~2°	2.1	1(d)
$HCl:H_2O_2:H_2O$	40:4:1	square, circular	sharp tip	2′–3′ [1][e]	10–30	1	weak[f]	2.1	1(e)
$HCl:H_2O_2:H_2O$	40:4:1	rectangular	sharp wedge	3′–4′ [1.1][e]	10–30	1.1	weak[f]	1.8	1(f)
$SiCl_4$(RIE)	···	circular dot	pillar	1h [1]	~20[g]		weak	0.05	2(a)
$SiCl_4$(RIE)	···	rectangular line	wall	1h [1]	~20[g]	12	4°	0.05	2(b)

[a] γ = height/(base/2).
[b] α = arctan[(base width-apex width)/2×height)].
[c] T_{sol} = 11.6 °C.
[d] Mask edges oriented at 45° to the natural cleavage planes.
[e] For a solution stabilized at room temperature and next heated at 40 °C (with and without stirring).
[f] Difficult to evaluate because of the strong verticality of the rounded forms at the apex.
[g] Corresponds to an equivalent radius equal to the width at the apex/2.

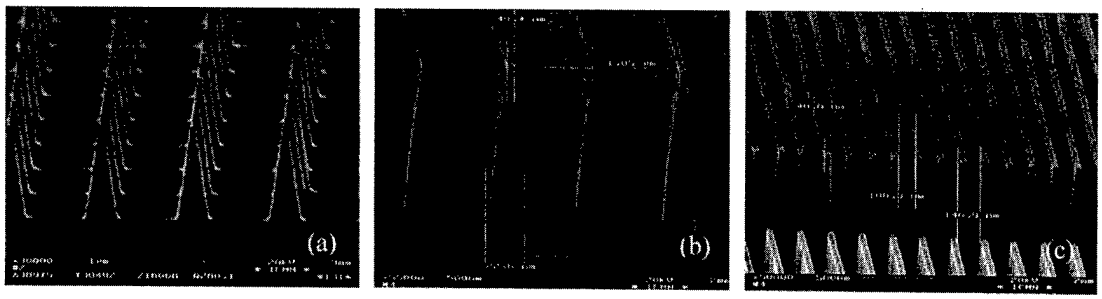


FIG. 2. Reactive ionic etching of GaAs emitters: (a) array of pillars 1 μm apart, (b) array of walls 1 μm apart, (c) high density array (wall to wall spacing: 150 nm).

the array is yet difficult to ensure. Moreover, the reduced bases of these structures may very well limit the heat dissipation at high emission levels, as was noticed during emissions testing. In any case, at moderate emission, such wall structures form an interesting geometry which could, for example, be integrated in a "knife on post" process.

Contrary to both previous systems, the $40HCl:4H_2O_2:1H_2O$ mixture presents very highly isotropic features for the etching of GaAs (Ref. 13) and provides roundish profiles in both cleavage directions. Wedge and tip emitters with a very sharp profile, an aspect ratio slightly higher to unity and a radius of curvature in the range of 10–20 nm have been typically fabricated using this etching mixture[14] [Figs. 1(e) and 1(f)]. However, the HCl-based solution is exothermic and subject to bubbling. In particular, when the etch is carried out a short while after initially mixing the constituents (under these conditions, the etch rate is ~2 μm/min and the solution temperature ~60 °C), etch pits are observed at large scale on the surface of the sample. These etch pits can be greatly reduced by letting the chemical solution settle at room temperature. But the significant reduction of the etching speed in this case is also found to favor the growth of an oxide layer on the surface leading to a detrimental effect on the emission. An optimized procedure limiting both etch pits and surface oxide is obtained by heating the solution and regulating it around 40 °C. These different operating conditions do not affect the profiles of the emitter structures and provide a good uniformity over the array.

We conclude that these wet chemical etching methods are able to shape a large variety of emitter geometries. Nevertheless, as shown in Table I, the required space for the mask patterns inevitably constitutes a severe limitation for the fabrication of very high density emitter arrays.

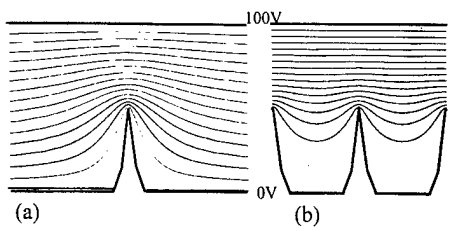


FIG. 3. 2D simulation of equipotential lines: (a) single emitter, (b) infinite number of emitters.

B. Dry etching process

In contrast to wet underetching, the RIE process is able to provide both a high resolution transfer of deep submicron masks and highly vertical profiles in view of enlarging microstructure integration. Electronic beam lithography and a lift-off processes were used to pattern 40 nm×10 μm metal lines and 40-nm-diam circular dots. For the lift-off, a high sensitivity bilayer resist system, composed of a bottom layer of 4% copolymer (80 nm) and a top layer of 3% 950 K polymethyl(methacrylate) (50 nm) has been applied, followed by a deposition of 30–35 nm thick metal (Ti–Au). The etching of GaAs was performed at 5 mTorr with a $SiCl_4$ plasma under an Ar flow to produce smooth surfaces and high vertical sidewalls. 1 μm high pillars [Fig. 2(a)] and walls [Fig. 2(b)] with less than 4° sidewall angle, high aspect ratio and 35–40 nm width at the apex were fabricated. An additional sharpening process can be performed on these structures using gentle wet chemical etching. A digital etching procedure[10] (based on cycles of successive steps of oxidation and formed oxide etching) was found to be well adapted when a highly accurate control is required (~1.5 nm per cycle). Interesting profile improvement was also observed using a $3H_3PO_4:1H_2O_2:50H_2O$ mixture. In this case, an etch speed of 35 nm/min is obtained at 3.2 °C on a beforehand HF deoxidized structure.

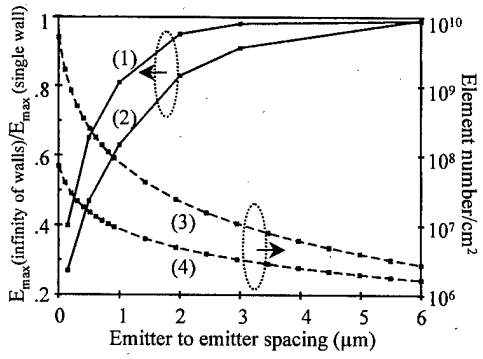


FIG. 4. 2D simulation of the electric field at the apex as a function of the emitter to emitter spacing. Case of 1 μm height wall structure (cf. Fig. 3) for two anode–cathode distances: (1) 2 μm; (2) 6 μm. Equivalent elements for an array of 1 cm^2 area: (3) 40 nm diam pillars, (4) walls of 40 nm width at the apex.

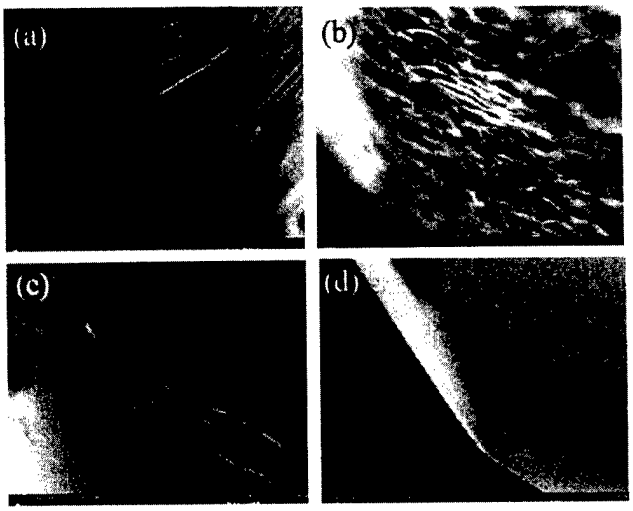


FIG. 5. Effect of a deoxidation procedure on the roughness of the GaAs surface preliminary etched by a $40HCl:4H_2O_2:1H_2O$ solution at room temperature: (a) as etched (reference), (b) +BOE, (c) +HF, (d) $+1NH_4OH:1H_2O$. In order to intentionally amplify effects in order to allow observation by SEM, the samples were dipped 30 min in the deoxidation solution (instead of 30 s to few minutes for the FEAs) and then annealed at 300 °C for 1 h.

The RIE process yields uniform and reproducible results with high packing densities. A reduction of the line to line spacing down to 150 nm was experimentally achieved [Fig. 2(c)], leading to a potential integration of 6×10^7 walls (or 6×10^9 pillars) per cm^2. However, as a consequence of high integration, the resulting proximity effects are expected to induce a lowering of the electric field at the apex, as illustrated in Fig. 3. These effects were assessed by simulating the evolution of the electric field as a function of the emitter to emitter distance. Simulations were made with a commercial software[15] using a finite element method applied to any geometrical structure. The electrostatic potential and electric field were computed at each node between the emitters and an anode by solving the Laplace equation (only the case of metal emitters was considered). The boundary conditions

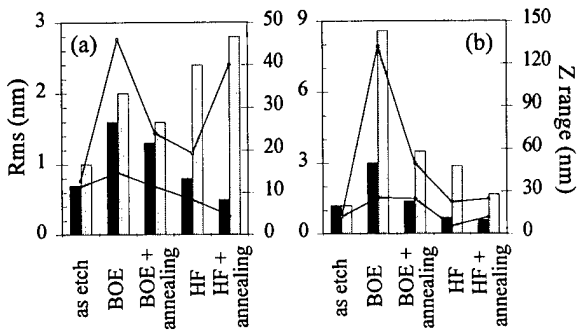


FIG. 6. Influence of the deoxidation treatment on the surface roughness after etching GaAs using a $40HCl:4H_2O_2:1H_2O$ mixture in two operating conditions: (a) etch without letting the solution settle beforehand, (b) etch at room temperature after the solution was settled. Histograms correspond to the rms value of the roughness measured by atomic force microscopy full line curves to the Z range value. The dark (light) color refers to a sample area of 1 μm^2 (25 μm^2).

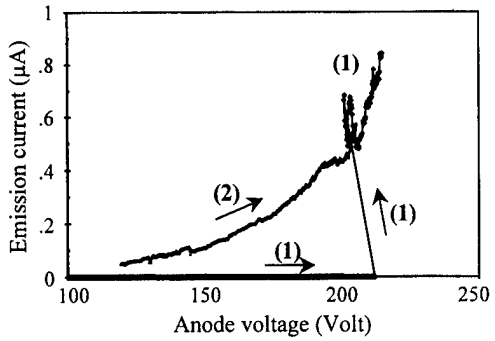


FIG. 7. Evidence of the "oxide breakdown" phenomenon: (1): first voltage sweep, (2): further voltage sweeps.

were set to have the conductor anode at voltage V and the emitter grounded. Simulation results (Fig. 4) show that the decrease of the electric field still remains relatively moderate (~50%–60% when the wall to wall spacing is reduced from 1 μm to 150 nm) compared to the potential gain on the emission current. Indeed, considering a constant electric field, this gain is expected to increase in the same proportion as the emitter number.

III. EXPERIMENT

The emission characteristics of FEAs are carried out in a vacuum chamber at a pressure of 3×10^{-8} Torr. The ungated devices are tested in a diode configuration in which the conducting anode is kept at a 6 μm distance above the cathode by an insulating film. In order to enhance the electrical emission, the samples are deoxidized just before being introduced into the vacuum chamber. As previously reported,[8] for structures etched using a $HF:HNO_3:H_2O$ mixture, As_2O_3 surface oxide is partially removed and smoothness and crystallinity improved by dipping the sample in a buffered oxide etchant (BOE) solution for a few seconds, followed by an annealing at 300 °C for 1 h. In contrast, a similar process applied to structures etched with a $HCl:H_2O_2:H_2O$ solution leads to a marked increase of the surface roughness [Fig. 5(b)]. This effect is noted whatever the etching operating conditions used and increases with the initial oxide thickness (Fig. 6). Therefore, in this case, a deoxidation procedure using a pure or dilute HF solution or a $1NH_4OH:1H_2O$ mixture, which are both found to improve the surface conditions, is preferred [Figs. 5(c) and 5(d)]. In particular, an enhancement of the emission is observed even on samples with a thick initial oxide layer after being submersed in $1NH_4OH:1H_2O$ for a few minutes. However, this process needs to be performed with some care, since it can create areas of instability where the surface is no longer passivated by a thin oxide layer during electron emission.

The current–voltage ($I-V$) characteristics are recorded by computer. The anode voltage is typically incremented in steps of 0.2 V. At each step, the emission current is measured with a frequency of 0.5 Hz for 1 min using an electrometer (Keithley 617) and the results numerically averaged. In order to stabilize the emission, the bias voltage is progressively

increased until the current reaches a specific limit. The voltage sweep is then repeated several times prior to going to the next current limit. During these cycles of measurement, a shift of the curves with the voltage is observed. The direction and amplitude of the shift appear strongly dependent on the surface conditions of the sample. In particular, when the oxide layer is expected to be greatly reduced or quasicompletely eliminated, it seems that the curves preferentially shift to the lower voltages, in agreement with previously reported observations.[4] This is, for example, the case for samples etched using $HCl:H_2O_2:H_2O$ (without stabilization), $NH_4OH:H_2O_2:H_2O$ or using a RIE process.

In contrast, a slighter shift to the high voltages is generally noticed on samples etched by $HF:HNO_3:H_2O$ or $HCl:H_2O_2:H_2O$ (process with stabilization and heating of the solution at 40 °C), for which the surface oxide layer is assumed to be thicker. Moreover, in these cases, during the first voltage sweep, instead of a progressive increase of the current as observed on the former samples, the emission turn-on occurs very suddenly (Fig. 7). This effect is not reversible and could be related to an "oxide breakdown" at the apex of the emitter. The further voltage sweeps show correct emission and the stabilization of the characteristics is generally rapidly achieved. The temporal evolution of the current at a fixed voltage is found to be very stable with fluctuations reduced to a few percent over several hours.

IV. FIELD EMISSION CHARACTERISTICS

A. Single emitter

The parameter extraction from the experimental $I-V$ curves was carried out using the planar Fowler–Nordheim (FN) formalism[16]

$$I(V)=A\times S\times(\beta V)^2\times\exp\left(-\frac{B}{\beta V}\right) \tag{1}$$

with

$$A=\frac{1.4\times10^{-6}}{\phi}\times\exp\left(\frac{9.81}{\sqrt{\phi}}\right),$$

$B=6.49\times10^9\times\phi^{3/2}$, S is the total emitting surface ($[S]=m^2$), β is the field enhancement factor ($E=\beta\times V$, $[\beta]=m^{-1}$) and ϕ is the emitter work function (ϕ is set to 4.07 eV). The main limitation of this formulation lies in the fact that a constant electric field along the emitter surface is assumed. For example, we can show that a linear regression performed on simulated data taking into account the variation of the electric field along a conical structure is found to lead to an emission area underestimate by around 300% and to an extracted average field enhancement factor β below its maximum value at the apex by 5%–7%. An improved parameter extraction procedure[17] was to consider the integration of the current density over the emitter surface using the ideal metal emitter floating sphere model for which the field distribution on the surface of the sphere varies as a function of the polar angle θ:[18]

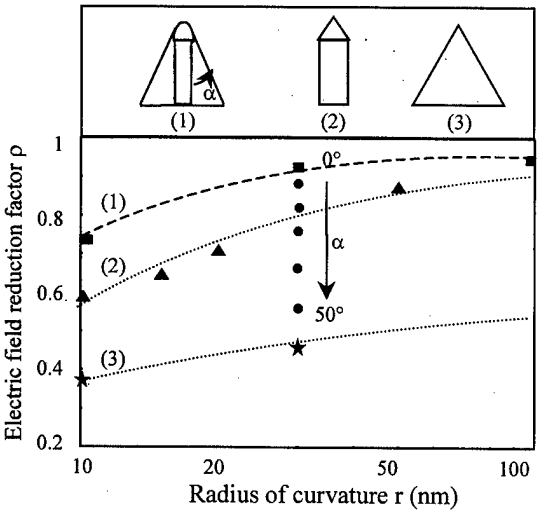


FIG. 8. Evaluation of the electric field reduction factor ρ as a function of the radius of curvature for several emitter geometry, the dashed lines are guides for eyes: (1) rounded whisker, (2) tip-on-post, (3) pyramid. At $r=30$ nm, the arrow indicates the evolution of ρ with the angle of a conical structure (from $\alpha=0$ (rounded whisker) to $\alpha=50°$ by step of 10°).

$$E(\theta)=\frac{V}{d}\times\left[\frac{h_r}{r}+3\times\cos(\theta)\right], \tag{2}$$

where r is the radius of curvature at the apex, d is the distance anode–cathode, h is the emitter height and $h_r=h-r$. However, the resulting expression does not systematically lead to a linear FN $I(V)$ plot and then nonlinear minimization techniques are needed to extract parameters. Indeed, a linear FN $I(V)$ plot is only obtained when E is taken equal to its maximal value. Moreover, it was shown that in any case, an accurate knowledge of the electric field at the apex of the emitter is required beforehand in order to achieve to a physical explanation. In particular, a reduction factor ρ needs to be introduced in the expression of the maximum electric field enhancement factor β deduced from the floating sphere model

$$\beta_{max}=\rho\times\frac{h_r}{d\times r}. \tag{3}$$

An evaluation of the reduction factor ρ was obtained from electric field simulations for various emitter geometries (Fig.

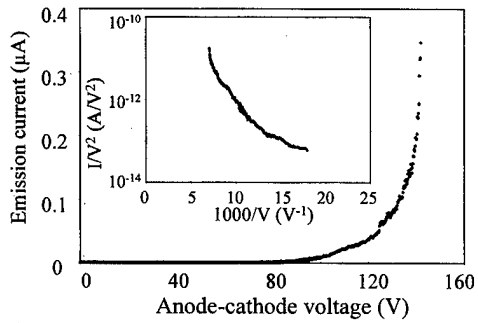


FIG. 9. Emission characteristics of a single GaAs tip. (Inset) The corresponding FN plot.

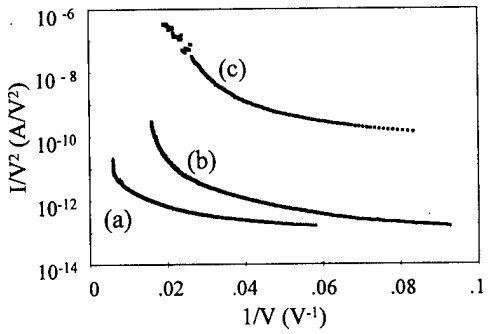


FIG. 10. Typical emission characteristics of GaAs field emitter arrays: (a) wedge array, (b) tip array, (c) wall array.

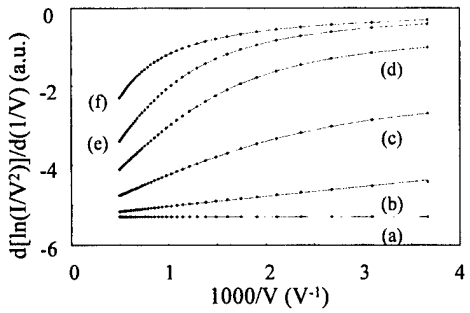


FIG. 11. Simulated data of the derivative of $\ln(I/V^2)=1/V$ considering a Gaussian distribution of the radius of curvature for several standard deviation values σ: (a) 0, (b) 2 nm, (c) 4 nm, (d) 6 nm, (e) 8 nm, (f) 18 nm. The maximum of the distribution is set at $r_0=20$ nm, the lower limit of the integration is $r_{min}=0.1$ nm.

8). Close to unity for blunted or very high aspect ratio structures, ρ is reduced down to 0.35–0.4 for wide angle structures. A similar investigation was conducted for two dimensional (2D) emitters using an approximate floating cylinder model. In this case, the field distribution on the surface can be expressed by

$$E(\theta)=\frac{V}{d}\times\left[\frac{h_r}{r\times\ln\left(\frac{2\times h_r}{r}\right)}+2\times\cos(\theta)\right] \tag{4}$$

and the maximum electric field enhancement factor by

$$\beta_{max}=\rho\times\frac{V}{d}\times\frac{h_r}{r\times\ln\left(\frac{2\times h_r}{r}\right)}. \tag{5}$$

The variations of ρ as a function of the structure angle and radius of curvature follow a similar tendency to those observed for 3D structures, but with a reduced amplitude.

Experimental emission characteristics of a single GaAs tip are shown in Fig. 9. A radius of curvature of 8 nm is extracted from the FN $I(V)$ plot (inset of Fig. 9) using a linear regression. This value is quite in agreement with scanning electron microscope (SEM) measurements. Even if corrected, the emission surface (around 2×10^{-20} m^2) remains much smaller than expected if the whole upper part of the tip is assumed to participate in emission. It could be suggested that a greatly reduced emission area may be generated by the oxide breakdown phenomenon, often observed on these tips. Moreover, this effect should lead to an electron accumulation at the apex and an increase of the local Joule heating,[19] which could explain the premature extinction of the emission which is noticed on these samples. On some damaged samples, SEM observations show that a tip melting occurs at the apex.

B. Field emitter arrays

On the investigated GaAs field emitter arrays, a curvature effect on the stabilized emission current FN plot at low voltages is generally observed (Fig. 10). This effect is attributed to the electric field distribution among the elements of the array and is well predicted using various statistical models.[19–22] In particular, it has been shown that considering a normal distribution of the emitter sharpness, an analytical expression of the FN plot can be derived[20]:

$$\ln\left(\frac{I}{V^2}\right)=A_0-\frac{B_0}{V}+\frac{\sigma^2}{2\times V^2} \tag{6}$$

with $B_0=B/\beta$ and $A_0=\ln(A\times S\times\beta^2)$. Compared to the linear FN expression, it is clearly seen from Eq. (6) that the curvature is caused by the addition of a third term, proportional to the square of the standard deviation σ of the emitter geometry, which represents the excess current introduced by the sharper emission sites. If a good agreement between this model and the experimental data is obtained for low values of the standard deviation (for example, for characteristics obtained during the first steps of the formation procedure), no correct fit could be achieved for most of the "well stabilized" characteristics. The limitation of the model is relative to the inevitable truncature effect of the Gaussian distribu-

TABLE II. Error estimation on the determination of curvature radii compared to nominal value r_0 for different parameter extraction procedures as a function of the degree of truncature of the Gaussian distribution: $\Delta r_0^{poly}/r_0$ and $\Delta r_0^{der}/r_0$ are the relative errors on the average curvature radii determined respectively from a polynomial regression of $\ln(I/V^2)=f(1/V)$ plot [Eq. (6)] and from a linear regression of the derivative [Eq. (7)]. $\Delta r_{min}^{parab}/r_0$ and $\Delta r_0^{parab}/r_0$ are the relative errors on the minimal and average values of the curvature radii determined from the linear regression of $I/V^4=f(1/V)$ plot [Eq. (9)].[a]

r_{min}	$\frac{\Delta r_0^{poly}}{r_0}$ [b]	$\frac{\Delta r_0^{der}}{r_0}$ [b]	$\frac{\Delta r_{min}^{parab}}{r_0}$ [c]	$\frac{\Delta r_0^{parab}}{r_0}$ [c]
$r_0-10\times\sigma$	−1%	−1%	...	...
$r_0-5\times\sigma$	−10%	−2%	...	...
$r_0-3\times\sigma$	−30%	−3%	...	...
$r_0-2.5\times\sigma$	−50%	−10%	<−70%	+45%
$r_0-2\times\sigma$	−60%	−20%	−20%	+20%
$r_0-1.5\times\sigma$	−70%	−30%	−15%	+10%
$r_0-\sigma$	<−80%	−40%	−10%	+10%

[a]Minus sign(plus sign) is arbitrary used in case of an underestimation (overestimation) with respect to the nominal value.
[b]See Fig. 11 for data simulation conditions.
[c]See Fig. 12 for data simulation conditions.

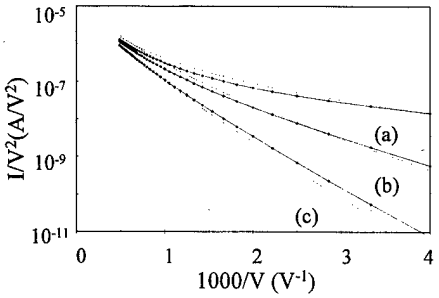


FIG. 12. Comparison of simulated FN plots using a truncated Gaussian distribution (full lines) and a parabolic distribution (dashed lines) model of the radius of curvature for different values of truncature r_{min}: (a) 1 nm, (b) 5 nm, (c) 10 nm. The distribution maximum is set at $r_0=20$ nm. For the Gaussian distribution, the standard deviation is fixed to 8 nm. For the parabolic distribution, the integration limits are $r_1=r_{min}$ and $r_2=2r_0-r_1$.

tion as the standard deviation increases due to the fact that the emitter radii cannot be physically smaller than some threshold value r_{min}. The effect is best illustrated using the linear formulation of the derivative of Eq. (6):

$$\frac{\partial[\ln(I/V^2)]}{\partial(1/V)}=-B_0+\frac{\sigma^2}{V}. \tag{7}$$

Increasing discrepancies to the expected linear behavior are observed as high values of σ are considered (Fig. 11), and therefore as the Gaussian is truncated in a significantly contributive region of the distribution. Despite the restriction of excess noise introduced by derivative techniques on experimental data, this method would be able to improve the validity of extracted parameters, in particular when high standard deviation values are involved. From simulated data, it can be shown that the determination of the mean curvature radius r_0 using a full Gaussian model can be highly erroneous as soon as $r_{min}>r_0-3\times\sigma$ (Table II).

Another approach was to consider a parabolic distribution.[21] As the contribution of the emitters in the upper tail of the Gaussian distribution does not significantly contribute to the emission current, the truncated Gaussian distribution can successfully be replaced by a parabolic distribution (Fig. 12). It was demonstrated that using the floating sphere model and a symmetrical parabolic distribution function (with a maximal at r_0 and lower and upper integration limits at, respectively, $r_1=r_{min}$ and $r_2=2\times r_0-r_1$), the following approximate FN expression can be derived:[21]

$$\frac{I}{V^2}=A\times S\times\beta^2\times\frac{6\times r_1^3}{(r_2-r_1)^3}\times\left(\frac{\beta(r_1)V}{B}\right)^3 \times\left(\frac{B}{\beta(r_1)V}\frac{(r_2-r_1)}{r_1}-2\right)\times\exp\left(-\frac{B}{\beta(r_1)V}\right). \tag{8}$$

Taking into account the typical parameter values of the investigated FEA and for sufficiently broad distributions (and low emission turn-on voltages), Eq. (8) can be simplified as

$$\frac{I}{V^4}=A\times S\times\beta^2\times\frac{6\times r_1^2}{(r_2-r_1)^2}\times\left(\frac{\beta(r_1)}{B}\right)^2\times\exp\left(-\frac{B}{\beta(r_1)V}\right) \tag{9}$$

providing an expression from which the parameters r_{min} and r_0 can be evaluated from simple linear regression. Extrapolations of r_{min} and r_0 from simulated data using Eq. (9) effectively show that coherent results are obtained as soon as a broad and highly truncated Gaussian distribution ($r_{min}>r_0-2\times\sigma$) is considered (Table II).

From experiments, on most of the investigated FEAs a consistent evaluation of the parameters is obtained using a comparison of both statistical models. The average r_0 and minimal r_{min} radii of curvature are typically found ranging from 1 to 10 nm and from 0.7 to 2 nm, respectively, in agreement with SEM observations. The geometry scattering (ratio of standard deviation to mean emitter sharpness) is evaluated between 15% and 30% for wet chemical etched arrays and between 15% and 20% for RIE ones. It is worth noting that this analysis only considers the geometric features of the emitters, although other factors can be sources of uncertainty in the parameter estimates, such as an inaccurate determination of the work function or the presence of interface states or surface oxide which can greatly modify the emission properties.[22] For a few characteristics of RIE etched wall arrays, this extraction procedure gives values of curvature radius (r_0 from 0.05 nm to 0.1) far lower than expected from SEM observations. This observation appears to be consistent with relatively low turn-on voltages and higher emission currents. As this effect does not systematically occur on all these structures, an enhancement of the factor β due to the angular geometry of the wall edges seems to be an insufficient assumption to satisfactorily explain this behavior. In contrast to wet chemical etching, RIE processing is supposed to limit the formation of oxide and therefore more readily expose the surface to perturbations. In particular, it can be suggested that a high density of positive surface states introduced during the latter steps of the fabrication process can contribute to enhance the emission, first by lowering the potential barrier, and second by reducing the depletion region at the emitter apex. The last condition also has the additional advantage of delaying Joule heating.[23]

V. CONCLUSION

Various ungated GaAs FEAs have been fabricated using wet and dry etching methods, and their emission capabilities were demonstrated. In most cases, a formation procedure was found necessary to stabilize the $I-V$ characteristics. This procedure leads to a blunting of the sharpest emitters and also to an increase of the standard deviation of the structure geometry, which leads us to suggest that after formation, the geometry distribution is well represented by a truncated Gaussian model. It was also shown that in the case of a highly truncated Gaussian distribution, a more coherent parameter extraction procedure is obtained using a parabolic distribution model. Nevertheless, a comparison between the full Gaussian and parabolic distribution is needed to get a correct idea of the geometry distribution. It appears that in some cases, the geometric features themselves are insufficient to explain the experimental results, and highly modified electrical surface properties need to be reviewed. Indeed,

several experimental evidences, such as current instabilities, oxide breakdown phenomena, or inconsistent parameter extraction all emphasize the influence of the surface properties at the emitter apex on the emission characteristics.

ACKNOWLEDGMENTS

This work was supported by DGA-DRET (Contract No. 94-078). The authors would like to thank Professor E. Constant for fruitful discussions and the technical staff of IEMN for their help in FEA's fabrication, and more particularly M. François and M. Muller for electronic beam lithography.

[1]R. K. Parker, K. L. Jensen, and R. H. Abrams, Proceedings of the 10th International Vacuum Microelectronics Conference, Kyongju, Korea, 17–21 August 1997, pp. 92–97.
[2]F. Charbonnier, Proceedings of the 10th International Vacuum Microelectronics Conference, Kyongju, Korea, 17–21 August 1997, pp. 7–13.
[3]R. B. Marcus, K. K. Chin, Y. Yuan, H. Wang, and W. N. Carr, IEEE Trans. Electron Devices **37**, 1545 (1990).
[4]J. L. Shaw, N. Papanicalau, and H. F. Gray, Le Vide, Les Couches Minces **271**, 124 (1994).
[5]H. F. Gray and J. L. Shaw, Proceedings of the 10th International Vacuum Microelectronics Conference, Kyongju, Korea, 17–21 August 1997, pp. 220–225.
[6]R. Williams, *GaAs Processing Techniques* (Artech House, Boston, 1990), pp. 95–114.
[7]S. Adachi and K. Oe, J. Electrochem. Soc. **130**, 2427 (1983).
[8]F. Ducroquet, P. Kropfeld, O. Yaradou, and A. Vanoverschelde, J. Vac. Sci. Technol. B **16**, 787 (1998).
[9]B. Tuck, J. S. K. Mills, and A. J. Hartwill, J. Mater. Sci. **11**, 847 (1976).
[10]G. C. DeSalvo *et al.*, J. Electrochem. Soc. **143**, 3652 (1996).
[11]T. Takebe, T. Yamamoto, M. Fujii, and K. Kobayashi, J. Electrochem. Soc. **140**, 1169 (1993).
[12]S. H. Jones and D. K. Walker, J. Electrochem. Soc. **137**, 1653 (1990).
[13]D. W. Shaw, J. Electrochem. Soc. **128**, 874 (1981).
[14]O. Yaradou, F. Ducroquet, P. Kropfeld, and A. Vanoverschelde, Proceedings of the 21th International Semiconductor Conference, CAS'98, Sinaia, Romania, 6–10 October 1998, pp. 361–364.
[15]"CATIA ® solutions", developed by S. A. Dassault.
[16]C. A. Spindt, I. Brodie, L. Humphrey, and E. R. Westerberg, J. Appl. Phys. **47**, 5248 (1976).
[17]D. Nicolaescu and V. Avramescu, J. Vac. Sci. Technol. B **12**, 749 (1994).
[18]T. Utsumi, IEEE Trans. Electron Devices **38**, 2276 (1991).
[19]R. J. Harvey, R. A. Lee, A. J. Miller, and J. K. Wigmore, IEEE Trans. Electron Devices **38**, 2323 (1991).
[20]J. D. Levine, J. Vac. Sci. Technol. B **13**, 553 (1995).
[21]D. Nicolaescu, J. Vac. Sci. Technol. B **12**, 759 (1994).
[22]S. W. Barry and M. H. Weichold, J. Vac. Sci. Technol. B **11**, 379 (1993).
[23]J. L. Shaw and H. F. Gray, Proceedings of the 11th International Vacuum Microelectronics Conference, Asheville, NC, 19–24 July 1998, pp. 146–147.

Novel lateral field emission device fabricated on silicon-on-insulator material

Minhee Yun, Anthony Turner, Ronald J. Roedel,[a] and Michael N. Kozicki
Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6206

(Received 3 September 1998; accepted 9 April 1999)

A field emission device which shows very low turn-on voltage with large electrode spacing has been fabricated by applying KOH orientation-dependent etching to single crystal silicon. Aqueous KOH solutions create sharp angular structures bounded by (111) planes at an angle 54.74° from the (100) surface. Using this technique, "self-sharpening" tips were formed on silicon-on-insulator wafers. This process can be used to form lateral sharp edges that can act as both the cathode and the anode in a field emission device. The shapes of the tips were determined using transmission electron microscopy and the smallest emitter tip radius in this study was estimated to be 3 nm. These sharp emitters enable significant electron emission at low turn-on voltage even for large spacing between the emitter and the collector. The lowest turn-on voltage was found to be 27 V with emission currents of 10 μA at 1 μm spacing. Electron emission was observed even at an electrode spacing larger than 2 μm. The field enhancement factor (β) and field emission area (α) have been calculated using the Fowler-Nordheim equation. As expected, the field enhancement factor decreases with increasing electrode spacing. Failure mechanisms for the devices associated with high electric field are also discussed in this article. © *1999 American Vacuum Society.* [S0734-211X(99)00904-X]

I. INTRODUCTION

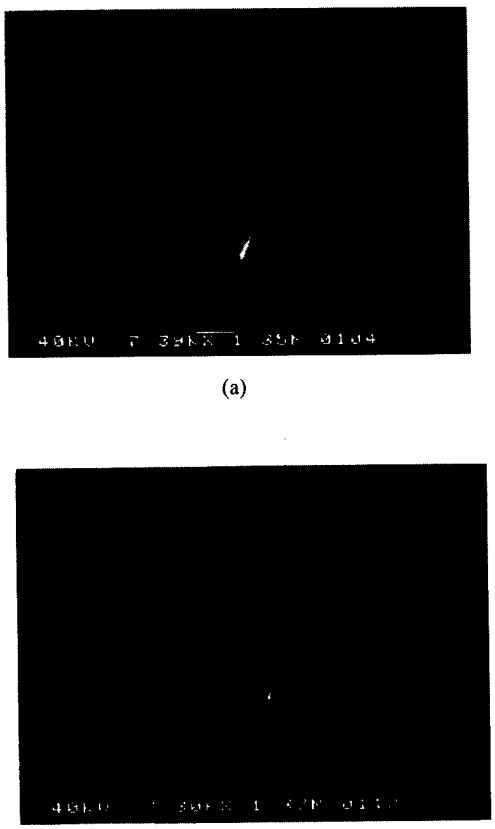
Since the first microelectronic field emission devices were fabricated by Shoulders in 1961,[1] significant developments have been achieved by others such as Spindt *et al.*[2] and Gray and Greene.[3] Since then, field emission devices have found many applications in free electron systems such as electron microscopes, flat panel displays, and high power amplifiers.[4–6] Their operation is based on high electric field induced transfer of electrons through a vacuum from a cathode to an anode.

Busta[7] reported a lateral field emission structure called a "nonself-aligned device" in 1989. This device employed a triangular shaped NiCr emitter of 800 Å thickness. It was fabricated on glass substrates over which a 1 μm thick Al gate was fabricated using a 0.3 μm silicon nitride film as a sacrificial layer. A major advancement came in 1991 when groups from Honeywell and the Naval Research Labs reported a thin film edge emitting vacuum microelectronic diode.[8] They obtained a high electric field on the emitter surface by applying up to 80 V across electrodes between the thin emitter (200 Å) and the anode. The device is essentially an emitter wedge with a curvature of less than 200 Å. Gray and co-workers proposed a vacuum microelectronic triode.[9] Oro and Ball[10] fabricated lateral tantalum multitip diodes with emitter to anode spacing ranging from 0.03 to 0.50 μm using high resolution electron beam lithography. Mason and co-workers[11] studied an improved monolithic diode structure. This device employs simple fabrication steps and is called a "table-top" emitter because of its shape. The process makes use of the conformal deposition of layers, followed by patterning and etching. This is a semiconductor industry-wide standard approach to fabrication. However, the fabrication of these devices is relatively difficult. The most recent work has been done by Park *et al.*,[12] who investigated lateral field emission diodes fabricated on separation by implantation of oxygen (SIMOX) wafers. The electrode spacing ranges from 0.3 to 0.8 μm and has a turn-on voltage of 22–25 V and emission current of 7 μA. However, the emitter was constructed by allowing a bridge between the anode and the cathode to separate by fracturing. The spacing was not well controlled.

Fabrication of these lateral field emission devices typically uses photo- or electron-beam lithography to define the distance between the emitter and the collector. This dimension determines the lowest operating voltage as it defines the electric fields at the emission tips. However, it is not easy to fabricate an electrode spacing of less than 1 μm using the typical lithography technique and wet etching. On the other hand, the electric field at the emission tip at a given voltage can be increased by choosing appropriate materials and by increasing the sharpness of the emission tip. Recently, various researchers have reported the development of suitable materials for electron emitters. These materials include Mo,[13] Au,[14] TiW/W/Al,[15] and chemically vapor deposited (CVD) diamond.[16] The advantages of metal and diamond emitters are a low work function, very small electron affinity, physical stability of the material, and high thermal conductivity. However, the fabrication of these emitters requires complicated processing and higher cost compared to single crystal silicon emitters.

In this study, we used silicon-on-insulator (SOI) to fabricate the electron emitters and focused our work on increasing the sharpness of the emitter. Etching of the top Si with aqueous KOH solution was used to form the emitters since this exposes facets at an angle 54.74° from Si(100) surface because of the different etch rates between (100) and (111)

[a]Electronic mail: r.roedel@asu.edu



(a)

(b)

FIG. 1. SEM picture of emitter sharpness based on KOH etching conditions of (a) 30% KOH at 70 °C reaction temperature and (b) 30% KOH at 90 °C reaction temperature.

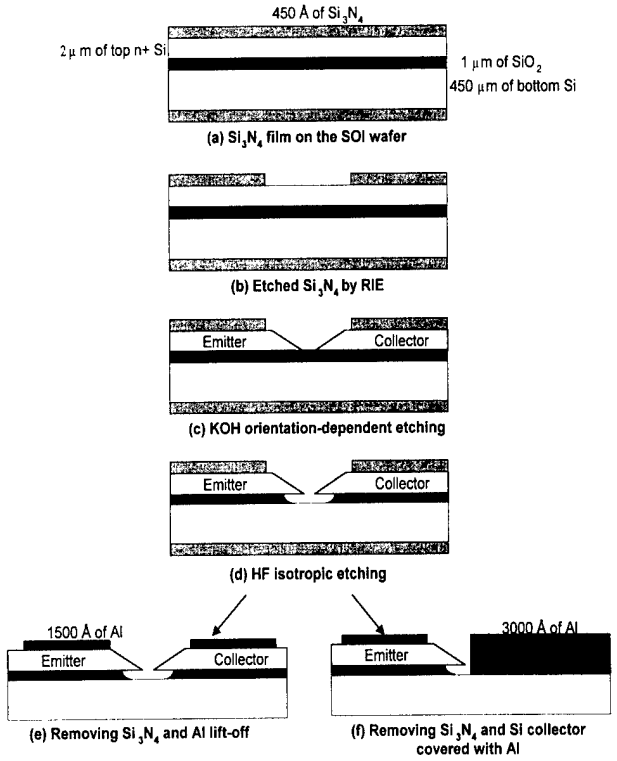


FIG. 2. Schematic diagram of the process sequence for the fabrication of field emission devices.

planes on Si(100) wafers. Although rectangular masks were used to define the cathodes, due to edge effects the KOH etches more rapidly at the corners and, consequently, extremely sharp tips were produced. Figure 1 shows a scanning electron microscopy (SEM) top view of the emitter tips on the SOI material under different etching conditions. These tips are formed with a 30% KOH solution at 70 and 90 °C etching reaction temperatures. Transmission electron microscopy (TEM) was used to observe the etched shape and sharpness of the tip and revealed that a tip diameter of 1–5 nm was possible. Many different types of tips can be formed by controlling the concentration of the KOH solution and other etching conditions.[17] The etching condition for tip sharpness was optimized by the "design of experiment approach."

Our results indicate that lower temperatures produce decreased tip sharpness because of the formation large H_2 bubbles which impede the etch process. Interestingly, the sharpness of the silicon tip shows a similar trend as the etch rate in that both the sharpness and etch rate increased to a critical point and then decreased at very high temperature and KOH concentration. The sharpest tips fabricated in this study were made at 30% KOH and 70 °C reaction temperature and have radius ≤ 3 nm. Although different types of tips can be formed by controlling the concentration of the KOH solution and other etching conditions, only the optimum tip formed by KOH etching will be described in this study.

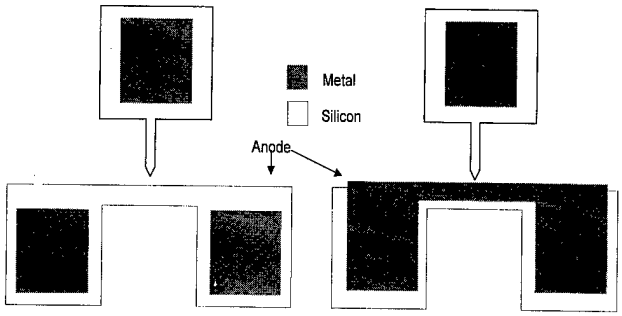
Using these fabricated field emission devices, the relationship between the field enhancement factor and electrode spacing will be described, and the electron field emission area will be reviewed using transmission electron micrographs of the device. Finally, possible device failure mechanisms due to the high electric field will be briefly discussed.

II. FABRICATION OF SOI FIELD EMISSION DEVICES

The process sequence of the fabrication of SOI field emission devices is shown in Fig. 2. The wafers used in this research were thermally bonded SOI wafers with (100) orientation. The thicknesses of the top Si and buried SiO_2 layer were 2 and 1 μm, respectively, with $\pm 10\%$ variation. Processing begins with standard RCA cleaning of the SOI wafers. The top Si layer was then highly doped using a phosphorus solid source at 970 °C to convert the p-type top Si layer to an n^+ layer. A 450 Å layer of Si_3N_4 [Fig. 2(a)], which will act as a KOH etching mask, was deposited on the wafer using low-pressure chemical vapor deposition (LPCVD). Silicon oxide can be used as an etch mask for short periods in the KOH solution but, for long periods, silicon nitride is a better etch mask because it etches more slowly. The SOI wafer was then cut into 10×10 mm sq pieces for use for discrete field emission devices. Positive photoresist was used to pattern the silicon nitride for the KOH etching mask. The nitride was etched with a CF_4 plasma [Fig. 2(b)] and the top Si on the SOI wafer was then etched with 30% KOH solution at 70 °C for 3 min [Fig. 2(c)]. Isotropic etching with HF was used to remove 440 nm



(a)



(b)

FIG. 3. (a) Cross-sectional TEM image of the field emitter tip and (b) top view of the field emission devices.

of SiO_2 around the field emission tip after the anisotropic etching [Fig. 2(d)]. This was done to allow the Si tip to overhang the isolation. Figure 3(a) shows a TEM cross-sectional view of the emitter after HF isotropic etching. Finally, electrical contacts to the n^+ Si layer were formed by electron beam evaporation of 1500 Å of aluminum which was subsequently patterned by lift off [Fig. 2(e)] or etching [Fig. 2(f)]. To decrease the electrode spacing and increase the physical stability of the anode, 3000 Å of Al was deposited on the Si anode by e-beam evaporation [Fig. 2(f)]. Figure 3(b) shows top-view sketches of the completed field emission devices on the SOI wafers. The spacing between the cathode and anode can be controlled from less than 1 to 5 μm by the KOH etch process condition and by moving the anode metal mask relative to the emitter. A select metal etchant, a mixture of HNO_3, HPO_3, CH_3COOH, and H_2O (1:4:4:1), was used to etch the Al in this research. Note that only two lithography steps are necessary for the fabrication of SOI field emission devices in this research.

Very sharp emitters can easily achieve the high electric fields necessary for field emission of about 4×10^7 V/cm. The tip diameter is one of the most significant factors determining the emission current because it affects the electric field at the emitter tip. In order to obtain high fields for electron emission, field emission devices must have a high curvature, resulting in a very high applied field at the emitting surface. Etching with KOH solutions is a process that can produce extremely sharp emitters.

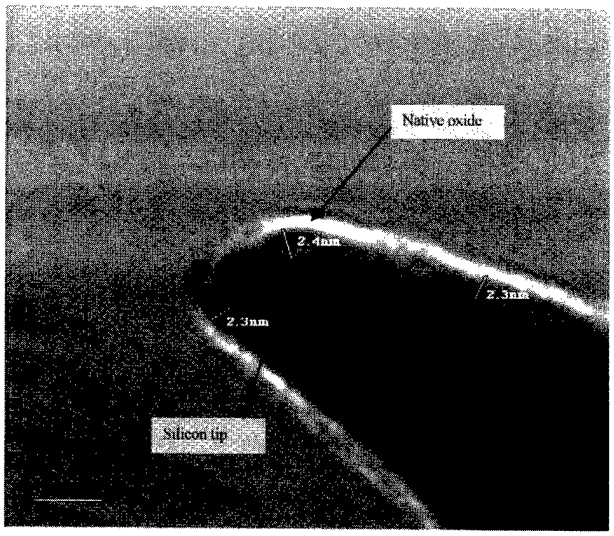


FIG. 4. TEM image of the emitter tip with native oxide around the tip.

A high vacuum system with a contact stage was built for testing of the field emission devices in this research. The fabricated devices were tested under a vacuum of approximately 10^{-7} Torr. The contact stage inside the high vacuum system was connected via coaxial feedthroughs to a Hewlett Packard model HP 4145B or Tektronix 372 semiconductor parameter analyzer for electrical characterization. The testing was performed by applying a positive voltage at the anode while holding the cathode at the measurement unit ground. A scanning electron microscope (ISI-DS 130) and a field emission gun source TEM (Philips CM2000) were used to determine the tip sharpness and diameter.

III. DISCUSSION

The tip diameter and sharpness are the most significant factors determining the electric field at the emitter tip and thus the emission current. This research indicated that single crystal silicon field emission tips formed using a KOH solution have a much better capability of electron emission compared to other previous methods.[7–12] TEM was used to observe the etched shape and sharpness of the tip and revealed that a tip diameter of 1–5 nm was possible. Such sharp emitters enable significant electron emission at low turn-on voltage even for relatively large spacing between the emitter and the collector.

Figure 4 shows the shape of the tip that was determined using TEM. The emitter diameter was estimated to be 3 nm. The thickness of the oxide layer around the emitter tip was found to be 2.3 nm. This layer formed around the emission tip as a result of exposing the silicon to air. Yang *et al.*[18] have reported the reduction of emission current density from an emitter covered with an oxide layer a few monolayers thick. In order to increase the efficiency of emission tip and electron emission, the native oxide should be minimized.

Figure 5 shows the *I-V* characteristics of SOI field emission devices for two prototype diodes with turn-on voltages of 27 and 60 V. The device with the 27 V turn on (B3381)

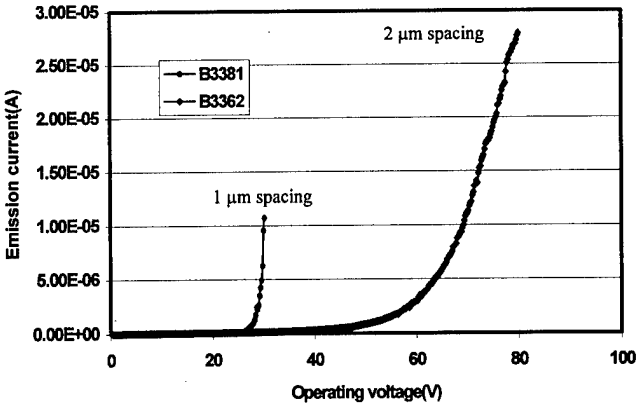


FIG. 5. *I-V* characteristics of field emission showing the dependence on electrode spacing.

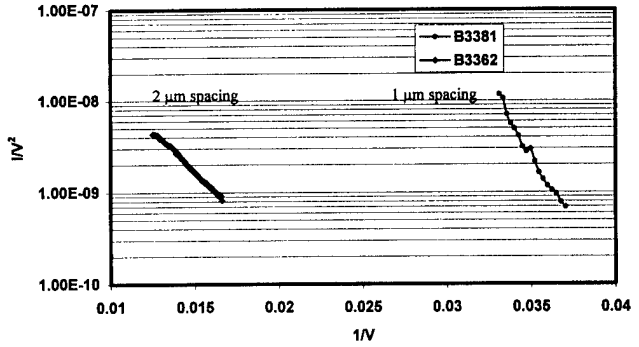


FIG. 6. Fowler-Nordheim plots showing the dependence on electrode spacing.

had a metal anode [Fig. 2(f)] and 1 μm spacing between electrodes. The emission current was 10 μA at 30 V of operating voltage in this device.

The silicon anode devices [Fig. 2(e)] generally showed higher turn-on voltages than the metal anode devices because of their larger spacing. For example, device B3362 had a spacing between electrodes of about 2 μm with a turn-on voltage of 60 V and an emission current of 27 μA at 80 V of operating voltage. The above devices were fabricated with the same etching condition (30% KOH concentration, 70 °C reaction temperature) but different electrode spacing. Different emission currents are expected because of different tip geometries and emission areas caused by processing conditions. However, it is clear that the turn-on voltage is also strongly dependent on electrode spacing. In this research, most of the devices had turn-on voltages of 100 V or less, but some were as high as 180 V when the spacing exceeded 3 μm.

Different geometrical shapes of emitter tips have been studied by many researchers.[2,19–22] For adequate emission at a reasonable voltage, the radius of a field emitter tip may vary from 10 to 200 nm but it is typically less than 100 nm. Measurement of current from such points can confirm the relevance of the Fowler-Nordheim plot. Typically, large radius tips show steeper lines than small radius tips.

We believe that our emission tip fulfills these conditions very well. Figure 4 shows the geometry and shape of the field emitter tip used in this research. It is assumed that electrons are emitted with uniform radial distribution because the tips have point geometries and the anodes were destroyed in a radial pattern, as shown in Fig. 7.

The Fowler-Nordheim equation expresses the field emission current density as a function of electric field at the surface. This equation with evaluated constants can be written as follows:

$$J=1.4\times10^{-6}\frac{E^2}{\phi}\exp\left(-\frac{6.87\times10^{7}\phi^{3/2}v(k)}{E}\right), \tag{1}$$

where J and E are the current density and the electric field, ϕ is the work function of the emitter material (4.5 eV), and $v(k)$ is a function dependent on the wave number (k) whose value can be found in the literature.[23] For a comparison of theory and experiment, the Fowler-Nordheim (FN) equation can be expressed in terms of the field enhancement factor (β) and field emission area (α): $J=I/\alpha$, $E=V\beta$. The constants β and α can be determined from the plot of $1/V$ vs I/V^2 (which is called the FN plot) as

$$\beta=\frac{6.526\times10^{7}\phi^{3/2}}{\text{slope}}\quad(\text{cm}^{-1}), \tag{2}$$

$$\alpha=\frac{\phi\exp\left(\text{intercept}-\dfrac{9.868}{\sqrt{\phi}}\right)}{1.4\times10^{-6}\beta^{2}}\quad(\text{cm}^{2}). \tag{3}$$

A FN plot was used to calculate the field enhancement factor (β) and field emission area (α) of the tips in our structures. These field enhancement factors and field emission areas can also be determined by the tip geometry. Smaller electrode spacing and sharper emitter tips result in a high enhancement factor and lower operating voltage. Figure 6 shows Fowler-Nordheim plots dependent on electrode spacing of the devices in Fig. 5. Table I shows the field enhancement factor and field emission area calculated from Fig. 6 and Eqs. (2) and (3), and includes results from other devices as well. As expected, the field enhancement factor decreases with increased electrode spacing. We assume that there will be no change in the field enhancement factor at a certain large distance. We conclude that our field enhancement factor is larger than that of other researchers because of the combination of small spacing and the sharpness of the emission tip.[7–12]

From the data listed in Table I, the highest field emission enhancement factor of $3.29\times10^{6}\,\text{cm}^{-1}$ and the smallest field

TABLE I. Field enhancement factors and field emission areas.

Device	β (cm^{-1})	α (cm^{2})	Spacing (μm)	Anode type
D829t13	2.88×10^{6}	3.70×10^{-9}	<1	Metal
SOI19	3.29×10^{6}	2.84×10^{-9}	$\ll1$	Metal
B3381	8.53×10^{5}	4.23×10^{-8}	$\leqslant1$	Metal
B3362	8.30×10^{5}	4.50×10^{-8}	$\gg2$	Silicon

TABLE II. Comparison between the tip diameter and the field enhancement factor.

Reference	Tip diameter (nm)	Field enhancement factor (cm^{-1})
19	18 000	1.28×10^{4}
20	50	2.17×10^{5}
2	25	2.10×10^{5}
This work	3	3.29×10^{6}
21	0 (very sharp)	2×10^{7}

emission area of $2.84\times10^{-9}\,cm^{2}$ have been achieved with device SOI 19. It is clear that only a few atoms at the end of the tip contribute to electron emission, which is consistent with our TEM studies. However, the calculated emission areas do not exactly match our TEM pictures. We assumed that the calculated emission area might be different due to the atomically sharp or nonplanar emitter. Cutler *et al.*[24] have proposed that the use of the conventional FN equation can lead to erroneous values for the emitting area and fields because the equation was derived using the planar model. In addition, the emission might be reduced because of the native oxide layer.

Table II summarizes a comparison of tip diameters investigated by other researchers.[2,19–21] Pang and co-workers[21] reported results of emitter tip simulations, and they calculated that the field enhancement factor of the sharpest tip would be $2\times10^{7}\,cm^{-1}$. Our result shows that a 3 nm diam tip has the highest experimental field enhancement factor, $3.29\times10^{6}\,cm^{-1}$. The emitter tip fabricated in this research shows the smallest emission area ($2.84\times10^{-9}\,cm^{2}$), leading to low turn-on voltage (25 V) and high emission current.

Reliability is one of the most important factors in field emission device operation. It is frequently found that ultrasharp tips become severely degraded and even melt when high current densities are produced. To prevent tips from melting, it is necessary to minimize the current density unless good heat sinking to the substrate is employed.[25]

In this research, device testing frequently destroyed tips. The emitters have a short lifetime because of the high current density, greater than $5\times10^{3}\,A/cm^{2}$, at the end of the tip. Figure 7 shows a SEM picture of a melted silicon emitter tip and electrode. The anode is melted in an arc pattern, clearly showing the effect of electron emission and bombardment. The exact cause of the destruction is currently under analysis. However, when this type of tip is operated at less than 2 nA, it does not suffer from failure by melting or ablation.

IV. CONCLUSION

Using KOH etching, silicon emitter tips were fabricated to a high degree of sharpness. Our research indicated that single crystal silicon field emission tips formed using KOH solution show a much better capability of electron emission compared to previous attempts.[7–10] Transmission electron microscopy was used to observe the etched shape and sharpness of the tips in this study and revealed that a tip diameter



FIG. 7. SEM picture of thermomechanical stress on the emission tip which has been destroyed at 200 V operating voltage.

of 3 nm was possible. The lowest attainable turn-on voltage was 27 V, and 10 μA emission current at 32 V operating voltage from 1 μm electrode spacing device was possible. Due to the sharpness of the tip, field emission could be achieved with electrode spacing larger than 2 μm. A high field emission enhancement factor of $3.29\times10^{6}\,cm^{-1}$ has been determined using Fowler-Nordheim plots. It is believed that further increases in the efficiency of the emitter tip can be obtained by minimizing the native oxide layer around the tip. The devices may possibly fail catastrophically because of thermomechanical stress. Methods to avoid thermomechanical stress are under investigation.

ACKNOWLEDGMENTS

The authors would like to thank the Electric Power Research Institute (EPRI) for support of this project and also thank R. E. Dunin-Borkowski and B. E. Kardynal for valuable insight and TEM inspection.

[1]K. R. Shoulders, Adv. Comput. **2**, 135 (1961).
[2]C. A. Spindt, I. Brodie, L. Humphrey, and E. R. Westerberg, J. Appl. Phys. **47**, 5248 (1976).
[3]H. F. Gray and R. F. Greene, U.S. Patent No. 4,307,507 (filed 29 December 1981).
[4]L. Reimer, *Scanning Electron Microscopy* (Springer, Berlin, 1985).
[5]C. A. Spindt, C. E. Holland, A. Rosengreen, and I. Brodie, IEEE Trans. Electron Devices **38**, 2355 (1991).
[6]J. P. Calame, H. F. Gray, and J. L. Sham, J. Appl. Phys. **73**, 1485 (1993).
[7]H. H. Busta, Inst. Phys. Conf. Ser. **99**, 29 (1989).
[8]H. F. Gray, IEEE International Electron Device Meeting, Washington, DC, December 1991.
[9]A. I. Akinwande, P. E. Buahahn, D. K. Arch, and H. F. Gray, 6th International Conference on Vacuum Microelectronics, July 1992.
[10]J. A. Oro and D. D. Ball, J. Vac. Sci. Technol. B **11**, 464 (1993).
[11]J. D. Legg, M. E. Mason, R. T. Williams, and M. H. Weichold, J. Vac. Sci. Technol. B **12**, 666 (1994).
[12]J.-H. Park, H.-I. Lee, H.-S. Tae, J.-S. Huh, and J.-H. Lee, IEEE Trans. Electron Devices **44**, 1018 (1997).
[13]J. H. Jung, B. K. Ju, Y. H. Lee, J. Jang, and M. H. Oh, IEEE Electron Device Lett. **18**, 197 (1997).
[14]J.-L. Shieh, R.-J. Ren, T.-J. Shieh, D. P. Klemer, and C.-Y. Chen, J. Vac. Sci. Technol. B **11**, 501 (1993).
[15]J. M. Kim, W. N. Carr, and R. J. Zeto, J. Vac. Sci. Technol. B **11**, 459 (1993).
[16]Z. Feng, I. G. Brown, and J. W. Ager III, J. Mater. Res. **10**, 1585 (1995).

[17]M. H. Yun, V. A. Burrows, and M. N. Kozicki, J. Vac. Sci. Technol. B **16**, 2844 (1998).
[18]G. Yang, K. K. Chin, and R. B. Marcus, IEEE Trans. Electron Devices **38**, 2373 (1991).
[19]W. P. Dyke and J. K. Trolan, Phys. Rev. **89**, 799 (1953).
[20]D. Hong, M. Aslam, M. Feldmann, and M. Olinger, J. Vac. Sci. Technol. B **12**, 764 (1994).
[21]M. R. Rakhshandehroo, F. Sukardi, and S. W. Pang, J. Vac. Sci. Technol. A **14**, 1832 (1996).
[22]T. Utsumi, IEEE Trans. Electron Devices **38**, 2276 (1991).
[23]R. E. Burgess, H. Kromer, and J. M. Houston, Phys. Rev. **90**, 515 (1953).
[24]P. H. Cutler, J. He, N. M. Miskovsky, T. E. Sullivan, and B. Weiss, J. Vac. Sci. Technol. B **11**, 387 (1993).
[25]M. G. Ancona, J. Vac. Sci. Technol. B **13**, 2206 (1995).

Field emission properties of diamondlike carbon films deposited by ion beam assisted deposition

Jing Wang,[a] Wen-Zhi Li, and Heng-De Li
Department of Materials Science and Engineering, Tsinghua University, Beijing 100084, People's Republic of China

(Received 14 December 1998; accepted 28 May 1999)

The ion beam sputtering technique was employed to prepare diamondlike carbon (DLC) films at nearly room temperature. Simultaneous ion beam bombardment during film growth was also conducted in order to study the bombardment effects. Raman spectroscopy was used to evaluate the structure property. The sp^3 fraction was found to strongly depend on the bombarding ion energy, giving a highly sp^3 bonded DLC at ion beam bombarding energies around 800 eV. The emission characteristics of these amorphous DLC films were compared. Field emission from DLC deposited with higher bombarding energies exhibit enhanced emission properties. High electron emission current, about 370 μA over a 1 cm^2 cathode area, was obtained with an electric field of 18.5 V/μm. The mechanism for electron emission from DLC films is also discussed. © *1999 American Vacuum Society.* [S0734-211X(99)09504-9]

I. INTRODUCTION

Field emission displays (FEDs) have the potential to be a low cost, high performance alternative to the currently dominant cathode ray tube and liquid crystal display technologies for flat panel displays. One of the key issues in FEDs is to develop a reliable and efficient cold cathode material for electron emission. Earlier field emitters were typically metals (such as Mo or W) or semiconductors (such as Si) with nanometer-sized sharp tips.[1,2] These emitters typically require complicated fabrication steps and need high control voltages for emission because of the high work functions associated with these materials. In addition, the fabrication of uniform sharp tips over a large area has been proven to be a difficult and expensive process.

One of the methods to solve this problem is to employ low work function materials, such as diamond, as the cold cathode material. Diamond films possessing negative electron affinity (NEA) characteristics are viewed as a possible cold cathode emitter. Excellent field emission characteristics with low turn-on field and high current density have been observed in chemical vapor deposition (CVD) diamond.[3–6] However, the CVD technique usually requires high growth temperatures, which are not compatible with the semiconductor technology.

Field emission has also been measured on several other carbon containing materials. It has been reported in recent years that diamondlike carbon (DLC) films possess excellent field emission properties.[7–12] Furthermore, compared to diamond, DLC films can be easily prepared using low substrate temperature (<200 °C) and over large substrate areas at low cost. The mechanism for electron emission from DLC films is still unclear and needs to be studied further.

Several techniques such as mass selected ion beam,[13,14] laser ablation of graphite,[15–18] and filtered catholic vacuum arc[19–25] have been used to deposit DLC. In the present work, an ion beam assisted deposition (IBAD) system was used to synthesize DLC films. IBAD is a combination of ion implantation and physical vapor deposition. In IBAD, most experimental parameters are independent of each other and are easily controlled. The deposition temperature can be as low as the ambient temperature. This greatly enlarges the choice of films and substrate materials. Film structure and property can also be easily tailored by adjusting the bombarding energy and species of the bombarding ion beam. In this article, the Raman spectroscopy and field emission properties of the DLC films deposited with a bombarding energy of 100, 200, and 800 eV are reported.

II. EXPERIMENT

DLC films were prepared using an innovative multifunction IBAD system[26,27] equipped with water-cooled rotatable target holders, a water-cooled rotatable sample holder, and Kaufman ion sources. Gas was introduced into the ion source chamber, and ionized due to impact of electrons emitting from the cathode filament. The ions were extracted by an electric field and used to sputter the target or bombard the growing films with varying energies. The base vacuum of the equipment was 1×10^{-3} Pa and the working pressure was about $(1.0–1.2)\times10^{-2}$ Pa. Highly doped p-Si (0.01–0.001 Ω cm) was used as a substrate for the films. The Si was cleaned with acetone and methanol, and then etched in a 20% HF solution prior to loading inside the chamber. Prior to deposition, the substrates were cleaned by Ar^+ bombardment with an ion energy of 3 keV and beam density of 127 μA cm^{-2} for 10 min. During deposition, an argon ion beam with an energy of 3 keV was employed to sputter a graphite target. The growing films were simultaneously bombarded by a 5 mA CH^{n+} beam with energies of 100, 200, and 800 eV.

The field emission property was measured with a simple field emission device. Prior to insertion, the samples were

[a]Author to whom correspondence should be addressed; electronic mail: zjzdms@mail.tsinghua.edu.cn

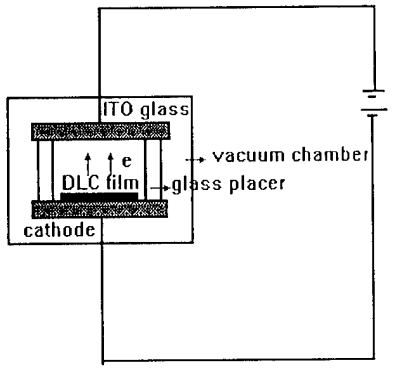


FIG. 1. Schematics of the apparatus used for measuring I–V characteristics of the DLC films.

rinsed in acetone, followed by ethanol, and finally dried under a N_2 stream. The DLC film was mounted as the cathode. A schematic diagram of the FED and the circuit used in the present work are shown in Fig. 1.

III. RESULTS AND DISCUSSION

Raman spectroscopy was performed to determine the diamondlike bonding character in the films. The Raman spectra of the DLC films deposited on Si substrates are shown in Fig. 2. In this figure, the Raman bands of the DLC films bombarded with different bombarding energy are shown. The spectra consist of a broad peak at about 1550 cm^{-1}, indicating a significant fraction of the carbon films is DLC.[28,29] In order to analyze the Raman spectra quantitatively, they are fitted with a Breit–Wigner–Fano (BWF) line shape[30] and a linear background using a least-squares computer program. The BWF line shape is described by

$$I(\omega)=\frac{I_0[1+2(\omega-\omega_0)/Q\Gamma]^2}{1+[2(\omega-\omega_0)/\Gamma]^2}, \tag{1}$$

where $I(\omega)$ is the intensity as a function of frequency, I_0 is the maximum peak intensity, ω_0 and Γ are the peak position and full width at half maximum, respectively, and Q is the BWF coupling coefficient. In the limit as $1/Q$ approaches zero, the Lorentzian line shape is recovered. For these films, a reasonable fit cannot be obtained by using a single BWF line shape, and it is necessary to include in the fit a second Lorentzian centered at about 1350 cm^{-1} (i.e., the D peak). This indicates the presence of small graphitic crystallites in these films.[31]

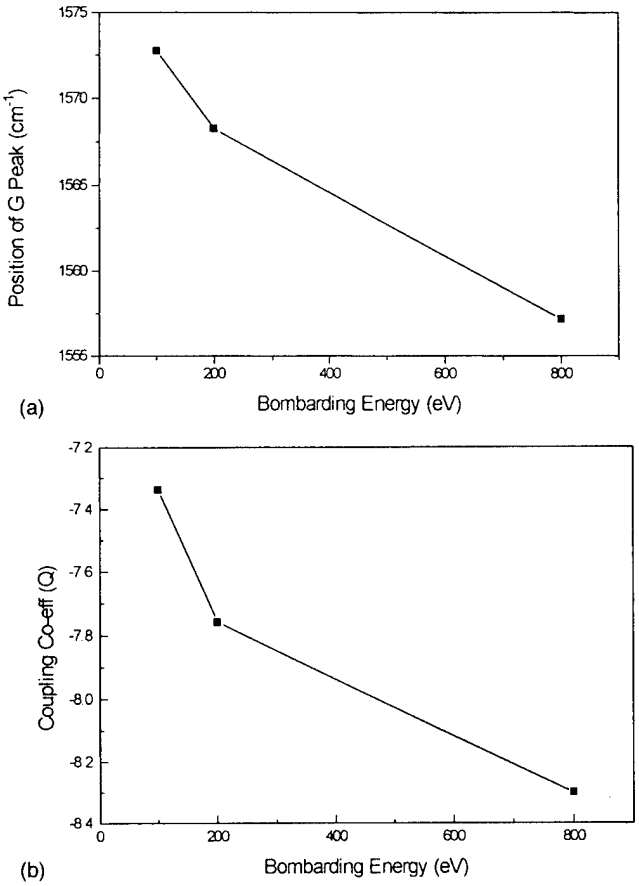


FIG. 3. (a) G peak position and (b) the coupling coefficient Q as a function of the bombarding energy.

In Fig. 3, the BWF line position and the coupling coefficient Q are shown as a function of the bombarding energy. It can be seen that the BWF line position and the coupling

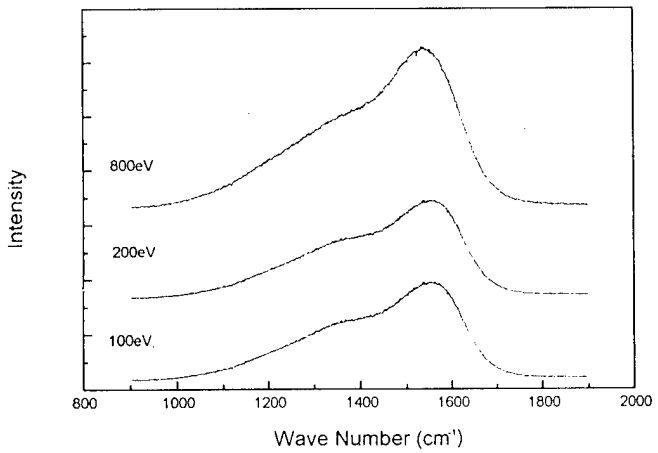


FIG. 2. Raman spectra of DLC films bombarded by different energy.

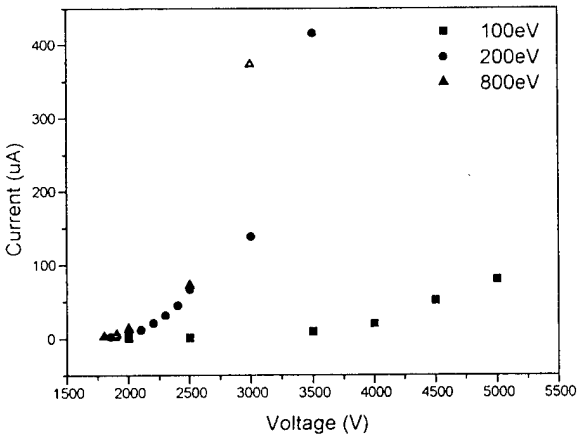


FIG. 4. I–V plot of the DLC films.

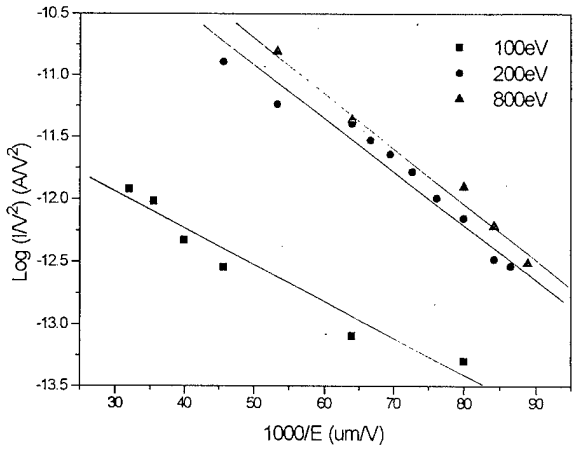


FIG. 5. FN plot of the DLC films.

coefficient Q decrease with bombarding energy. According to Prawer's theory, as the sp^2 fraction decreases, the BWF line position moves down in energy and the coupling parameter decreases. Thus, for the DLC films bombarded by a CH^{n+} beam with an energy of 100, 200, and 800 eV, the sp^2 fraction reached a minimum at an ion energy of 800 eV.

The current–voltage *(I–V)* plots are shown in Fig. 4. It can be seen that the emission current increases with bombarding energy. Figure 5 shows the Fowler–Nordheim (FN) plot for the DLC films. The field threshold for emission decreases from 12.5 V/μm for the film grown with 100 eV ions to 11.56 V/μm for the film grown with 200 eV ions to 11.25 V/μm for the film grown with 800 eV ions. Control experiments were carried out using an uncoated Si substrate and reversing the polarity of the anode–cathode terminals. No current could be detected in the range of experimental field strength.

From the above experimental observations, we conclude that the emission current from our DLC films was enhanced when the sp^3 fraction was increased by the higher bombarding energy. To understand the origins of the emission process from DLC films, we consider that the DLC films consisting of sp^2 and sp^3 bonded carbons have no long-range ordering, but may possess medium- or short-range ordering.[32,33] According to the theory of energy banding of semiconductors, the energy band does not depend on the long-range ordering, but on the short-range ordering of the materials. The structure of energy band in the sp^3 banding area is similar to that of the diamond. Since diamond possesses NEA characteristics, the DLC film with higher sp^3 fraction should have better field emission characteristics.

IV. CONCLUSIONS

In summary, we have investigated the field emission behavior of IBAD synthesized DLC films containing different amounts of sp^3 bonded carbon. Increasing the ion beam bombarding energy caused the sp^3 fraction in the films to increase and the emission current to increase. It is noted that the enhanced field emission is linked to the larger sp^3 fraction in the higher bombarding energy film.

ACKNOWLEDGMENT

The work was supported by State Key Laboratory of Tribology at Tsinghua University.

[1]C. A. Spindt, C. E. Holland, I. Brodie, J. B. Mooney, and E. R. Westerberg, IEEE Trans. Electron Devices **36**, 1 (1989).
[2]C. A. Spindt, I. Brodie, L. Humphrey, E. R. Westerberg, J. Appl. Phys. **47**, 5248 (1976).
[3]A. A. Talin, L. S. Pan, K. F. McCarty, and T. E. Felter, Appl. Phys. Lett. **69**, 3842 (1996).
[4]K. Okano and K. K. Gleason, Electron. Lett. **31**, 74 (1995).
[5]F. Lacher, Diamond Relat. Mater. **6**, 1111 (1997).
[6]M. W. Geis, J. C. Twichell, J. Macaulay, and K. Okano, Appl. Phys. Lett. **67**, 1328 (1995).
[7]B. S. Satyanarayana, A. Hart, W. I. Milne, and J. Robertson, Appl. Phys. Lett. **71**, 1430 (1997).
[8]S. Lee, B. Chung, and T.-Y. Ko, J. Phys. IV **6**, 91 (1996).
[9]G. A. J. Amaratunga and S. R. P. Silva, Appl. Phys. Lett. **68**, 2529 (1996).
[10]N. Missert, T. A. Friedmann, and J. P. Sullivan, Appl. Phys. Lett. **70**, 1995 (1997).
[11]Y. K. Hong, 97'IVMC.
[12]K. C. Park, J. H. Moon, and S. J. Chung, J. Vac. Sci. Technol. B **15**, 428 (1997).
[13]Y. Lifshitz, S. R. Kasi, and J. W. Rabalais, Phys. Rev. Lett. **68**, 620 (1989).
[14]Y. Lifshitz, S. R. Kasi, J. W. Rabalais, and W. Eckstein, Phys. Rev. B **41**, 10468 (1990).
[15]D. L. Pappas, K. L. Saenger, J. Brueley, W. Krakow, J. J. Cuomo, T. Gu, and R. W. Collins, Appl. Phys. Lett. **67**, 984 (1995).
[16]F. Xiong, Y. Y. Chang, and R. P. H. Chang, Phys. Rev. B **48**, 8016 (1993).
[17]A. A. Voevodin, S. J. P. Laube, S. W. Walck, J. S. Solomon, M. S. Donley, and J. S. Zabinski, J. Appl. Phys. **78**, 4123 (1995).
[18]I. I. Aksenov, S. I. Vakula, V. G. Padalka, R. E. Strelnitski, and V. M. Khoroshikh, Sov. Phys. Solid State **25**, 1164 (1980).
[19]D. R. Mckenzie, D. Muller, and B. A. Pailthorpe, Phys. Rev. Lett. **67**, 773 (1991).
[20]P. J. Fallon, V. A. Veerasamy, C. A. Davis, J. Robertson, G. A. J. Amaratunga, W. I. Milne, and J. Koskinen, Phys. Rev. B **48**, 4777 (1993).
[21]B. F. Coll and M. Chhowalla, Surf. Coat. Technol. **79**, 76 (1996).
[22]R. Lossy, D. L. Pappas, P. A. Roy, J. J. Cuomo, and V. M. Sura, Appl. Phys. Lett. **61**, 171 (1992).
[23]S. Falabella, D. B. Boercker, and D. M. Saunders, Thin Solid Films **209**, 165 (1992).
[24]A. Anders, S. Anders, I. G. Brown, M. R. Dickinson, and R. A. MacGill, J. Vac. Sci. Technol. B **12**, 815 (1994).
[25]H. J. Scheibe and B. Shultrich, Thin Solid Films **246**, 92 (1994).
[26]H. D. Li and X. M. He, MRS Bull. **17**, 1415 (1994).
[27]X. M. He, W. Z. Li, and H. D. Li, J. Vac. Sci. Technol. A **11**, 2964 (1993).
[28]P. J. Fallonk, V. S. Veerasamy, C. A. Davis, J. Robertson, F. A. J. Amaratunga, W. I. Milne, and J. Koskinen, Phys. Rev. B **48**, 4777 (1993).
[29]V. S. Veerasamy, G. A. J. Amaratunga, W. I. Wilne, P. Hewitt, P. J. Fallon, D. R. McKenzie, and C. A. Davis, Diamond Relat. Mater. **2**, 782 (1993).
[30]M. Yoshikawa, Mater. Sci. Forum **52–53**, 365 (1989).
[31]S. Prawer, K. W. Nugent, Y. Lifshitz, G. D. Lempert, and E. Grossman, Diamond Relat. Mater. **5**, 433 (1996).
[32]D. R. McKenzie, D. A. Muller, E. Kravtchinskaia, D. Segal, D. J. H. Cockayne, G. Amaratunga, and R. Silva, Thin Solid Films **206**, 198 (1991).
[33]A. Grill and B. S. Meyerson, in *Synthetic Diamond: Emerging CVD Science and Technology*, edited by K. E. Spear and J. P. Dismukes (Wiley, New York, 1994).

High aspect ratio all diamond tips formed by focused ion beam for conducting atomic force microscopy

A. Olbrich,[a] B. Ebersberger, and C. Boit
Siemens Semiconductor Division HL FA, 81730 Munich, Germany

Ph. Niedermann and W. Hänni
Swiss Center for Electronics and Microtechnology, Incorporated, CH-2007 Neuchâtel, Switzerland

J. Vancea and H. Hoffmann
Institute for Experimental and Applied Physics, University of Regensburg, 93040 Regensburg, Germany

(Received 14 December 1998; accepted 14 May 1999)

Conductive (boron doped) all diamond tips are best suited for use as electrical probes for scanning probe experiments due to their hardness, chemical inertness, and resistivity against wear. In order to overcome the problems of image distortion induced by the tip shape the aspect ratio of the diamond tips was increased to about 7:1 by focused ion beam milling, maintaining a tip radius of typically 30 nm at the apex. The application of these sharpened tips to conducting atomic force microscopy for local electrical characterization of thin metal–oxide–semiconductor dielectrics is demonstrated. Excess current flow was detected at the transition region between the gate oxide and the bordering field oxide due to oxide thinning introduced by the local oxidation of silicon process. © *1999 American Vacuum Society.* [S0734-211X(99)06504-X]

I. INTRODUCTION

Since its invention in 1985 atomic force microscopy (AFM) has become an important tool for three-dimensional imaging of surfaces.[1] The development of further AFM-based techniques like conducting-AFM (C-AFM), scanning capacitance microscopy (SCM), electric force microscopy (EFM), and surface potential microscopy (SPM) allows the electrical characterization of samples on a nanometer length scale in addition to its topography by detecting local conductivity, capacitance, and surface potential.[2–7] The continuously increasing integration density of integrated circuits (ICs) of very large scale integration (VLSI) and ultra-large scale integration (ULSI) areas on the one hand and the lack of adequate conventional analysis tools for nanoscale semiconductor characterization on the other have opened up a wide field of applications in the semiconductor industry for these new scanning probe techniques, especially in failure analysis, research, and development.[8] One of the most promising tools seems to be C-AFM, because of its inexpensive setup and versatility for various extremely important device problems, like the local electrical characterization of thin dielectrics of metal–oxide–semiconductor (MOS) structures and the delineation of doping profiles.[9,10]

All the AFM-based techniques above require conductive tips and thus must have tips with sufficient conductivity and long term resistance against wear which are essential for reliable, reproducible, and quantitative topographical and electical measurements.

Usually conductive tips are made by coating a Si or Si_3N_4 tip with a metallic (Pt, PtIr, CoCr, Au, etc.) or other well-conducting film (e.g., boron doped diamond) with thicknesses ranging from 25 to 100 nm. The problems encountered with metallic films are usually the fast mechanical wear during scanning and hence the loss of conductivity at the tip apex.[9,11] This problem is especially severe for C-AFM and SCM, because both techniques are based on contact mode AFM, where higher lateral forces act on the tips in contrast to tapping or noncontact mode as used for EFM and SPM. Also, the surface structure plays a crucial role—rough surfaces with steep features lead to faster degradation of the tip and coating than flat surfaces. Often the coating is worn off the apex and the sidewalls of the tip in the scan direction after a short period of imaging. Additionally, tip sample currents in the nA range and higher can lead to melting and evaporation of the metallic coating.[12] All hard metallic coatings (PtIr, Ti, CoCr, a-C, TiN) we tested showed these problems. Due to surface oxides they usually show a metal–isolator–metal (MIM) like *I-V* curve on Au samples. The only coating which exhibits good conductivity (linear *I-V*) and resistance against wear was a 100 nm thick film of boron doped diamond (0.1–0.2 Ω cm).[9–14] These tips have been successfully used in C-AFM for microscopic investigation of thin oxides and for nano-scanning resistance profiling (nano-SRP). Unfortunately, the tip radius at the apex is on the order of 100 nm, making the tip less useful to resolve rough topographies. A way to overcome the problem of diminishing conductivity during wear is to use tips made entirely of conductive material. Doped Si tips are not suitable because of the native oxide, which forms under ambient conditions and the possible growth of anodic oxide underneath the tip.[11] All metal tips also show a surface oxide (except Au and Pt tips) and usually the tip apex gets blunt very fast making these tips not useful for high resolution imaging.

AFM measurements of semiconductor structures, either deprocessed from fully processed wafers to expose the surface of certain parts or layers of the IC or pulled directly after the processing steps of interest, often show structured

[a]Electronic mail: Alexander.Olbrich@siemens-scg.com

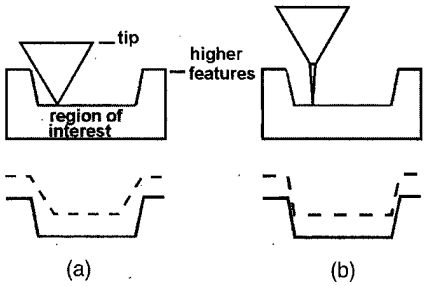


FIG. 1. Image distortion encountered in scanning probe microscopy. (a) Effect of a pyramidal shaped tip: The region of interest (bottom: solid line) is traced only partially (dashed line). (b) A high aspect ratio tip gives access to the full region of interest. The measured trace (dashed line) follows the topography (solid line) exactly.

surfaces and the regions that are to be characterized may be embedded between higher (tens to hundreds to nanometers) features, as schematically shown in Fig. 1.

In order to get complete electrical and topographical information, tips with a high aspect ratio and small tip radius are of primary importance. Using conventional pyramidal tips, shown in Fig. 1(a), leads to distortion of the image induced by the shape of the tip.[15] High aspect ratio tips can help to solve this problem and can get complete information on the region of interest. This is illustrated in Fig. 1(b). The dashed line corresponds to the measured topography and, ideally, traces the real topography (solid line) exactly.

In this article, we report on conductive all diamond tips which can overcome the problems stated above. These tips have been modified by focused ion beam (FIB) milling, which leads to a high aspect ratio and a small tip radius. This makes them very suitable for topographical and electrical characterization of flat as well as rough semiconductor surfaces. The application of these tips to conducting-AFM measurements for nanoscale electrical characterization of thin oxides is demonstrated.

II. EXPERIMENT

All diamond cantilevers and tips were fabricated as described in Ref. 14. In short, the structures were made by depositing a 1 μm chemical vapor deposited (CVD) layer on a silicon wafer containing (111) plane bounded pyramidal molds and the cantilevers were structured by reactive ion etching. The CVD diamond was *in situ* boron doped, resulting in a resistivity of 0.02 Ω cm. The diamond levers and supporting membranes were underetched using KOH. The backside of the lever was gold coated to provide sufficient optical reflectivity for the AFM cantilever deflection detector.

The AFM cantilevers were then introduced into the working chamber of a focused ion beam apparatus (Micrion 2100). The sample holder is capable of carrying a full 3 in. wafer of micromachined cantilevers. The *x-y* stage is programmable and allows batch fabrication of the tips. The principle of tip sharpening by FIB milling is schematically shown in Fig. 2. The liquid gallium ion source produces a 30 keV Ga^+ ion beam with a spot of about 10 nm. Several

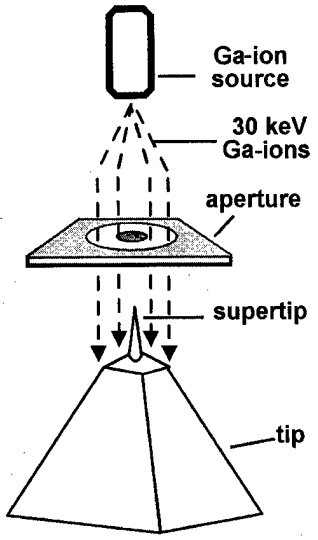


FIG. 2. Principle of focused ion beam tip sharpening. The software design and adjustable aperture allow the Ga ion beam to mill only at predefined parts of the tip, i.e., at the white angular region of the aperture.

mechanical apertures can be selected to choose a proper beam current between 1 pA and 8 nA. The cantilever is adjusted perpendicular to the ion beam in order to maintain a conformal milling of the sidewalls of the tip. A software programmable aperture is chosen to define the area scanned by the beam. In principle, the focused ion beam traces the field of view of the tip and the beam passes very fast over those locations so that the software aperture is not allowed to mill. In Fig. 2, for example, the beam only mills at the white area of the aperture, but does not markedly affect the gray areas. We use a circular disk as the aperture, which only allows the beam to mill the sidewalls, i.e., the white area of the aperture in Fig. 2, and the beam leaves out the very end of the tip and the outer side of the disk (gray regions). This configuration leads to an effective sharpening of the tip and yields a tip with a high aspect ratio. The shape and the aspect ratio of the tip are mainly adjustable by the size and shape of the aperture, the beam current, and the etching time. Usually three milling steps are performed. First, the rough shape of the tip is defined by a fast mill at a relatively large beam current of 100 pA. To protect the tip apex the unexposed region at the tip apex has a diameter of about 500 nm. The next step uses reduced current and aperture size. For the final etch a current of 4 pA and an inner aperture diameter of about 100 nm show the best results and yield high aspect ratio tips with tip radii on the order of 30–40 nm.

Complete milling of a tip in the way described above requires about 15 min. Batch fabrication and a further improved milling process would further reduce this time. Due to the good conductivity of the diamond film, electrostatic charging of the probes did not cause any problems. After modification by the FIB the membranes with the cantilevers were glued on to support chips. Milling of the tip leaves a 10–20 nm thick nonconducting amorphous carbon layer which must be etched by pure HNO_3 (99.9%) to expose the diamond surface. AFM and C-AFM measurements were per-

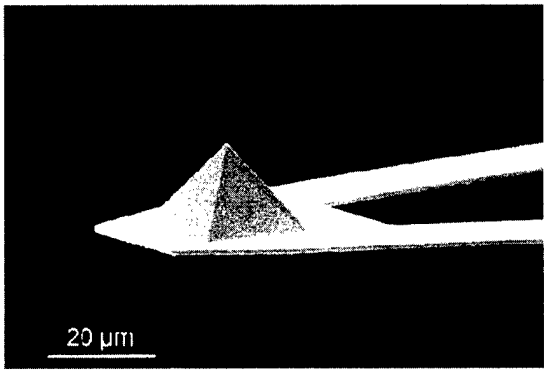


FIG. 3. Pyramidal all diamond probe with an integrated cantilever (tip height 15 μm).

formed on a modified contact mode AFM.[16] Details of the experimental setup have been described previously (see, for example, Refs. 3 and 9).

III. RESULTS

A. All diamond tips

Diamond as a tip material has several advantages. Because of the hardness of diamond (Knoop hardness=6000–9000) it is used as a reference for nearly all material hardness tests. Diamond is highly hydrophobic and reduces capillary forces acting on the tip under ambient conditions. Due to the chemically inert nature of the material, diamond tips can be used as electronic probes at weak contact forces in AFM experiments. Figure 3 shows a cantilever with a pyramidal tip made in the fashion described above. Three V-shaped levers with lengths between 150 and 500 μm and spring constants ranging from 0.08 to 6 N/m, corresponding to resonance frequencies of 11–110 kHz, are placed on one side of a chip. The tip height is 15 μm. At relatively low contact forces ($\sim 10^{-7}$ N) contact resistances of less than 1 MΩ were measured and they can be further decreased by higher contact forces.[10] Figure 4 shows a high resolution scanning electron microscope (SEM) image of a pyramidal all diamond tip. The tip radius is about 10 nm. Tip radii on the order of 10–20 nm are routinely obtained.

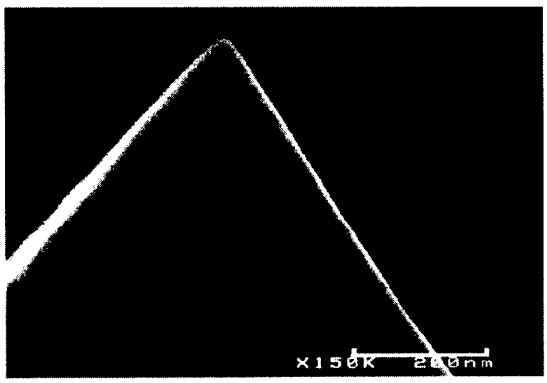


FIG. 4. High resolution SEM image of the tip apex of the pyramidal CVD all diamond tip. The tip radius is about 20 nm.

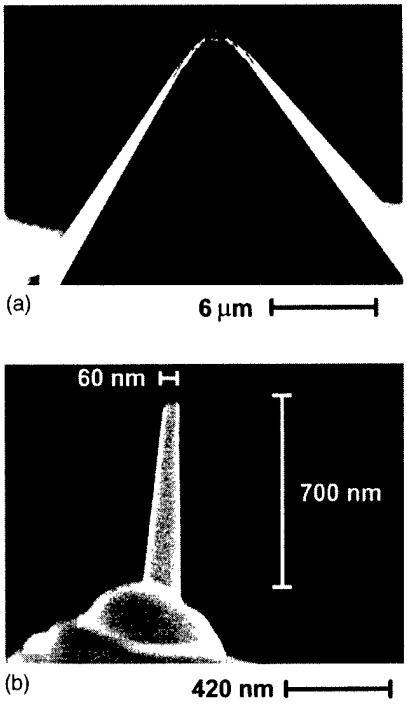


FIG. 5. SEM images of the FIB modified all diamond tip. (a) SEM image showing the supertip on top of the pyramidal tip (b) supertip at high resolution revealing an aspect ratio of 1:7.

Taking these tips as the source material, high aspect ratio tips are fabricated by FIB milling. The typical shape of these tips is shown in the SEM image of Fig. 5(a). In order to measure in the deep grooves of the IC samples approximately 1 μm of the sidewalls of the pyramidal tip is cut away, leaving a high aspect ratio tip. A SEM image of this tip is shown in Fig. 5(b). The tip radius is about 30 nm, the height 700 nm, and the aspect ratio is 1:7. Still higher values of the aspect ratio can also be obtained. The upper limit is mainly determined by the diamond film thickness. Since a FIB can also be used for imaging, one can easily measure the tip radius after milling and decide whether or not the aspect ratio or the tip radius is satisfactory. In the latter case the milling procedure for the tip can be continued to fulfill the properties desired.

The tips show very good long term resistance against wear. Contact AFM measurements with forces up to several 10^{-7} N on surfaces with steep features did not lead to breaking of the end of the tip. Figures 6(a) and 6(b) show the same structure, a gate oxide with an extension of about 1 μm×0.5 μm surrounded topographically ~300 nm higher field oxide, imaged with a pyramidal tip and the FIB modified tip, respectively. The structure was deprocessed from a fully processed IC structure by lapping it parallel to the wafer surface until the polysilicon (thickness ~500 nm), which acts as the gate of a MOS transistor, is reached. Afterwards the polysilicon was selectively etched by KOH (10%) at room temperature to expose the gate oxide surface beneath the gate (selectivity >2000:1). The distortion of the image due to the shape of the tip is clearly visible in Fig. 6(a). A large part of the gate oxide near the edge is inaccessible and cannot be inves-

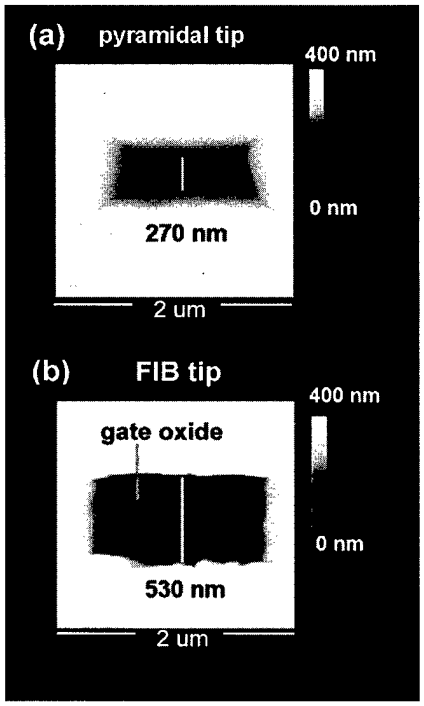


FIG. 6. Topography of a MOS gate oxide embedded in the bordering field oxide after deprocessing. (a) Imaged with a pyramidal tip, and (b) imaged with the FIB modified high aspect ratio tip.

tigated. The width of the imaged gate oxide area is about 270 nm. In contrast, the high aspect ratio tip traces the topography much more accurately and images the entire gate oxide surface (width: 530 nm) for structural and electrical characterization.

B. Application: Conducting atomic force microscopy measurements

With the FIB modified diamond tips C-AFM measurements are performed to locally characterize device grade MOS gate and tunneling oxides. The quality and homogeneity of these dielectrics is of paramount importance for device reliability. Weak spots, like oxide thinning, induced by the local oxidation of silicon process (LOCOS), lead to locally enhanced electric fields and Fowler-Nordheim (FN) currents, which degrade the gate oxide. Finally, intolerable leakage currents are generated and/or dielectric breakdown occurs. In order to prevent early device failure, an analysis tool for the local electrical characterization of the oxide quality is of primary importance. C-AFM seems to be the most adequate technique for a nanoscale investigation of these oxides.[9,17] C-AFM offers lateral resolution on the order of 10 nm.[9] The absolute and relative accuracies for oxide thickness determination are ± 0.3 and ± 0.15 nm, respectively.[3] In Fig. 7 the principle of C-AFM for oxide characterization is shown schematically.

A constant voltage is applied between the conductive tip and the Si substrate across the oxide and Fowler-Nordheim tunneling currents are measured with a pA amplifier simultaneously with the oxide topography. This results in a two-

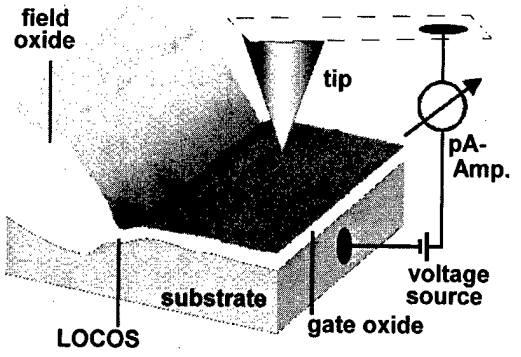


FIG. 7. Experimental setup for C-AFM on MOS gate oxides. A voltage is applied across the oxide and the corresponding Fowler-Nordheim current is measured with a pA amplifier.

dimensional map of the topography and the FN current. The tip is positively biased with respect to the substrate.[9] As already shown in Fig. 6, the shape of the tip, i.e., the tip radius and aspect ratio, is essential for reliable measurements on deprocessed as well as on partially processed IC structures.[3,9]

Ill-processed gate oxides often exhibit significant oxide thinning at the LOCOS groove, i.e., transition region between the gate oxide and the bordering thick field oxide, as indicated in Fig. 7. C-AFM is used to detect and quantitatively measure the oxide thinning. This is demonstrated in Fig. 8. The structure was deprocessed from a fully processed test capacitor and consists of a 20 nm thick gate oxide. The topographical image of Fig. 8(a) reveals the characteristic

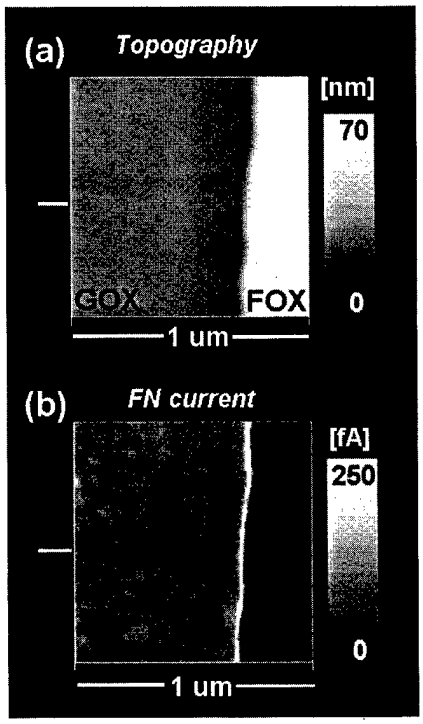


FIG. 8. C-AFM measurement on a 20 nm thick MOS gate oxide. (a) Top view of the oxide topography (gray: gate oxide, white: field oxide) and (b) simultaneously measured FN current image at a tip sample voltage of 18.7 V. The white marks on the left side indicate the linescans of Fig. 9.

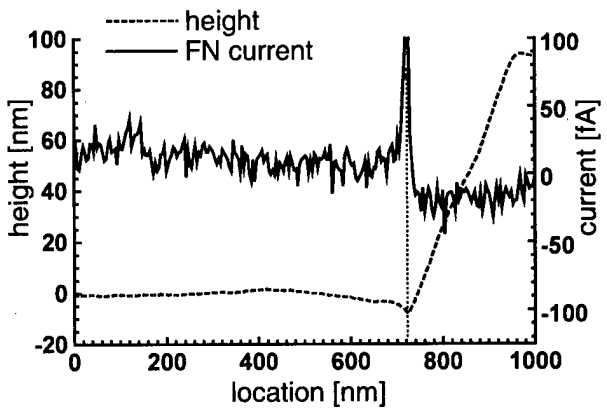


FIG. 9. Example of a linescan through Figs. 8(a) and 8(b) at the same location. The FN current (line) shows a peak value of ~100 fA at the minimum of the height trace (dashed curve).

shape of the LOCOS groove between the gate oxide and the field oxide. The simultaneously recorded FN current image is shown in Fig. 8(b). The tip voltage was 18.7 V and the scan frequency was 0.2 Hz. Clearly an enhanced current flow at the LOCOS groove is detected, whereas on the gate and the field oxide the current is below the detection limit. Local *I-V* curves obtained at the LOCOS groove and on the gate oxide and their corresponding fits to the FN current equation yield an oxide thickness decrease of up to 3.3 nm at the LOCOS groove with respect to the gate oxide thickness.[3]

Linescans through Figs. 8(a) and 8(b) taken at the same location are shown in Fig. 9. The FN current (line) shows a sharp increase to a peak value of ~100 fA within the 20 nm extension at the topographically deepest value, i.e., at the minimum of the LOCOS groove (dashed curve). Small currents are already detectable on the gate oxide. The current noise has a root mean square (RMS) value of about 10 fA. Also, a very small negative current is measured on the field oxide. It is due to a change in tip-sample capacitance (~0.5 pF) when scanning the steep topographical step at the field oxide, which results in a small displacement current.

There also have been several attempts with conventional tips and metallic coatings to measure the oxide thinning on this sample, but without success. Up to now only the FIB modified bulk diamond tips exhibit the overall good conductivity, small tip radius, high aspect ratio, and mechanical stability to reproducibly measure the oxide thinning at the LOCOS groove.

IV. CONCLUSION

Focused ion beam modification of all diamond tips has been demonstrated to yield high aspect ratio tips with small tip radii, which can be used as stable electrical probes for measurements of semiconductor IC structures. This was demonstrated by the application of conducting-AFM on a 20 nm MOS gate oxide. The Fowler-Nordheim current image reveals weak spots due to oxide thinning at high lateral resolution (~10 nm) and allows a direct correlation with oxide surface features. The tips also have been successfully tested for the application to scanning capacitance microscopy. Moreover, the high aspect ratio will be advantageous, either due to the mechanical stability or the good conductivity and low stray capacitance of the tips, for other AFM-based techniques for electrical as well as structural characterization of semiconductor devices. Further useful applications include nanolithography, adhesion investigations, and all techniques that require a stable tip with precise knowledge of the tip shape. The last can be prepared by focused ion focused ion beam milling in a wide variety of methods.

ACKNOWLEDGMENT

This work was partially supported by the German Bundesministerium für Forschung und Technik.

[1]G. Binnig, C. F. Quate, and Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986).
[2]S. J. O'Shea, R. M. Atta, M. P. Murrell, and M. E. Welland, J. Vac. Sci. Technol. B **13**, 1945 (1995).
[3]A. Olbrich, B. Ebersberger, and C. Boit, Appl. Phys. Lett. **73**, 3114 (1998).
[4]P. DeWolf, T. Clarysse, W. Vandervorst, L. Hellemans, Ph. Niedermann, and W. Hänni, J. Vac. Sci. Technol. B **16**, 355 (1998).
[5]K. M. Mang, Y. Khang, Y. J. Park, Y. Kuk, S. M. Lee, and C. C. Williams, J. Vac. Sci. Technol. B **14**, 1536 (1996).
[6]Y. Martin, D. W. Abraham, and H. K. Wickramasinghe, Appl. Phys. Lett. **52**, 1103 (1988).
[7]N. Nonnenmacher, M. P. O'Boyle, and H. K. Wickramasinghe, Appl. Phys. Lett. **58**, 2921 (1991).
[8]J. A. Dagata, A. C. Diebold, C.K. Shih, and R. J. Colton, M.D., *Workshop Summary Report: Industrial Application of Scanned Probe Microscopy*, National Institute of Standards and Technology, Gaithersburg, Report No. 5550 (1994).
[9]A. Olbrich, B. Ebersberger, and C. Boit, *36th Annual 1998 IEEE International Reliability Physics Proceedings, Reno, NV, 1998* (IEEE, New York, 1998), p. 163.
[10]J. Snauwaert, N. Blanc, P. DeWolf, W. Vandervorst, and L. Hellemann, J. Vac. Sci. Technol. B **14**, 1513 (1996).
[11]S. J. O'Shea, R. M. Atta, and M. E. Welland, Rev. Sci. Instrum. **66**, 2508 (1995).
[12]M. A. Lantz, S. J. O'Shea, and M. E. Welland, Rev. Sci. Instrum. **69**, 1757 (1998).
[13]Ph. Niedermann, W. Hänni, N. Blanc, R. Christoph, and J. Burger, J. Vac. Sci. Technol. A **14**, 1233 (1996).
[14]Ph. Niedermann, W. Hänni, D. Morel, A. Perret, N. Skinner, P. -F. Indermühle, N. F. de Rooij, and P. -A. Buffat, Appl. Phys. A: Mater. Sci. Process. **66**, 31 (1998).
[15]G. Reiss, J. Vancea, H. Wittmann, J. Zweck, and H. Hoffmann, J. Appl. Phys. **67**, 1156 (1990).
[16]Digital Instruments, 112 Robin Hill Road, Santa Barbara, CA 93103.
[17]B. Ebersberger, C. Boit, H. Benzinger, and E. Günther, *1996 IEEE International Reliability Physics Proceedings, 34th Annual 1996* (IEEE, New York, 1996), p. 126.

Sensitivity analysis of the field emitter

Lei Wei,[a] Wang Baoping, and Yin Hanchun
Department of Electronic Engineering, Southeast University, Nanjing 210096, People's Republic of China

Guohui Li
Acroloop Motion Control System, Inc., Chaska, Minnesota 55318

(Received 7 July 1998; accepted 30 April 1999)

Many types of field emitters have been studied up to now to improve the emission current. However, only the ideal parameters are taken into account in some designs. In the fabrication process, the structural parameters of emitters may change. Therefore, it is useful to analyze the parameter sensitivity of the field emitter. This sensitivity should also be included in the defect function of the optimum design. In this article we propose a method to calculate the derivative of the electric field and emission current density versus geometrical parameters of emitter. This derivative can be used to represent the parameter sensitivity of the field emitter. Using this method, the sensitivities of some parameters in the diode and triode emitters are analyzed. © *1999 American Vacuum Society.* [S0734-211X(99)02604-9]

I. INTRODUCTION

With the development of vacuum microelectronics, many scholars have studied different types of field emitters to obtain high current density.[1–4] In the fabrication of field emitters, the practical parameters are usually somewhat different than the values in the design. Sometimes, these differences will greatly influence the performance of the field emitter.[5] Therefore, it is important to analyze the sensitivity of the parameters in the field emitter. If the performance of the emitter is very sensitive to the parameters during fabrication, the emission current may change considerably in the practical application. In the optimization design of a field emitter, the defect function should include the influence of parameter sensitivity.

In this article the propose a method to analyze the parameter sensitivity of field emitters. First, the derivative of potential versus geometrical parameters that we are interested in is calculated. Then, the change in electric field can be obtained. Finally, we calculate the sensitivity of the emission current density according to the Fowler–Nordheim (FN) equation. Using this method, we studied the parameter sensitivities of both the diode emitter and the triode emitter.

II. SIMULATION MODEL

As mentioned in Sec. I, an analysis of the parameter sensitivity is important in the design of the field emitter. Usually, we can use the derivative of current density versus the geometrical parameter to represent the sensitivity T_s,

$$T_s = dJ/dp, \tag{1}$$

where J is the current density of field emitter, and p is the geometrical parameter of the field emitter which has a small perturbation during fabrication.

The parameter sensitivity can be obtained by the finite difference at various points of parameter p,

$$dJ/dp \approx \frac{J_{p+\Delta p} - J_p}{\Delta p}. \tag{2}$$

Because the perturbation of parameter p is usually small, this calculation may introduce serious numerical errors in the result.[6,7]

In this article we use another way to predict dJ/dp. First, the derivative of the potential versus parameter p is obtained from the unperturbed result. Because the emission activity takes place in a very tiny area of the emitter, the finite difference method with an irregular mesh is employed to calculate the distribution of the potential of the field emitter. The potential of the grid can be obtained from the Laplace equation if the space charge effect is ignored.

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = 0. \tag{3}$$

The irregular mesh is shown in Fig. 1.

By ignoring the tremendous procedure, the potential of the grid (i,j,k) collapses into the simple form

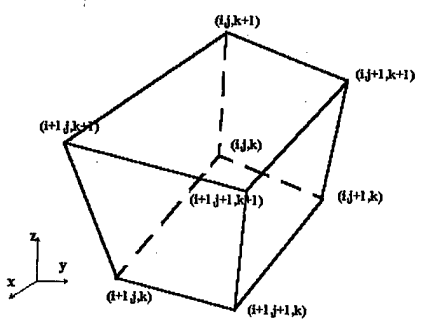


FIG. 1. Irregular mesh in the calculation.

[a]Electronic mail: lw@seu.edu.cn

$$\varphi_{(i,j,k)}=\frac{-1}{T_0}[T_1\varphi_{(i+1,j,k)}+T_2\varphi_{(i+1,j-1,k)}+T_3\varphi_{(i+1,j,k-1)}+T_4\varphi_{(i,j+1,k)}+T_5\varphi_{(i-1,j+1,k)}+T_6\varphi_{(i,j+1,k-1)}$$
$$+T_7\varphi_{(i-j,k+1)}+T_8\varphi_{(i-1,j,k+1)}+T_9\varphi_{(i,j-1,k+1)}+T_{10}\varphi_{(i-1,j,k)}+T_{11}\varphi_{(i,j-1,k)}+T_{12}\varphi_{(i,j,k-1)}], \quad (4)$$

where coefficients T_0-T_{12} are some parameters that are determined by the coordinates of the grids.

From Eq. (4), we can obtain

$$\frac{\partial\varphi_{(i,j,k)}}{\partial p}=\frac{1}{T_0^2}[T_1\cdot\varphi_{(i+1,j,k)}+T_2\cdot\varphi_{(i+1,j-1,k)}+T_3\cdot\varphi_{(i+1,j,k-1)}+T_4\cdot\varphi_{(i,j+1,k)}+T_5\cdot\varphi_{(i-1,f+1,k)}+T_6\cdot\varphi_{(i,j+1,k-1)}$$
$$+T_7\cdot\varphi_{(i-j,k+1)}+T_8\cdot\varphi_{(i-1,j,k+1)}+T_9\cdot\varphi_{(i,j-1,k+1)}+T_{10}\cdot\varphi_{(i-1,j,k)}+T_{11}\cdot\varphi_{(i,j-1,k)}+T_{12}\cdot\varphi_{(i,j,k-2)}]\cdot\frac{\partial T_0}{\partial p}$$
$$-\frac{1}{T_0}\left[\frac{\partial T_1}{\partial p}\cdot\varphi_{(i+1,j,k)}+\frac{\partial T_2}{\partial p}\cdot\varphi_{(i+1,j-1,k)}+\frac{\partial T_3}{\partial p}\cdot\varphi_{(i+1,j,k-1)}+\frac{\partial T_4}{\partial p}\cdot\varphi_{(i,j+1,k)}+\frac{\partial T_5}{\partial p}\cdot\varphi_{(i-1,j+1,k)}\right.$$
$$+\frac{\partial T_b}{\partial p}\cdot\varphi_{(i,j+1,k-1)}+\frac{\partial T_7}{\partial p}\cdot\varphi_{(i,j,k+1)}+\frac{\partial T_8}{\partial p}\cdot\varphi_{(i-1,j,k+1)}+\frac{\partial T_9}{\partial p}\cdot\varphi_{(i,j-1,k+1)}+\frac{\partial T_{10}}{\partial p}\cdot\varphi_{(i-1,j,k)}$$
$$\left.+\frac{\partial T_{11}}{\partial p}\cdot\varphi_{(i,j-1,k)}+\frac{\partial T_{12}}{\partial p}\cdot\varphi_{(i,j,k-1)}\right]-\frac{1}{T_0}\left[T_1\cdot\frac{\partial\varphi_{(i+1,j,k)}}{\partial p}+T_2\cdot\frac{\partial\varphi_{(i+1,j-1,k)}}{\partial p}+T_3\cdot\frac{\partial\varphi_{(i+1,j,k-1)}}{\partial p}\right.$$
$$+T_4\cdot\frac{\partial\varphi_{(i,j+1,k)}}{\partial p}+T_5\cdot\frac{\partial\varphi_{(i-1,j+1,k)}}{\partial p}+T_6\cdot\frac{\partial\varphi_{(i,j+1,k-1)}}{\partial p}+T_7\cdot\frac{\partial\varphi_{(i,j,k+1)}}{\partial p}+T_8\cdot\frac{\partial\varphi_{(i-1,j,k+1)}}{\partial p}$$
$$\left.+T_9\cdot\frac{\partial\varphi_{(i,j-1,k+1)}}{\partial p}+T_{10}\cdot\frac{\partial\varphi_{(i-1,j,k)}}{\partial p}+T_{11}\cdot\frac{\partial\varphi_{(i,j-1,k)}}{\partial p}+T_{12}\cdot\frac{\partial\varphi_{(i,j,k-1)}}{\partial p}\right]. \quad (5)$$

In Eq. (5), the derivative of T_0-T_{12} can be calculated from the derivative of the coordinate of the grid versus parameter p. Therefore, the $d\varphi/dp$ can be obtained with Eq. (5) after we know the $d\mathbf{r}/dp$ of every grid.

It is easy to determine the derivative of electric field $d\mathbf{E}/dp$ from $d\varphi/dp$. In the calculation of the field emitter, the current density is given by the FN equation,[8]

$$J=154\times10^{-6}\frac{E^2}{\phi t^2(y)}\exp\left(-6.87\times10^7\frac{\phi^{3/2}v(y)}{E}\right)\ \text{A/cm}^2, \quad (6)$$

where E is the electric field at the tip in V/cm, $v(y)$ and $t^2(y)$ are electric field dependent elliptical functions, ϕ is the work function of the emitter material in eV, and y, the image charge lowering contribution to the work function, is given by $y=3.79\times10^{-4}E^{1/2}/\phi$. For simplicity, it is generally assumed that $t^2(y)=1.1$ and $v(y)=0.95-y^2$. The parameter sensitivity can be obtained from Eq. (6) as

$$dJ/dp=1.54\times10^{-6}\left\{\frac{2E}{\phi\cdot t^2(y)}\cdot\frac{dE}{dp}+\frac{E^2}{\phi\cdot t^2(y)}\right.$$
$$\left.\times\left[-6.87\times10^7\left(\frac{\phi^{3/2}}{E}\cdot\frac{dv(y)}{dp}-\frac{\phi^{3/2}\cdot v(y)}{E^2}\cdot\frac{dE}{dp}\right)\right]\right\}$$
$$\cdot\exp\left(-6.87\times10^7\cdot\frac{\phi^{3/2}\cdot v(y)}{E}\right), \quad (7)$$

where $dv(y)/dp=-2ydy/dp$ and $dy/dp=3.79\times10^{-4}\times(1/2)(dE/dp/E^{1/2}\phi)$.

III. SENSITIVITY ANALYSIS OF THE FIELD EMITTER

A. Analysis of the diode model

The parameter sensitivity of the diode field emitter is analyzed using the method mentioned above. We only studied the cone-shaped structured emitter here. The simulation model is shown in Fig. 2. In the simulation, the radius of the tip is 10 nm, the height of the emitter is 4.1 μm, the distance between tip and anode is 3.0 μm, the emitter angle is 70°, the voltage of anode is 200.0 V and the work function is 4.5 eV.

We suggest that the apex of the emitter, z_a, the radius of the tip, r, the height of tip, h_c, and the angle of emitter, α,

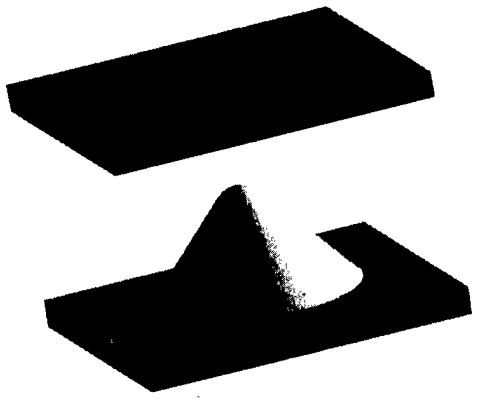


FIG. 2. Simulation model of the diode emitter.

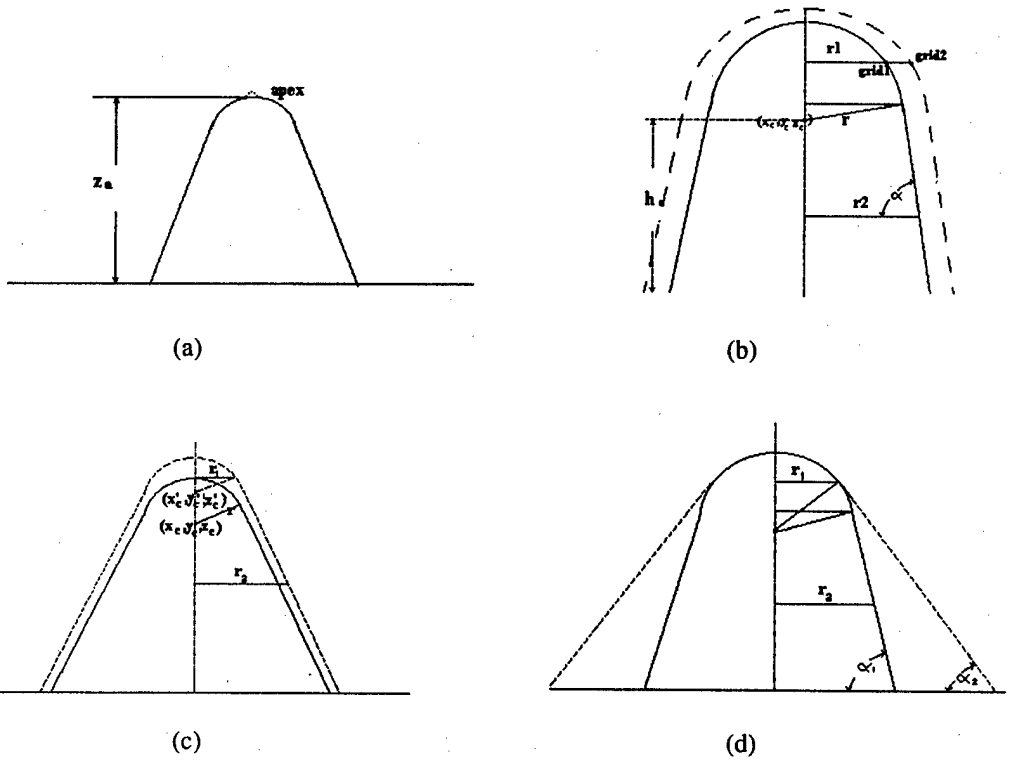


FIG. 3. Different emitter perturbations. (a) The apex has a displacement in the Z direction. (b) The radius of the tip has a perturbation. (c) The height of the tip has a displacement. (d) The angle α has a perturbation.

have perturbations. Figure 3 shows these variations.

In Fig. 3, parameters z_a, r, h_c, and α are used to describe the perturbations. After the perturbation, grid 1 moved to grid 2. From Fig. 3, we can obtain the following equations for Figs. 3(a)–3(b), respectively:

$$dx/dz_a=0.0 \quad dx/dz_a=0.0$$

$$dy/dz_a=0.0 \text{ at the apex}, \quad dy/dz_a=0.0 \text{ at other grids};$$

$$dz/dz_a=1.0 \quad dz/dz_a=0.0$$

$$\frac{dr_1}{dr}=\frac{r}{\sqrt{r^2-(z-z_r)^2}}, \quad \frac{dr_2}{dr}=1/\sin\alpha;$$

$$dr_1/dz_r=\frac{z-z_r}{\sqrt{r^2-(z-z_r)^2}}, \quad dr_2/dz_r=1/\tan\alpha; \tag{8}$$

$$dr_1/d\alpha=0.0, \quad dr_2/d\alpha$$

$$=-\frac{r\cdot\cos\alpha}{\sin^2\alpha}-(z-z_r)\left(1+\frac{\cos^2\alpha}{\sin^2\alpha}\right).$$

Using Eqs. (8), we can obtain the derivative of the coordinate of the grid from dr_1/dp and dr_2/dp. After calculation of the derivative of the coordinate, we use Eq. (5) to

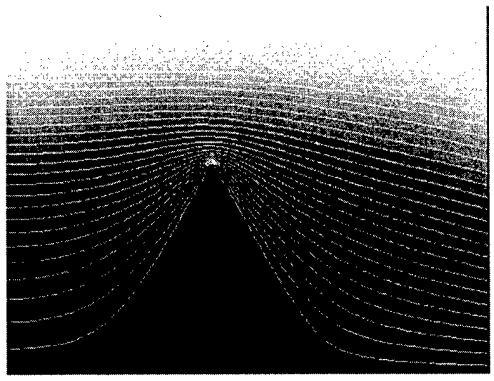


FIG. 4. Distribution of the unperturbed potential.

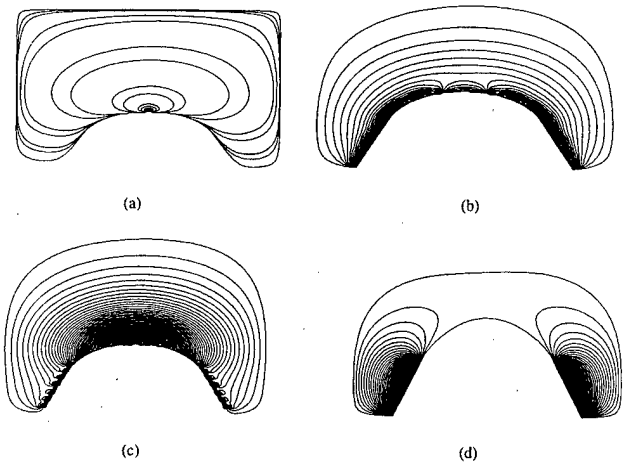


FIG. 5. Distribution of $d\varphi/dp$ near the tip of the emitter.

TABLE I. Derivative of the electric field and current density; z_a: the position of the apex; r: the radius of the tip; h_c: the height of the emitter; α: the angle of the emitter.

Parameter that has a perturbation	z_a	r	h_c	α
de/dp (V/μm^2)	6.5789×10^6	6.3280×10^6	6.3650×10^6	21.5042
dJ/dp (A/cm^3)	2.4462×10^{13}	2.3529×10^{13}	2.3667×10^{13}	7.9958×10^7

compute $d\varphi/dp$. The distribution of the unperturbed potential is shown in Fig. 4. The equal values of the curves of the $d\varphi/dp$ near the tip are shown in Fig. 5.

The derivatives of the electric field and current density are calculated using Eqs. (6) and (7). The results are listed in Table I.

In Table I, the unperturbed electric field is 4.6667×10^7 V/cm. The current density is 1.1304×10^5 A/cm^2. If the geometrical parameter p has a small perturbation, Δp, variations in the electric field and current density can be obtained by first order approximations $de/dp\cdot\Delta p$ and $dj/dp\cdot\Delta p$. As is shown in Table I, the electric field and current density are very sensitive to the position of the apex. When the apex has a displacement in the Z direction, it is close to the anode. Therefore, the electric field and current density increase greatly. In Figs. 3(b) and 3(c), the distance between the apex and the anode is decreased. Hence, the electric field and current density also increase. Nevertheless, the emitter becomes somewhat flat in Fig. 3(b), so dJ/dr is smaller than dJ/dh_r. In Fig. 3(d) the emitter becomes sharper and the angle α increases. The electric field changes greatly near the edge of the cone. In the center of the tip, the variation of the electric field is small. Because the angle α does not influence the radius of the tip and the distance between the apex and the anode, $dE/d\alpha$ is much smaller than dE/dz_a, dE/dr, and dE/dh_c. The current density is not very sensitive to angle α.

B. Sensitivity analysis of the triode model

In this article, we have also calculated the triode model (Fig. 6). It is assumed the apex of the emitter, the radius of the gate, angle α and the radius of the emitter r_b have small perturbations. In the calculation of the triode model, the radius of the tip is 10 nm, the radius of the cylinder is 0.05 μm, the radius of the gate is 0.3 μm, the height of the emitter is 0.605 μm, the angle of the cone is 31.8°, the thickness of the grid is 0.2 μm, the distance between the grid and the anode is 1.685 μm, the voltage of the grid is 60 V, the voltage of the anode is 200.0 V and the work function is 4.0 eV.

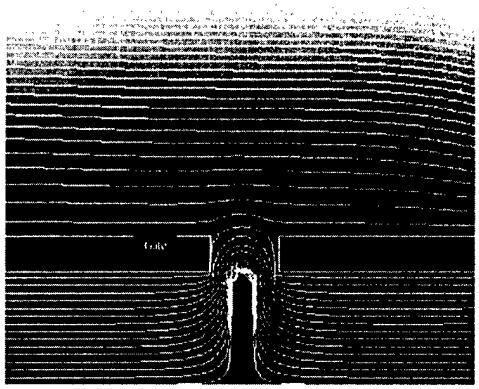


FIG. 7. Distribution of the unperturbed potential of the triode emitter.

The distribution of the unperturbed potential is shown in Fig. 7. The contour curves of $d\varphi/dp$ are shown in Fig. 8.

In the calculation, we suggest that there is a free region on the right side of emitter, so the contour curves are not symmetric in Fig. 8(a). From Figs. 7 and 8 we can determine in which part the potential and electric field change the most. The derivatives of the electric field and current density at the tip are shown in Table II. In Table II, the unperturbed electric field is 2.3×10^7 V/cm, and the current density is 3.56 A/cm^2.

From Table II, it can be seen that the current density increases greatly when the apex has a displacement in the Z direction. This is the same as what we found in the calculation of the diode model. Because the potential of the gate is 60 V, there is an electric field between the gate and the tip of the emitter. When the radius of the gate increases, the distance between the tip and the gate electrode increases. Therefore, the electric field at the tip becomes weaker and the current density also decreases. In Fig. 8(c), the angle α increases, so the emitter becomes sharper and the electric field

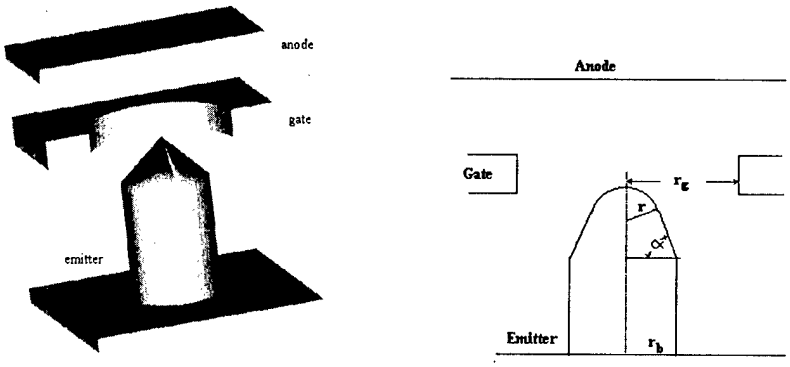


FIG. 6. Simulation model of the triode emitter.

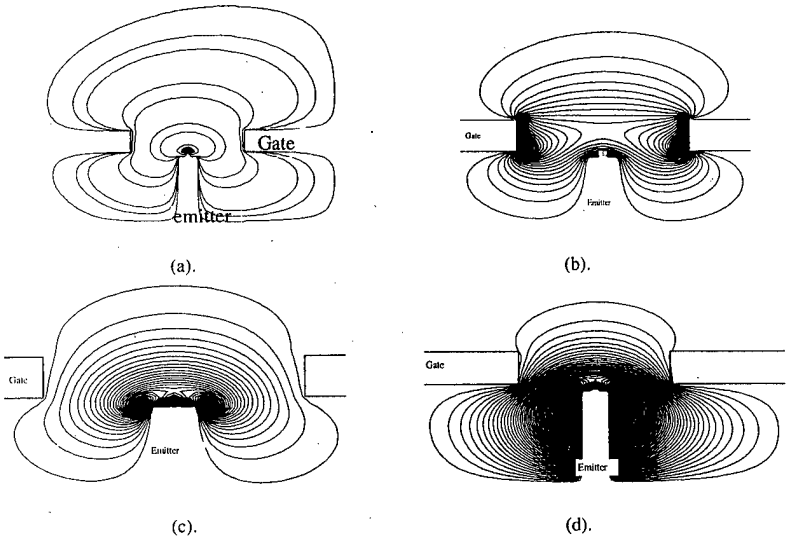


FIG. 8. Distribution of $d\varphi/dp$ near the tip of the triode emitter. (a) Apex with a small displacement in the Z direction. (b) Radius of the gate with a perturbation. (c) Angle α with a perturbation. (d) Radius of emitter r_b with a perturbation.

becomes stronger. However, variations of the electric field are concentrated at the edge of the cone. The influence of α on the peak electric field at the tip is small. If the radius of emitter r_b increases, the emitter becomes flat. Consequently, the electric field and current density are reduced.

IV. CONCLUSION

In the fabrication process of a field emitter, the physical parameters may have some small variations. Therefore, it is important to analyze the parameter sensitivities of the field emitter. In this article we proposed a way to predict the geometrical change effects on the electric field and emission current density of the emitter. Using this method, the derivatives of the electric field and current density versus the parameters which have perturbations are used to represent the parameter sensitivity. First, the derivative of the potential versus the displacement vector is calculated from the unperturbed result. Then, the derivative of electric field can be obtained. Finally, the derivative of the current density versus the displacement vector is calculated in terms of the FN equation.

In this article, both the diode model and the triode model were analyzed. In the calculation of the diode model, the position of the apex, the radius of the tip, the height of the emitter and the emitter angle are assumed to have perturbations. From the results obtained in the simulation, it was found that the emission current density is most sensitive to the position of the apex. The variation in the angle is relatively weaker.

The influence of the apex, radius of the gate, emitter angle and radius of the cylinder was studied in the calculation of the triode model. As shown in the simulation results, the influence of different parameters is not the same. The position of the apex is still the most sensitive parameter that influences the current density.

TABLE II. Derivative of the electric field and current density in the triode model; z_a: the position of the apex; r_g: the radius of the gate; α: the angle of the emitter; r_b: the radius of the emitter.

Parameter that has a small perturbation	z_a	r_g	α	r_b
de/dp (V/μm^2)	1.1273×10^5	-670.54	277.9215	-4351.254
dJ/dp (A/cm^3)	4.3104×10^7	-2.5639×10^5	1.0627×10^5	-1.6638×10^6

ACKNOWLEDGMENTS

The authors would like to thank Jean Coolean and Gerard van Poppel for their encouragement and support.

[1]R. L. Hartman, W. A. Mackie, and P. R. Davis, J. Vac. Sci. Technol. B **12**, 754 (1994).
[2]R. B. Marcus, K. K. Chin, Y. Yuan, H. Wang, and W. N. Carr, IEEE Trans. Electron Devices **37**, 1545 (1990).
[3]Z. Cui and L. Tong, IEEE Trans. Electron Devices **40**, 448 (1993).
[4]D. Hong and M. Aslam, J. Vac. Sci. Technol. B **12**, 764 (1994).
[5]T. Asano, IEEE Trans. Electron Devices **38**, 2392 (1991).
[6]J. L. Coulomb and G. Meunier, IEEE Trans. Magn. **20**, 1894 (1984).
[7]S. Gitosusastro, J. L. Coulomb, and J. C. Sabonnadiere, IEEE Trans. Magn. **25**, 2834 (1989).
[8]C. A. Spindt, I. Brodie, L. Humphrey, and E. R. Westerberg, J. Appl. Phys. **47**, 5248 (1976).

High field characteristics of dielectric spacers in thin-film electrode vacuum gaps

Xianyun Ma and T. S. Sudarshan[a)]
Electrical and Computer Engineering Department, University of South Carolina, Columbia, South Carolina 29208

(Received 20 July 1998; accepted 28 May 1999)

This article reports a test setup designed to study the prebreakdown and breakdown characteristics of spacers relevant to field emission displays (FEDs). The electrical characteristics of different spacer structures, bridging vacuum gaps formed by two thin-film electrodes, were studied in an attempt to determine the maximum field to which such gaps can be stressed. An extra low light detection three dimensional imaging system using an intensified charge coupled device camera was developed to diagnose and record the different prebreakdown phases and the breakdown phenomenon. This system enabled us to obtain spatially resolved images of light activities emanating from the electrically stressed surfaces. The experimental results showed that the investigated insulation materials did not introduce significant leakage current under high field stress. Breakdown voltages of fiber, disk-shaped and thick-wall spacers (750 μm high) were 12–14, 13–16, and 15–20 kV, respectively. The insulation capability of the thick-wall structure was limited by the edge effect of the thin-film electrodes and the quality of the thin-film electrodes rather than processes intrinsic to the spacer material or its geometry. This work provides some guidance to the design of spacers for FEDs. © *1999 American Vacuum Society.* [S0734-211X(99)09204-5]

I. INTRODUCTION

In large area field emission displays (FEDs), solid-insulating spacers must be used to provide mechanical support between the cathode and anode electrodes. Different-structure spacers, such as spheres for low-voltage FEDs to high-aspect-ratio rails and strips (thin wall) for high-voltage FEDs, are used for this purpose.[1] These spacers must be thin, mechanically strong, vacuum compatible, exhibit low current leakage, be hidden between pixels and not affect the electron beamlets, while withstanding the high voltage necessary in a high-voltage display. Another consideration for the spacer is its electrical resistance. The spacer resistivity should be high enough to minimize the leakage current between the anode and cathode, minimizing the power consumption and it must also be low enough to dissipate the charge buildup around the spacer surface, avoiding the breakdown caused by the spacer surface charge buildup. The insulation consideration for FED spacers is complicated in light of the operational environment of FEDs: electron bombardment, outgassing, and the elevated operational temperature of the device. In fact, the reliability and performance of FEDs are strongly dependent on spacer performance. The higher the voltage a spacer can withstand, the higher will be the efficiency of the phosphor, and the higher will be the brightness and the contrast of the FED.

Without considering outgassing or sputtering from the phosphor anode and electron bombardment from the cold cathode in FEDs, the above insulation structure can be electrically simulated using different spacer geometries bridging two thin-film electrodes. Although insulators in vacuum have been extensively studied for a long time, the investigated insulators were generally confined to circular cylindrical disks and conical frustums bridging two solid electrodes, and no flashover studies have been reported for insulator-bridged gaps shorter than 2 mm or thin-wall structured spacers, which bridge two thin-film electrodes.[2–5] In the context of developing advanced spacer technology for FEDs, this article is the first reported study to address issues related to the prebreakdown and breakdown characteristics of spacers that bridge vacuum gaps (750 μm) formed between thin-film metal electrodes.

In this study, Teflon is mainly used as a spacer material. Although Teflon is not a practical material for FED spacers, it is used in this initial work as a reference since it is known to be a good electrical insulating material that can be fabricated easily into different structure spacers. Teflon material is somewhat flexible, therefore it is possible to form well-defined triple junction contacts (without voids) between spacers and thin-film electrodes. The triple junction is the region where the spacer, electrodes and vacuum meet. If there are some voids between the spacer and the electrodes, the electrical field in those voids will be enhanced, which results in a low breakdown voltage.[5] Hence the flashover values of Teflon spacers perhaps represent the upper range for the thin-wall spacers made of different materials that are relevant to FEDs—specifically hard, brittle, ceramic materials. These initial studies are necessary, since based on the knowledge gained from the high-field characteristics of these spacer structures, we will be in a good position to investigate high-aspect-ratio spacers made of ceramic materials with specific designs in the near future, in addition to taking into account the FED environment. The results presented here will provide the basis for high-aspect-ratio spacer design in FEDs.

[a)]Electronic mail: Sudarsha@ece.sc.edu

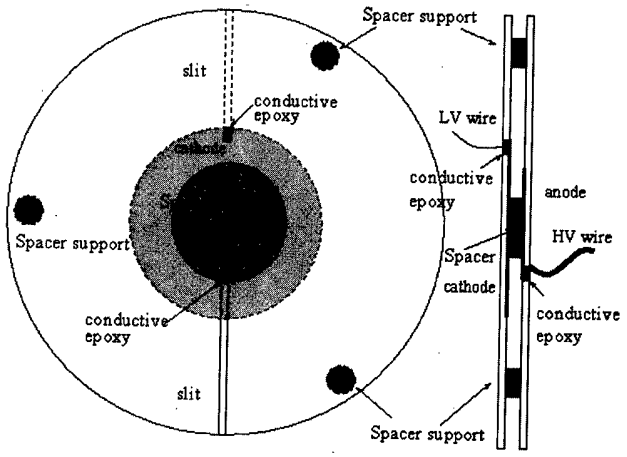


FIG. 1. Spacer setup.

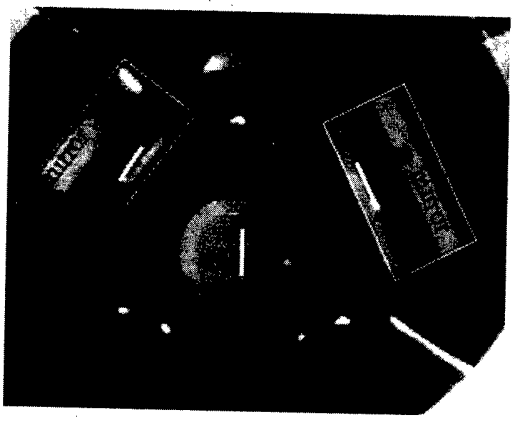


FIG. 2. Extra low light detection 3D imaging system. The white strip indicates the spacer location in the test gap and its reflections in the two mirrors.

II. EXPERIMENTAL SETUP

Three different spacer test structures were used: (1) a V-shaped thick-wall structure with 500 μm width and 750 μm thickness, made of Teflon, (2) a circular disk-shaped spacer with ~5 mm diameter and 750 μm thickness, made of Teflon, and (3) a plastic fiber with 750 μm diameter and 10 mm long, made of fluorine polymer. The V-shaped thick-wall spacers were used only for their ease in setting up the bridged vacuum gap.

The spacer structures in this study bridged a vacuum gap of 750 μm between two thin-film metal electrodes deposited on optically flat glass disks. Figure 1 shows the setup of the V-shaped thick-wall spacer. In order to simulate closely the electrode gap in FEDs, a special electrode structure was developed as shown in Fig. 1. To avoid the high field stress at the edge of the cathode, the diameter of the cathode (25 mm in diameter) was larger than that of the anode (17 mm in diameter) and a special means of attaching metallic leads to the metallized films was developed. The wire for each electrode was glued using conductive silver epoxy (EPO-TEK H20E) through a slit cut in each glass substrate. The tip of the metal lead was flush with the substrate surface on which the metal thin film was deposited. Since this method of attaching a metallic lead to the substrate unavoidably resulted in some surface defects on the electrode surface, the defect region on the cathode surface was designed to be away from the area opposite to the smaller diameter anode working area. The whole structure was also supported by three spacer supports outside the metal thin-film area, which had the same thickness and material as the spacer (Teflon).

Ni/Cr thin films were sputter deposited on optical glass plates via shadow masks to produce the electrode shapes as shown in Fig. 1. The substrates were heated to 150–200 °C in vacuum at 10^{-5} Torr. Argon gas was introduced into the reactor and the magnetron was excited by a dc power supply. The plasma reactor was operated at 140 W for 7 min to deposit a Ni/Cr metal film of 0.8–1.0 μm thickness on the glass plates.

After the deposition of the thin-film electrodes, no electrode pretreatment, e.g., baking, was applied. During the setup of the spacer, the thin-film electrodes were handled using ultralow-lint wipes. The test chamber was evacuated by a turbomolecular pump backed by an oil-free diaphragm pump and the final pressure in the chamber was ~1 $\times 10^{-6}$ Torr.

In order to diagnose and record the different prebreakdown phases and the breakdown phenomenon, and obtain spatially resolved images of light activities emanating from the electrically stressed spacers, an extra-low light detection three dimensional (3D) imaging system using an intensified charge coupled device (CCD) camera (Xybion ISG-750) was developed as shown in Fig. 2. The camera was positioned directly above the test gap such that the camera axis was parallel to the axis of the plane electrode gap. In order to obtain 3D information with the intensified CCD camera, two mirrors were used to reflect the two side views of the gap and the spacer in the direction of the camera's axis. Three white lines (two in the mirror and another in the work area) are drawn to indicate the relative position of the spacer (Fig. 2). For extra low light detection, the support for the entire setup was made of polyacetal (Delrin), a black material to avoid reflections, in addition to elimination of any background light. The spectral range of the Xybion ISG-750 CCD camera is 550–920 nm and sensitivity was better than 6 $\times 10^{-9}$ fc (7×10^{-15} W/mm^2). Furnished with a video system, we watched and recorded very weak prebreakdown light activity, then viewed it frame by frame. Due to the low frequency of the video signal, light activity shorter than the duration between two video frames could be missed.

The electrical prebreakdown and breakdown diagnostics and extra low noise high-voltage test setup were discussed in detail in our earlier article.[6]

III. RESULTS AND DISCUSSIONS

Thin-film (~1 μm thick) electrodes made of Ni/Cr were bridged by spacers 750 μm in height. Figure 3 shows their breakdown voltages and calculated average electrical fields, together with those values of the thin-film plain vacuum gap at the same gap distance.[7] Compared to the breakdown and

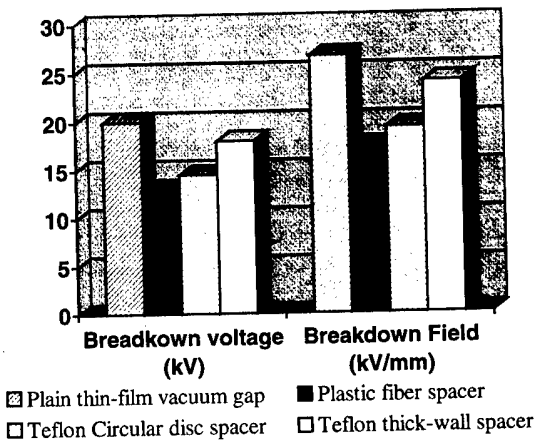


FIG. 3. Breakdown characteristics of different spacer structures vs a plain vacuum gap at 750 μm.

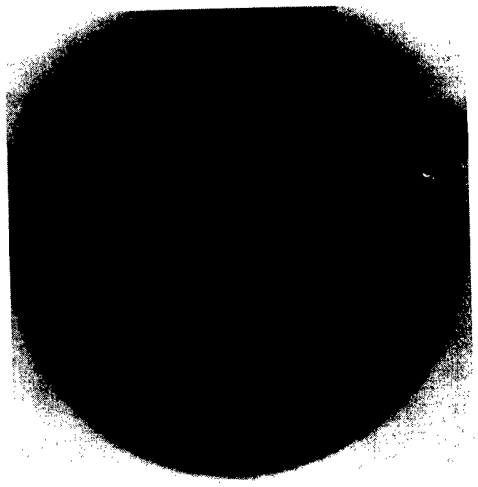


FIG. 4. Damaged area of the fiber spacer after catastrophic breakdown.

prebreakdown characteristics of a Ni/Cr thin-film plain vacuum gap at 750 μm, the bridged gaps at 750 μm gap distance have the following features.

A. Spacer leakage current

The vacuum gap leakage currents with the spacer materials were almost the same as that in a plain vacuum gap.[7] The gap equivalent resistance in both were on the order of 10^{14} Ω, implying that the surface leakage through the spacer surface with the selected materials was not significant. Leakage current consideration is important in spacer material selection. As stated above, the spacer resistivity must be high enough to minimize leakage current between anode and cathode, and it must also be low enough to dissipate any charge buildup around the spacer surface, avoiding breakdown caused by the spacer surface charge buildup. According to our present experimental results and previous experience with large dimension spacers, the surfaces of all traditional good insulating materials, such as Teflon, alumina (Al_2O_3) of different compositions and zirconia, will introduce very low leakage current. Hence the surface charging effect will be the main concern in our further investigations of FED spacers.

B. Fiber spacer

The breakdown voltage of the plastic fiber spacer is in the 12–14 kV range, corresponding to a gap spacing of 750 μm. The above breakdown voltages are below the threshold for the onset of prebreakdown current in the plain vacuum gap.[7] Breakdown occurred only at the end (tip) of the fiber with associated prebreakdown light activity. Figure 4 shows the damage produced by such a breakdown event at the tip of the fiber. In some cases, a spark breakdown occurred simultaneously at both ends of the spacer (Fig. 5). Hence, the weakest region in a fiber spacer is at its ends. It should be noted, however, that the fiber ends were not rounded. Attempts to round off the fiber spacer ends were not successful due to unavailability of proper tools. Since the shortest distance of the investigated spacer structures are the same, and the triple junction effect is the least likely cause of breakdown in the studied spacers, the low breakdown voltage was mainly due to the insulation material itself or its geometrical structure.

C. Disk-shaped spacer

For the disk-shaped spacer, 5 mm in diameter and 750 μm high, the breakdown voltage is in the 13–16 kV range. The characteristics were similar to those observed for the fiber spacers, i.e., breakdown occurred suddenly before the onset of a prebreakdown current. Considering that the disk shaped spacer results in a quasiuniform field, it is rather intriguing that the disk-shaped spacer did not exhibit superior behavior compared to the thick-wall structure (in Sec. III D). Since the spacer surface distance between the two electrodes is the same as in the other two cases (fiber and thick wall), one obvious explanation is the area effect or the larger triple-junction perimeter effect, since the disk-shaped spacer covers a larger area of the electrode.[5]

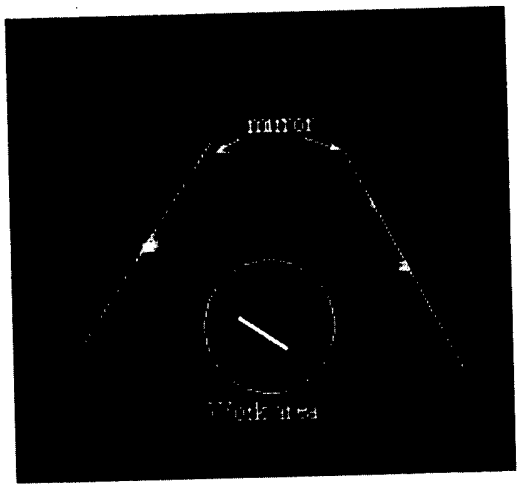
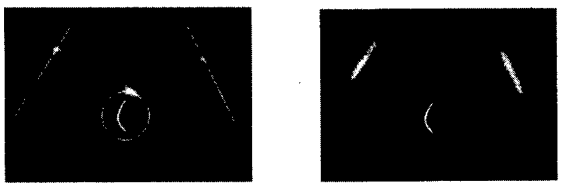


FIG. 5. Spark breakdown occurred at both ends of the plastic fiber spacer (the white line indicates the relative position of the fiber spacer).



(a) (b)

FIG. 6. Two different breakdown modes for Teflon thick-wall spacers; (a) breakdown at the edge of the thin-film electrodes; (b) breakdown across the spacer surface.

D. Thick-wall spacer

The width of the investigated thick-wall structure was 500 μm. For the Teflon thick-wall spacer, breakdown occurred in the 15–20 kV range, which was within the range of voltages at which the onset of prebreakdown current occurred in plain vacuum gaps (Fig. 3). Two distinct modes of breakdown were observed: flashover across the spacer surface or breakdown at the edge of the thin-film electrode (Fig. 6). Figure 6(a) shows breakdown at the edges of thin-film electrodes which were bridged by a thick-wall Teflon spacer. Figure 6(b) shows a flashover image observed across the two faces of the spacer wall. However, it is not clear if the flashover occurs at just one location on the spacer wall or if it is spread along the entire lateral length of the spacer. Another interesting result is that the prebreakdown characteristics followed Fowler–Nordheim (FN) behavior (Fig. 7). The FN plots are obtained in order to determine if the observed prebreakdown current is a result of field emission from the cathode.[5,8]

The phenomena of FN behavior and breakdown at the electrode edges indicate that the flashover voltage of thick-wall Teflon spacers approaches the breakdown strength of the thin-film plain vacuum gap. Hence it is imperative to address the problem of edge breakdown from the electrodes in these test structures. The insulation capability of the investigated structure can be improved by suppression of electrode edge breakdown via some thin-film edge treatment, perhaps even without any treatment of the spacer surface. Also, the above results imply that careful considerations should be given to the edge termination of the phosphor anode and the gate of the field emission arrays (FEAs) for higher voltage operation in FEDs.[7]

Because of the requirements imposed by the FED's vacuum and device structure, only ceramic insulating materials can be considered as spacers; and most likely the thin-wall structured spacers will be used to satisfy the optical invisibility requirement. According to the present study, having eliminated the triple junction process as the source of breakdown initiation, the insulation capability of the wall-structured spacer is mainly limited by the edge effect of the thin-film electrodes and the quality of the thin-film electrodes rather than processes intrinsic to the spacer material or its geometry. However, a conservative breakdown field of ~20.0 kV/mm (15.0 kV/0.75 mm) achieved for thick-wall spacers appears quite adequate to provide the electrical insulation needed for a 1 mm high vacuum gap spacing for FEDs. The main insulation limitation will most likely arise due to electron emission from the field emission arrays. Further studies involving the effect of outgassing from the phosphor anode and the influence of electron bombardment on the spacer surface are presently underway. Investigations are also underway on thin-wall ($<150\,\mu$m) ceramic spacer structures.

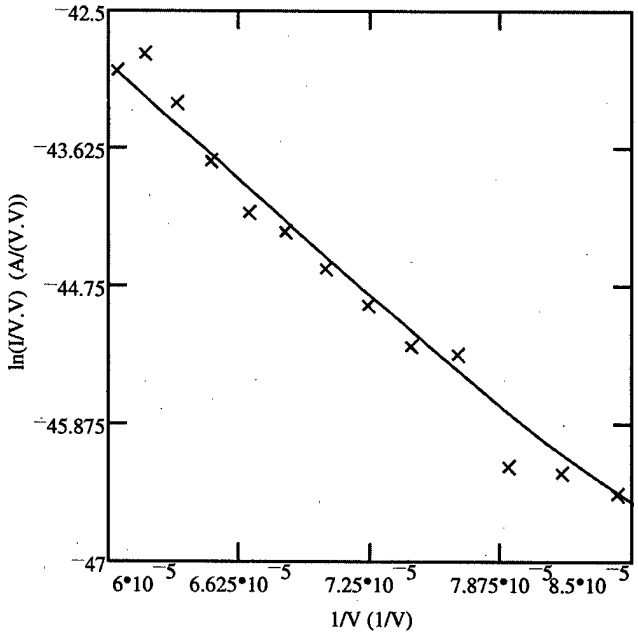


FIG. 7. FN plot of the prebreakdown current of a vacuum gap bridged by a thick-wall Teflon spacer.

IV. CONCLUSIONS

The traditionally good insulating materials investigated in this study did not introduce significant leakage current in the vacuum gap. Breakdown voltages of fiber, disk-shaped and thick-wall spacers were 12–14, 13–16, and 15–20 kV, respectively, across a 750 μm vacuum gap. Compared to the breakdown and prebreakdown characteristics of a Ni/Cr thin-film plain vacuum gap at 750 μm, the breakdown of the fiber and disk-shaped spacers occurred abruptly before the onset of any prebreakdown current, and their breakdown voltages were relatively low. However, the thick-wall structure spacer provided relatively good electrical insulation among the investigated spacer structures, and breakdown occurred in the range of voltages at which the onset of prebreakdown current occurred in plain vacuum gaps. Two distinct modes of breakdown were observed: flashover across the spacer or breakdown at the edge of the thin-film electrode.

Since the insulator flashover voltage of a thick-wall structure spacer approached the breakdown strength of the thin-film plain vacuum gap, the electrode edge effect was the main limiting factor. Hence it can be concluded that apart from the dielectric spacer, the other factors that could limit the insulation capability of FEDs arise from the edge termi-

nation of the phosphor anode and the gate of the FEAs for high voltage operation in FEDs.

ACKNOWLEDGMENTS

This research was supported from DARPA Grant No. N000149610845 managed by ONR. The authors are grateful to Dr. C. Wood for his interest and support of this research. They thank M. McLester and P. G. Muzykov for preparing the thin-film electrodes. They also gratefully acknowledge the discussions with Dr. D. Cherednichenko and Dr. V. P. Madangarli.

Presented at the 11th International Vaccum Microelectronic Conference, Asheville, NC, 19–23 July 1999.

[1]B. R. Chalamala, Y. Wei, and B. E. Gnade, IEEE Spectr. **April**, 42 (1998).
[2]H. C. Miller, IEEE Trans. Electr. Insul. **28**, 521 (1993).
[3]H. C. Miller, IEEE Trans. Electr. Insul. **24**, 765 (1989).
[4]R. Hawley, Vacuum **24**, 383 (1968).
[5]R. Latham, *High Voltage Vacuum Insulation* (Academic, London, 1995).
[6]X. Ma and T. S. Sudarshan, J. Vac. Sci. Technol. B **16**, 1174 (1998).
[7]X. Ma, P. G. Muzykov, and T. S. Sudarshan, 11th International Vacuum Microelectronics Conference, 1998, pp. 59–60.
[8]R. H. Fowler and L. Nordheim, Proc. R. Soc. London, Ser. A **119**, 173 (1928).

BRIEF REPORTS AND COMMENTS

This section is intended for the publication of (1) brief reports which do not require the formal structure of regular journal articles, and (2) comments on items previously published in the journal.

Mechanism of highly preferred (002) texture of Ti films sputter deposited on water-absorbed borophosphosilicate glass films

Tomoyuki Yoshida,[a] Koyu Aoki, and Yasuichi Mitsushima
Toyota Central Research and Development Laboratories, Inc., Nagakute, Aichi 480-1192, Japan

(Received 1 December 1998; accepted 16 April 1999)

The texture of Ti films sputter deposited at 350 °C on water-immersed borophosphosilicate glass (BPSG) films were investigated as a function of the boron content and immersion time of the BPSG films. The Ti film texture was the most strongly (002) preferred for the BPSG films that absorbed a certain amount of water, regardless of their boron content and immersion time. The following mechanism is proposed for the improved texture. The water-absorbed BPSG films increase H_2O partial pressure near the BPSG surface during the heating up to 350 °C prior to Ti sputtering; the increased H_2O pressure increases the density of silanol groups on the BPSG surface to reduce surface free energy; this results in Ti film growth toward highly preferred (002) texture by self-assembly of Ti atoms. © *1999 American Vacuum Society.* [S0734-211X(99)01104-X]

I. INTRODUCTION

Electromigration (EM) failures in Al-alloy interconnects used for very large scale integrated circuits (VLSIs) have become more serious as device feature sizes have shrunk. It is well known that the EM resistance of Al-alloy interconnects is strongly influenced by their microstructure, such as grain size and crystallographic texture. For example, it has been reported that the EM resistance of Al-alloy interconnects can be effectively improved by enhancing Al(111) texture.[1–5] For Al-alloy interconnects stacked with Ti, TiN/Ti, or Ti/TiN/Ti underlayers, which are widely used in VLSIs, the Al texture is strongly (111) preferred if the texture of the bottom Ti films is (002) preferred.[3,4] This can be explained by the epitaxial transfer of the crystallographic information from the Ti(002) plane to the Al(111) plane, because Ti(002), TiN(111), and Al(111) planes have very similar atomic spacings.[3] Therefore, improvement in the (002) texture of the bottom Ti films can make layered Al-alloy interconnects more highly reliable against EM.

We have recently reported that the air exposure or water immersion of borophosphosilicate glass (BPSG) underlayers drastically improves the (002) texture of bottom Ti films and thus Al(111) texture and EM lifetime in AlSiCu/Ti/TiN/Ti interconnects.[6] The improved Ti(002) texture was related to the water absorption of the BPSG films. In another article,[7] we have also reported that the (002) texture of Ti films deposited on thermal SiO_2 films is drastically improved by increasing the H_2O partial pressure of a Ti sputtering chamber from 1×10^{-9} to 3×10^{-8} Torr. In both the above cases, we attributed the mechanism of the improved Ti(002) texture to the self-assembly of Ti atoms on silicate surfaces, which had a low free energy owing to the formation of surface OH groups (silanol groups). Although the silanol formation is probably caused by water adsorption on silicate surface, in the former case the role of the water absorbed in the BPSG films in forming silanol groups has not been fully understood yet.

In this work, we have investigated the texture of Ti films sputter deposited on water-immersed BPSG films as a function of the boron content of BPSG films and water immersion time. Also, we examined the change in the profile of the water absorbed in a BPSG film by heating in vacuum. From the results of these experiments, the role of the water absorbed in the BPSG films in forming silanol groups is discussed.

II. EXPERIMENT

Three kinds of BPSG films with boron contents of 2.0, 3.6, and 4.4 wt. % and a phosphorus content of 3 wt. % and thermal SiO_2 films were formed on Si (100) wafers, and were then used as substrates for Ti films. The thickness of these dielectric films was 800 nm. The BPSG films were formed by plasma enhanced chemical vapor deposition, and annealed at 900 °C in an O_2 ambient for 30 min. These substrates were immersed in hot (95 °C) water for 0.2–5 h, and dried with a spin dryer. Then, Ti (50 nm) films were deposited on the substrates at 350 °C by direct current (dc) magnetron sputtering from a Ti target using Ar gas. Sputtering conditions were a dc power of 1.3 kW and an Ar pressure of 2.4 mTorr. The sputtering was done under an H_2O partial pressure of about 1×10^{-9} Torr.

[a]Electronic mail: tom-yoshida@mosk.tytlabs.co.jp

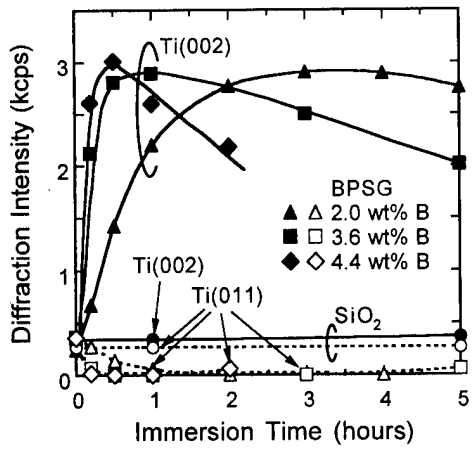


FIG. 1. Ti(002) and Ti(011) diffraction intensities for Ti films as a function of immersion time for BPSG (2.0, 3.6, and 4.4 wt. % boron) and SiO_2 films in hot water.

The crystallographic texture of as-deposited Ti films was examined by x-ray diffraction (XRD) with a conventional θ–2θ method using Co$K\alpha$ radiation. The distribution of hydrogen, boron, and phosphorus atoms in water-immersed BPSG films were evaluated by secondary ion mass spectrometry (SIMS) depth profiling.

III. RESULTS AND DISCUSSION

The texture of Ti films sputter deposited on water-immersed BPSG/Si and SiO_2/Si substrates was examined by XRD. All the XRD patterns showed either two peaks of Ti(002) and Ti(011) or only one peak of Ti(002). Figure 1 shows the Ti(002) and Ti(001) diffraction peak intensities as a function of the immersion time. On the substrates without water immersion, all the films show nearly the same Ti(002) and Ti(011) intensities, as plotted at the immersion time of 0 h. When the substrates are immersed in the hot water, the Ti(002) and Ti(011) intensities on the BPSG films change with increasing immersion time, though those on the SiO_2 film remain unchanged. On the BPSG films, as the immersion time increases, the Ti(002) intensity at first increases rapidly, showing a maximum, and then decreases, while the Ti(011) intensity decreases to approach nearly zero. This indicates that the (002) preferred texture of the Ti films changes with the water immersion of the BPSG films. The maximum Ti(002) intensities for the three kinds of BPSG films are almost the same (near 3000 cps) and the immersion time giving the maximum Ti(002) intensity becomes shorter with increasing boron content of the BPSG films.

It is known that BPSG films become more hygroscopic with increasing boron content,[8] while thermal SiO_2 is non-hygroscopic. Thus, the change in Ti(002) texture on the water-immersed BPSG films can be associated with the water absorption of the BPSG films. To investigate the mechanism of this phenomenon, two kinds of BPSG/Si substrates with boron contents of 2.0 and 4.4 wt. % were immersed in hot (95 °C) water for various times, and the distribution of water absorbed in the BPSG films were then evaluated by

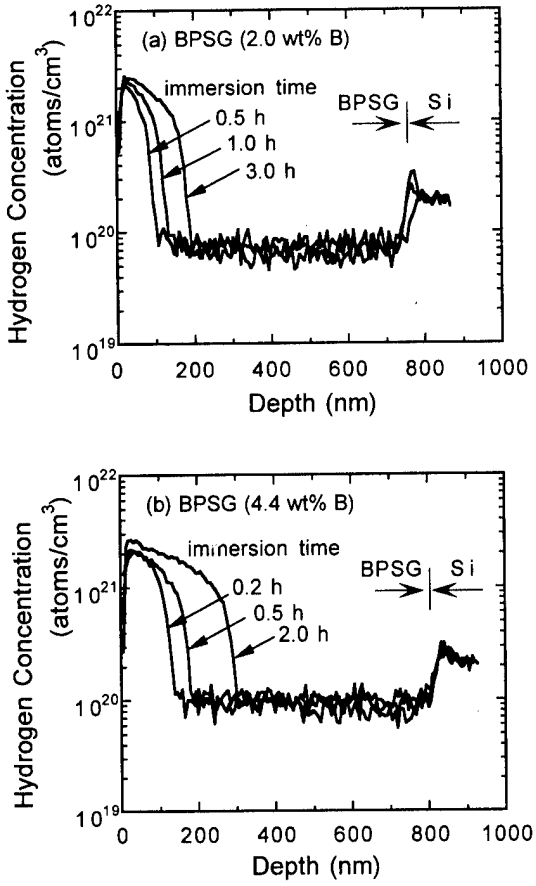


FIG. 2. SIMS depth profiles of hydrogen concentration for BPSG films immersed in hot water for various times. The boron contents of the BPSG films were (a) 2.0 wt. % and (b) 4.4 wt. %.

SIMS depth profiling of hydrogen atoms. At the same time, depth profiles of boron and phospshorus concentrations were also measured.

Figure 2 shows the depth profiles of hydrogen concentration for the two kinds of substrates that were immersed in the hot water for various times. In all the profiles, the hydrogen concentrations much higher than the background level (about 10^{20} atoms/cm^3) are observed near the surfaces. These increased hydrogen concentrations are attributed to absorbed water. From these profiles, it is found that the peak concentration of the absorbed water near the surfaces is independent of both boron content and immersion time, and that the penetration depth of the absorbed water increases with increasing immersion time for both of the two kinds of substrates. Moreover, it is found that the water penetration rate becomes greater with increasing boron content of the BPSG films.

Figure 3 shows the typical dependence of boron and phosphorus depth profiles on the immersion time for the BPSG/Si substrate with 4.4 wt. % boron. Both boron and phosphorus concentrations are almost constant throughout the film thickness and are independent of the immersion time. The BPSG/Si substrate with 2.0 wt. % boron also showed essentially the same result. Thus, the hot-water immersion does not affect both boron and phosphorus distributions in the BPSG films.

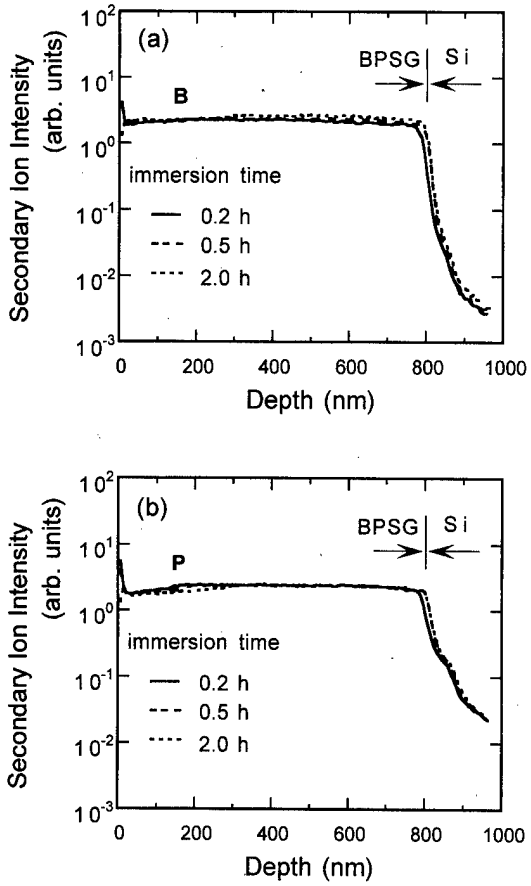


FIG. 3. SIMS depth profiles of (a) boron and (b) phosphorus concentrations for BPSG films immersed in hot water for various times. The boron content of the BPSG films was 4.4 wt. %.

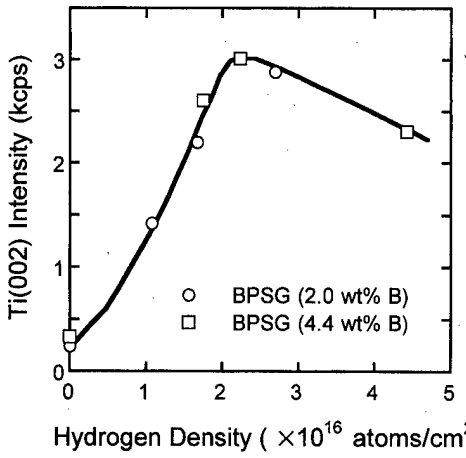


FIG. 4. Relation between Ti(002) intensities and hydrogen densities (per unit area) for BPSG (2.0 and 4.4 wt. % boron) films immersed in hot water. The hydrogen densities were calculated from the depth profiles of hydrogen concentration shown in Fig. 2.

From the above results, it is evident that only the water penetration depth in the BPSG films depends strongly on the water immersion time. Therefore, the change in Ti(002) texture shown in Fig. 1 can be related with the amount of the water absorbed in the BPSG films. Thus, to estimate the amount of the water absorbed in the BPSG films shown in Fig. 2, we calculated hydrogen densities per unit area by integrating the hydrogen concentration profiles against the depth. Figure 4 shows the relation between the Ti(002) intensities and the hydrogen densities for the two kinds of substrates. All the data well fit a single curve, which has a maximum Ti(002) intensity of near 3000 cps at the hydrogen density of 2.2×10^{16} atoms/cm^2, indicating that the Ti(002) texture is determined by the amount of the water absorbed in the BPSG films, regardless of the boron content.

Next, we discuss the reason why the Ti(002) texture is determined by the amount of the water absorbed in the BPSG films. It has been reported that a water-absorbed BPSG film has a water desorption peak at about 280 °C.[8] This suggests that when a water-absorbed BPSG film is heated up to 350 °C in vacuum prior to Ti sputtering, some of the absorbed water desorbs during the heating. To confirm this, the hydrogen profile of the 0.5 h water-immersed BPSG film with 4.4 wt. % boron was measured by SIMS after heating in vacuum at 350 °C for 80 s. Figure 5 compares the hydrogen depth profiles of the BPSG film measured before and after the heating. After the heating, the hydrogen profile broadens toward to the deeper region and the hydrogen concentration near the surface decreases. Hydrogen densities (per unit area) calculated from the two hydrogen profiles indicated that an about 15% of the water absorbed in the substrate desorbed during the heating. Practically, since the period of the heating at 350 °C prior to Ti sputtering is only about 8 s, the amount of the water desorbing from the substrate should be much less than 15%. However, it is probably large enough to increase H_2O partial pressure near the substrate surface, because the initial background H_2O partial pressure is very little (about 1×10^{-9} Torr). Moreover, the H_2O partial pressure near the substrate surface probably increases with increasing amount of the absorbed water. We previously demonstrated that the Ti(002) texture of Ti films sputter deposited on a SiO_2 surface at 350 °C is improved drastically by increasing the H_2O partial pressure of a Ti sputtering chamber from 1×10^{-9} to 3×10^{-8} Torr.[7] Fur-

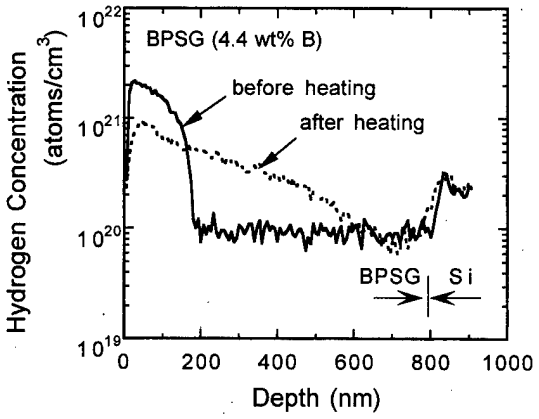


FIG. 5. Hydrogen concentration profiles for 0.5 h water-immersed BPSG (4.4 wt. % boron) film measured before and after heating in vacuum at 350 °C for 80 s.

thermore, we attributed the mechanism of the improved Ti(002) texture to the self-assembly of Ti atoms on the SiO_2 surface, which had a low free energy owing to the formation of silanol groups by water adsorption. This mechanism can be applied to the improved Ti(002) texture on the water-absorbed BPSG films, because the absorbed water plays a role in increasing H_2O partial pressure near the BPSG surfaces, as mentioned above. In Fig. 4, the Ti(002) intensity increases with increasing hydrogen density, reaching a maximum at 2.2×10^{16} atoms/cm^2. This improved Ti(002) texture is attributed to the increased H_2O partial pressure near the BPSG surface. At the hydrogen density of 2.2×10^{16} atoms/cm^2, the H_2O partial pressure near the BPSG surface is thought to reach 3×10^{-8} Torr, because the Ti(002) intensity is in good agreement with that obtained for a Ti film sputter deposited on a SiO_2 surface under a background H_2O partial pressure of 3×10^{-8} Torr.[7] At the hydrogen densities larger than 2.2×10^{16} atoms/cm^2, the Ti(002) intensity decreases with increasing hydrogen density. This indicates that the excess H_2O partial pressure near the BPSG surface restrains the growth of Ti films toward highly preferred (002) texture.

IV. CONCLUSION

The Ti film texture is the most strongly (002) preferred for the BPSG films that absorb a certain amount of water, regardless of their boron content and water immersion time. The following mechanism is proposed for the phenomenon. While the water-absorbed BPSG films are heated up to 350 °C in vacuum prior to Ti sputtering, the water in the BPSG films desorbs partially to increase H_2O partial pressure near the BPSG surface. The increased H_2O partial pressure increases the density of silanol groups on the BPSG surface to reduce surface free energy; this results in Ti film growth toward highly preferred (002) texture by self-assembly of Ti atoms. However, the excess H_2O partial pressure near the BPSG surface restrains self-assembly of Ti atoms.

[1]S. Vaidya and A. K. Sinha, Thin Solid Films **75**, 253 (1981).
[2]D. B. Knorr, D. P. Tracy, and K. P. Rodbell, Appl. Phys. Lett. **59**, 16 (1991).
[3]H. Shibata, M. Murota, and K. Hashimoto, Jpn. J. Appl. Phys., Part 1 **32**, 4479 (1993).
[4]H. Onoda, T. Narita, K. Touchi, and K. Hashsimoto, *Proceedings of the 34th International Reliability Physics Symposium* (IEEE, New York, 1996), p. 139.
[5]H. Toyoda, T. Kawanoue, M. Hasunuma, H. Kaneko, and M. Miyauchi, *Proceedings of the 32nd International Reliability Physics Symposium* (IEEE, New York, 1994), p. 178.
[6]T. Yoshida, S. Hashimoto, H. Hosokawa, T. Ohwaki, Y. Mitsushima, and Y. Taga, J. Appl. Phys. **81**, 7030 (1997).
[7]T. Yoshida, S. Hashimoto, Y. Mitsushima, T. Ohwaki, and Y. Taga, J. Vac. Sci. Technol. B **16**, 2751 (1998).
[8]M. Yoshimaru and H. Matsuhashi, J. Electrochem. Soc. **143**, 3032 (1996).

Microfabrication and testing of suspended structures compatible with silicon-on-insulator technology

A. A. Ayón,[a] K. Ishihara, R. A. Braff, H. H. Sawin, and M. A. Schmidt
Microsystems Technology Laboratories, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 16 February 1999; accepted 21 May 1999)

The footing or notching effect arises during the dry overetching of silicon layers on top of dielectric films. The visible consequence of this effect is the resulting etch that propagates along the interface between the underlying dielectric films and the silicon layer. Footing is usually considered an undesirable artifact during etching. Thus, the vast majority of efforts made to date have been oriented towards reducing or eliminating the aforementioned effect. There is, however, another alternative that has not been fully exploited: the application of the notching effect in the microfabrication of released structures. Furthermore, with deep reactive ion etching (DRIE) tools it is also feasible to deposit fluorocarbon films for electrical isolation purposes *in situ*. Thus, it is possible to microfabricate suspended structures by combining the footing effect with the capabilities offered by DRIE. For this purpose, we have developed, built, and tested suspended electrostatic actuators applying this new microfabrication scheme. The process is well suited for applications involving silicon-on-insulator wafers. Electrostatic actuators microfabricated with this passivation film subsequently underwent 10^5 pull-in cycles without failure. © *1999 American Vacuum Society.* [S0734-211X(99)07904-4]

I. INTRODUCTION

The footing or notching effect, which arises due to the differential charging of sidewalls and bottom of features, is observed when overetching polysilicon or single crystal silicon layers on top of dielectric films[1] (see Fig. 1). This effect is usually negligible in conventional reactive ion etching (RIE) of submicron features due to the low current fluxes and small dimensions involved. However, the etching of large features (several microns or more), using currently available high-density plasma etchers, can exhibit a notching effect that extends laterally several microns. The extent of footing depends on many variables such as the electron temperature,[2–4] the ion energy,[5] the ion/electron current to the surface, the size of the feature and the amount of sidewall passivation.

In particular, the dependence of this effect on etching conditions in a time multiplexed deep etcher (TMDE) has been described elsewhere.[5] In this exercise we used a deep reactive ion etching (DRIE) tool from Surface Technology Systems.[6] This tool multiplexes between a passivating cycle that employs a C_4F_8 plasma and an etching cycle with a SF_6 plasma. In addition to the 13.65 MHz rf power applied to the electrodes, the tool has another rf power source connected to a single-turn coil around the processing chamber. This dual rf power arrangement permits the generation of high-density plasmas for deep silicon etching applications. A complete description of this piece of equipment as well as process optimization techniques can be found elsewhere.[7–9]

The notching effect also has a strong dependence on the aspect ratio of etched features, as can be seen in Fig. 2. Aspect ratio is the ratio of depth to horizontal opening for a particular trench. For aspect ratios of the order of 1 or less, there is enough electron bombardment at the lower part of etched structures, and on the bottom surface, to preclude the charge buildup associated with the footing effect. It can be further observed in the scanning electron microscopy (SEM) micrographs (a) and (b) in Fig. 2 that the notching is less than 1 μm deep. However, as the aspect ratio increases the footing depth also increases. It then achieves a depth of 1.84 μm for an aspect ratio of 1.95, Fig. 2(c), and even 5.63 μm when the aspect ratio reaches 3.30 in Fig. 2(d). As the aspect ratio increases even further, the footing depth decreases to 3.96 μm for an aspect ratio of 5.16 [see Fig. 2(e)], and 1.89 μm when the aspect ratio reaches 7.25 as can be seen in Fig. 2(f). For narrow trenches, there is a significant drop in the etching rate due to shadowing and the concomitant ion flux reduction reaching the bottom surface. The reduction in notching depth observed in Figs. 2(e) and 2(f) is then related to the reduced time the bottom surface is exposed to ion bombardment, and the decreased ion current at the bottom of a trench. The preceding observations suggest that notching can be controlled by tailoring the aspect ratio of etched trenches.

Another variable of interest that concerns notching is the amount of overetching time. The relevance of this variable is shown in Fig. 3, which includes an array of trenches similar to those of Fig. 2 but with an additional 10 min etching time. Footing remained small in those cases involving low aspect ratio trenches as can be seen in the micrographs (a) and (b) in Fig. 3. However, in all higher aspect ratio cases, and compared to similar trenches in Fig. 2, the notching effect became deeper and wider as shown in the micrographs (c) through (e) in Fig. 3.

The observations indicate that feature charging is the

[a]Electronic mail: ayon@mtl.mit.edu

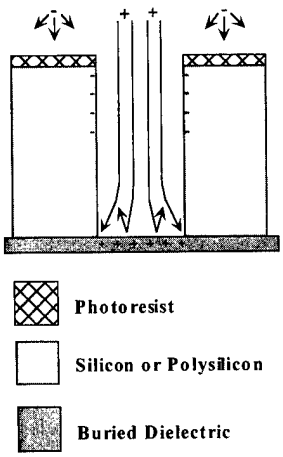


FIG. 1. Footing originates from the isotropic flux of electrons on the sidewalls of the feature that creates a negative charge near the feature opening. The flux of electrons to the lower sidewalls and bottom of the feature is limited by the reduced view factor to the plasma and the repulsive negative charging near the feature opening. The bottom regions of narrow spaces charge positively due to the greater monodirectional ion flux.

main mechanism for explaining the footing effect,[1–3,5] and that the charging potential of etched features on a wafer is influenced by their electrical connection.[4,10–12] Thus, for a given trench width, different notch depths can be expected across a wafer on features charging at different potentials. Similarly, notch symmetry can be predicted for features charging at the same potential. The micrograph shown in Fig. 4 is an example where both symmetric and asymmetric notches are observed on features with dissimilar electrical connections.

It is also relevant to ponder the influence of etching species availability on footing, considering that upon reaching the buried dielectric layer, the concentration of fluorine is expected to increase. It has been reported that conditions that promote the availability of fluorine or decrease the concentration of etching byproducts, such as higher chamber pressure settings or higher SF_6 flow rates, decrease[5] the extent of notching with respect to the overetching time. Furthermore, the excess supply of fluorine cannot explain the observed asymmetrical variations in notching shown in Fig. 4, or the absence of notch formation for trenches when low aspect ratios are involved (see Figs. 2 and 3), where etching species are predictably more abundant. Also, the shape of the notches appears to be line of sight with the deflection of the ions leaving the silicon dioxide near the interface, instead of being circular with the center based upon the sidewall. Thus, the observations suggest that surface charging controls footing, although the availability of etching species could have a synergistic effect in notch formation.

The footing or notching effect described in the previous paragraphs can be utilized for releasing silicon structures during processing of silicon-on-insulator (SOI) wafers. For these applications it is also necessary to emphasize the rel-

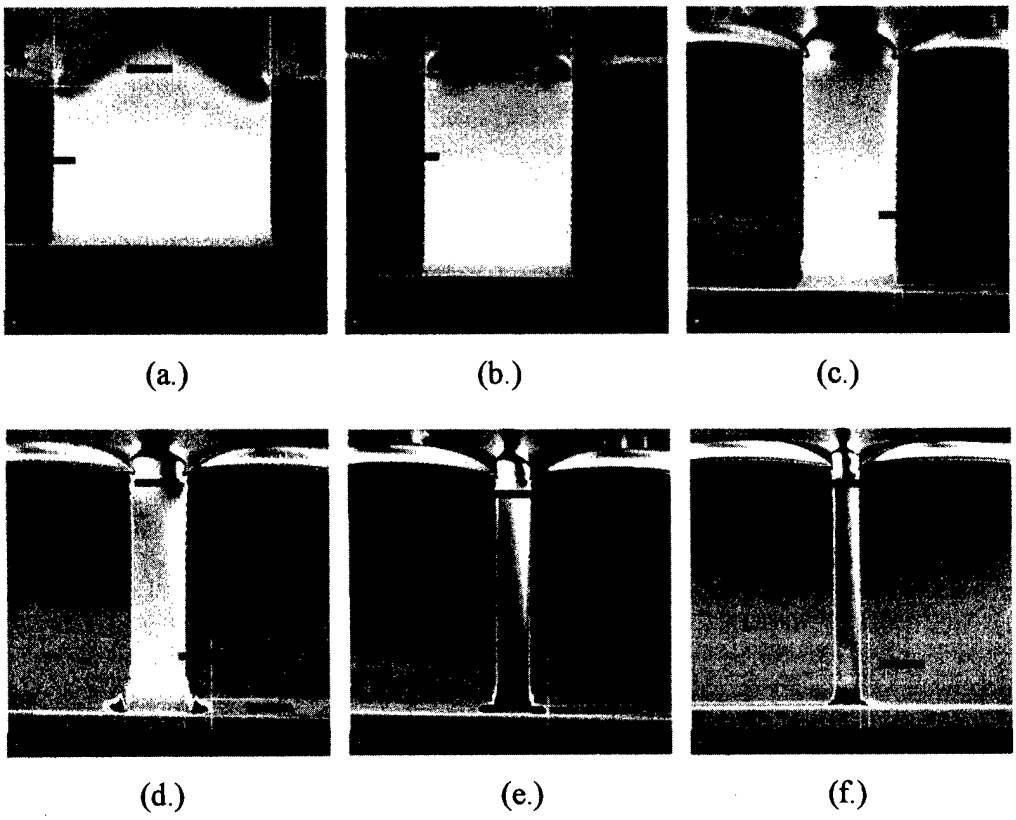


FIG. 2. Footing effect dependence on the aspect ratio of etched structures. Aspect ratio is defined as the ratio of depth to opening of a particular trench. Thus, the respective aspect ratios for the SEM micrographs shown are 0.56 in (a), 1.07 in (b), 1.95 in (c), 3.30 in (d), 5.16 in (e), and 7.25 in (f).

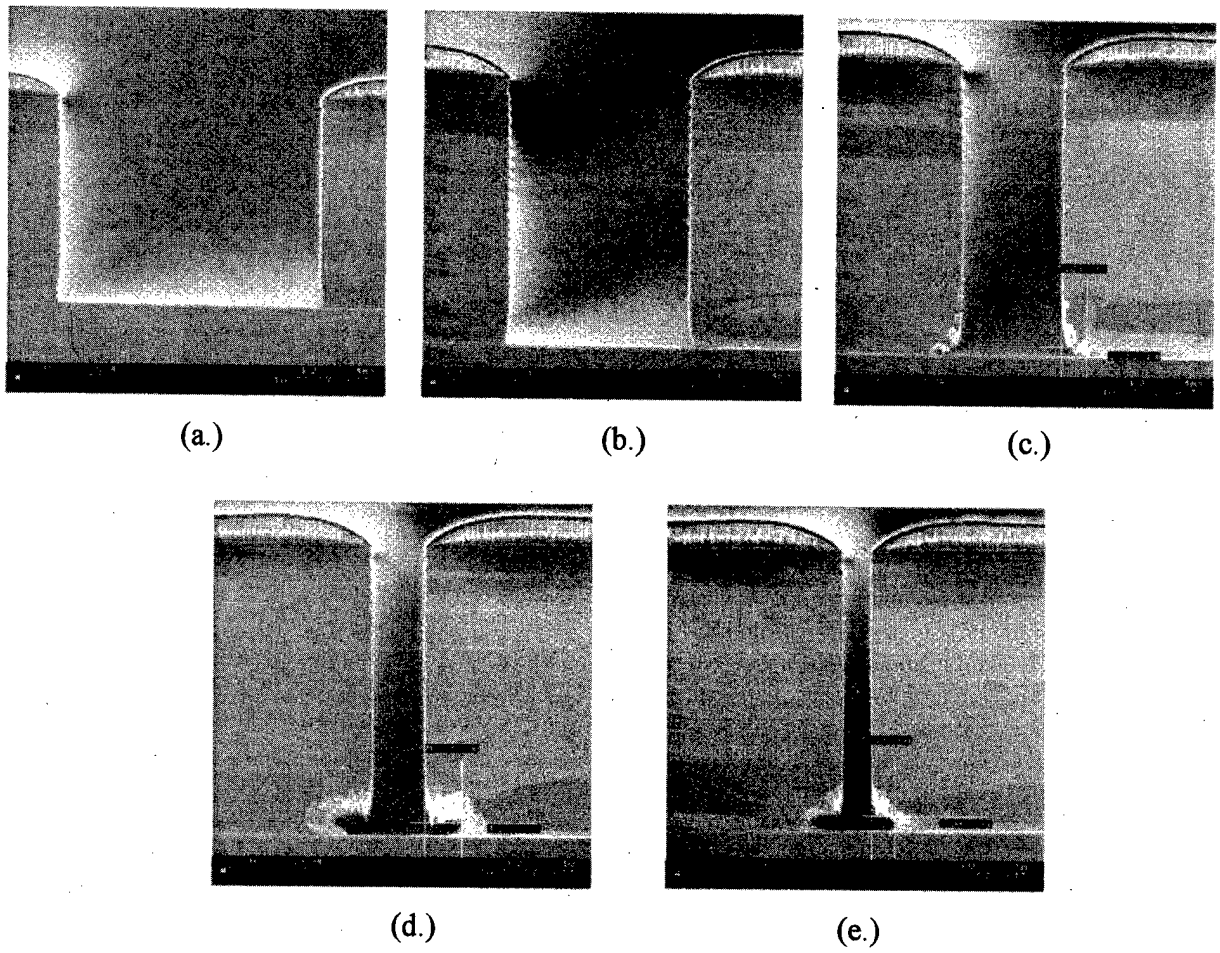


FIG. 3. Dependence of footing on overetching time. The SEM micrographs show trenches similar to those in Fig. 2, that have been etched 10 min more compared to those in Fig. 2.

evance of charging and ion bombardment during the microfabrication of cantilevered beams. In the case of a simple fluorinated, isotropic glow discharge, when releasing a silicon beam that has its sidewalls protected by a film such as silicon dioxide, the spontaneous etch of silicon causes the quick loss of the silicon in the beams,[13,14] leaving behind only the protective shell. By comparison, this deleterious effect is small when exploiting the footing effect. The explanation underlying this observation is that during the release of structures the tool is still operating in TMDE mode, therefore, protective films are also being deposited within the notch. Thus, ion bombardment is still required to proceed with the etch.

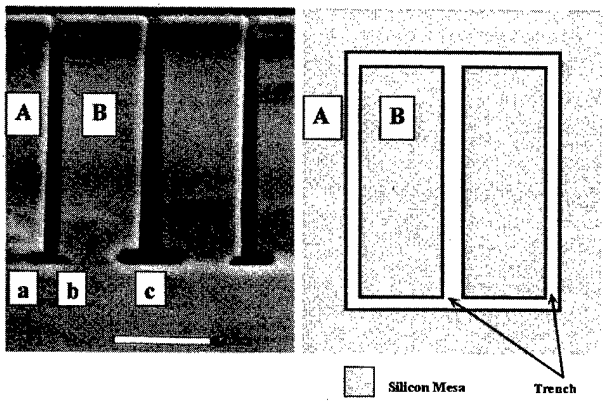


FIG. 4. Electrical connection between different etched features influences the extent of footing. Thus the different depths achieved by notches a and b can be explained in terms of differences in the charging potential of the respective features. Similarly, the observed symmetry in the notches identified with c is expected for features charging at the same potential.

II. ELECTROSTATIC ACTUATORS

The observed extent of footing can, therefore, be used to microfabricate high aspect ratio suspended structures for such applications as microvalves, micropumps, and microrelays. However, devices using electrostatic actuation also require electrical isolation to avoid failure by stiction. Several techniques have been proposed to solve this problem, usually involving additional film depositions such as silicon dioxide.[13,14]

The suggested approach contained in this article is the utilization of *in situ* deposited fluorocarbon films employing the same TMDE piece of equipment. It has been previously reported[15] that the characteristics of those fluorocarbon films

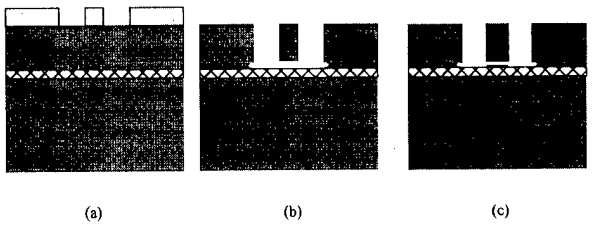


FIG. 5. Sequence for the microfabrication of released structures. After the photolithographic step (a), a TMDE tool can perform all the remaining steps *in situ*: silicon etch, (b) release of structures during overetch, ash photoresist, conformational deposition of a fluorocarbon film, and (c) use of an oxygen plasma to remove the fluorocarbon film from all horizontal surfaces.

have a strong dependence on deposition conditions, and that the deposition rate can exceed 0.18 μm/min. Therefore, the application of the footing effect in combination with the *in situ* deposition of electrical isolation films enables a new microfabrication sequence detailed in the following paragraphs.

Step (i). Transfer the topography to photoresist using a standard photolithography step. See Fig. 5(a).

Step (ii). Etch the silicon substrate until the underlying dielectric film is reached. Etching conditions and optimization of operating points during TMDE can be found elsewhere.[7,8] The TMDE processing conditions used in this exercise were as follows: during the passivating cycle 40 sccm of C_4F_8 were flown 11 s, with the electrode power set at 6 W and the coil power at 600 W; during the etching cycle 105 sccm of SF_6 were flown 14 s (with an additional overlap over the passivating cycle of 0.5 s) with the electrode power set at 12 W and the coil power set at 750 W. The position of the throttle valve was fixed at 65° during operation.

Step (iii). During the overetching time, charging produces the footing effect,[1] and the structures are released. This time can be tailored to release the widest feature in the design. As was previously mentioned, the deleterious effect of losing height control in narrower beam due to spontaneous silicon etching reported in other schemes is very small during this step.

Step (iv). Photoresist can be removed *in situ* with an oxygen glow discharge. See Fig. 5(b).

Step (v). Operating the equipment in continuous mode, and flowing only C_4F_8, a fluorocarbon film is deposited on all surfaces.

Step (vi). Additional *in situ* oxygen plasma will preferentially remove the passivation film from all horizontal surfaces exposed to ion bombardment, while preserving the film on sidewalls.[5] Auger spectroscopy was used to inspect the horizontal and vertical surfaces after exposure to oxygen plasma. Figure 5(c) shows the cross section of a completed device.

III. TESTING

An electrostatic actuator was built to validate the preceding approach. The design (see Fig. 6) included beams 3.8 μm wide, 1000 μm long etches using SOI wafers supplied by Motorola with a single crystal silicon ⟨100⟩ film 38 μm thick, and of 5–20 Ω cm resistivity. Upon transferring the topography to photoresist, the devices were prepared following the process outlined in Fig. 5. After photolithography, all necessary steps were performed in the same DRIE tool including silicon etching, structure release, photoresist ashing, fluorocarbon film deposition as well as its removal from all horizontal surfaces. After ashing the photoresist (step iv), the wafer were cleaved and half of each wafer was tested with-

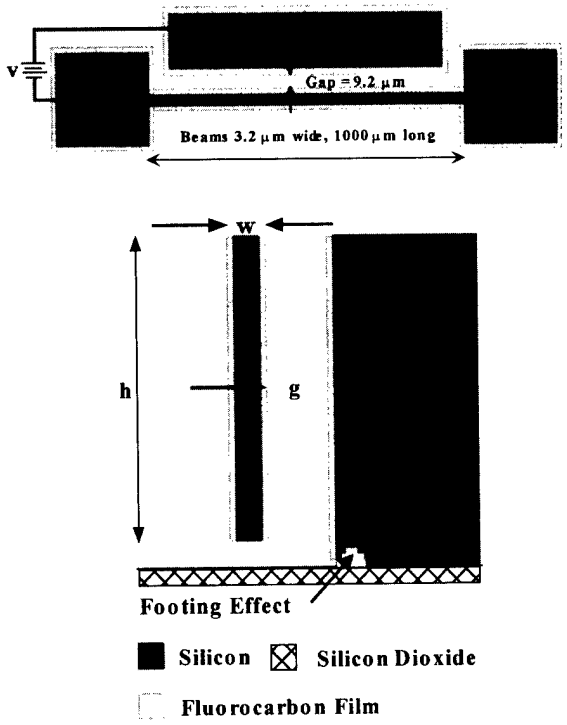


FIG. 6. Drawing of the electrostatic actuator described herein.

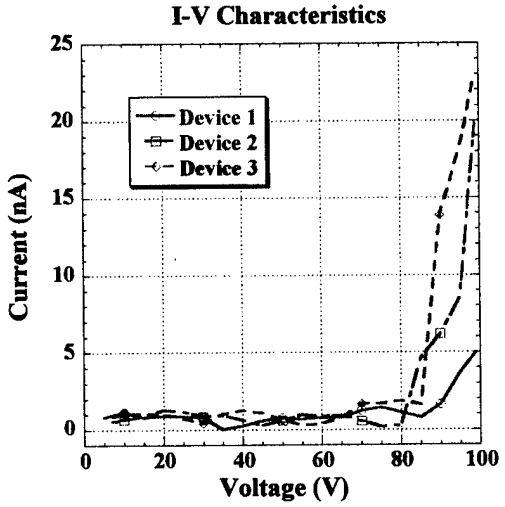


FIG. 7. *I–V* characteristics of the electrostatic actuator described in this article. Fluorocarbon films for electrical isolation purposes were deposited *in situ*.

out the additional fluorocarbon film deposited, while the other half continued with steps (v) and (vi) outlined in the previous paragraph.

Testing was done by applying a ramping voltage in steps of 5 V, between the points illustrated in Fig. 6. Although the pull-in voltage is of the order of 85 V, it was decided to apply up to 100 V to further stress the passivation film. Actuators without the additional passivating film failed by stiction within 50 pull-in cycles. Actuators that followed the sequence described herein (see Fig. 5), continued operating without failure or noticeable performance degradation after 10^5 pull-in cycles. Figure 7 shows the I–V characteristics for three typical devices obtained with a semiconductor parameter analyzer. Even though there is a noticeable increase in current drained beginning at pull-in voltage, the leakage current only increased from $\approx$1 nA to $<$25 nA even for an applied 100 V.

IV. CONCLUSIONS

We have, therefore, demonstrated the application of the footing effect in the microfabrication of suspended structures using a process compatible with SOI technology. The measured performance of the microfabricated electrostatic actuators demonstrates the usefulness of the *in situ* deposition of fluorocarbon films for electrical isolation purposes. Furthermore, after photolithography, all steps required to microfabricate suspended structures (i.e., silicon etching, release of devices, photoresist removal, deposition of electrical isolation films and clearing of fluorocarbon films from all horizontal surfaces) are done in the same high-density plasma, DRIE tool. This process is, therefore, compatible with a large variety of microelectromechanical systems (MEMS) applications.

ACKNOWLEDGMENTS

The U.S. Army Research Office, Dr. R. Paur, Technical Manager, and DARPA, Dr. R. Nowak, Program Manager, supported this project. One of the authors (R.A.B.) was supported by a NSF Fellowship. The cooperation of the staff of the Microsystems Technology Laboratories at MIT and the processing help of Kurt Broderick are also appreciated.

[1]J. C. Arnold and H. H. Sawin, J. Appl. Phys. **70**, 5314 (1991).
[2]S. Tabara, Jpn. J. Appl. Phys., Part 1 **35**, 2456 (1996).
[3]G. S. Hwang and K. P. Giapis, J. Vac. Sci. Technol. B **15**, 70 (1997).
[4]N. Fujiwara, T. Maruyama, and M. Yoneda, Jpn. J. Appl. Phys., Part 1 **34**, 2095 (1995).
[5]A. A. Ayón, K. Ishihara, R. A. Braff, H. H. Sawin, and M. A. Schmidt, 45th International Symposium of the American Vacuum Society, Baltimore, MD, 2–6 November 1998.
[6]Surface Technology Systems USA Inc., Redwood, CA.
[7]A. A. Ayón, C. C. Lin, R. Braff, R. Bayt, H. H. Sawin, and M. Schmidt, J. Electrochem. Soc. **146**, 339 (1999).
[8]A. A. Ayón, C. C. Lin, R. Braff, R. Bayt, H. H. Sawin, and M. Schmidt, 1998 Solid Sate Sensors and Actuator Workshop, Hilton Head, SC, 8–11 June 1998.
[9]J. Bhardwaj, H. Ashraf, and A. McQuarrie, Annual Meeting of the Electrochemical Society, Symposium on Microstructures and Microfabricated Systems, Montreal, Quebec, 4–9 May 1997, p. 118.
[10]T. Nozawa, T. Kinoshita, T. Nishizuka, A. Narai, T. Inoue, and A. Nakaue, Jpn. J. Appl. Phys., Part 1 **34**, 2107 (1995).
[11]T. Kinoshita, M. Hane, and J. P. McVittie, J. Vac. Sci. Technol. B **14**, 560 (1996).
[12]K. K. Chi, H. S. Shin, W. J. Yoo, C. O. Jung, Y. B. Koh, and M. Y. Lee, Jpn. J. Appl. Phys., Part 1 **35**, 2440 (1996).
[13]Z. L. Zhang and N. C. MacDonald, J. Micromech. Microeng. **2**, 31 (1992).
[14]A. A. Ayón, N. Kolias, and N. C. MacDonald, in Proceedings of the 1995 Asia-Pacific Microwave Conference, Taejon, South Korea, 10–13 October 1995, p. 147.
[15]A. A. Ayón, K. Ishihara, R. Braff, H. H. Sawin, and M. Schmidt, Fall Meeting of the materials Research Society, Boston, MA, 30 Nov.–4 Dec. 1998.

RAPID COMMUNICATIONS

This section is reserved for short submissions which contain important new results and are intended for accelerated publication.

Low-energy electron-beam lithography using calixarene

A. Tilke, M. Vogel, F. Simmel, A. Kriele, R. H. Blick,[a] H. Lorenz, D. A. Wharam,[b] and J. P. Kotthaus

Center for NanoScience and Sektion Physik, Ludwig-Maximilians-University München, D-80539 München, Germany

(Received 19 November 1998; accepted 9 April 1999)

Low-energy electron-beam lithography using calixarene as a negative electron resist has been investigated in the energy range between 0.5 and 20 keV. The suitability of electron energies down to 2 keV with a writing resolution of about 10 nm is clearly demonstrated. At low electron energies the required electron dose is drastically reduced. Moreover, irradiation damage during the exposure of a high-mobility two-dimensional electron gas using calixarene plays no significant role in the low-energy regime. © *1999 American Vacuum Society.* [S0734-211X(99)00804-5]

I. INTRODUCTION

Low-energy electron-beam lithography in the range of 1–10 keV offers significant advantages over the use of high electron energies for the exposure of sensitive positive electron resists such as PMMA. Since the penetration depth of electrons is smaller for lower energies the proximity effect is strongly suppressed due to a reduction in the number of backscattered electrons from the substrate.[1] The majority of the electrons are inelastically scattered in the resist film and it is their spatial range which limits the lateral resolution. Furthermore, the irradiation damage of the underlying substrate is substantially lower, making low-energy electron-beam lithography an attractive prospect for lateral structuring of high-mobility semiconductor devices.

Recently Fujita *et al.* investigated a new high-resolution negative resist called calixarene.[2–4] Because this resist has a low sensitivity, it is especially important to use low electron energies. Reducing the electron energy will significantly reduce the electron dose required for exposure as is immediately seen using a simplified Bethe equation where x denotes the distance from the sample surface and $W(x)$ the electron energy:[5]

$$dW(x)/dx \sim -1/W(x). \tag{1}$$

Moreover, in the case of negative resists the electron beam usually exposes the active area of the defined device structure so reducing the electron energy is important for the elimination of radiation damage. Low-energy electron-beam lithography on high-resolution negative electron resists is therefore especially well suited for defining nanostructures such as single-electron tunneling transistors.[6]

[a]Electronic mail: robert.blick@physik.uni-muenchen.de

[b]Permanent address: Institut für Angewandte Physik, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany.

Here, we investigate the exposure parameters of a calixarene resist at electron energies down to 0.5 keV. We determine both the resolution of the resist as well as the radiation damage of a high electron mobility transistor (HEMT) structure in the energy range between 1 and 20 keV. In addition we investigate the differences in the exposure parameters of nanometer scale and large scale structures caused by proximity effects.

II. EXPERIMENTS

A. Exposure parameters

In order to study the influence of the electron energy on the exposure dose, writing resolution and proximity effects of the calixarene resist hexaacetate *p*-methylcalixarene (MC6AOAc), various silicon samples were coated with a thin calixarene film (~40–50 nm). These films were prebaked at 170 °C for 30 min and then exposed at different electron energies and electron doses. After exposure, the samples were developed for 30 s in xylene and the development was then stopped by immersion for 30 s in isopropanol.[2] The exposure was performed by a scanning electron microscope (SEM) with a thermally assisted field-emission electron source using a commercial beam and stage control system.

The relationship between the electron energy and the required electron dose was determined by exposing patterns consisting of large and small scale structures. The required dose for small patterns (of the order of 10 nm) can be much larger than for large structures. In order to investigate this difference, the test patterns consisted of large areas connected to small lines with widths varying between 3 and 21 nm (as will be shown in detail later). These lines were written with a meander scan and a pixel separation of 3 nm. This structure also enables us to investigate proximity effects occurring at higher electron energies. The electron dose for the

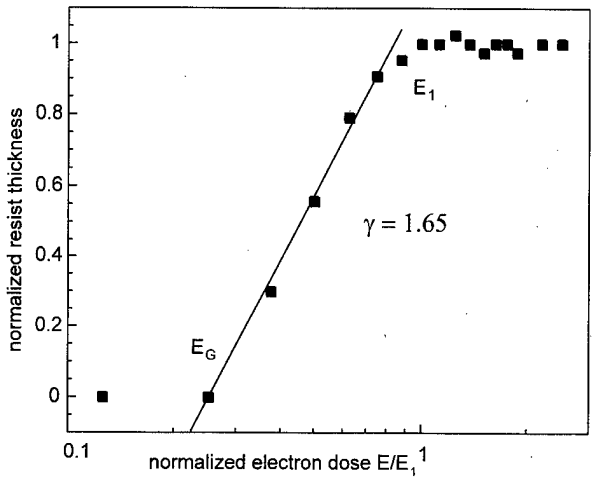


FIG. 1. Resist contrast curve for exposure at an electron energy of 2 keV. The minimum gel dose is $E_G = 200\ \mu C/cm^2$, and the saturation or maximum dose is $E_1 = 800\ \mu C/cm^2$. The contrast is measured to be $\gamma = 1.65$.

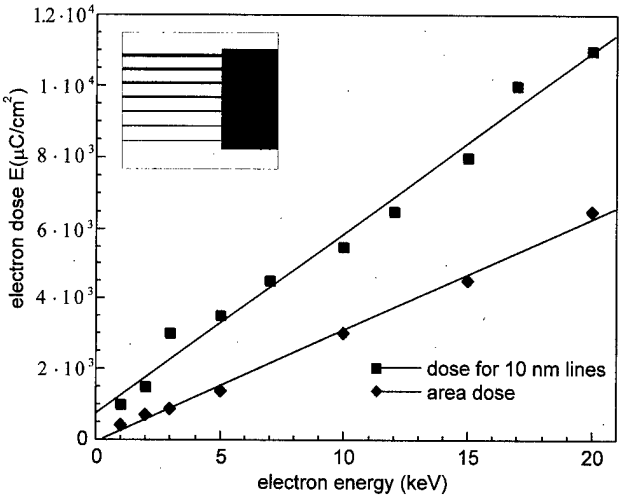


FIG. 2. Electron dose for 9–20 nm lines and for large scale patterns. The saturation dose for the large patterns was determined by measuring the resist thickness with an AFM, whereas the electron dose for the small features was controlled with a SEM. The inset shows the test structure used, where lines with different widths are connected to a large area with a lateral extension of $5 \times 20\ \mu m$. The nominal linewidth of the thinnest line is 3 nm (1 pixel line) and increases to 21 nm (7 pixels) for the thickest one.

small lines was determined by measuring the exposed linewidths in the developed resist with the SEM. The 100% dose was determined by requiring that the exposed linewidth be identical to the intended linewidth.

In the case of large patterns we determined the saturation dose required for a given resist thickness after exposure. For lower electron doses only a fraction of the resist film remained on the substrate after development.[7] This was confirmed by studies performed with an atomic force microscope (AFM) operating in the so-called tapping mode. The step height at the edges of the pattern as well as the slope of the film edge were also determined.

B. Irradiation damage

In order to investigate the influence of the electron beam at different energies on the quality of the underlying semiconductor layers, various Hall bar geometries were defined on a high-mobility $GaAs/Al_xGa_{1-x}As$ heterostructure grown by molecular beam epitaxy, where the two-dimensional electron gas is located 85 nm from the top surface. The Hall bar structures were defined using standard optical lithography and a wet chemical etching process. On the defined Hall bar structures a 42 nm thick calixarene film was subsequently exposed with different electron energies varying between 1 and 20 keV. The resulting electron density and electron mobility at 4.2 K were extracted from the period of the Shubnikov–de Haas oscillations in high magnetic fields and the resistance at zero magnetic field, respectively.[8]

III. RESULTS AND DISCUSSION

A. Exposure parameters

Figure 1 shows the resist contrast curve for an electron energy of 2 keV. The minimum gel dose is $E_G = 200\ \mu C/cm^2$, the saturation or maximum dose is $E_1 = 800\ \mu C/cm^2$. The contrast γ is defined as the slope of the line from E_G to E_1 and turns out to be $\gamma = 1.65$, in agreement with the value obtained previously for the exposure of calixarene films with 25 keV electrons.[7] Determining the maximum dose E_1, where the resist film achieves its maximum thickness, leads to the dependence of this saturation dose on the electron energy as shown in Fig. 2. Also shown in Fig. 2 is the required electron dose for the narrow lines. The inset in Fig. 2 shows an overview of the test structure used. Clearly, the electron dose for small structures exceeds the one for the large patterns by a factor of about 1.7 over the whole energy range considered.

Both plots in Fig. 2 indicate an almost linear relationship between the electron energy and the electron dose as predicted by Eq. (1) when the penetration depth of the incident electrons is much larger than the resist thickness.

The slope of the resist edge for different electron doses at an electron energy of 2 keV is shown in Fig. 3. For a dose of 500 $\mu C/cm^2$, just below the saturation dose, the maximum

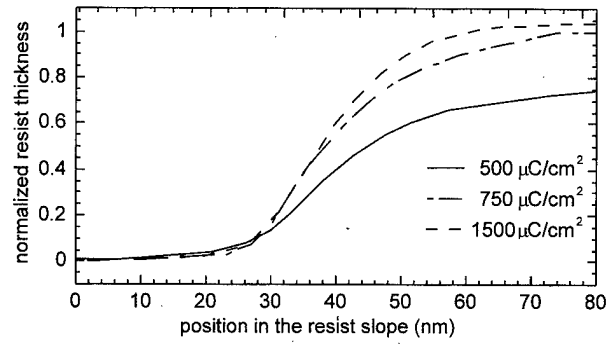


FIG. 3. Resist slope for electron doses of 500, 750, and 1500 $\mu C/cm^2$. Although the maximum resist thickness is almost achieved at $E = 750\ \mu C/cm^2$, the slope still gets steeper for higher electron doses.

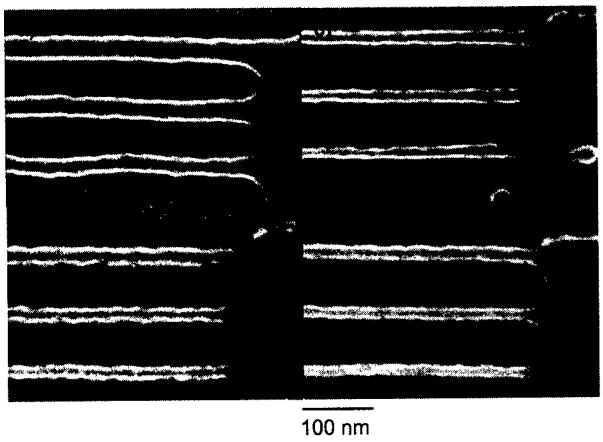


FIG. 4. (a)–(d) SEM picture of the test structure used for electron energies of 1, 2, 7, and 20 keV, respectively. Only a small influence of proximity effects can be seen at an electron energy of 20 keV, where the spatial range of the proximity induced line broadening is about 40 nm. At 1 keV (a) weak adhesion leads to curved lines.

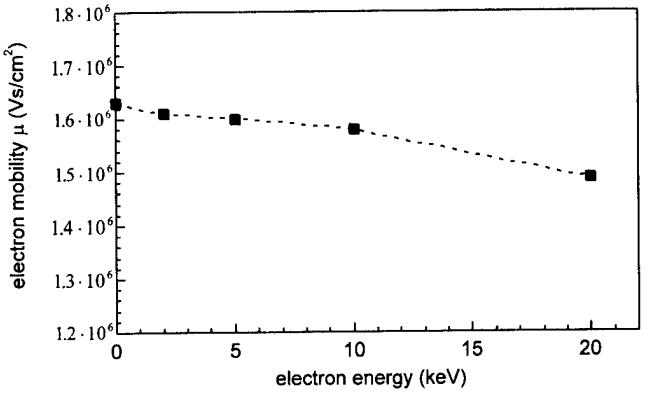


FIG. 5. Electron mobility in a high-mobility two-dimensional electron system after exposure of a 42-nm-thick calixarene film at different energies. The electron dose was chosen according to the saturation dose depicted in Fig. 2.

resist thickness is not yet achieved. Increasing the electron dose above the saturation dose of 800 μC/cm^2 to the dose required for the small lines of 1500 μC/cm^2 leads to a steeper resist slope. The width of the resist slope at the saturation dose allows one to estimate the range in which electron scattering is important to be about 40 nm. As has been found previously the scattering range is expected to follow a power law as a function of the incident beam energy:

$$\beta[\mathrm{nm}] = c \cdot W[\mathrm{keV}]^{1.7}, \tag{2}$$

where c is of the order of 10.[9] For an incident beam energy of 2 keV the measured scattering range in our experiments is comparable to the value expected.

One advantage of low-energy electron-beam lithography is the reduction of proximity effects due to the reduced scattering range of the incident electrons. In addition, the range for direct electrodynamic interaction of the incident electron beam with the resist is smaller for lower electron energies.[10] In order to investigate the influence of proximity effects in calixarene, we determined the width of the thin lines in the vicinity of the large patterns in our test structure and found only small broadening due to proximity effects even at the highest electron energy of 20 keV (Fig. 4). The broadening turns out to be about 40 nm for all electron energies considered. We therefore conclude that proximity effects only play a role within a range of about 40 nm in the low-energy regime of electron-beam lithography with calixarene and are mainly caused by electrons scattered in the resist film, which has a thickness comparable to the observed scattering range. Backscattered electrons from the substrate could, in principle, lead to additional background exposure in the vicinity of the large patterns, especially for high beam energies. However, we do not observe any significant variation in the broadening for the beam energies considered and presumably this background exposure is smaller than the minimum gel dose E_G, making calixarene relatively insensitive to such proximity effects.

In our investigations we found a resolution limit of 10 nm for the lines used in our test structure even at accelerating voltages as low as 2 kV. Moreover, by connecting short lines at both ends, even at an electron energy of 2 keV, minimum structure sizes down to 6 nm were realized, suggesting that adhesion problems are responsible for the resolution limit. This is in good agreement with recent data of Fujita *et al.*, who found no change in resolution of a calixarene resist down to an electron energy of 5 keV, where the smallest possible structure size is about 10 nm. From Monte Carlo simulations they also concluded that the resolution limit of calixarene does not depend on the electron scattering but on weak adhesion to the substrate for structures smaller than 10 nm.[11]

Lowering the electron energy below 1 keV leads to reduced adhesion of the developed resist [Fig. 4(a)]. Calculations and measurements for a PMMA electron resist of 50 nm thickness indicate that below an electron energy of 1 keV the resist cannot be penetrated completely by the incident electrons.[5] In the case of positive resists this leads to a thin resist layer remaining after development, the thickness of which can easily be measured. For calixarene the whole thickness of about 42 nm was maintained even for an electron energy below 1 keV. Nevertheless, an unexposed thin film between the substrate and the exposed upper part of the resist layer leads to the observed reduced adhesion. A further reduction of the electron energy down to 500 eV results in the removal of the exposed film surface during the development process by dissolving the unexposed resist.

The dose reduction in low-energy electron-beam lithography can cause serious problems for electron resists with high sensitivity like PMMA due to shot noise in the electron beam. Since the shot noise limit is around a minimum number of about 200 electrons the limiting dose is about 150 μC/cm^2 for a resolution of 10 nm.[12] For calixarene, where the electron dose is 800 μC/cm^2 at an energy of 2 keV, the influence of shot noise is presumably already vis-

ible [Fig. 4(b)]. Thus, in addition to adhesion problems, shot noise turns out to be another limiting factor for high-resolution lithography at very low energies with calixarene.

B. Irradiation damage

We have measured the variation of the electron mobility in a high-mobility two-dimensional electron gas ($\mu = 166\,m^2/V\,s$) upon the electron energy in the lithographic process (Fig. 5).

A slight decrease of the electron mobility with increasing acceleration voltage is observed. Nevertheless, this effect plays only a small role in the energy regime considered. Since the total process of irradiation damage in electron-beam lithography is not yet fully understood, this result will be a topic of further research.

IV. SUMMARY

We have investigated the use of low-energy electron-beam lithography for the negative electron resist calixarene. The resolution limit at energies as low as 2 keV is found to be about 10 nm. The range of electron scattering in the resist film that leads to proximity effects was determined to be about 40 nm. Irradiation damage to a high-mobility two-dimensional electron gas in the low-energy regime evidently plays no significant role. Lowering the exposure energy of the electrons below 1 keV leads to incomplete vertical exposure of the resist film. We conclude that the most suitable energy regime for high-resolution low-energy electron-beam lithography with calixarene is in the range between 2 and 10 keV.

ACKNOWLEDGMENTS

The authors would like to thank L. Pescini, W. Frank, B. Irmer, and A. Erbe for useful discussions. The semiconductor heterostructure was kindly grown by Dr. W. Wegscheider of the Walter-Schottky-Institut of the Technical University, Munich, Germany. The authors acknowledge financial support from the Bundesministerium für Bildung und Forschung (Contract No. 01M2413C6) and from the Deutsche Forschungsgemeinschaft (SFB No. 348).

[1]P. A. Peterson, Z. J. Radzimski, S. A. Schwalm, and P. E. Russell, J. Vac. Sci. Technol. B **10**, 3088 (1992).
[2]J. Fujita, Y. Ohnishi, Y. Ochiai, and S. Matsui, Appl. Phys. Lett. **68**, 1297 (1996).
[3]Y. Ohnishi, J. Fujita, Y. Ochiai, and S. Matsui, Microelectron. Eng. **35**, 117 (1997).
[4]J. Fujita, Y. Ohnishi, S. Manako, Y. Ochiai, E. Nomura, T. Sakamoto, and S. Matsui, Jpn. J. Appl. Phys., Part 1 **36**, 7769 (1997).
[5]K.-D. Schock, F. E. Prins, S. Sträle, and D. P. Kern, J. Vac. Sci. Technol. B **15**, 2323 (1997).
[6]E. Leobandung, L. Guo, Y. Wang, and S. Y. Chou, Appl. Phys. Lett. **67**, 938 (1995).
[7]F. E. Prins, J. Pfeiffer, S. Raible, D. P. Kern, and V. Schurig, Microelectron. Eng. **41/42**, 359 (1998).
[8]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 41 (1982).
[9]P. M. Mankiewich, L. D. Jackel, and R. E. Howard, J. Vac. Sci. Technol. B **3**, 174 (1985).
[10]W. Langheinrich and H. Beneking, Microelectron. Eng. **23**, 287 (1994).
[11]J. Fujita, Y. Ohnishi, S. Manako, E. Nomura, and S. Matsui, Microelectron. Eng. **41/42**, 323 (1998).
[12]J. Greeneich, J. Trotel, and B. Fay, in *Electron Beam Technology in Microelectronic Fabrication,* edited by G. Brewer (Academic, New York, 1980), p. 245.

New method to prepare W–B^+–N ternary barrier to Cu diffusion by implanting BF_2^+ ions into W–N thin film

Dong Joon Kim
Department of Metallurgical Engineering, Hanyang University, 17 Haengdang-dong, Seongdong-ku, Seoul 133-791, Korea

Yong Tae Kim[a)]
Semiconductor Materials Laboratory, Korea Institute of Science and Technology, P.O. Box 131, Cheongryang, Seoul, Korea

Jong-Wan Park
Department of Metallurgical Engineering, Hanyang University, 17 Haengdang-dong, Seongdong-ku, Seoul 133-791, Korea

(Received 21 January 1999; accepted 23 April 1999)

Implanting 1×10^{17} BF_2^+ ions/cm^2 with as low an energy as 40 keV into W–N thin films, W–B^+–N thin layer was formed for the region near the surface of the W–N thin film. Experimental results reveal that thermal stability of the W–B^+–N/W–N thin film and its barrier performance against Cu diffusion were improved compared to those of the W–N thin films after annealing at 600–800 °C for 30 min. These excellent properties of the W–B^+–N/W–N barrier are due to the B^+ ions to prevent nitrogen out-diffusion and to keep the W–B^+–N/W–N thin film in an amorphous phase after annealing at 800 °C because the grain growth of W or W–N and the Cu diffusion were suppressed by the B and N impurities in the amorphous thin film. © *1999 American Vacuum Society.* [S0734-211X(99)02204-0]

Throughout the evolution of very large scale integrated circuits, the scale down of active devices eventually became less profitable since the limitations of the circuit speed and functional density came to depend more on the characteristics of multilevel interconnection than on the scaled device.[1,2] To implement a multilevel interconnection for high speed and high density new materials such as conductors and interlevel dielectrics are required. In addition to new materials, several difficulties and failure modes related to etching and annealing processes, for example, anisotropy, thermal stability, and electromigration might be encountered. A metal–semiconductor contact structure especially needs a thermally stable barrier layer to prevent interaction between a metal and a semiconductor. If we use Cu films as a contact or interconnect metal, the barrier layer has to block the fast diffusive Cu atoms during post processing. Previous works reported that binary diffusion barriers such as TiN, Ta–N, and W–N were superior to pure refractory metal films.[3–5] Recently the amorphous ternary metal film is expected to show higher thermal stability than the binary film because grain boundaries offering a diffusion path do not exist in the amorphous ternary films.[6–9] However, it is not easy to control the composition, electrical resistivity, and anisotropic etching of the ternary metal thin film compared to the binary thin film. In this work, to solve these problems, we have proposed a new ternary metal-Si contact structure by implanting boron ions after patterning the binary W–N thin film on the Si contact window. This boron ion implantation method enables us to modify the binary W–N film into the ternary W–B–N thin film barrier and to precisely control the concentration and depth profile of B impurity. In addition, this method offers the self-aligned contact structure as follows: For the contact window, we have deposited the W–N or W thin film and delineated the contact metal. After that, we have implanted B^+ ions into the diffusion barrier metal film. Thus, we can avoid a difficult etching process of the ternary metal film.

We use *p*-type wafers as starting materials which orientation and resistivity were (100) and 5–10 Ω cm, respectively. W–N thin films of about 1100 Å in thickness were deposited on Si contact windows by the plasma enhanced chemical vapor deposition method using WF_6, NH_3, and H_2 reactant gases. The flow ratio of $WF_6/NH_3/H_2$ was fixed at 2/1/50. The system was evacuated to 2×10^{-6} Torr before the deposition and the working pressure was maintained at 0.5 Torr. Radio frequency power density and substrate temperature were 0.7 W/cm^2 and 300 °C, respectively. Detail deposition processes and experimental results regarding the W–N thin film were already published.[10,11] To estimate the maximum projected range of the BF_2^+ ion in the W–N thin film, we used a transport ion mass (TRIM) simulation of boron implantation at 40 keV and 1×10^{17} ions/cm^2. By the simulation of BF_2^+ implantation, we could estimate that the projected range of boron ions would be distributed in the range of 100–200 Å depth from the top surface of the 1100-Å-thick W–N thin film. After the implantation of BF_2^+ ions, the W–B^+–N/W–N thin films were annealed at temperatures in the range of 600–800 °C for 30 min in Ar ambient to determine phase transition during the annealing processes. The thickness and resistivity of W–B^+–N/W–N thin films were measured by Rutherford backscattering spectrometry

a)Author to whom the correspondence should be addressed; electronic mail: ytkim@kistmail.kist.re.kr

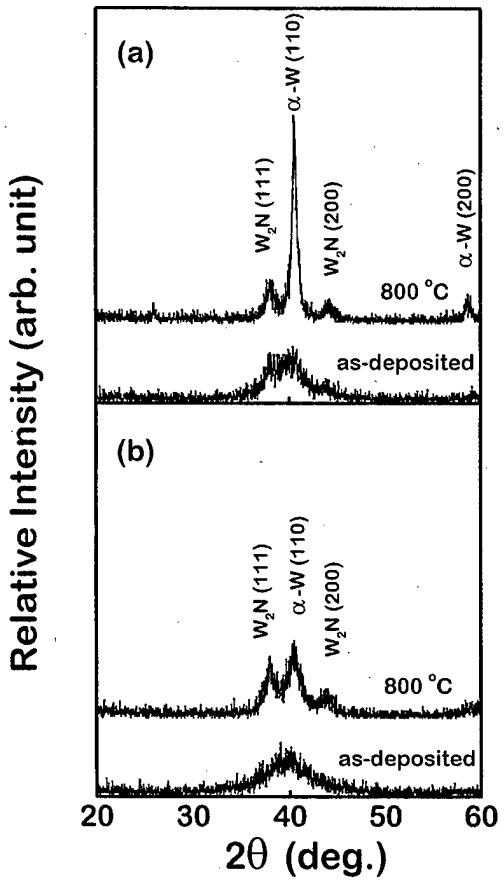


FIG. 1. XRD patterns of (a) W–N and (b) W–B$^+$–N/W–N thin films before and after annealing at 800 °C.

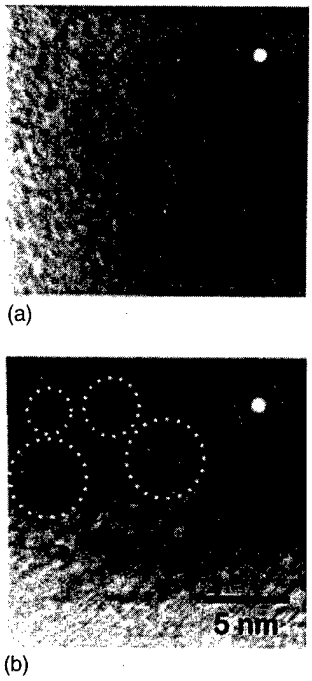


FIG. 2. TEM micrographs and electron diffraction patterns of the W–B$^+$–N/W–N thin film (a) before and (b) after annealing at 800 °C for 30 min.

(RBS) and four point probe, respectively. The B and N profiles in the W–B$^+$–N/W–N thin films were determined by secondary ion mass spectroscopy (SIMS) and RBS. We investigated the crystalline structures and surface morphologies of as-deposited and annealed W–B$^+$–N/W–N and W–N thin films by x-ray diffraction (XRD) and atomic force microscopy (AFM), respectively.

Figure 1 shows XRD patterns of the W–B$^+$–N/W–N and W–N thin films as a function of the annealing temperature. The crystal structure of the as-deposited W–N thin film has a microcrystalline state of W_2N and W phases as shown in Fig. 1(a). However, after boron implantation, the crystallinity of the W–B$^+$–N/W–N thin film appears more broad than that of the W–N thin film as shown in Fig. 1(b). This means that the crystallinity of the W–B$^+$–N/W–N thin film is closer to the amorphous state than the W–N film due to the damage of the ion bombardment. After annealing at 800 °C, the W–N film shows a sharp and strong (110) oriented W phase and (111), (200) oriented W_2N phases. In contrast, there is no sharp and strong (110) W peak in the W–B$^+$–N/W–N thin film. This result suggests that the boron implantation into the W–N thin film modifies the surface of the W–N film, having a tendency to prevent crystallization of the whole W–N layer during annealing even at high temperatures.

Figures 2(a) and 2(b) show transmission electron microscopy (TEM) micrographs of the W–B$^+$–N/W–N thin films before and after the annealing at 800 °C for 30 min. We can tell that the crystal structure of the W–B$^+$–N/W–N before the annealing is a mixture of amorphous and microcrystalline phases from the electron diffraction pattern in Fig. 2(a). The amorphous phase remained as it is even after the annealing at 800 °C in all the area except some local region where some degree of grain growth occurred as shown in Fig. 2(b). Generally, it is known that the grain growth takes place by reducing the grain boundary energy among neighboring homogeneous grains. However, B^+ ions in the W–B$^+$–N thin film might play an important role to keep the amorphous phase such that the microcrystalline grains are embedded in the amorphous phase. Therefore, these microcrystalline grains cannot be free to crystallize into large grains because there is no neighboring grain boundary available in the surrounding amorphous regions consisting of W–B, B–N, and W–B–N. The amorphous phases perform a barrier against the propagation of grain boundaries of the microcrystalline grains, preventing grain growth during the post-annealing.

To determine the boron and nitrogen redistributions in the W–B$^+$–N/W–N thin film, the depth profiles of B and N atoms are investigated by SIMS before and after annealing at 800 °C as shown in Figs. 3(a) and 3(b). Figure 3(a) shows that the majority of B atoms is located near the surface and this result is well consistent with the TRIM simulation. In this figure, it is found that a B-rich W–B$^+$–N thin layer is formed at the surface after the implantation at 40 keV and 1×10^{17} ions/cm^2. After annealing at 800 °C for 30 min as shown in Fig. 3(b), boron atoms seem to be redistributed into the inside of thin film. The N profile indicates that nitrogen impurity remains constant in the W–B$^+$–N/W–N thin film

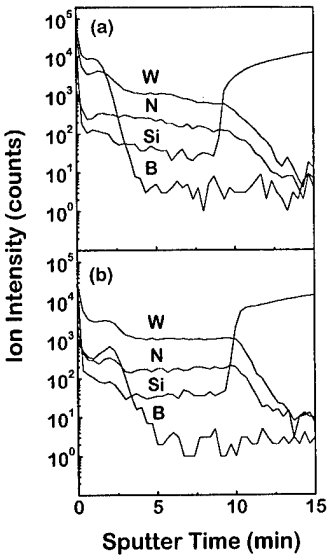


FIG. 3. SIMS depth profiles of W–B$^+$–N/W–N thin film (a) before and (b) after annealing at 800 °C for 30 min.

before and after the annealing. However, in the case of the W–N thin film, 5%–6% of N was diffused out after annealing at 800 °C for 30 min (not shown here). Figure 3(b) suggests that the B^+ ions are redistributed in the inside of the W–N thin film and prevent nitrogen out-diffusion during the annealing at 800 °C for 30 min. From the experimental results we can conclude that B^+ ions prevent the N out-diffusion and keep the W–B$^+$–N/W–N thin film in amorphous ternary phase because the crystallization temperature of W–B$^+$–N/W–N film becomes higher than 800 °C. In conclusion, implanting B^+ ions into W–N, we can improve the barrier property of the W–N thin film due to the stuffing effects of boron and nitrogen to inhibit the grain growth of W or W–N and the formation of W–Si at the contact window. Electrical resistivities of the W–B$^+$–N/W–N and W–N thin films were investigated corresponding to annealing temperature. The resistivity of the W–N thin film deposited at 300 °C is 174 $\mu\Omega$ cm. The resistivity of the W–B$^+$–N/W–N thin film increases to 200 $\mu\Omega$ cm and gradually decreases with increasing the annealing temperature from 500 to 800 °C. The resistivity of W–B$^+$–N/W–N film decreases to 150 $\mu\Omega$ cm, while that of the W–N thin film decreases to 75 $\mu\Omega$ cm after annealing at 800 °C. This result means that the increased resistivity of the W–N thin film after the implantation is due to the amorphous phase, as confirmed with XRD patterns, as well as impurity effect of the W–B$^+$–N/W–N film. Therefore, the decrease in the resistivity after the annealing can be ascribed to annihilation of disordered surface, partial crystallization, and impurity redistribution in the W–B$^+$–N/W–N thin film. The resistivity of the annealed W–N film is relatively lower than that of the annealed W–B$^+$–N/W–N thin film because in the W–N thin film W_2N and W phases could be more easily crystallized than in the W–B$^+$–N/W–N film.

Another interesting point of this work is morphological changes due to the implantation. We observed the surface of W–B$^+$–N/W–N and W–N thin films before and after annealing by AFM. Figure 4(a) reveals that roughness root mean square (rms) of the as-deposited W–N thin film is ~3 Å. However, Fig. 4(b) shows that after annealing at 800 °C the surface roughness greatly increases to 47.46 Å. This re-

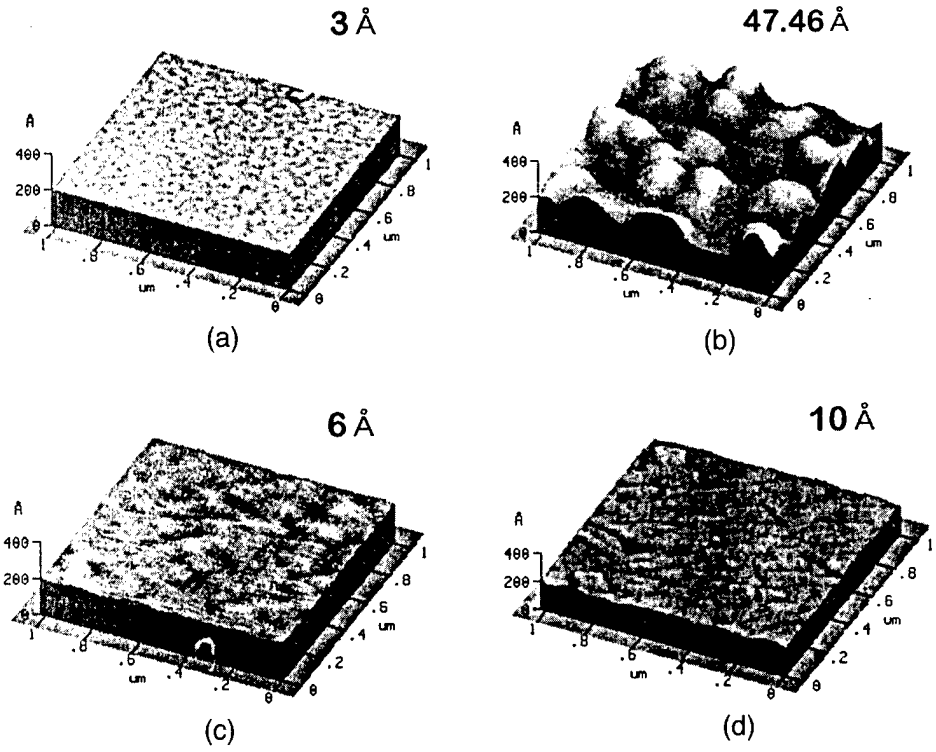


FIG. 4. AFM images of (a) and (b) W–N and (c) and (d) W–B$^+$–N/W–N thin films before and after annealing at 800 °C for 30 min.

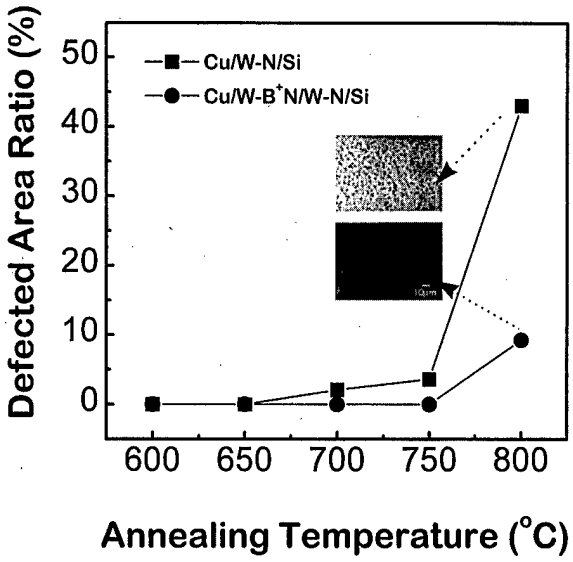


FIG. 5. Defected area ratio on Si surface as a function of the annealing temperature.

sult can be easily understood by considering that large W grain growth takes place in the W–N film during annealing. While, in the case of the W–B$^+$–N/W–N film, the rms of the as-implanted film slightly increases from ~3 to ~6 Å as shown in Fig. 4(c) due to the sputtering phenomena during the implantation process. However, Fig. 4(d) shows that this surface roughness remains at ~10 Å even after the annealing at 800 °C for 30 min. Comparing Fig. 4(d) with Fig. 4(b), we can tell that the smooth surface of the W–B$^+$–N/W–N film after the annealing process is due to the suppression of W grain growth. This result is in good agreement with the fact that the amorphous thin film has a very smooth surface.

Figure 5 shows that the barrier properties of the W–B$^+$–N/W–N and W–N thin films against the Cu diffusion after annealing at 600–800 °C for 30 min. The defected area ratio was determined by the rate of defected area to measured area. We determined the defected area by measuring the etch pit size in the 2×2 cm^2 and multiplying it by the number of etch pit since the etch pit size is quite different with each other. After that, we normalized the defected area by the measured area (2×2 cm^2). This figure indicates that the W–B$^+$–N/W–N barrier prevents the Cu diffusion up to the annealing temperature of 750 °C. In the case of W–B$^+$–N/W–N barrier, the defected area ratio is as low as 0.25 times that of the W–N barrier even at 800 °C.

In summary, we can modify the thermal stability and crystalline phase of the W–N thin film by the boron implantation. The W–B$^+$–N/W–N thin film seems to have the following advantages: We can prepare the excellent barrier against the Cu diffusion without another etching process of the W–B$^+$–N layer and reduce the resistivity of ternary refractory metal film to 200 $\mu\Omega$ cm. The thermal stability is improved as high as 750 °C and the smooth surface roughness can be obtained after annealing at 800 °C. These advantages are mainly due to the thermal behavior of B$^+$ ions to prevent the grain growth of W, W–Si, and W–N in the W–B$^+$–N/W–N thin film: i.e., boron and nitrogen impurities in the amorphous W–B$^+$–N/W–N thin film to inhibit the reaction with W and Si at the contact window as well as preventing the Cu diffusion.

This work was financially supported by the Ministry of Science and Technology under Contract No. 2N15742.

[1]S. P. Murarka, Mater. Sci. Eng., R. **19**, 87 (1997).
[2]M. E. Thomas, Mater. Res. Soc. Symp. Proc. **337**, 13 (1994).
[3]N. Cheung, H. von Seefeld, and M.-A. Nicolet, *Proceedings of the Symposium on Thin Film Interface and Interactions* (Electrochemical Society, New York, 1980), p. 323.
[4]K. Holoway, P. M. Fryer, C. Cabral, Jr., J. M. E. Harper, P. J. Bailey, and K. H. Kelleher, J. Appl. Phys. **71**, 5433 (1992).
[5]Y. T. Kim, C. W. Lee, and S. K. Min, Jpn. J. Appl. Phys., Part 1 **32**, 6126 (1993).
[6]D. J. Kim, Y. T. Kim, and J. W. Park, J. Appl. Phys. **82**, 4847 (1997).
[7]Y. T. Kim, D. J. Kim, C. W. Lee, and J. W. Park, Proc. SPIE **3214**, 48 (1997).
[8]J. S. Reid, R. Y. Liu, P. M. Smith, R. P. Ruiz, and M.-A. Nicolet, Thin Solid Films **262**, 218 (1995).
[9]Y. T. Kim, C. W. Lee, and D. J. Kim, Appl. Phys. Lett. **72**, 1507 (1998).
[10]C. S. Kwon, D. J. Kim, C. W. Lee, Y. T. Kim, and I. H. Choi, Mater. Res. Soc. Symp. Proc. **355**, 441 (1995).
[11]C. W. Lee, Y. T. Kim, and J. Y. Lee, Appl. Phys. Lett. **64**, 619 (1994).

Growth of a near-atomic protrusion on molybdenum field emitter tips under argon ion bombardment

F. Okuyama,[a] H. Sugie,[b] and M. Sato
Department of Environmental Technology, Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showaku, Nagoya 466-8555, Japan

(Received 20 July 1998; accepted 28 May 1999)

Ar^+ ion bombardment of a molybdenum field emitter is shown to grow a single protrusion with a near-atomic dimension at the very tip of the emitter. This kind of protrusion is generated from a thick amorphous layer covering the monocrystalline emitter tip, and lengthens toward the incoming ions. The protrusion apex is roughened, resulting in atomic projections on its surface. This surface phenomenon may be developed into a technique to prepare metal tips with atomically small dimensions. © *1999 American Vacuum Society.* [S0734-211X(99)09304-X]

One of the key concerns in focused-electron beam techniques is how to produce field emission tips with atomically small dimensions. Although one technique of tip preparation is to polish metal wires electrolytically, such electropolishing results in heavy contamination of the tip surface by the electropolishing chemical agents. While surface contaminants can be easily removed by either heating[1] or field evaporation,[2] these simple methods generally entail tip blunting, which depends on the melting point of the metal and temperature required. An alternate technique is therefore required to clean electropolished tips without losing their original sharpness.

One candidate for the above purpose is ion milling or sputtering in an ultrahigh vacuum. In ion sputtering, the atoms at the outermost surface are ejected into a vacuum through a collisional cascade in the subsurface.[3] This process etches away the contaminated surface layer, thereby exposing the noncontaminated subsurface to the vacuum. The etching rate is a function of the ion incidence angle, the maximum being around 60°; its inevitable consequence is a conical or pyramidal shaping of surface projections.[4,5] Sputtering also reshapes field emission tips into a cone or pyramid with a very sharp apex, irrespective of their original geometry. For example, hemispherical tips of refractory metals are sharpened into cones a few nm in radius of curvature by bombardment with inert-gas ions at low energies ($\lesssim$10 keV).[6–9] Recently, we found an anomaly, i.e., a single protrusion with a near-atomic dimension induced to grow on an ion-bombarded field emitter tip of molybdenum (Mo). This phenomenon, or the "growth of near-atomic protrusion" (GNAP), may be developed into a technique to prepare "atomic-scale" metal tips, which show promise as coherent electron sources and scanning microprobes.

The experimental system was a JAMP-7100 scanning Auger microprobe (SAM), the analyzing chamber of which was slightly modified.[10] In brief, the sample was mounted on a eucentric goniometer stage, which enabled us to direct the sample tip toward the focused-beam ion gun. Since the ion-current distribution across the beam cross section is Gaussian, the tip axis must be aligned with the ion-beam axis as accurately as possible; otherwise, the resulting tip geometry would be asymmetrical. This axial alignment was accomplished by carefully operating the goniostage while observing the sample using SAM's scanning electron microscope (SEM) capability. The sputtering conditions involved 3 keV Ar^+ ions focused into a circle $\sim$300 μm in diameter with a current density of $\sim 1\times10^3$ μA/cm^2, equivalent to an ion-dose rate of $\sim 5\times10^{17}$ ions/cm^2 min. The sample was kept at room temperature during sputtering.

To visualize the tip-sharpening process, the tip geometry was examined by using a field emission (FE) SEM (S-5000) as a function of sputtering time. To do so, the sample was repeatedly transferred back and forth in air from the SAM to the FE SEM.

Polycrystalline Mo wires spot welded on W loops were electropolished in KOH solution to serve as samples. To smooth their surfaces, the original sample tips were flashed at $\sim$2000 °C in an auxiliary vacuum chamber. The analyzing chamber was modified to install a counterelectrode for monitoring the field electron current emitted from the sample tip. This chamber modification was based on our experience that the current–voltage (I–V) curves reflect the tip sharpness (cf. Fig. 2). In I–V measurements, the tip and the counterelectrode were positioned opposite each other at a distance of $\sim$100 μm with the aid of the goniometer stage and the counterelectrode manipulator. Since the FE SEM was not able to fully resolve the sputtered tips, some samples were analyzed by high-resolution (HR) transmission electron microscopy (TEM). For this purpose, the emitter was mechanically removed from the supporting loop with the greatest care.

Figure 1 illustrates the tip-sharpening process observed by FE SEM. Note that the tip dimension was dramatically reduced by the bombarding Ar^+ ions. Indeed, the final sample tip was too fine to be perfectly resolved with the FE SEM. The vertical lines in Figs. 1(a) and 1(b) indicate the same area with a cross-sectional diameter of $\sim$120 nm, and thus served as a "marker" to estimate the sputter-removal rate. It is clear by comparing Figs. 1(a) and 1(b) that sputtering proceeded at a rate of $\sim$9 nm/min along the center axis.

[a]Electronic mail: okuyama@system.nitech.ac.jp
[b]Also with Department of Electrical and Computer Engineering.

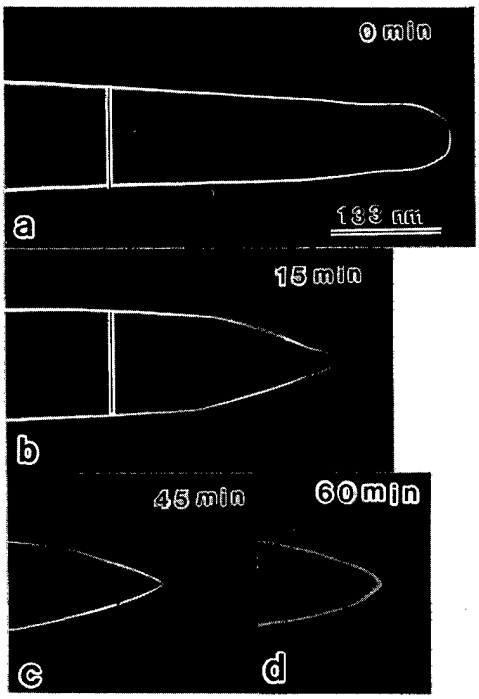


FIG. 1. Typical tip-sharpening process due to Ar^+ sputtering, observed by FE SEM. The total sputtering time is indicated in the respective images. The vertical lines in (a) and (b) correspond to a cross-sectional diameter of ~120 nm.

The $I-V$ curves (a)–(d) in Fig. 2 correspond to the FE SEM images (a)–(d) in Fig. 1, respectively. The "onset voltage," at which the electron current from the tip began to be detectable, was around 900 V before sputtering, decreasing to ~150 V after a 60 min sputtering. (Due to the microampere sensitivity of the ammeter, emission currents in the nanoampere region could not be detected.) This dramatic decrease in onset voltages is directly related to the increase in tip sharpness illustrated in Fig. 1. Since the onset voltage, designated as V_0, depends on the distance between the emitter and counterelectrode, the distance had to be kept constant

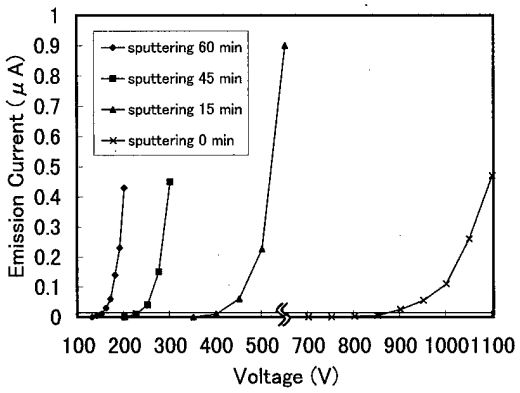


FIG. 2. $I-V$ curves for the tip geometries in Fig. 1. Curves (a), (b), (c), and (d) correspond to the FE SEM images (a), (b), (c), and (d), respectively. The horizontal line roughly indicates the lowest level of detectable current emitted, the intersections of which with the respective curves give a measure of onset voltage.

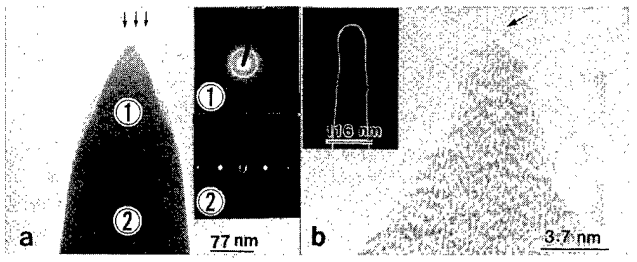


FIG. 3. (a) TEM image of a sputtered Mo tip taken at a medium magnification. The inset in the image shows the ED patterns from areas 1 and 2, which prove that the tip area was entirely amorphous, with the shank area maintaining monocrystallinity. (b) Enlarged image of the apex area, revealing a near-atomic protrusion grown at the apex. The inset in (b) is the original emitter geometry (FE SEM image). The corresponding onset voltage was around 150 V. The arrows in (a) indicate the ion-incidence direction, while the arrow in (b) indicates an atomically small projection. Instrument: JEM-3010.

for repetitive measurements in order to use V_0 as a measure of tip sharpness. Although this required careful operation of the goniostage and the counterelectrode manipulator, the method was very convenient in that it was *in situ* and hence time saving. (The sample transfer needed for FE SEM was enormously difficult and time consuming.)

Ion sputtering modifies the target surface structurally as well as topographically.[11] Figure 3(a) shows the medium magnification TEM image of a sputtered emitter, in which the emitter is seen to be entirely amorphous at the tip area. [Note also the halo electron diffraction (ED) pattern inset in the image.] By energy dispersive x-ray analysis, the amorphous area was confirmed to be solely composed of Mo (data not shown). The shank area, on the other hand, was monocrystalline, as corroborated by the corresponding ED pattern. The amorphized depth was around 200 nm along the emitter axis, which is two orders of magnitude larger than the ion-penetration depth computed for 3 keV Ar^+ ions normally incident on polycrystalline Mo (~3 nm). Monocrystalline semiconductor surfaces are known to be amorphized by bombarding ions, but this "ion-induced amorphization" does not occur on metal surfaces. The physical process underlying this tip amorphization thus remains unknown.

In the above TEM image, the very tip (or apex) was not discerned distinctly, implying that the emitter apex was extremely sharp. To obtain a clear image, this emitter tip was re-observed in the HR mode. The result is shown in Fig. 3(b), which reveals a single protrusion at the apex. The center axis of the protrusion was parallel to the ion incidence direction, meaning that the protrusion was oriented toward the incident ions. If the protrusion's tip is replaced with a sphere, its radius is around 1.2 nm; the protrusion was of near-atomic dimensions. Actually, however, the surface of the protrusion's apex was not smooth but roughened to accommodate an atomic projection (see arrow).

Figure 4 shows another example of near-atomic protrusion. This protrusion was ~17 nm in length, and thus the protrusion itself was like a "nanometric" field emitter. The tip of the original emitter before sputtering was slightly sharpened at the apex (see the FE SEM image inset). The

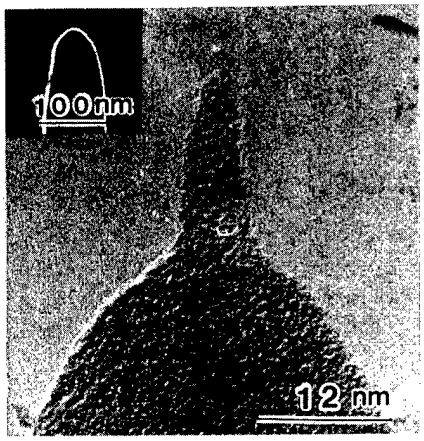


FIG. 4. Another example of near-atomic protrusion (HR TEM image). The inset shows the sample before sputtering (FE SEM image).

apex of the original sample in Fig. 3, by contrast, was flat. The fact that near-atomic protrusions grew independently of the original tip geometry is proof of its reproducibility.

Until now, no atomic-scale protrusion has been detected on Ar^+-sputtered W emitters. However, W tips sputtered at the optimal ion doses were generally around 5 nm in radius. Since such a nanometric W tip has a wide range of applications including high performance electron sources,[9,12] the sputtering behavior of W tips that we observed may be of practical significance. The details will be reported in a forthcoming article.

As noted earlier, the nature of GNAP is still unclear. This issue will be elucidated by successively following the growth dynamics of near-atomic protrusions, with the aid of real-time TEM.[13]

Sincere thanks are expressed to Professor H. Morikawa and Y. Fujimoto for their technical assistance. The authors are also grateful to Y. Katada for the figure drawing. This work was supported by the Ministry of Education, Science, Sports and Culture, Japan, through Grant-in-Aid for Developmental Research A No. 05505001.

[1]W. P. Dyke, J. K. Trolan, W. W. Dolan, and G. Barnes, J. Appl. Phys. **24**, 570 (1953).
[2]E. W. Müller and T. T. Tsong, in *Field Ion Microscopy* (Elsevier, Amsterdam, 1969), Chap. III.
[3]P. Williams, in *Applied Atomic Collision Physics,* edited by S. Datz (Academic, Orlando, 1983), Vol. 4, Chap. 7.
[4]A. D. G. Stewart and M. W. Thompson, J. Mater. Sci. **4**, 56 (1967).
[5]O. Auciello, J. Vac. Sci. Technol. **19**, 841 (1981).
[6]D. K. Biegelson, F. A. Ponce, and J. C. Tramontana, Appl. Phys. Lett. **54**, 1223 (1989).
[7]J. A. Kubby and B. M. Siegel, J. Vac. Sci. Technol. B **4**, 120 (1986).
[8]S. Morishita and F. Okuyama, J. Vac. Sci. Technol. A **9**, 167 (1991).
[9]C. Schiller, A. A. Koomans, T. L. Van Rooy, C. Schönenberger, and H. B. Elswijk, Surf. Sci. **339**, L925 (1995).
[10]M. Sato (unpublished).
[11]R. Kelly, in *Ion Bombardment Modification of Surfaces,* edited by O. Auciello and R. Kelly (Elsevier, Amsterdam, 1984), Chap. 3.
[12]H. W. Fink, IBM J. Res. Dev. **30**, 460 (1986).
[13]F. Okuyama and J. Kato, Surf. Sci. **338**, L857 (1995).

SHOP NOTES

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.

Multilayer resist films applicable to nanopatterning of insulating substrates based on current-injecting scanning probe lithography

Hiroyuki Sugimura[a] and Osamu Takai
Department of Materials Processing Engineering, Graduate School of Engineering, Nagoya University, Nagoya 464-8603, Japan

Nobuyuki Nakagiri
1st R&D Department, R&D Headquarters, Nikon Corporation, Tokyo 140-8601, Japan

(Received 26 February 1999; accepted 23 April 1999)

Scanning probe lithography based on localized current injection using the probe tip of atomic force microscopy (AFM) has been applied to nanoprocessing of an insulating substrate. An electrically conductive resist film composed of triple layers was developed for this current-injecting AFM lithography. The bottom layer of the resist, which served as a current pass during patterning, consisted of amorphous silicon (a-Si) with 20 nm in thickness prepared by ion-beam sputtering. An organosilane monolayer, that is, octadecylsilyl self-assembled monolayer (ODS-SAM) of 2 nm in thickness, was used as the top layer of the resist, therefore, as the imaging layer in which nanoscale patterns were drawn by AFM. In order to bind the a-Si and the ODS-SAM together, the intermediate layer of the resist, that is, Si oxide of 2 nm in thickness, was prepared by photooxidation of the a-Si layer. Through an AFM-lithographic process using this multilayered resist, nanofabrication of fine grooves on a Si oxide substrate was demonstrated. The minimum feature size about 50 nm was successfully fabricated. © *1999 American Vacuum Society.* [S0734-211X(99)02004-1]

Scanning probe microscopes (SPMs) have shown high potential as tools for nanolithography which is of primary importance in order to attain higher integration density of future electronic microdevices.[1–3] Scanning probe lithography based on atomic force microscopy (AFM) is particularly promising, since this technique provides us with versatility in materials to be patterned, and with high line-drawing rates in the range of thousands of μm/s.[4–6] Further important advantages of AFM-based scanning probe lithography are compatibility with other high throughput lithographies such as photolithography[7,8] and that with *in situ* electrical characterization of fabricated nanopatterns.[9,10] At present, most of the successful results have been obtained using the AFM lithography in which patterning was conducted with injecting current or applying electric field to a substrate. Hence, substrate materials patterned through such AFM lithographies had to be electrically conductive so that the application of electrically conductive thin films to scanning probe lithography was explored.[11–13]

Organic thin films have been widely applied as resists for various types of lithography. Patterning of organic materials by AFM is of particular interest as well. Monolayers formed through chemisorption of organosilane molecules onto an oxide surface have successfully served as resist films for high-resolution lithography, e.g., electron beam lithography[14] and scanning probe lithography.[15,16] Due to hydrophobic and van der Waals interactions between the alkyl chains of the organosilane molecules, the molecules are spontaneously organized into a thin layer of monomolecular thickness in which the molecules are closely packed together. Such a film belongs to a class of material known as self-assembled monolayer (SAM). The organosilane SAMs, therefore, fulfill the requirements for high-resolution resist films including thickness, uniformity, patternability and compatibility to various pattern transfer processes and, therefore, are promising for scanning probe lithography.

In this article, we demonstrate the capability of organosilane SAMs in the patterning of an insulator, that is, SiO_2 which is a key insulating material for Si-based microdevices, based on scanning probe lithography. However, the SAMs have not been applicable so far to patterning of insulating substrates since they lack electrical conductivity similar to most organic materials. In order to overcome such a shortcoming, we have combined the SAM and an electrically conductive film and served this film as a resist for current-injecting AFM lithography. Before selecting the conductive film, we considered the following conditions. First, it can be covered with an organosilane SAM. Second, it can be used as a mask for chemical etching of SiO_2. Third, it can be formed as a thin and uniform film of a few tens of nm in

[a]Electronic mail: sugimura@numse.nagoya-u.ac.jp

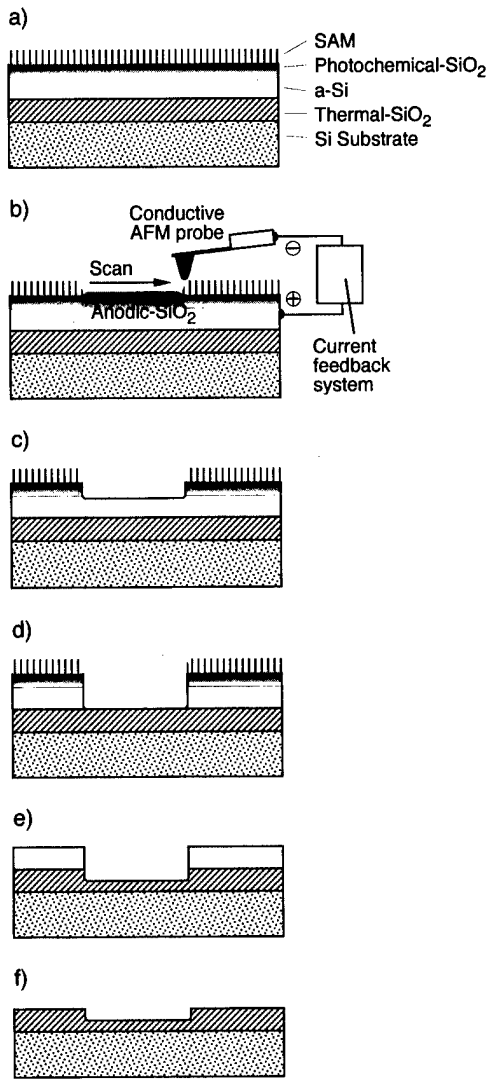


FIG. 1. Schematic illustration of AFM lithography for nanopatterning of SiO_2 using the multilayered resist system. (a) Cross section of a sample coated with a resist film. (b) Resist exposure. Drawing patterns into the imaging layer (ODS-SAM) by current injection from an AFM probe in constant current mode. (c) Resist development Step I. HF etching in order to remove the SiO_2 layer of the resist in the exposed area. (d) Resist development Step II. TMAH etching in order to remove the *a*-Si layer in the exposed area. (e) Pattern transfer. HF etching in order to remove the thermal SiO_2 in the exposed area. (f) Resist removal. TMAH etching for the whole *a*-Si layer.

thickness. Finally, it would be trouble free and compatible to the fabrication process of Si-based microelectronic devices. Accordingly, amorphous silicon (*a*-Si) was chosen as a candidate and a multilayered resist film was prepared. In this resist system, an organosilane SAM and an *a*-Si film were served as as an imaging layer and a current-passing layer, respectively.

Figure 1 shows a process chart of AFM lithography for SiO_2 nanopatterning using the multilayered resist film. A sample substrate was a single crystal Si plate with a thermally grown oxide layer of ~20 nm in thickness. A triple-layered resist film was prepared on this Si substrate [Fig. 1(a)]. A thin Si film of 20 nm in thickness, which was confirmed to be amorphous by electron microscopy and diffractometry, was deposited on the substrate by ion-beam sputtering. Next, the sample was cleaned by ultraviolet light-generated ozone cleaning using an excimer lamp of $\lambda = 172$ nm (Ushio Electric, UER20-172 V). By this cleaning, organic contamination on the *a*-Si layer was removed and a clean surface oxide photochemically prepared on the layer. The oxide surface was hydroxylated and, therefore, completely hydrophilic with its water contact angle less than 5°. Finally, the cleaned sample substrate was exposed to *n*-octadecyltrimethoxysilane (OTMS) vapor. Through the chemical reaction between the surface hydroxyl groups and the OTMS vaporized molecules, the molecules immobilized onto the substrate resulting in the formation of an octadecylsilyl (ODS) monolayer. The thickness of this monolayer was estimated by ellipsometry to be 1.5–2 nm. The monolayer surface was hydrophobic showing a water contact angle of ~105°. Details of this chemical vapor deposition were described elsewhere.[17,18] Note that our multilayered resist film consists of triple layers, that is, the ODS-SAM, the photochemically grown SiO_2 layer and the *a*-Si film from the top surface of the resist. This intermediate SiO_2 layer is crucial for this resist system, since, without this, the SAMs are not formed. The SAM formation proceeds due to the chemical reaction between precursor organosilane molecules and hydroxyl sites on the oxide surface.

The sample was patterned using an AFM (Park Scientific Instruments, Autoprobe LS) as shown in Fig. 1(b). In order to inject current into the ODS monolayer, a bias voltage was applied between the AFM probe and the *a*-Si layer of the resist which was positively polarized. A conductive probe (Park Scientific Instruments, Ultralever, a heavily doped Si probe) was used. When a SPM is operated in the presence of atmospheric water vapor, the SPM probe-sample junction is connected through a water column created by capillarity of the adsorbed water and can serve as a minute electrochemical cell.[19,20] As a result of electrochemical reactions induced in this cell, the ODS monolayer became degraded in the region where the probe had been passed. Hence, the probe-scanning pattern was printed into the ODS monolayer on the top of the multilayered resist film. It was confirmed that anodization of the underlying Si occurred in the probe-scanned region.[21] As schematically illustrated in Fig. 1(b), the oxide film in the probe-scanned region became thicker than the photochemical oxide due to the growth of anodic oxide.

In current-injecting AFM lithography, current flowing through the probe-sample junction is a key factor in order to control it. The dependency of the junction current on the lithography has been studied by applying a constant bias voltage.[22] We have also demonstrated that line drawing with faster probe-scan rates could be achieved by increasing the junction current through applying a higher bias voltage. However, the controllability of the current was not satisfactory in using the constant bias mode. The relationship between the junction current and the bias voltage depended on several factors, such as the age and idiosyncrasies of the particular probe used. The current could not be precisely

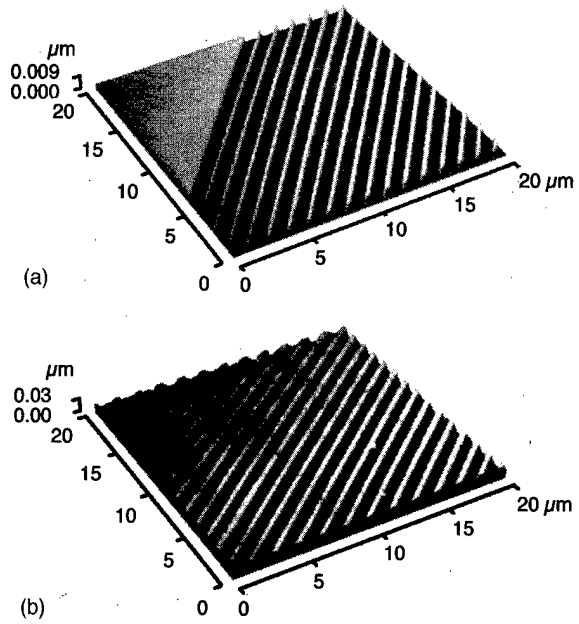


FIG. 2. AFM patterning of a resist at a probe scan rate of 10 μm/s and a probe current of 5 nA. (a) AFM image acquired after development Step I. (b) AFM image acquired after development step II. The etching duration times were 0.5 and 3 min, respectively.

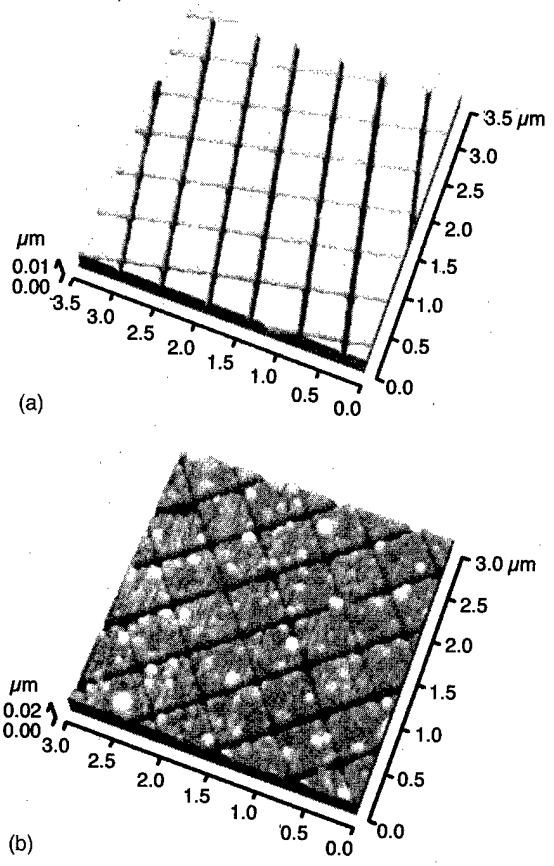


FIG. 3. AFM images of a nanopatterned SiO_2 surface. AFM patterning was conducted at a probe scan rate of 20 μm/s and a probe current of 1 nA. (a) AFM image of the exposed and developed resist film through Steps I and II prolonged for 0.5 and 3 min, respectively. (b) AFM image of the SiO_2 substrate etched for 5 min in HF using the developed resist as an etching mask. The etch depth was estimated to be ~10 nm. Prior to imaging, the resist film was removed by TMAH etching for 4 min.

controlled by simply applying a defined bias voltage. Thus, a current feedback system, which enables to perform AFM lithography in constant current mode, is crucial to the reliable AFM lithography.[6,23] In the present study, the AFM was equipped with a constant current source (Keithley, SMU Model 236).

The developing process for the exposed resist film was a two-step chemical etching conducted at room temperature. Here we demonstrate this resist development using rather wide lines that were drawn by the use of a dull-tip probe for ease of etch depth measurement. At the first step of the process [Step I, Fig. 1(c)], the exposed resist film was etched in an aqueous solution of 0.5 wt. % hydrofluoric acid (HF). Figure 2(a) shows an AFM image[24] of the resist film after Step I. In the current injected region, i.e., where the probe had been passed, both the photochemical and anodic oxides were selectively etched while the unscanned region remained unetched due to the protection by the ODS-SAM. The etch depth, estimated to be ~5 nm from the cross section of the AFM image, was slightly greater than the sum of the thicknesses of the ODS-SAM and the photochemical oxide, that is, 4 nm at the thickest. This result indicates that the bottom *a*-Si layer was also etched to the depth of at least 1 nm and, thus, had further oxidized due to anodization in the probe scanned region. Although organosilane SAMs are gradually degraded in HF solution, we confirmed that the ODS-SAM on Si endured at least for 5 min in 0.5 wt. % HF solution.

At the second step of the resist developing process [Step II, Fig. 1(d)], the HF-etched resist was further treated in an aqueous solution of 25 wt. % tetramethylammonium hydroxide (TMAH). Since, in this solution, Si is effectively dissolved while SiO_2 is rarely etched,[25] the *a*-Si is expected to be locally etched in the region where its surface oxide had been removed. In addition, this *a*-Si etching actually stops when the whole *a*-Si layer is etched and the underlying substrate, i.e., thermal SiO_2, is exposed. As clearly seen in the AFM image shown in Fig. 2(b), the current-injected region became further etched compared with Fig. 2(a). Its depth was almost equal to the thickness of the *a*-Si layer in the multilayered resist. The probe-scanning pattern printed in the imaging layer, that is, ODS-SAM, on top of the resist was successfully transferred into the *a*-Si layer. Thus, the developing process of the exposed multilayer resist was concluded to be completed.

Nanostructures were fabricated on the substrate SiO_2 through chemical etching using the developed resist film as an etching mask. Figure 3(a) shows an AFM image of a developed resist film with a grid pattern at horizontal and vertical intervals of ~500 nm. This sample was further etched in the HF solution. The grid pattern on the resist film was thus transferred to the thermal SiO_2 [Fig. 1(e)]. After this pattern transfer step had been completed, the resist film was etched again in the TMAH solution. Since the ODS-SAM and the photochemical oxide had been removed

through the step shown in Fig. 1(e), the remaining *a*-Si layer was etched and degraded and removed from the sample [Fig. 1(f)]. The result is shown in Fig. 3(b). This AFM image demonstrates that fine grooves less than 50 nm in width were successfully fabricated on the thermal SiO_2 substrate, although some part of the *a*-Si film, particulate features seen on the surface, remained after the final TMAH etching.

Pattern widening must occur during the pattern transfer processes since wet chemical etchings employed in the present experiments proceed isotropically. It becomes more significant when an etch depth increases. The minimum linewidth which can be written into the multilayer resist is thus determined by the thickest layer, that is, the *a*-Si layer. The resolution limit is thought to be larger than its thickness of 20 nm in the present case.

In summary, we have demonstrated nanostructuring of insulator based on current-injecting AFM lithography. A multilayered resist film with electrical conductivity was employed. This resist consisted of triple layers, that is, *a*-Si, its photochemical oxide and ODS-SAM with thicknesses of 20, 2, and 2 nm, respectively. Nanoscale patterns were first defined by AFM in the ODS-SAM. These patterns were then transferred into the *a*-Si layer through the two-step resist developing process. Finally, the substrate SiO_2 was nanostructured using the developed resist as an etching mask.

This research has been supported in part by a Grant-in-Aid for Scientific Research on the Priority Area of "Electrochemistry of Ordered Interface" and a Grant-in-Aid for Scientific Research (No. 10555247) by the Ministry of Education, Science, Sports, and Culture, Japan.

[1]K. Matsumoto, S. Takahashi, M. Ishii, M. Hoshi, A. Kurokawa, S. Ichimura, and A. Ando, Jpn. J. Appl. Phys., Part 1 **34**, 1387 (1995).
[2]S. C. Minne, H. T. Soh, Ph. Flueckiger, and C. F. Quate, Appl. Phys. Lett. **66**, 703 (1995).
[3]M. S. Hagedorn, D. D. Litfin, G. M. Price, A. E. Gordon, and T. K. Higman, J. Vac. Sci. Technol. B **14**, 4153 (1996).
[4]S. W. Park, H. T. Soh, C. F. Quate, and S. Park, Appl. Phys. Lett. **67**, 2415 (1995).
[5]J. P. Bourgoin, R. V. Sudiwala, and S. Palacin, J. Vac. Sci. Technol. B **14**, 3381 (1996).
[6]H. Sugimura and N. Nakagiri, Nanotechnology **8**, A15 (1997).
[7]A. Born, C. Hahn, M. Löndorf, A. Wadas, Ch. Witt, and R. Wiesendanger, J. Vac. Sci. Technol. B **14**, 3625 (1996).
[8]H. Sugimura and N. Nakagiri, Jpn. J. Appl. Phys., Part 2 **36**, L968 (1997).
[9]E. S. Snow and P. M. Campbell, Science **270**, 1639 (1995).
[10]T. Yasue, H. Koyama, T. Kato, and T. Nishioka, J. Vac. Sci. Technol. B **15**, 614 (1997).
[11]E. S. Snow, P. M. Campbell, and B. V. Shanabrook, Appl. Phys. Lett. **63**, 3488 (1993).
[12]H. Sugimura, T. Uchida, N. Kitamura, and H. Masuhara, J. Vac. Sci. Technol. B **12**, 2884 (1994).
[13]N. Kramer, J. Jorritsma, H. Birk, and C. Schonenberger, J. Vac. Sci. Technol. B **13**, 805 (1995).
[14]M. J. Lercel, R. C. Tiberio, P. F. Chapman, H. G. Craighead, C. W. Sheen, A. N. Parikh, and D. L. Allara, J. Vac. Sci. Technol. B **11**, 2823 (1993).
[15]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).
[16]H. Sugimura and N. Nakagiri, Langmuir **11**, 3623 (1995).
[17]H. Sugimura and N. Nakagiri, J. Photopolym. Sci. Technol. **10**, 661 (1997).
[18]A. Hozumi, K. Ushiyama, H. Sugimura, and O. Takai, Langmuir (in press).
[19]R. L. MacCarley, S. A. Hendricks, and A. J. Bard, J. Phys. Chem. **96**, 10089 (1992).
[20]H. Sugimura, T. Uchida, N. Kitamura, and H. Masuhara, J. Phys. Chem. **98**, 4352 (1994).
[21]H. Sugimura and N. Nakagiri, J. Vac. Sci. Technol. A **14**, 1223 (1996).
[22]H. Sugimura, K. Okiguchi, N. Nakagiri, and M. Miyashita, J. Vac. Sci. Technol. B **14**, 4140 (1996).
[23]K. Wilder, H. T. Soh, A. Atalar, and C. F. Quate, J. Vac. Sci. Technol. B **15**, 1811 (1997).
[24]Image SXM version 1.61 was used for data processing of the AFM images in this article. This software is an extention of NIH Image (public domain software) for images acquired by SEM, SPM, SAM, etc., and is written by S. Barret.
[25]O. Tabata, R. Asahi, H. Funabashi, K. Shimaoka, and S. Sugiyama, Sens. Actuators A **34**, 51 (1992).

Papers from the 26th Conference on the Physics and Chemistry of Semiconductor Interfaces

17–21 January 1999
Catamaran Resort Hotel
San Diego, California

Sponsored by
The American Vacuum Society
The Army Research Office
The Office of Naval Research

Editor for the Conference:
J. E. Rowe



Published by the American Vacuum Society through
the American Institute of Physics, New York, 1999

ORGANIZING COMMITTEE

CONFERENCE CHAIR
Karen Kavanagh
Univ. California, San Diego

PROGRAM COMMITTEE

D. E. Aspnes (North Carolina State Univ.)
L. Cooper (Office of Naval Research)
J. D. Dow (Arizona State Univ.)
R. M. Feenstra (Carnegie Mellon Univ.)
G. Lucovsky (North Carolina State Univ.)
C. J. Palmstrom (Univ. of Minnesota)
J. E. Rowe (Army Research Office)
H. W. M. Salemink (IBM Res. Labs, Zurich)
J. S. Speck (Univ. California, Santa Barbara)
D. J. Wolford (Iowa State Univ.)

CONFERENCE ADMINISTRATION
Carole and John Dow

PREFACE

The Proceedings of the 26th Annual Conference on the Physics and Chemistry of Semiconductor Interfaces (PCSI-26) are contained in this volume. The Conference was held at the Catamaran Resort Hotel in Pacific Beach, San Diego, California from Sunday evening, 17 January 1999 to noon on Thursday 21 January 1999. A total of 18 invited and 69 contributed papers were presented, with 113 attendees.

This year's Conference saw a large number of papers relating to nitride semiconductor interfaces and piezoelectric effects (41) reflecting the increasing interest in that field worldwide. Besides the traditional emphasis on growth and characterization of other compound semiconductor interfaces (35) a significant number of papers addressed the chemistry and physics of SiO_2/Si interfaces (23). There were also a small but significant number of papers dealing with magnetic materials, and spin transport (6), an area that the PCSI Committee believes will increase in the future. Carol Ashby presented some interesting new results on the role of As in oxidation of AlGaAs layers. Steve Streiffer gave an overview of recent work on Perovskite structure ferroelectric films for possible gate dielectrics in transistors. Other exciting advances reported at the Conference include: silicon substrate-based optoelectronic structures predicted by John Joannopoulos from MIT, MBE studies of GaN growth, impurity and surfactant effects in GaN total energy calculations, and real-time x-ray scattering studies of MOCVD growth of GaN films. The Tuesday "Rump" session's theme this year was piezoelectric effects at semiconductor interfaces and included stimulating talks by D. L. Smith and a number of others.

The meeting successfully brought together researchers from a variety of specialties that would not normally interact. The quality of the presentations was very high, and lively discussion occurred following each talk. All participants deemed this year's meeting a great success. The Conference succeeded through the efforts of many persons. We would like to thank particularly the many anonymous reviewers who performed their work carefully and thoroughly, Jack Dow for conference 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 *JVSTB* is also appreciated. The Program Committee would particularly like to thank Carole Dow for handling much of the conference organizational details, including hotel negotiations, abstracts, and abstract booklets as well as conference registration. The staff at the Catamaran Resort Hotel including Lisa Marie McNorton are thanked for making our stay there very enjoyable. We also greatly appreciate the financial support provided by the Office of Naval Research (Larry Cooper) and the Army Office of Scientific Research (Jack Rowe). The conference was held under the sponsorship of the American Vacuum Society (Electronic Materials and Processing Division). Finally, we thank our conference mascot the Grey Pelican for making frequent appearances on the beaches nearby.

Jack E. Rowe
Conference Proceedings Editor

Karen L. Kavanagh
Conference Chairperson

Deliberately designed interfaces for monolithic integration in optoelectronics

Tairan Wang,[a] N. Moll, Kyeongjae Cho,[b] and J. D. Joannopoulos
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 20 January 1999; accepted 30 March 1999)

A novel class of semiconductors is introduced, based on computational design, to solve the long-standing problem of lattice and polarity mismatch in heteroepitaxial growth of III–V alloys on silicon substrates. *Ab initio* total-energy calculations and quasiparticle GW calculations are used to investigate the physical properties of these new semiconductors. One particular configuration is designed to match lattice constant and polarity with the Si(100) surface and to possess a direct band gap of 1.59 μm, which is close to the canonical frequency used by the optoelectronics industry. These results could pave the way for eventual monolithic integration of optical materials on silicon.
© *1999 American Vacuum Society.* [S0734-211X(99)04104-9]

I. INTRODUCTION

For many years, the issue of monolithic integration of Si technology with optoelectronics has been widely investigated, and yet it still remains largely unresolved. Various different approaches have been suggested.[1–6] For example, one approach is to use porous Si as the optoelectronic material.[1,2] In porous Si, the lack of complete diamond-structure symmetry allows electron transitions to occur which can emit light. Unfortunately, they do not appear to produce the needed intensities. Another approach is to use heteroepitaxy of optically active III–V materials on Si.[3–6] However, the III–V alloys that have band gaps matching the operating wavelength of optical fibers[7] also have large lattice constant mismatches (approximately 8% larger) as compared to crystalline Si. Therefore, during growth many defects are created that propagate from the interface. Attempts to reduce the number of defects inovlve the use of buffer materials to try and slowly increase the lattice constant.[8] In this article, we describe a totally new approach to solve this long-standing problem. We introduce a novel class of deliberately designed optical semiconductor alloys which should, at least in principle, be amenable to layer-by-layer heteroepitaxial growth.

This novel class of semiconductors exhibits the following properties: First of all, it resolves the polarity-mismatch problem since the growth sequence guarantees that each bond has two and only two valence electrons. Second, there are various combinations of elements to choose from, so that the lattice mismatch between the new material and Si could be minimized. Finally, by choosing different combinations of elements, the material could be tuned so that it exhibits a direct band gap, with a band gap value matching the operating wavelength of optical fibers. The goal of this work is to find an appropriate configuration from this new class of materials that satisfies these specifications.

Instead of an experimental search for the optimum configuration, which would clearly be a formidable task given anything beyond binary alloys, we exploit the predictive and tractable power of *ab initio* calculations. For different configurations of our semiconductor alloys we perform *ab initio* calculations to obtain their structural properties and quasiparticle band structures. Before we discuss our results in Sec. IV, we will first describe the general construction of this class of semiconductors and their properties in Sec. II. And in Sec. III we give an overview of the computational details.

II. NOVEL CLASS OF SEMICONDUCTORS

Let us begin by considering the polarity mismatch problem and how one might resolve this issue. The typical growth orientation of Si in molecular beam epitaxy (MBE) or metalorganic chemical vapor deposition (MOCVD) is the (100) orientation. In the (100) direction, a III–V compound has a alternating layering of group-III and group-V elements. Growing a III–V compound on a Si(100) substrate will either create an interface with a group-III element layer and the Si substrate or a group-V element layer and the Si substrate. The interface of group-V element layer and Si substrate is shown in left panel of Fig. 1. In the diamond or zinc-blende lattice, each atom has four bonds with its four neighbors. The electron counting rule has to be satisfied which means that each bond is filled exactly with two electrons. In order for the III–V and Si interface to fulfill this, half an electron per group-V interface atom has to be redistributed to the outer surface of the structure. This creates a long-range electric field that leads to the polarity mismatch. In the same way, a group-III layer at the interface would create a long-range electric field because of the lack of an half electron per group-III atom to satisfy the electron counting rule.

To solve this polarity mismatch, we generalize the concept of III–V materials. A III–V material is a group-IV material where the group-III and the group-V atoms alternate and together have the same property as group-IV atoms. Every group-III atom provides 3/4 of an electron to a bond, whereas every group-V atoms provides 5/4 of an electron. Therefore, the III–V material satisfies the electron counting rule. Following this concept we create a class of pseudo III–V materials with two possibilities:

[a]Electronic mail: tairan@mit.edu
[b]Present address: Mechanics and Computation Division, Department of Mechanical Engineering, Stanford University, Stanford, CA 94305-4040.

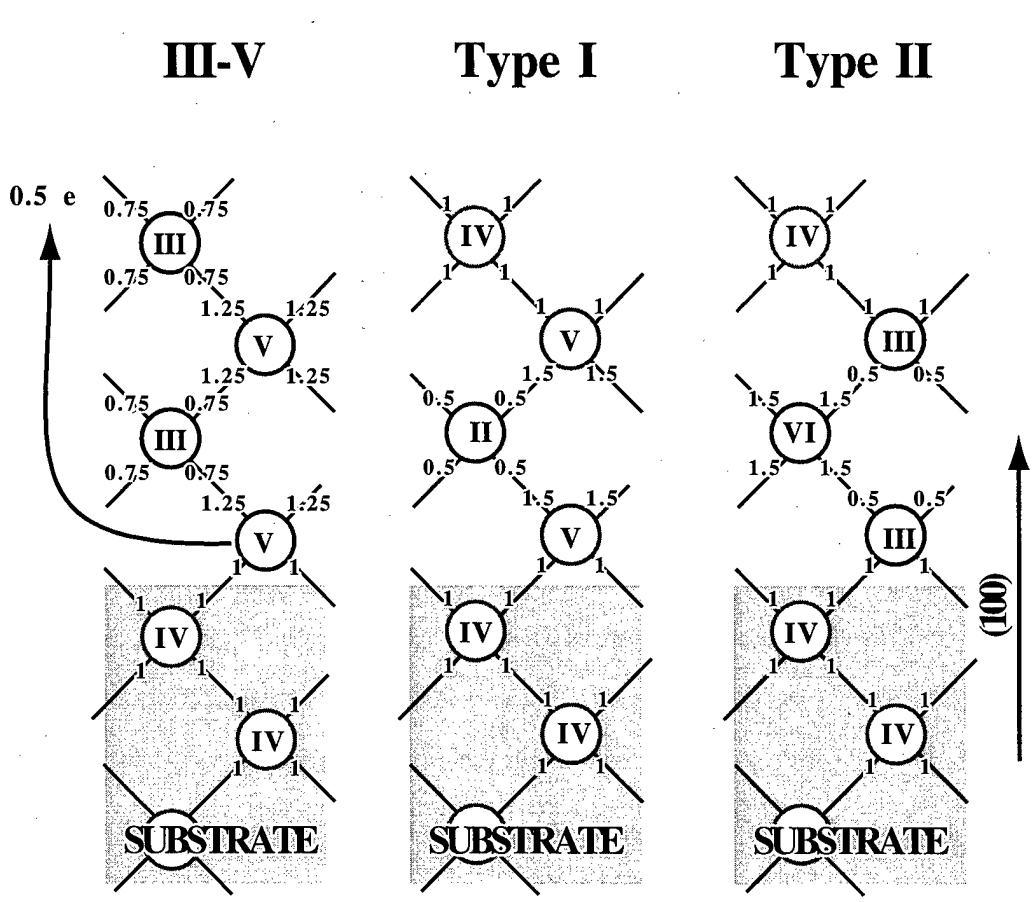


FIG. 1. Schematic illustration of the bond saturation for the interface of a Si(100) substrate and a III–V, a type I, and a type II material. The number of electrons contributed by each bond atom is indicated along each bond.

$$\text{III–V}\begin{cases}\nearrow & (\text{II–IV})_{1/2\text{–V}} \quad \text{Type I}\\ \searrow & \text{III–(IV–VI)}_{1/2} \quad \text{Type II}\end{cases}$$

One possibility which is denoted as type I materials is constructed by replacing two group-III atoms with a group-II and a group-IV atom. The other possibility which is denoted as type II materials is constructed replacing two group-V atoms with group-IV and a group-VI atom.

As shown in Fig. 1 both types of the new material are constructed by having complete layers of the different elements perpendicular to the (100) direction. Both types contain a complete layer of group-IV atoms. The electron compensation of the different layers is satisfied locally. Therefore, the polarity mismatch is resolved at the interface. At each group-IV layer in the growth sequence the growth can be continued by only depositing group-IV material. No excess or missing electrons are created at the interface.

Another way to look at the resolution of the polarity mismatch is to view the new material as a stacking of three layers of different elements acting effectively together as group-IV layers. As shown in Fig. 1 the two group-V layers and the one group-II layer are grouped between two group-IV layers. The three layers function as pseudo group-IV layers. They fulfill the electron counting rule locally and it is easy to understand that no polarity mismatch is created. The same argument is true for the type II materials.

The type I materials have the same chemical formula as naturally occurring Chalcopy-rite materials. However, the structure of both materials differ slightly, even though both materials are in a distorted zinc-blende structure. The alternating group-II and group-IV layers of the Chalcopyrites are intermixed. This would make a layer-by-layer heteroepitaxial deposition very difficult and would create the same polarity mismatch as for the III–V materials.

For both types of the new material, there are many possible configurations because there are many different elements to choose from for the different layers. For example, for the group-II layer Zn can be used, for the group-IV layer Si and for the group-V layer P. This combination of elements would make $(ZnSi)_{1/2}P$. However, any other elements of the different groups can also be used. One can even combine different type I materials, or type I and type II materials, to form new material alloys. From this tremendous amount of different configurations, it is very probable that one exists which has the desired properties.

Given that the polarity-mismatch problem is now resolved in principle, we proceed to identify the configurations that have a matching lattice constant to Si. Since the amount of possible configurations is enormous we begin with the simplest cases. As a starting point we use the tetrahedral-covalent radii of elements from Shay[9] to approximate the

TABLE I. Estimated lattice mismatch to Si for type I and type II materials.

Type I	$\Delta a/a_{Si}$ (%)	Type II	$\Delta a/a_{Si}$ (%)
⋮		⋮	
$(BeSi)_{1/2}As$	−3.2	$Al(CSe)_{1/2}$	−5.7
$(BeSn)_{1/2}P$	−3.1	$B(SnTe)_{1/2}$	−3.5
$(BeGe)_{1/2}As$	−2.2	$Ga(CTe)_{1/2}$	−3.1
$(ZnSi)_{1/2}P$	−1.9	$Al(CTe)_{1/2}$	−2.8
$(ZnGe)_{1/2}P$	−0.8	$In(SiO)_{1/2}$	−1.6
$(MgSi)_{1/2}P$	−0.3	$In(GeO)_{1/2}$	−0.7
$(ZnC)_{1/2}Sb$	0.5	$In(CS)_{1/2}$	0.1
$(MgGe)_{1/2}P$	0.8	$Ga(SiS)_{1/2}$	1.5
$(BeSn)_{1/2}As$	1.1	$Al(SiS)_{1/2}$	1.7
$(CdSi)_{1/2}P$	1.7	$In(CSe)_{1/2}$	1.8
$(MgC)_{1/2}Sb$	1.8	$In(SnO)_{1/2}$	2.0
$(ZnSi)_{1/2}As$	2.2	$Ga(GeS)_{1/2}$	2.6
$(ZnSn)_{1/2}P$	2.9	$Al(GeS)_{1/2}$	2.8
$(CdGe)_{1/2}P$	2.9	$Ga(SiSe)_{1/2}$	3.6
$(ZnGe)_{1/2}As$	3.3	$Al(SiSe)_{1/2}$	3.8
⋮		⋮	

lattice constants of various configurations of the type I and type II materials. These estimates are not too accurate (having errors up to 4%) but are still very valuable in narrowing the which are closest to the Si lattice constant are shown in Table I. The lattice mismatch varies from −3.2 to 3.3% for the type I and from −5.7 to 3.8% for the type II materials. To gain more accurate structural properties, we employ *ab initio* total energy calculations to the configurations which have the smallest lattice mismatch. Before doing this we will describe our computational method in the next section.

III. COMPUTATIONAL DETAILS

To determine the structural properties we carry out total energy calculation using density functional theory.[10–12] We apply the local-density approximation to the exchange-correlation functional, choosing the parameterization by Perdew and Zunger[13] of Ceperley and Alder's[14] data for the correlation energy of the homogeneous electron gas.

We use *ab initio* norm-conserving pseudopotentials. The pseudopotentials for the first row elements are created with scheme of Rappe *et al.*[15] or of Troullier and Martins.[16,17] The other pseudopotentials are generated using Hamann's scheme.[17,18] For the group-II elements, we employ nonlinear core corrections to improve the transferability of these potentials. These corrections take the nonlinear exchange of core and valence electrons into account.[19] The semilocal pseudopotentials are further transformed into fully separable Kleinman–Bylander pseudopotentials,[20] with the *d* potential chosen as the local potential.

The wave functions are expanded into plane waves[21] with a kinetic energy up to 20 Ry. For structures which include first row elements, we used 40 Ry. The electron density is calculated from special **k**-point sets,[22] their density in reciprocal space being equivalent to 384 **k** points in the whole Brillouin zone of a 8-atom cell.

TABLE II. Lattice constant mismatch and band gaps from *ab initio* LDA and GW calculations. The abbreviations *d, i, s, mi* refer to direct, indirect, semimetallic, and marginally indirect (<0.1 eV) band gaps, respectively.

Structure	$\Delta a/a_{Si}$ (%)	E_g^{LDA} (eV)	E_g^{GW} (eV)	Directness
$(CdSi)_{1/2}P$	0.9	0.41	⋯	*mi*
$(ZnGe)_{1/2}P$	−1.6	0.33	⋯	*i*
$(ZnSi)_{1/2}As$	0.7	0.32	0.68	*mi*
$(ZnSi)_{1/2}P$	−3.1	0.76	1.56	*d*
$Al(SiSe)_{1/2}$	6.6	0.40	⋯	*i*
$Ga(SiS)_{1/2}$	4.0	0.40	⋯	*mi*
$Al(SiS)_{1/2}$	4.9	0.20	⋯	*i*
$In(CSe)_{1/2}$	3.4	⋯	⋯	*s*
$In(CS)_{1/2}$	2.3	0.41	⋯	*d*
$Ga(CTe)_{1/2}$	−4.9	⋯	⋯	*s*

The equilibrium lattice parameters for different configurations of the novel class are found by minimizing the total energy. For each set of lattice parameters, the ions are relaxed until the forces are smaller than 50 meV/Å. For the equilibrium structures, we compute the band structures to investigate if they are optically active. However, it is not only of interest to have structures with a direct band gap but also to possess the correct band gap size. Since local density approximation (LDA) methods give inaccurate band gap results we use a quasiparticle GW scheme to obtain accurate band gap information for the most promising configurations.

We employ the GW approximation[23,24] for the electron self-energy to calculate the quasiparticle band structure. This method which is fully based on first principles has been shown to yield band structures in excellent agreement with experiment for a large class of materials including semiconductors.[25] Gap energies are usually given within an uncertainty in the order of 0.1 eV. Using the obtained LDA wave functions and energies, we determine the quasiparticle band structure using a procedure that has become state of the art in band-structure calculations.[25] We construct the electron self-energy operator within the GW approximation. The difference between the GW self-energy operator and the LDA exchange-correlation potential constitutes the quasiparticle corrections to the LDA band structure, thus leading to the quasiparticle band structure.

IV. RESULTS

The results for lattice constant mismatch and band gaps from the *ab initio* LDA and GW calculations are shown in Table II. The two configurations which have lattice constants within 1% of that of Si are $(ZnSi)_{1/2}As$ and $(CdSi)_{1/2}P$. However, all these configurations have an indirect or marginally indirect gap. Indeed, all configurations except for $(ZnSi)_{1/2}P$ listed in Table II have an indirect or marginally indirect gap. However, $(ZnSi)_{1/2}P$ has a −3.1% too small lattice constant. Therefore, no pure type I or type II configuration has the required properties.

The natural next step is to alloy two configurations from Table II. The best candidates for this are $(ZnSi)_{1/2}P$ and $(ZnSi)_{1/2}As$. $(ZnSi)_{1/2}As$ has a marginally indirect gap smaller than 0.76 eV which is the desired band gap value.

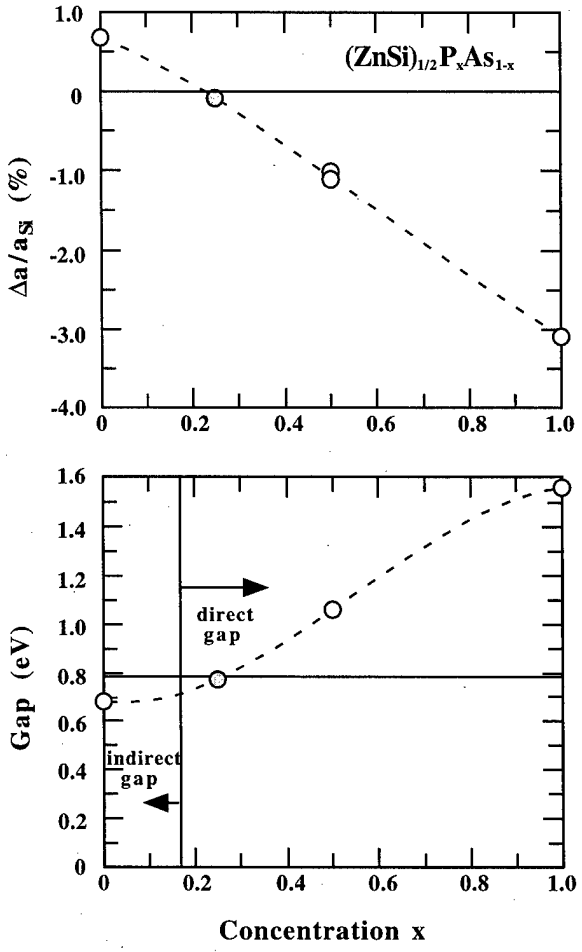


FIG. 2. Lattice constant mismatch (top panel) and quasiparticle band gap (bottom panel) for $(ZnSi)_{1/2}P_xAs_{1-x}$ as a function of the P concentration x.

This desired value corresponds to the canonical optoelectronics wavelength of 1.55 μm. Also the lattice constant of $(ZnSi)_{1/2}As$ is 0.7% larger than that of Si. Therefore, mixing $(ZnSi)_{1/2}As$ with $(ZnSi)_{1/2}P$ which has a too small lattice constant and a larger direct gap than 0.76 eV could lead to material which has desired properties. Both lattice constant and band gap should be closer to the preferred values.

We denote this combination of the two type I materials as $(ZnSi)_{1/2}P_xAs_{1-x}$, where x indicates the relative percentage of phosphorus used. This material can easily be achieved because the epitaxial layering scheme allows to use different elements for the different group-V layers. By choosing different numbers of layers for the two group-V elements we perform computations for different values of x. Both the lattice constant mismatches and band gap are calculated, whereby the band gap is computed using the quasiparticle GW scheme. The results are shown in Fig. 2. Clearly, for $x = 0.25$ which corresponds to the material $(ZnSi)_{1/2}P_{1/4}As_{3/4}$, the lattice mismatch is very small and only has a value of -0.08%. This material also has a direct band gap of 0.78 eV which is very close to the desired 0.76 eV for optoelectronics. The precise value of this gap is not as important as the possibility of tuning it by the appropriate phosphorus concentration.

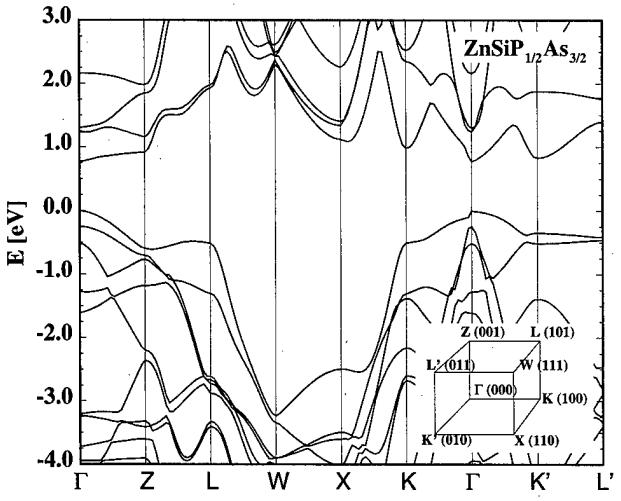


FIG. 3. Quasiparticle band structure for $(ZnSi)_{1/2}P_{1/4}As_{3/4}$.

The quasiparticle band structure is shown in Fig. 3. It is not feasible to do a complete GW band structure calculation for these configurations. The quasiparticle energies can only be calculated for certain symmetry points. Therefore, we obtain the band structure in Fig. 3 by using the LDA band structure and distorting it linearly so that it matches the GW band energies at the symmetry points. The band structure shows that the material is marginally direct. The K' point competes with the Γ point, but is still 0.06 eV higher in energy.

The growth of this new material on Si(100) can be achieved by the layer-by-layer deposition which is shown schematically in Fig. 4. After growing the last Si substrate layer the structure is begun by growing a layer of As on the Si(100) surface. Next, a layer of Zn has to be grown and so forth. The material with deposition layering as illustrated in Fig. 4 will have a slight dipole moment, due to the difference in chemical properties of phosphorus and arsenic. However,

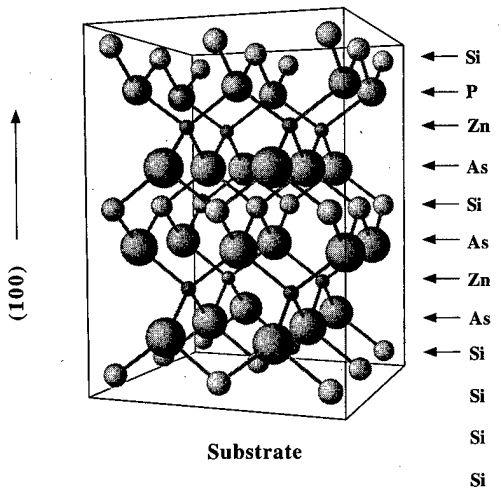


FIG. 4. Atomic model of the crystal structure of $(ZnSi)_{1/2}P_{1/4}As_{3/4}$ and its layering for growth.

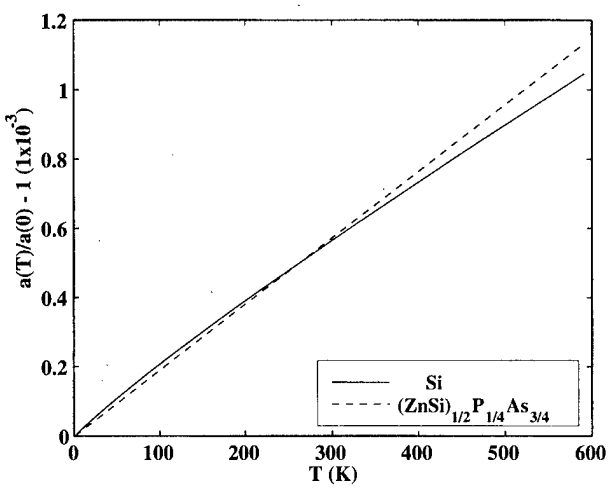


FIG. 5. Relative lattice expansion $a(T)/a(0)-1$ for Si and $(ZnSi)_{1/2}P_{1/4}As_{3/4}$ as a function of the temperature T.

this dipole moment can be eliminated by inverting every other cell in growth direction. The supercell would be then twice as long in the growth direction.

It is also important to examine if the lattice mismatch of $(ZnSi)_{1/2}P_{1/4}As_{3/4}$ and Si continues to be small at nonzero temperatures. For this, we estimate the thermal expansion coefficients from the total energy surfaces for both materials using the anharmonicities of the elastic energies. The expectation value of the lattice constant at a temperature T is given by

$$\langle a\rangle_T=\frac{1}{Z}\int a\exp^{-E_{\text{tot}}(a)/kT}da, \tag{1}$$

where Z is the partition function. From this expectation value we calculate the relative lattice expansion for both materials which are shown in Fig. 5. The difference of the relative lattice expansion of both materials is less than 0.01% from 0 to 600 K, and actually vanishes at 350 K. This suggests that thermal expansion will not cause additional lattice mismatch.

To accurately predict if this new material can be grown under actual experimental conditions without defects is, of course, impossible until tried. However, we can discuss if the new material is thermodynamically stable once it is fabricated. To ensure that the new material does not segregate into equilibrium phases of the elements, we calculate the formation enthalpies for the different configurations. Indeed, all the formation enthalpies are positive, indicating that the materials are stable against segregation. In the case of $(ZnSi)_{1/2}As$, for example, it is 0.33 eV per atom. This is only half the value of formation enthalpy of the corresponding Chalcopyrite. The Chalcopyrite is therefore the equilibrium structure and our new material is only meta stable. To convert our structure to the Chalcopyrite structure requires that two second-nearest-neighbor atoms have to be exchanged. We expect the diffusion barrier for this process to be very high because interstitials have to be created. Therefore, this diffusion process is rather unlikely to occur and we do not expect our structure to dissolve into Chalcopyrite.

V. SUMMARY AND CONCLUSION

We have introduced a novel class of compound semiconductors that should be intriguing candidates for creating monolithic integration with Si using epitaxy. The different possible combinations of elements provide a wide range of lattice constants and band gaps. Different configurations of this novel class are investigated using *ab initio* total energy calculations and quasiparticle GW computations. The subclass $ZnSiP_xAs_{1-x}$ spans over the Si lattice constant and possesses a direct band gap that covers the optical fiber wavelength. The particular configuration $(ZnSi)_{1/2}P_{1/4}As_{3/4}$ has the smallest lattice mismatch and a band gap closest to the desired size. From the estimated thermal expansion coefficients, we conclude that the lattice mismatch should continue to be small for nonzero temperatures.

ACKNOWLEDGMENTS

The authors would like to thank Professor S. Louie for his generosity in providing us with GW computer codes and B. Králik for helping to run these codes. This work was supported in part by the Office of Naval Research Contract No. N0001-94-1-0591.

[1]L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[2]L. T. Canham, T. I. Cox, A. Loni, and A. J. Simons, Appl. Surf. Sci. **102**, 436 (1996).
[3]*Heteroepitaxy in Silicon II, MRS Symposia Proceedings No. 91*, edited by J. Fan, J. Phillips, and B.-Y. Tsaur (Materials Research Society, Pittsburgh, PA, 1987).
[4]R. F. Davis, Proc. IEEE **79**, 702 (1991).
[5]J. H. Edgar, J. Mater. Res. **7**, 235 (1992).
[6]J. C. R. Eddy, T. D. Moustakas, and J. Scanlon, J. Appl. Phys. **73**, 448 (1993).
[7]E. Desurvire, *Erbium-doped Fiber Amplifiers: Principles and Applications* (Wiley, New York, 1994).
[8]R. M. Sieg, S. A. Ringel, S. M. Ting, S. B. Samavedam, M. Currie, T. Langdo, and E. A. Fitzgerald, J. Vac. Sci. Technol. B **16**, 1471 (1998).
[9]L. J. Shay, *Ternary Chalcopyrite Semiconductors: Growth, Electronic Properties, and Applications* (Pergamon, Oxford, 1975).
[10]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[11]W. Kohn and L. J. Sham, Phys. Rev. A **140**, A1133 (1965).
[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. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[14]D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[15]A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990).
[16]N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[17]M. Fuchs and M. Scheffler, http://www.fhi-berlin.mpg.de/th/fhi96md/code.html (unpublished).
[18]D. R. Hamann, Phys. Rev. B **40**, 2980 (1989).
[19]S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).
[20]L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[21]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).
[22]H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[23]L. Hedin, Phys. Rev. A **139**, A796 (1965).
[24]L. Hedin and S. Lundqvist, in *Solid State Physics, Advances in Research and Application*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1969), Vol. 23, p. 1.
[25]M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).

Pyroelectronics: Novel device concepts based on nitride interfaces

G. Zandler, J. A. Majewski, and P. Vogl[a)]
Physik Department and Walter Schottky Institut, Technische Universität München, Am Coulombwall, D-85748 Garching, Germany

(Received 20 February 1999; accepted 20 April 1999)

A key property of nitrides is their large spontaneous and piezoelectric polarization fields that allow significant tailoring of the carrier dynamics and optical properties of nitride devices. In this article, based on first-principles calculations of structural and electronic properties of bulk nitrides and their heterostructures, we investigate the potential of this novel material class for modern device applications by performing self-consistent Monte Carlo simulations. We demonstrate that the internal electric fields have a significant and favorable influence on the transistor characteristics.
© *1999 American Vacuum Society.* [S0734-211X(99)03004-8]

I. INTRODUCTION

The strong internal electric fields in nitrides, resulting from the charge induced at the nitride interfaces by the divergence of their strong pyro- and piezoelectric polarizations,[1] have a dramatic effect on the electronic and optical properties in these systems, but only relatively few systematic theoretical predictions of this have been published so far.[2]

II. THEORY OF INTERNAL POLARIZATION FIELDS IN GaN AND AlN INTERFACES

The present systematic theoretical studies of structural and electronic properties of heteroepitaxial AlN/GaN interfaces and homoepitaxial stacking faults[3] are based on a well-established first-principles total-energy pseudopotential method within local-density-functional formalism.[4] The band offsets, the charge accumulation at the polar interfaces of the junctions, and the interface electronic states have been investigated, taking fully into account the effects of lattice relaxation and electric polarization.

A. Polarity of interfaces

In the following, we consider strained GaN layers that are grown pseudomorphically on a cubic or wurtzite AlN substrate. For cubic heterostructures, we consider [001], [110], and[111], and for wurtzite heterostructures [0001], $[0\bar{1}10]$, and $[\bar{1}2\bar{1}0]$ growth directions, respectively. In the case of stacking fault interfaces, we assume the growth axis to be the hexagonal c axis ([0001] in wurtzite, [111] in zinc blende).

The space group $P6_3mc(C_{6v}^4)$ of wurtzite is compatible with spontaneous polarization along the hexagonal c axis. Therefore, the polarization can be of both pyroelectric and piezoelectric origin in wurtzite AlN and GaN, but only piezoelectric in cubic phases. Whenever the polarization lies parallel to the growth direction, its change (divergence) across the interface is equivalent to an interface charge. Among the interfaces between GaN and AlN that we have studied, the [111] GaN/AlN, the [0001] GaN/AlN, and the stacking fault interface between the wurtzite and zinc blende GaN interface are of this type and may be termed "polar." In cases such as $[\bar{1}2\bar{1}0]$ GaN/AlN, $[0\bar{1}10]$ GaN/AlN, or [110] GaN/AlN, on the other hand, the polarization lies parallel to the interface and therefore does not give rise to charge accumulation. We term those interfaces "nonpolar." A limiting case of this type is the cubic [001] GaN/AlN interface.

Computationally, we have modeled all of these interfaces by supercells containing up to 40 atoms (for $[0\bar{1}10]$ and $[\bar{1}2\bar{1}0]$ wurtzite structures). All atomic positions in the unit cell have been optimized by minimizing the total energy via Hellmann–Feynman forces. The length of the supercell was determined by minimizing the stress tensor component along the growth direction. Based on these first-principles calculations, we find that the charges induced at the interfaces and valence band offsets VBOs are insensitive to the lattice constant along the growth direction, but highly sensitive to the atomic relaxation at the interface (e.g., the VBO for the relaxed [0001] GaN/AlN interface is 0.4 eV smaller than that for the unrelaxed one). The lattice relaxation near the interface reduces the interface charge by approximately a factor of 2.

B. Valence band offsets and interface charges

The valence band offset ΔE_V at an interface between two semiconductors (or between two phases of one material) can be conveniently split into two terms[5] $\Delta E_V = \Delta \bar{V} + \Delta E_{\mathrm{BS}}$, where $\Delta \bar{V}$ is the asymptotic difference between the laterally and vertically averaged electrostatic potential $\bar{V}(z)$, and ΔE_{BS} is the bulk band structure contribution. For neutral interfaces, $\Delta \bar{V}$ equals the dipole moment of the electrostatic charge density $\bar{\rho}(z)$ across the interface. In the case of charged interfaces, the electrostatic charge density $\bar{\rho}(z)$ contains an additional monopole term that is proportional to the difference between the polarizations in the two adjacent materials. The monopole charge density causes the macroscopically averaged electrostatic potential to change linearly with distance from the interface. This contribution can be filtered out[6] and the intrinsic $\Delta \bar{V}$ can be obtained based on the remaining dipolar contribution to $\bar{\rho}(z)$. By integrating the

a)Electronic mail: vogl@wsi.tum.de

TABLE I. Valence and conduction band offsets, ΔE_V (eV) and ΔE_C (eV), potential lineups $\Delta \bar{V}$ (eV), and band structure contributions ΔE_{BS} (eV) for GaN/AlN heterostructures and stacking fault interfaces. For the GaN/AlN heterostructures, all values are given relative to the band edges in AlN, whereas for stacking faults they are given relative to the band edges of the wurtzite phase. The conversion factor for σ (C/m^2) to cm^{-2} is 6.241×10^{14}.

	Wurtzite GaN/AlN			Zinc blende GaN/AlN			Stacking faults	
	[0001]	$[0\bar{1}10]$	$[\bar{1}2\bar{1}0]$	[001]	[110]	[111]	AlN	GaN
ΔE_{BS}	−0.26	−0.15	−0.17	−0.15	−0.15	0.13	−0.25	−0.046
$\Delta \bar{V}$	0.93	0.88	0.97	0.89	1.07	0.77	0.27	0.081
ΔE_V	0.67	0.73	0.80	0.75	0.92	0.90	0.02	0.04
ΔE_C	−1.92	−1.72	−1.67	−0.62	−0.51	−0.53	−1.30	−0.17
σ	0.010	0	0	0	0	0.006	0.009	0.003

monopole charge density across the interface, we have been able to calculate the charge induced at the interface σ. Finally, the conduction band offset ΔE_C is defined by $\Delta E_C = \Delta E_V + \Delta E_{gap}$, where ΔE_{gap} is the difference between the fundamental band gaps of two constituent bulk crystals.

The calculated band offsets and charges induced at polar interfaces are given in Table I, together with the individual contributions to the VBOs. These predicted VBOs agree very well with recent experimental data,[7,8] but not with some core-level photoemission data that suggest an extremely high VBO of 1.36 eV.[9]

It is interesting to note that all of the calculated VBOs are of the order of 0.7–0.9 eV and are insensitive to the polar/nonpolar character of the interface. Thus, the macroscopic electric fields do not grossly alter or modify the intrinsic band offsets. The large difference between conduction band offsets in cubic and wurtzite GaN/AlN heterostructures is a consequence of the indirect energy gap ($\Gamma \rightarrow X$) in cubic AlN that is 1.4 eV smaller than the direct gap in wurtzite AlN.

All the GaN/AlN interfaces and stacking faults studied are of type I, with the valence band lying higher and the conduction band being lower in GaN and the zinc blende phase of the stacking fault, respectively. The large charges accumulated at the stacking fault interfaces result mainly from the spontaneous dielectric polarization of the hexagonal phase, whereas the strain-induced piezoelectric contribution is negligible. The induced charges are seen to be reduced by a factor of 2 by relaxation of the atomic positions at the interface. The charges calculated compare well (the difference amounts to 15%–20%) with results estimated previously from bulk calculations.[1] The stacking faults can generate persistent photoconductivity and they can be tuned to act either as excitonic traps or luminescence centers, depending on the width of the cubic layer.

Our calculations reveal the existence of large internal electric fields. Their influence on the properties of the electronic devices will be presented.

III. MONTE CARLO CALCULATIONS

A. Bulk transport properties of AlGaN

In our Monte Carlo transport calculations, we have used band parameters[10] deduced from the electronic structure calculations described above and scattering rates for ionized impurity, intra- and intervalley phonon scattering, acoustic, piezoelectric, polar optical phonon and alloy scattering.

The resulting electron drift velocities (Fig. 1) in bulk GaN at both low and high fields are in good agreement with recently published theoretical results. At high fields, GaN as well as AlN shows approximately twice the saturation velocity of GaAs. For the $Al_xGa_{1-x}N$ material system, the heavier effective mass leads to considerably lower drift velocities at low fields, and to lower peak velocities with increasing Al contents. As a consequence of the comparable and high phonon frequencies, the saturation velocities, on the other hand, are predicted to be practically independent of the Al content, $\sim 1.5 \times 10^7$ cm/s at, namely, 1 MV/cm. These results indicate that one can, in principle, achieve much higher transit time frequencies in submicron nitride devices than in GaAs devices.

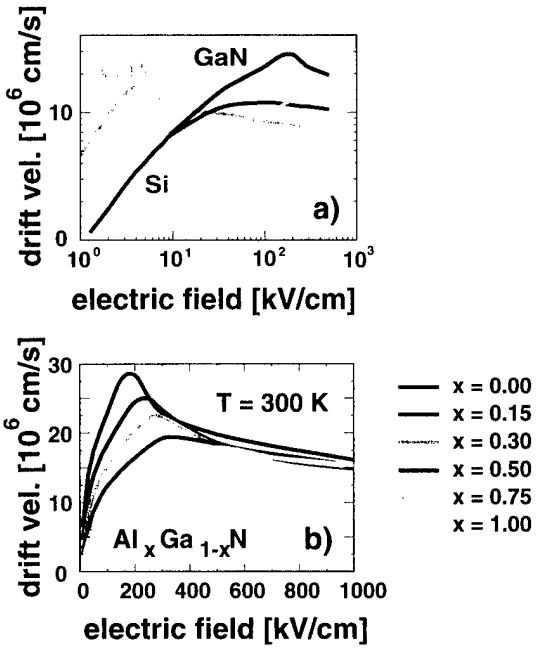


FIG. 1. (a) Calculated electron drift velocities of GaN and AlN in comparison with those of GaAs and Si. (b) Electron drift velocities in $Al_xGa_{1-x}N$ for various Al contents.

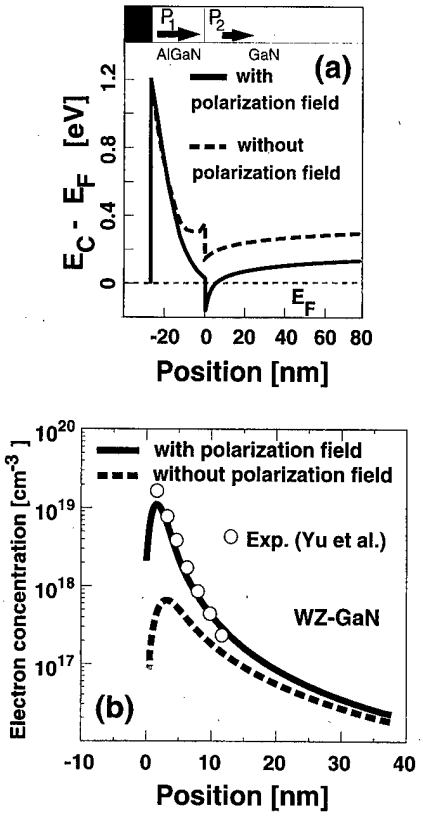


FIG. 2. Calculated results for a $Al_{0.15}Ga_{0.85}N$/GaN HFET with (solid lines) and without (dashed lines) spontaneous and piezoelectric fields. A homogeneous n-doping background of $10^{16}\,cm^{-3}$ throughout the structure has been assumed. (a) Self-consistently computed conduction band edge profile along the growth direction. (b) Calculated electron density profile in comparison with published experimental values.

B. Pyroelectric and piezoelectric fields in wurtzite AlGaN/GaN single heterostructures

As discussed in detail in Sec. II, wurtzite structure III nitrides possess pyroelectric polarization along the [0001] direction. Although the corresponding electric fields are of the order of some MV/cm,[1] such fields are normally neutralized by unavoidable charged states at the surface of bulk samples. However, in heterostructures, each interface between different materials carries a charge that equals the divergence in the spontaneous and/or strain-induced piezoelectric polarization[11] which cannot be compensated by charged states at the surface and substrate interface.

To study the effect of this polarization induced interface charge on carrier transport and heterostructure device characteristics, we focus on a single $Al_{15}Ga_{85}N$/GaN heterostructure (see Fig. 2) first. In a tensile strained 30 nm thick $Al_{15}Ga_{85}N$ layer on a GaN substrate, there is both pyro- and piezoelectrical polarization, while a relaxed 300 nm thick GaN layer contains only a pyroelectric moment. Assuming charge neutrality for the whole structure, there is a polarization induced charge density σ at the $Al_{15}Ga_{85}N$/GaN interface and a compensating charge of $-\sigma/2$ at the surface and at the nucleation layer. We take σ to be $6\times10^{12}\,cm^{-2}$ which interpolates between a recent experimental value[11] ($4\times10^{12}\,cm^{-2}$) and the theoretical value[1] ($8\times10^{12}\,cm^{-2}$).

In Fig. 2, we depict the self-consistent band edge and density profiles along the growth direction obtained from a coupled one-dimensional Schrödinger and a Poisson equation. The Poisson equation includes all interface charges that result from the polarization fields. The bulk conduction band is modeled by a nonparabolic isotropic band centered at wave vector $\mathbf{k}=0$ in the wurtzite structure of GaN.[10,12,13] We have included up to 40 electronic subband energies and wave functions. As is shown in Fig. 2, the polarization induced positive interface charge causes stronger confinement of a two-dimensional (2D) channel electron and an increase in the channel density by more than an order of magnitude. This is consistent with the experimental values for this layer structure shown in Fig. 2. Furthermore, we have performed ensemble Monte Carlo calculations for drift mobilities along the channel that forms parallel to the interface. The inter- and intrasubband-scattering rates[14] have been determined consistently with the confined electronic channel states. We find excellent agreement between the experimental and calculated temperature dependent mobilities,[14] consistent with the high channel densities reported for these devices.[11,15] We would like to stress that this type of agreement strongly supports the existence of strong pyroelectric polarization fields. For channel densities exceeding $10^{13}\,cm^{-2}$, we predict a phonon limited drift mobility for 2D electrons close to 2000 $cm^2/V\,s$ for high quality interfaces, in good agreement with recent data on such structures.[16]

C. Influence of polarization fields on device performance

To study the influence of polarization fields on device characteristics, we have performed self-consistent two-dimensional Monte Carlo simulations of various nitride based heterostructure devices.[10] At high bias voltages, it suffices to use a bulk description for the hot carrier transport in the channel layers.

In accordance with the high channel densities, the drain currents in these devices can become extraordinarily high. We have studied a 300 nm gate length single heterostructure device with the same layer sequence as that described above, but with additional doping densities of $10^{18}\,cm^{-2}$ in the supply layer. This results in channel densities of $8\times10^{12}\,cm^{-2}$ in the ungated structure. For this device, we find drain currents at open channel conditions up to 2000 A/m and predict a maximum transconductance of 430 S/m. Although the polarization induced interface charge in these devices suppresses parasitic substrate currents very efficiently, the magnitude of the threshold voltage is still fairly high (−6 V), mainly as a consequence of the high gate to channel separation of 30 nm.

Based on these bulk results, we have carried out self-consistent Monte Carlo simulations for planar short channel heterostructure field effect transistors (HFETs).[17] We predict the following layer sequence to be well suited for high speed power applications. Beneath the contacts, we assume a highly doped $Al_{0.16}Ga_{0.84}N$ supply layer of 10 nm (4 nm with

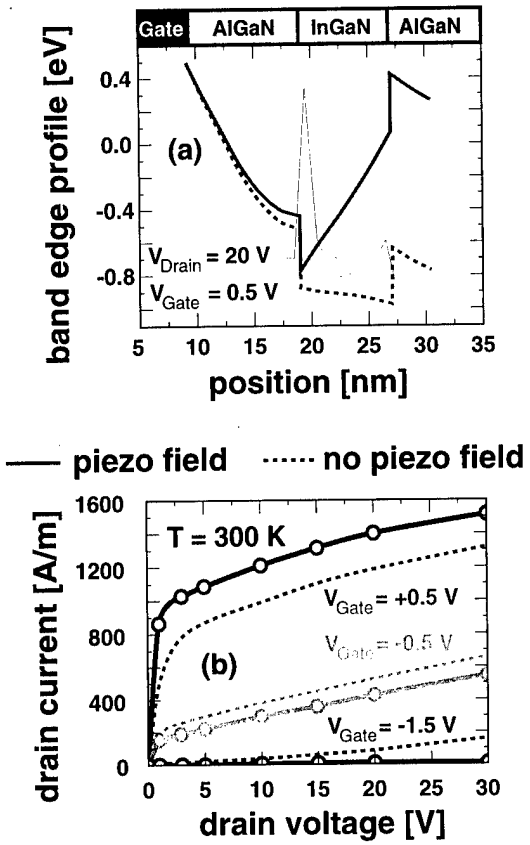


FIG. 3. Simulation results for a 300 nm gate length $In_{0.05}Ga_{0.95}N/Al_{0.16}Ga_{0.84}N$ HFET with (solid lines) and without (dashed lines) spontaneous and piezoelectric fields. (a) Self-consistent conduction band edge profile and electron density. (b) Calculated drain current vs drain voltage for different gate voltages.

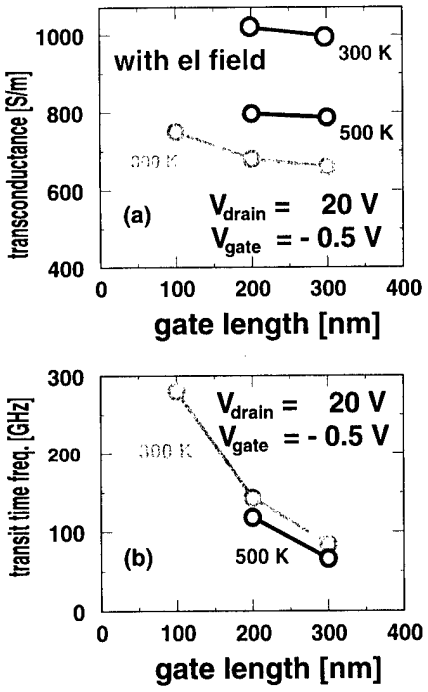


FIG. 4. (a) Calculated transconductance for different temperatures vs gate length of a $In_{0.05}Ga_{0.95}N/Al_{0.16}Ga_{0.84}N$ HFET with and without interface strain-induced piezoelectric fields. (b) Predicted transit time frequency as a function of gate length and temperature at a drain voltage of 20 V and open channel conditions.

$n=1\times10^{18}\,cm^{-3}$ and 6 nm with $n=1\times10^{19}\,cm^{-3}$). This is followed by an 8 nm $In_{0.05}Ga_{0.95}N$ channel with $n=3\times10^{18}\,cm^{-3}$ and a subsequent 400 nm buffer layer that consists of $Al_{0.16}Ga_{0.84}N$. We assume this buffer layer to be fully compensated/nearly intrinsic, except for the first 6 nm that we take to be n type with $n=2\times10^{18}\,cm^{-3}$. The intrinsic buffer is a prerequisite for obtaining acceptable off characteristics.

The lattice mismatch between the channel and the barrier material amounts to approximately 1%. This strain causes macroscopic polarization of the channel in the growth direction. The resulting electric field can be calculated from the piezoelectric tensor and the elastic constants. It amounts to 1.7×10^{8} V/m for the HFET layer sequence given above and is predicted to point towards the substrate. There is some experimental evidence for such polarization effects.[18]

The influence of this huge polarization field on the drain current characteristics of a 300 nm gate HFET is shown in Fig. 3. The piezoelectric field pushes the electrons closer to the gate contact and therefore increases the transconductance from 650 to 980 S/m (see Fig. 4). Furthermore, this confining effect improves the turn-off behavior of the device considerably.

Our calculations predict average channel velocities exceeding 2×10^{5} m/s for gate lengths less than 200 nm. The resulting transit time frequencies are depicted in Fig. 4(b) as a function of the gate length. The Monte Carlo calculations yield transit time frequencies of 120 GHz for gate lengths between 200 and 300 nm and up to 300 GHz for a 100 nm gate length. The velocity overshoot is found to remain small in all of these cases but the turn-off behavior is poor for very short gate lengths. These short channel effects become even more pronounced at elevated temperatures. In conlusion, we predict excellent high frequency and power performance for wide-gap nitride based HFETs with gate lengths larger than 200 nm.

Polarization induced internal fields can be effectively utilized to tailor devices with optimal potential barriers towards the substrate. Indeed, we have theoretically designed the following inverted HEMT structure that should yield high channel densities and excellent turn-off behavior. On top of a relaxed GaN buffer, there is a 12 nm thick strained $Al_{40}Ga_{80}N$ layer. The inner part (8 nm) of this layer is highly n doped with $1.9\times10^{19}\,cm^{-3}$. The remaining part of the barrier is lowly doped with $10^{17}\,cm^{-3}$ carriers. On top of this layer, there is a 30 nm GaN (n doped with $4\times10^{16}\,cm^{-3}$). The gate is assumed to be recessed in order to achieve a small gate to channel separation of 14 nm. In this HFET, we assume the opposite polarity of the structure to that in the previous example. The GaN is assumed to be field free, whereas the two GaN/AlGaN interfaces have opposite interface charges. We have estimated them from bulk calculations[1] and set them to $\pm2\times10^{13}\,cm^{-2}$. The calculated band edge profile, together with the confined electron density, is shown in Fig. 5(a) at the midgate position.

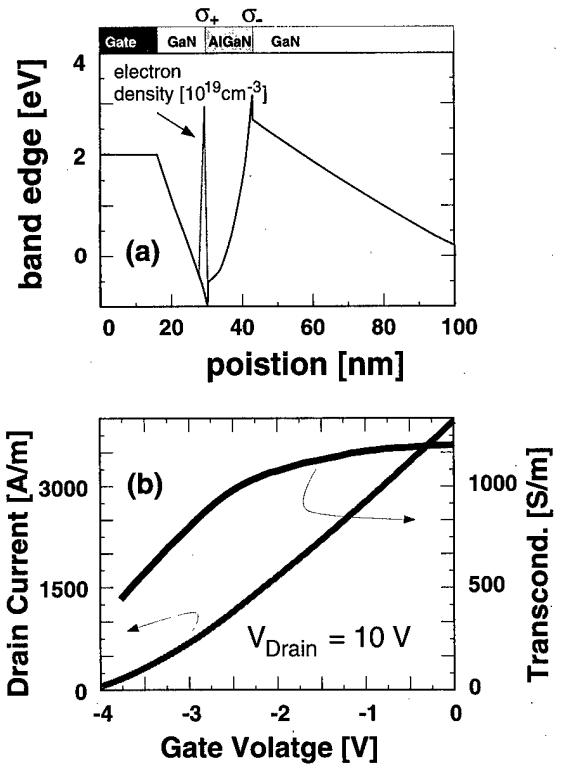


FIG. 5. Monte Carlo results for a 300 nm gate length inverted $GaN/Al_{40}Ga_{60}N$ HFET. (a) Self-consistent band edge profile for −1 V gate voltage and 10 V drain voltage. The channel electron density is indicated schematically. (b) Drain current and transconductance vs gate voltage at a 10 V drain voltage.

As can be seen from Fig. 5, the polarization induced field produces a high potential barrier of approximately 3 eV towards the substrate. This has two major consequences. First, the highly doped AlGaN barrier becomes completely depleted and all the electrons are strongly confined in the channel that forms at the interface with the top GaN layer. Second, the high barrier completely suppresses leakage current across the AlGaN barriers up to high voltages. We predict excellent electrical characteristics for such types of devices. In the present example, we find drain currents exceeding 3000 A/m and a transconductance that lies well above 1000 S/m. We note that the thickness of the AlGaN barrier layer and its doping level must be chosen carefully in order to avoid the formation of parasitic hole channels.

In conclusion, both pyro- and piezoelectrical polarizations must be taken into account in order to explain experimentally observed mobilities and densities in nitride heterostructures. They are crucial for device performance and offer a unique possibility to design novel high power high frequency heterostructure transistors.

ACKNOWLEDGMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft and by the Bayerischer Forschungsverbund (FOROPTO).

[1]F. Bernardini, V. Fiorentini, and D. Vanderbilt, Phys. Rev. B **56**, R10024 (1997).
[2]J. A. Majewski, G. Zandler, and P. Vogl, Semicond. Sci. Technol. **13**, 90 (1998).
[3]X. H. Wu, L. M. Brown, D. Kapolnek, S. Keller, B. Keller, S. P. DenBaars, and J. S. Speck, J. Appl. Phys. **80**, 3228 (1996); Z. Liliental-Weber, C. Kisielowski, S. Ruvimov, Y. Chen, J. Washburn, I. Grzegory, M. Bockowski, J. Jun, and S. Porowski, J. Electron. Mater. **25**, 1545 (1996).
[4]M. Städele, J. A. Majewski, and P. Vogl, Phys. Rev. B **56**, 6911 (1997); J. A. Majewski, M. Städele, and P. Vogl, MRS Internet J. Nitride Semicond. Res. **1**, 30 (1996).
[5]A. Baldereschi, S. Baroni, and R. Resta, Phys. Rev. Lett. **61**, 734 (1988).
[6]F. Bernardini and V. Fiorentini, Phys. Rev. B **57**, R9427 (1998).
[7]G. Martin, A. Botchkarev, A. Rockett, and H. Morkoç, Appl. Phys. Lett. **68**, 2541 (1996).
[8]G. Martin, S. Strite, A. Botchkarev, A. Agarwal, A. Rockett, H. Morkoç, W. R. L. Lambrecht, and B. Segall, Appl. Phys. Lett. **65**, 610 (1994).
[9]J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **68**, 2879 (1996).
[10]G. Zandler, J. A. Majewski, M. Städele, P. Vogl, and F. Compagnone, Phys. Status Solidi B **204**, 133 (1997).
[11]E. T. Yu, G. J. Sullivan, P. M. Asbeck, C. D. Wang, D. Qiao, and S. S. Lau, Appl. Phys. Lett. **71**, 2794 (1997).
[12]N. E. Christenson and I. Gorczyca, Phys. Rev. B **50**, 4397 (1994).
[13]K. Kim, W. R. L. Lambrecht, and B. Segall, Phys. Rev. B **53**, 16310 (1996).
[14]R. Oberhuber, G. Zandler, and P. Vogl, Appl. Phys. Lett. **73**, 818 (1998).
[15]Y. F. Wu, B. P. Keller, D. Kapolnek, P. Kozodoy, S. P. Denbaars, and U. K. Mishra, Appl. Phys. Lett. **69**, 1438 (1996).
[16]R. Gaska, J. W. Yang, A. Osinsky, Q. Chen, M. A. Khan, A. O. Orlov, G. L. Snider, and M. S. Shur, Appl. Phys. Lett. **72**, 707 (1998).
[17]M. S. Shur and M. Asif Khan, MRS Bull. **22**, 44 (1997).
[18]G. Martin, A. Botchkarev, A. Rockett, and H. Morkoç, Appl. Phys. Lett. **68**, 2541 (1996).

Selective area chemical vapor deposition of titanium oxide films: Characterization of $Ti(OC_3H_7)_4$ as an electron beam resist

William J. Mitchell and Evelyn L. Hu[a)]
Center for Quantized Electronic Structures, University of California, Santa Barbara, California 93106

(Received 26 February 1999; accepted 20 April 1999)

We have shown that condensed multilayer films of titanium isopropoxide [$Ti(-OC_3H_7)_4$] on cold GaAs(001) substrates are easily converted to nonvolatile titanium oxide deposits by exposure to a 10 keV electron beam. Using spatially resolved Auger electron spectroscopy, we have measured the initial kinetics of this electron beam induced decomposition reaction and have found it to be zeroth order (in precursor concentration) with an extremely high zeroth-order reaction cross section of $(1.5\pm0.6)\times10^{-14}$ cm^2/electron. Consequently, exposures as low as 23 μC/cm^2 result in titanium oxide films with thicknesses on the order of 5 Å. Moreover, the remaining unexposed precursor simply desorbs upon annealing to room temperature, ensuring selective area oxide deposition. The etch resistance of these titanium oxide patterns were characterized by etching in Cl_2 at a substrate temperature of 250 °C. We found that exposures above 23 μC/cm^2 produce robust, etch resistant oxide films which result in efficient pattern transfer to the underlying GaAs(001) substrate by etching. Moreover, clear undercut- and V-etch profiles were observed in the orthogonal $\langle 011\rangle$ and $\langle 01\bar{1}\rangle$ directions, respectively, characteristic of a crystallographic etch mechanism on the GaAs(001) surface. © *1999 American Vacuum Society.* [S0734-211X(99)03104-2]

I. INTRODUCTION

The phenomenon of electron-induced decomposition of adsorbed molecules on surfaces has been known for over 25 years.[1] Recently, technological implications of this phenomenon have been realized.[2–8] Localized electron irradiation, when coupled with a computer-controlled electron beam writer, has been used to deposit thin films in a spatially selective manner under ultrahigh vacuum (UHV) conditions without the need to first generate masks (via the usual multistep process).

The majority of the previous studies on electron-induced selective area deposition have focused on single metal or semiconductor film deposition from various metal halide, hydride, organic, and carbonyl precursors.[9] By virtue of the substrate to be patterned being at room temperature or higher (i.e., above the precursor desorption temperature), these depositions were effected in a single step by irradiating the sample surface with the electron beam while simultaneously maintaining a (high) background pressure of the precursor to produce a small, but nonzero, steady state concentration of precursor on the substrate. Under these conditions, deposition rates are expected to be precursor limited.

In the work to be presented here, we show how selective area deposition can be utilized in an *in situ* electron beam lithography scheme, a primary requirement of which is the identification of a suitable resist material, i.e., one that is sensitive to electrons, is UHV compatible, and is robust in an etch/pattern transfer process. Previous work in this area has focused on a thin semiconductor oxide layer as a positive resist in which the irradiated areas of the oxide are rendered less resistant to etching in Cl_2.[10–17] However, issues with these oxide resists remain, such as poor electron sensitivity (exposures as high as 10^4-10^6 mC/cm^2 are needed) and poor pattern transfer characteristics (e.g., in one case,[14] 90 min in 10^{-4} Torr Cl_2 at 70 °C produced an etch depth of only 1000 Å). Instead, we utilize a volatile metal alkoxide precursor as a negative-type resist material and use it to deposit robust, etch resistant metal oxide films via electron-beam induced decomposition of the alkoxide. Specifically in this study, we deposit titanium oxide films on GaAs(001) using a three-step lithography process. The first "spin-on resist" step involves condensing multilayers of the titanium isopropoxide precursor on the GaAs surface held at temperatures below the precursor desorption temperature (approximately −20 °C). To pattern the resist, we expose selected areas of the precursor-covered surface to a scanning electron beam which converts the volatile precursor to a nonvolatile titanium oxide deposit. Finally, we perform "resist development" where the substrate is annealed to room temperature desorbing the unexposed precursor and defining the deposited oxide pattern. Pattern transfer to the underlying GaAs substrate is then achieved by etching in Cl_2 at elevated substrate temperatures (approximately 250 °C).

In this article, we characterize both the electron beam induced deposition of titanium oxide from condensed titanium isopropoxide films using spatially resolved Auger electron spectroscopy (SR-AES) as well as the transfer of these titanium oxide patterns to the underlying GaAs substrate via Cl_2 etching [using scanning electron microscopy (SEM)].

II. EXPERIMENTAL METHODS

All deposition experiments were conducted in an ultrahigh vacuum (UHV) chamber (base pressure of 5×10^{-10} Torr) which is equipped with a commercial scanning Auger microscope (PHI 660 system). The microscope has

[a)]Electronic mail: hu@ece.ucsb.edu

the ability to perform both (computer-controlled) SEM and SR-AES experiments. In this study, the selective area deposition was effected using the SEM function, whereas SR-AES was used to monitor the deposition kinetics *in situ*.

The samples used in this study were 100 Å thick GaAs (undoped) films grown on n^+-GaAs(001) substrates by molecular beam epitaxy which were Sb capped to prevent air oxidation upon exposure to the atmosphere. The samples were mounted onto molybdenum holders via indium bonding and were loaded into the chamber via a turbopumped loadlock. The GaAs substrates could be cooled to −60 °C using liquid nitrogen in a remote cryostat or heated radiatively to 600 °C by three W filaments (from halogen light bulbs) located within a heater housing underneath the sample holder. Sample temperature was monitored directly by a type-K thermocouple attached to a movable arm which can either press the thermocouple on the sample surface during a measurement or lift and clear it from the sample area to ensure unimpeded sample transfer to and from the loadlock.

The titanium isopropoxide source (Aldrich, 99.999%) was transferred, in a nitrogen-purged glovebox, to a valved glass bulb which was then attached to a high-vacuum gas manifold system (evacuated by a 60 l/s turbopump). The bulb was mounted in a temperature-controlled water bath to ensure flux reproducibility. Repeated freeze–pump–thaw cycles under vacuum were used to further purify the precursor source. A directional tubular doser consisting of a $\frac{1}{4}$ in. tube mounted on a z-transfer manipulator was used to expose the precursor vapor directly onto the sample surface (background pressure remained below 4×10^{-9} Torr during exposure).

In all experiments reported here, the Sb-capped GaAs substrates was first annealed to 520 °C to desorb the Sb cap and produce a clean, oxygen- and carbon-free, GaAs(001) surface. After cooling to below −20 °C (which roughly corresponds to the precursor desorption temperature), the surface was then exposed to $Ti(OC_3H_7)_4$ vapor for 2 min. This exposure would ensure that a thick multilayer (>50 Å) had formed as measured by AES.[18] Nominal squares of a dimension of 70×70 μm^2 were then patterned by sweeping a 0.2 nA, 10 keV electron beam over the selected areas for different times. Upon completion of patterning, and prior to *in situ* analysis, the substrate was allowed to warm to room temperature to desorb the unexposed precursor and volatile reaction products, and define the patterned titanium oxide squares. After analysis, the patterned substrates were transferred (quickly through air) to a remote etch chamber (described in detail elsewhere[19]) for the Cl_2 etching experiments.

III. RESULTS AND DISCUSSION

Auger line scans recorded (using SR-AES) after patterning a series of 70×70 μm^2 oxide squares with electron exposures varying from 23 to 4000 $\mu C/cm^2$ are shown in Fig. 1. The spectra were collected by scanning lines (128 points/line) across all the squares after warming the substrate to room temperature and they show explicitly the spatial variation in the intensity of the Ga, O, Ti, and C Auger transitions

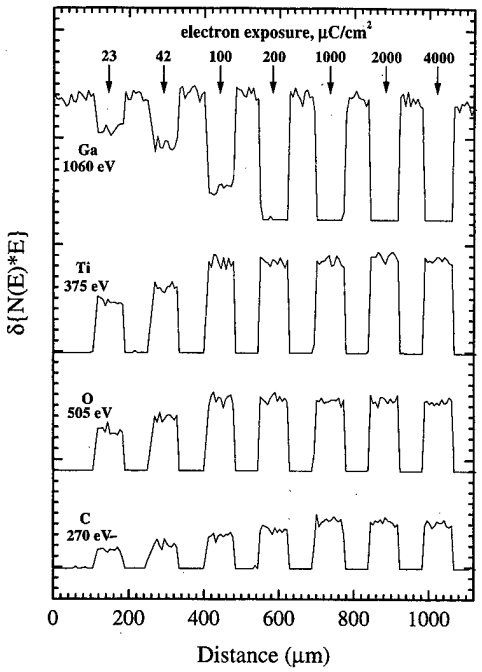


FIG. 1. *In situ* Auger line scans recorded across all the titanium oxide squares which were patterned by the indicated electron exposures upon the titanium isopropoxide multilayer at −60 °C. All scans were recorded after warming the substrate to room temperature. A beam energy of 10 keV and a beam current of 40 nA were used.

at 1060, 505, 375, and 270 eV, respectively. The buildup of the titanium oxide film with electron exposure (within the patterned areas) is clear from the increase in both the Ti and O signal intensities, as well as the decrease in the intensity of the Ga signal originating from the underlying GaAs substrate. Note that by virtue of a larger electron escape depth,[20] the Ga signal ($\lambda\approx20$–25 Å) goes to zero some time after saturation of the Ti and O signals ($\lambda\approx6$–8 Å) and, therefore, the Ga signal provides more information regarding the film growth kinetics. We utilize this fact below to determine effective reaction cross sections.

It is clear that C signal intensity also increases with electron exposure, indicative of increased carbon incorporation within the thicker oxide films. Furthermore, saturation of the C signal occurs at significantly higher exposures than that for the Ti and O signals (order of 10^3 $\mu C/cm^2$ versus order of 10^2 $\mu C/cm^2$, respectively), an effect that can be explained by simply assuming a larger cross section for O–C bond cleavage, relative to C–H bond cleavage, in the original precursor.[21] With this picture, the electrons prefer to initially attack the O–C bond, producing titanium oxide and volatile hydrocarbon (HC) species. These HC species can then be removed from the film by desorption, either during the post exposure anneal to room temperature or directly via electron stimulation, resulting in minimal C incorporation, or be decomposed by the less facile C–H cleavage reaction which produces a nonvolatile C deposit. As the electron exposures are increased, the balance between these two possible routes is tilted toward the decomposition channel and increased C incorporation in the thicker films results.

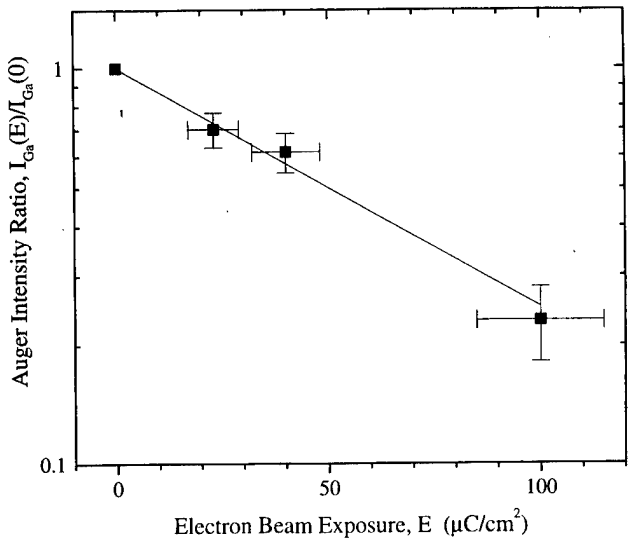


FIG. 2. Electron exposure dependence of the Ga Auger signal intensity (relative to the intensity from the unexposed substrate). The data were taken from Fig. 1.

Further analysis of the spectra in Fig. 1 show that the Ti, O, and C transitions (as monitors for the precursor in any form, intact or decomposed) maintain essentially zero signal intensity between the patterned oxide squares. This indicates that negligible decomposition of the precursor occurs outside the exposed areas and clearly demonstrates the excellent pattern selectivity achievable with titanium isopropoxide.

It is well known[22] that the intensity of the Ga transition from the underlying GaAs substrate, $I_{Ga}(E)$, can be directly related to the thickness of the growing oxide film, $h_{ox}(E)$, as

$$\frac{I_{Ga}(E)}{I_{Ga}(0)}=\exp\left(-\frac{h_{ox}(E)}{\lambda_{Ga}\cos\theta}\right), \tag{1}$$

where E is the total electron exposure on the condensed multilayer (in C/cm^2), λ_{Ga} is the electron escape depth for the Ga transition (22±3 Å), and θ is the electron energy analyzer acceptance angle (42° for a cylindrical mirror analyzer). Shown in Fig. 2 is the dependence of the measured Auger intensity ratio, $I_{Ga}(E)/I_{Ga}(0)$ (cf. Fig. 1), upon the electron exposure, E. It is clear from Fig. 2 that $\ln[I_{Ga}(E)/I_{Ga}(0)]$ is proportional to E (for E equal to 100 $\mu C/cm^2$ and below), so that we may write

$$\ln\left(\frac{I_{Ga}(E)}{I_{Ga}(0)}\right)=-mE, \tag{2}$$

where m, being the slope of the line in Fig. 2, is found to be 13 900±800 cm^2/C. Equating Eqs. (1) and (2) gives

$$h_{ox}(E)=mE\lambda_{Ga}\cos\theta, \tag{3}$$

implying that the deposited oxide thickness is also linear in electron exposure, E. Implications of this linear relation are derived below.

From the standard collision theory model of reaction rates, we can write for the deposition rate

$$-\frac{d\vartheta_{pre}(t)}{dt}=\sigma^1\frac{F}{q}\vartheta_{pre}(t), \tag{4}$$

where $\vartheta_{pre}(t)$ is the surface concentration of alkoxide precursor [in monolayers (ML)] at time t, σ^1 is the first-order reaction cross section (cm^2/electron), F is the incident electron flux (in $C/cm^2/s$), and q is the electronic charge (1.6×10^{-19} C/electron). Since $d\vartheta_{pre}=-d\vartheta_{ox}$ by stoichiometry (where ϑ_{ox} is the surface concentration of deposited oxide), $h_{ox}=\bar{h}\vartheta_{ox}$ (where $\bar{h}$ is the effective height of the oxide monolayer ≈2–3 Å/ML), and $E=Ft$, we can write Eq. (4) as

$$\frac{dh_{ox}(E)}{dE}=\frac{\sigma^1\bar{h}}{q}\vartheta_{pre}(E). \tag{5}$$

To ensure a linear relation between h_{ox} and E as determined above experimentally [cf. Eq. (3)], we must have $\vartheta_{pre}(E)=\vartheta_{pre}(0)$=constant in Eq. (5), i.e., the deposition kinetics must be zeroth order in precursor concentration.[23] With this understanding, we can integrate Eq. (5) to get the correct linear dependence of h_{ox} on E,

$$h_{ox}(E)=\frac{\sigma^0\bar{h}}{q}E, \tag{6}$$

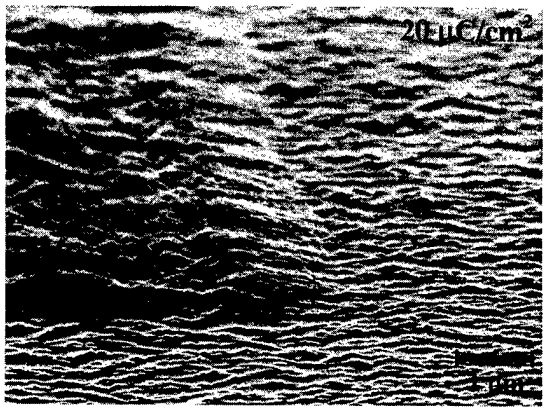
where we have defined an effective zeroth-order cross section as σ^0. Equating Eq. (1) with Eq. (6) then finally gives

$$\sigma^0=\frac{qm\lambda_{Ga}\cos\theta}{h}. \tag{7}$$

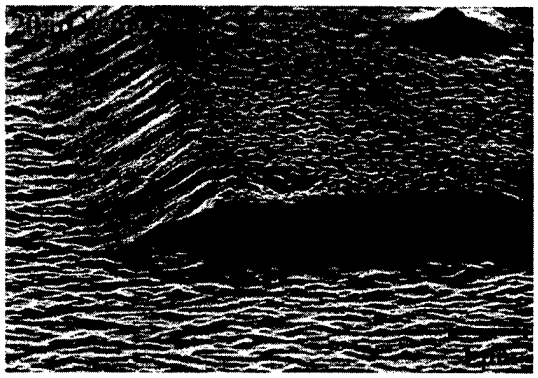
Consequently, using Eq. (7), the zeroth-order cross section for the electron beam induced decomposition of the titanium isopropoxide multilayers was estimated to be $(1.5\pm0.6)\times10^{-14}$ cm^2/electron. This value is extremely high, e.g., Matsui *et al.* find, under precursor-limited deposition conditions, a much lower σ^0 of 10^{-20}–10^{-21} cm^2/electron for the deposition of W and Si from WF_6 and $Si_2H_2Cl_2$, respectively.[3,5]

It is clear from the above that titanium isopropoxide exhibits the superior deposition properties, namely, excellent electron sensitivity and selectivity, required of a practical *in situ* resist. It now remains to investigate the etch transfer characteristics of these patterned oxide films. The results of a preliminary study are presented below.

Shown in Figs. 3–5 are SEM micrographs recorded after etching the patterned GaAs substrate at 250 °C in 2×10^{-4} Torr Cl_2 for 3 min. Figure 3 clearly shows that too low an electron exposure, i.e., 23 $\mu C/cm^2$ or lower, results in a poor quality mask—although a rough oxide surface is maintained in one corner of the patterned square [cf. Fig. 3(b)], the complete breakdown of the oxide at another corner [cf. Fig. 3(a)] is indicative of a film that is too thin. However, as can be seen in Fig. 4, increasing the electron exposure to only 40 $\mu C/cm^2$ is sufficient to generate a robust oxide film with good pattern transfer characteristics (crystallographic undercut- and V-etch profiles in the $\langle011\rangle$ and $\langle01\bar{1}\rangle$ directions, respectively). It should be pointed out that these exposure densities approach those used in patterning the standard

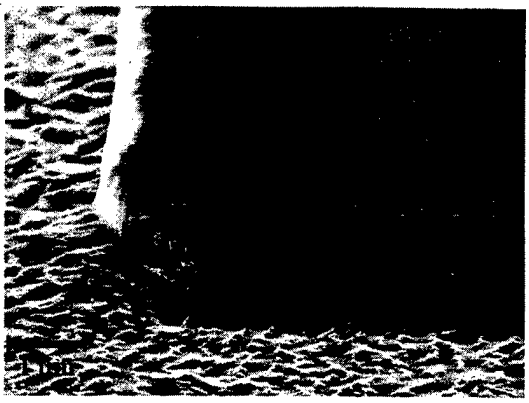


(a)

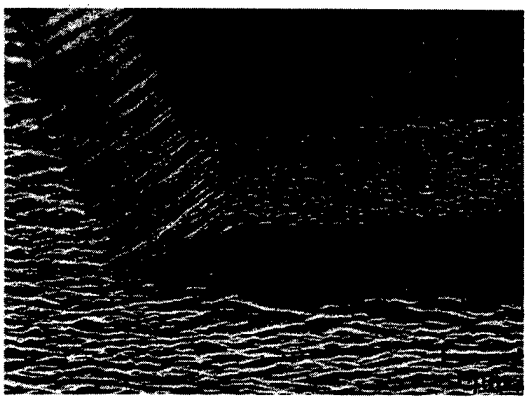


(b)

FIG. 3. *Ex situ* SEM micrographs recorded at different ends of the titanium oxide square patterned by a 20 $\mu C/cm^2$ electron exposure and after etching in 2×10^{-4} Torr Cl_2 for 5 min at 250 °C.



(a)



(b)

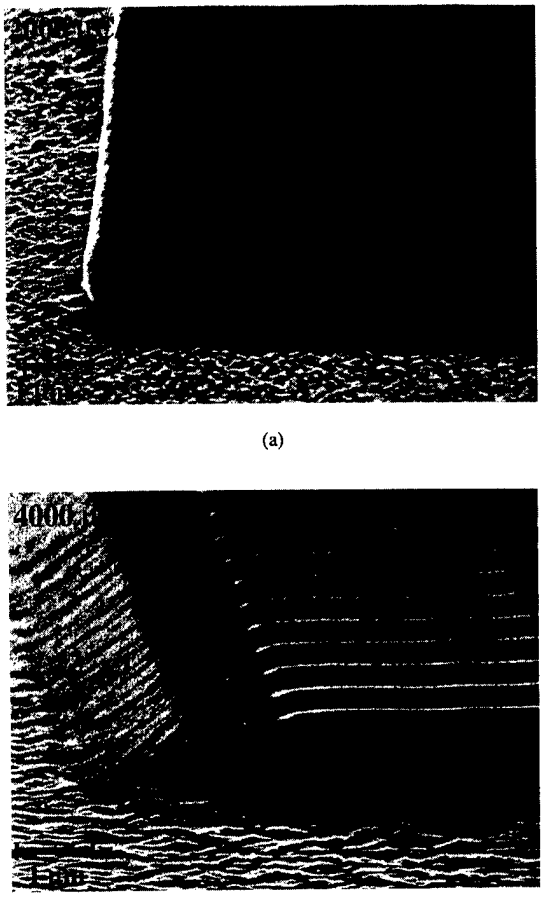
FIG. 4. *Ex situ* SEM micrographs recorded along the (a) $\langle 01\bar{1}\rangle$ and (b) $\langle 011\rangle$ directions after etching the patterned substrate in 2×10^{-4} Torr Cl_2 for 5 min at 250 °C. The titanium oxide square shown was written with an electron exposure of 40 $\mu C/cm^2$.

organic electron beam sensitive resists, such as PMMA. Similar robust behavior was also observed with the squares patterned by 100, 200, and 1000 $\mu C/cm^2$ exposures (micrographs not shown).

Issues that relate to the ultimate resolution of patterned features using titanium isopropoxide as an *in situ* resist are raised in Fig. 5. Shown in Fig. 5(a) is the etched oxide square patterned by a 2000 $\mu C/cm^2$ electron exposure which, as expected, exhibits the excellent etch resistance and pattern transfer properties of the film shown in Fig. 4. Of more interest, however, are the clearly resolved rows of closely spaced dots on the surface of the oxide film (aligned along the $\langle 011\rangle$ direction[24]) which appear only after these higher exposures. We believe that these raised oxide dots are an artifact of the (discrete) computer-controlled scan generator and represent the overexposed areas where the primary 10 keV beam repeatedly dwells during the fast patterning scans (at 5 frames/s). The presence of titanium oxide in the regions between these primary dots in Fig. 5(a) must, therefore, be a result of precursor decomposition induced by scattered electrons, either lower energy secondary electrons that are produced by inelastic collisions between the condensed precursor layer and the primary electron beam or backscattered electrons which are primary electrons that have scattered elastically in the precursor layer and from the substrate. These scattered electrons exhibit a broad distribution in angle and, consequently, will result in a poorer pattern resolution than that based solely on the incident spot size of the primary beam, e.g., in Fig. 5(a), the oxide pattern edge extends at least 1 μm beyond the closest row of primary dots that are submicron in size. Further implications of this effect will be discussed in more detail elsewhere.[25] Moreover, note that after doubling the electron exposure to 4000 $\mu C/cm^2$ [cf. Fig. 5(b)], the rows of discrete dots resolved in Fig. 5(a) converge into continuous rows. It appears that the increased exposure has produced bigger dots which, by virtue of their close spacing along the $\langle 011\rangle$ direction, merge to form a continuous row.

IV. CONCLUSIONS

We have investigated the interaction of electrons with multilayer films of titanium isopropoxide on cold (<-20 °C) GaAs substrates as well as the etch characteristics of the patterned titanium oxide films which result. Titanium isopropoxide is extremely sensitive to electron irradiation and is found to decompose to nonvolatile titanium oxide (with



(a)

(b)

FIG. 5. *Ex situ* SEM micrographs recorded after etching the patterned substrate in 2×10^{-4} Torr Cl_2 for 5 min at 250 °C. The titanium oxide squares shown were written with electron exposures of (a) 2000 and (b) 4000 $\mu C/cm^2$.

some C incorporation) with a high zeroth-order cross section of $(1.5\pm0.6)\times10^{-14}$ cm^2/electron (under excess precursor conditions). All unexposed precursor simply desorbs upon annealing to room temperature which clearly defines the patterned titanium oxide regions. Electron exposures greater than 20 $\mu C/cm^2$ result in titanium oxide films that are resistant to etching in Cl_2 and which exhibit excellent pattern transfer characteristics to the underlying GaAs substrate. We find that "scattered" electrons (secondary and/or backscattered) play an important part in the electron induced chemistry observed, an effect which will limit the ultimate pattern resolution possible.

It is clear that titanium isopropoxide makes an excellent candidate for a negative-type resist in an *in situ* electron beam lithography scheme.

ACKNOWLEDGMENTS

This work was supported by QUEST, a National Science Foundation (NSF) Science and Technology Center (Grant No. DMR91-20007). The *ex situ* SEM study made use of the MRL Central Facilities at U.C., Santa Barbara, which is supported by the NSF (Grant No. DMR91-23048).

[1]See, for example, H. Madden and G. Ertl, Surf. Sci. **35**, 211 (1973).
[2]L. Ng, J. Chen, P. Basu, and J. Yates, Langmuir **3**, 1161 (1987).
[3]S. Matsui and K. Mori, Appl. Phys. Lett. **51**, 646 (1987).
[4]H. Koops, R. Weiel, D. Kern, and T. Baum, J. Vac. Sci. Technol. B **6**, 477 (1988).
[5]S. Matsui and M. Mito, Appl. Phys. Lett. **53**, 1492 (1988).
[6]F. Bozso and P. Avouris, Appl. Phys. Lett. **53**, 1095 (1988).
[7]J. Swanson, F. Flitsch, and C. Friend, Surf. Sci. Lett. **215**, L293 (1989).
[8]F. Hirose and H. Sakamoto, Appl. Surf. Sci. **135**, 293 (1998).
[9]One exception was the study performed by Bozso and Avouris (Ref. 6) who effected reaction between coadsorbed Si_2H_6 and O_2, NO, or NH_3 to produce Si oxide, oxynitride, and nitride films, respectively.
[10]M. Taneya, Y. Sugimoto, H. Hikada, and K. Akita, J. Appl. Phys. **67**, 4297 (1990).
[11]M. Taneya, Y. Sugimoto, H. Hikada, and K. Akita, Jpn. J. Appl. Phys., Part 2 **29**, L182 (1990).
[12]M. Taneya, Y. Sugimoto, H. Hikada, and K. Akita, J. Appl. Phys. **68**, 3630 (1990).
[13]Y. Sugimoto, K. Akita, M. Taneya, and H. Hikada, Appl. Phys. Lett. **57**, 1012 (1990).
[14]H. Kawanishi, Y. Sugimoto, N. Tanaka, and T. Ishikawa, Jpn. J. Appl. Phys., Part 1 **32**, 4033 (1993).
[15]T. Ishikawa, N. Tanaka, M. Lopez, and I. Matsuyama, J. Vac. Sci. Technol. B **13**, 2777 (1995).
[16]E. Clausen, J. Harbison, C. Chang, P. Lin, H. Craighead, and L. Florez, Appl. Phys. Lett. **57**, 1043 (1990).
[17]S. Kohmoto, N. Takado, Y. Sugimoto, M. Qzaki, M. Sugimoto, and K. Asakawa, Appl. Phys. Lett. **61**, 444 (1992).
[18]It would not be possible to detect the high-energy GaAs transitions at 1070 and 1230 eV in the measured Auger spectra.
[19]D. Mui, T. Strand, B. Thibeault, L. Coldren, P. Petroff, and E. Hu, Proceedings of the 21st International Symposium Compound Semiconductors, edited by H. Goronkin and U. Mishra (1994), p. 291.
[20]I. Lindau and W. Spicer, J. Electron Spectrosc. Relat. Phenom. **3**, 409 (1974).
[21]Since there is no significant difference in their electron escape depths, we would expect that the Ti, O, and C signals would saturate at about the same time if the growth rate of each element within the film were equal.
[22]See, for example, O. Millo, A. Many, and Y. Goldstein, J. Vac. Sci. Technol. A **7**, 4033 (1993).
[23]This suggests that the precursor is so much in excess that its effective surface concentration during reaction remains constant.
[24]The alignment of the dots along this direction has no structural significance. The scanned squares were simply aligned relative to the cleaved edges of the GaAs(001) substrate initially.
[25]W. Mitchell and E. Hu (to be published).

Reduction and creation of paramagnetic centers on surfaces of three different polytypes of SiC

P. J. Macfarlane[a] and M. E. Zvanut
Department of Physics, University of Alabama at Birmingham, 310 Campbell Hall, Birmingham, Alabama 35294-1170

(Received 19 January 1999; accepted 27 May 1999)

SiC is of interest to create power metal–oxide–semiconductor field-effect transistors because it can be thermally oxidized to form a SiO_2 dielectric layer. Previously, we used electron paramagnetic resonance to identify centers in 3C–SiC epilayer samples and 4H–SiC and 6H–SiC wafer samples after oxidation and dry heat treatment [P. J. Macfarlane and M. E. Zvanut, Appl. Phys. Lett. **71**, 2148 (1997); Mater. Res. Soc. Symp. Proc. **513**, 433 (1998)]. The spectroscopic and thermal characteristics of these centers indicate that they are related to C. Because these centers are activated in a H_2O-poor ambient and are passivated in an ambient that is H_2O-rich, we suggest that the activation mechanism is release of a hydrogenous species. In this investigation, the effect of repeated oxidations on the concentration of heat-treatment-induced centers is studied. Samples are successively oxidized at 1150 °C in O_2 bubbled through de-ionized water for 1, 2, 4, 8, and 16 h. After each oxidation, the samples are heat treated in dry (<0.1 ppm H_2O) N_2. Prior to the next oxidation, the oxide is removed. Upon oxidation of the samples we observe an order of magnitude reduction in the concentration of centers that are present on the as-prepared substrates. After each oxidation centers are activated by dry heat treatment. We suggest that the centers present on the as-prepared substrates are related to surface damage and are removed during the oxidation as the surface SiC material is converted in the oxidation products. Two models are offered for the source of the centers generated by dry heat treatment. The centers could be activated from C–H bonds related to damage like micropipes, nanopipes, or Si vacancies distributed throughout the SiC substrate, or they could arise from C–H bonds that form during the oxidation. We will discuss the merits of both of these models. © *1999 American Vacuum Society.* [S0734-211X(99)09004-6]

I. INTRODUCTION

SiC, with its wide band gap and its high thermal conductivity, is proposed to replace Si for high power, high temperature microelectronic devices. In particular, for power metal–oxide–semiconductor field-effect transistors (MOSFETs), SiC is of interest because unlike the other wide band-gap semiconductors, it can be thermally oxidized similarly to Si in order to create a SiO_2 insulating layer. Because SiC-based MOSFETs can be created using methods that are similar to those used to manufacture Si-based MOSFETs, the transfer of manufacturing technology to produce SiC MOSFETs is expected to be less complicated than to produce similar devices based on other wide band-gap materials like GaN.

SiC-based devices suffer from unacceptable concentrations of electrically active defects that affect the expected reliability of these devices.[1] In the development of Si-based MOS structures, electrically active defects were identified using electron paramagnetic resonance (EPR). For example, electrically active interface states were found to be related in part to an interfacial Si dangling bond, called the P_b center,[2] and charge traps in the SiO_2 have been attributed to the E' center.[3]

In our previous studies of 3C-SiC, 4H-SiC, and 6H-SiC,[4,5] we have observed surface related centers the concentration of which can be reduced by wet thermal oxidation. However, it appears that the effect produced by oxidation can be reversed by dry heat treatments at temperatures greater than 800 °C. The g values of these centers, from 2.0025 to 2.0029, are within the range of g values that are typical of C-related centers, and the temperatures at which these centers are generated are significantly greater than those used to generate Si dangling bonds. Thus, we suggest that the centers are unpaired electrons located on C atoms. Because these centers can be activated in an ambient that does not contain moisture and passivated in an ambient that contains moisture, we suggest that these centers are generated by release of a hydrogenous species from C dangling bonds. Supporting the relation to hydrogen is an experiment in which we heat treat oxidized samples in dry (<0.6 ppm H_2O) O_2. The concentration of centers activated is similar to that activated by dry heat treatment in N_2. HF etching studies of the samples indicate that the centers are not located in the oxide.

In studies of Si-based MOSFETs, release of hydrogen has been shown to lead to the degradation of the oxide and the oxide/substrate interface.[6] Thus, as the SiC is processed to produce a MOSFET, it is necessary to understand how SiC surfaces are altered by oxidation and how hydrogen is incorporated in the structure. In this investigation of 3C–SiC, 4H–SiC, and 6H–SiC, we examine the effects of repeated oxidations, dry heat treatments, and HF etching on the concentration of centers in a single sample. While the hexagonal

[a]Electronic mail: pmacfarl@phy.uab.edu

polytypes of SiC are more applicable to current high power device research, we include 3C–SiC to gain a clearer scientific understanding.

II. EXPERIMENT

We examine 4H–SiC and 6H–SiC substrates, and 3C–SiC epilayers. Both SiC substrate wafers are p type and double-sided polished. They are cut so that their polished faces are oriented 3.5° off the (0001). The 3C–SiC epilayer samples are deposited by chemical vapor deposition (CVD) on both the polished and the unpolished (100) faces of a Si wafer. The 3C–SiC epilayers are approximately 1 μm thick. Samples are cut into strips 1.5 cm by 0.23 cm, a size suitable for the EPR microwave cavity. Prior to oxidation, we clean the samples by rinsing them in trichloroethane, xylenes, acetone, methanol, and de-ionized water for 5 min time intervals. The samples are then etched for 1 min in 9:1 H_2O:HF (50%). The samples are oxidized for 1, 2, 4, 8, and 16 h time intervals at 1150 °C in O_2 bubbled through de-ionized water. After each oxidation the samples are measured with EPR and then heat treated in dry (<0.1 ppm H_2O) N_2 for 200 min. After the dry heat treatment, the samples are measured with EPR. Before the samples receive their next oxidation, the oxide is removed using a 9:1 H_2O:HF (50%) solution and the samples are measured with EPR again.

All EPR measurements are conducted at room temperature using an X-band Bruker 200 spectrometer. EPR is a spectroscopic technique able to detect paramagnetic defects in solids. We used the dry heat treatment to activate the paramagnetic state of a defect already present in the material. For example, often after oxidation, an unpaired electron on a substrate Si atom at the Si/SiO_2 interface (Si P_b center), is passivated by H. Thus, it cannot be detected with EPR. Vacuuming annealing or dry heat treating the oxidized Si at temperatures greater than 500 °C breaks the Si–H bond, activates the paramagnetic state, and reveals the center. Unpaired electrons in a solid absorb microwaves according to the resonant condition:

$$h\nu = \mu_B g B_0, \tag{1}$$

where $h\nu$ is the energy of the microwave, μ_B Bohr magneton for the electron, B_0 is the value of the magnetic field where the EPR signal intensity crosses zero, and g is the g value. We use the g value as a characteristic defect marker for the center. The concentration of centers is found by double integration of the EPR spectrum and comparison with the double integral of a spectrum obtained from a known standard. Typically, we integrate the EPR spectrum with the largest amplitude and determine the concentration from the other spectra by scaling their amplitudes. All spectra are measured with 1 G peak-to-peak modulation amplitude and 1 mW microwave power.

After dry heat treatment, the oxide thickness is measured by surface profilometry. Before receiving the next oxidation, we thoroughly clean the samples with solvents and the samples are etched in HF to remove the SiO_2. We approximate the amount of SiC converted into oxidation products by the method used for Si oxidation. For an oxide of thickness d, the thickness of SiC removed from the substrate is $0.46d$. This provides a measure of the amount of SiC removed during oxidation.

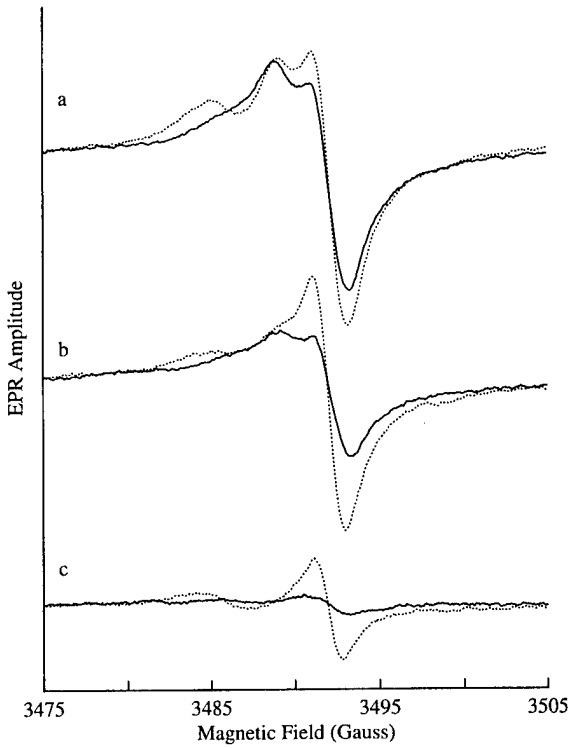


FIG. 1. EPR spectra of 3C–SiC epilayer samples measured at room temperature after 1 h oxidation (a), after 4 h oxidation (b), and after 16 h oxidation (c). The solid curves are spectra measured just after oxidation and the dotted curved are measured after dry heat treatment. Centers are observed at approximately $g=2.0060$ (3486 G), $g=2.004$ (3490 G), and 2.0029 (3492 G).

III. RESULTS

Prior to oxidation, there is a large signal present in the EPR spectrum of each of the SiC polytype samples. In the 6H–SiC, the center is located at $g=2.0026\pm0.0001$, in the 4H–SiC, the center is located at $g=2.0025\pm0.0001$, and in the 3C–SiC epilayer samples, centers are located at $g=2.0029\pm0.0001$, $g=2.004\pm0.001$, and $g=2.006\pm0.001$. The concentration of the centers present in the hexagonal polytypes and the center at $g=2.0029$ in the 3C–SiC epilayer samples is approximately 3×10^{13} cm^{-2}.

EPR spectra of the 3C–SiC epilayer samples are shown in Fig. 1 after 1 h [Fig. 1(a)], 4 h [Fig. 1(b)], and 16 h [Fig. 1(c)] oxidations. The solid lines indicate spectra measured after oxidation and the dotted curves are spectra measured after dry heat treatment. The solid curves indicate the presence of two centers located at approximately $g=2.004\pm0.001$ (3490 G) and $g=2.0029\pm0.0001$ (3492 G). The amplitudes of the signals remaining after oxidation decrease with increasing oxidation time. In particular, we note the center at $g=2.004$ is scarcely resolved from the center at $g=2.0029$ after 16 h oxidation. The dry heat treatments increase the amplitude of the signal at $g=2.0029$ but do not appear to strongly affect the concentration of centers at $g=2.004$. The dry heat treatments also produce a third center

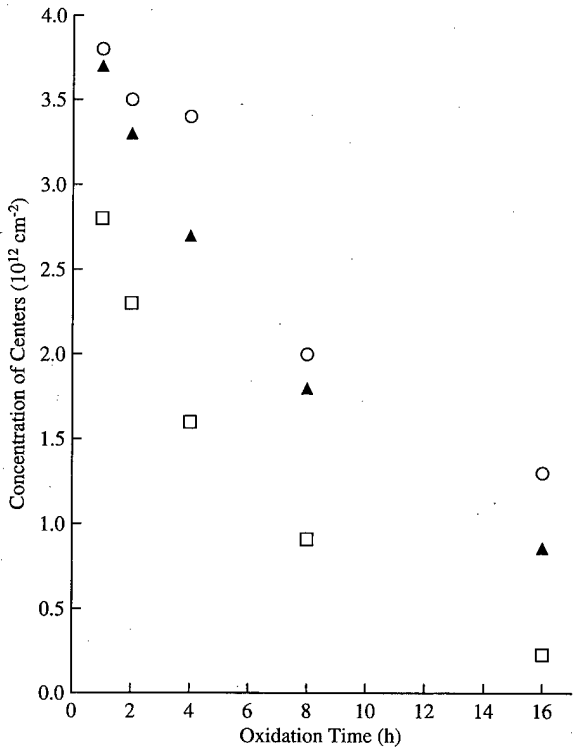


FIG. 2. Concentration of centers in the oxidized 3C–SiC epilayer samples located at $g=2.0029$ (3492 G) as a function of individual oxidation time. (□) represent the total concentration of centers measured after oxidation, (○) represent the total concentration of centers measured after dry heat treatment, and (▲) represent the total concentration of centers measured after the oxide is removed.

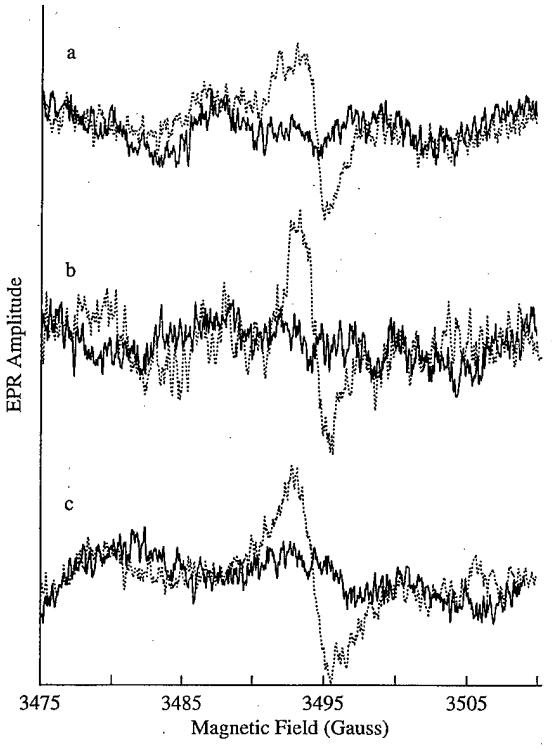


FIG. 3. EPR spectra of 6H–SiC measured at room temperature after 1 h oxidation (a), after 4 h oxidation (b), and after 16 h oxidation (c). The solid curves are spectra measured just after oxidation and the dotted curved are measured after dry heat treatment. A center is observed at approximately 3494 G (2.0026).

at approximately $g=2.006\pm0.001$ (3486 G), which we suggest is related to the Si P_b center.[7]

In Fig. 2, we plot the concentration of the center at $g=2.0029$ in the 3C–SiC epilayer samples after 1, 2, 4, 8, and 16 h oxidation times. Note that the x axis represents the time for a single oxidation, not the total time the sample was oxidized. The unfilled squares represent the concentration of centers after oxidation, the unfilled circles represent the concentration of centers after the dry heat treatment, and the filled triangles represent the concentration of centers after the oxide is removed by HF etching. The concentration of centers present after oxidation (□) decreases as the 3C–SiC epilayer samples are repeatedly oxidized. After each individual oxidation, we observe an increase in the total concentration of centers after dry heat treatment (○) and a slight reduction in center concentration after removal of the oxide by HF etching (▲). The variation of the concentration of centers induced by the heat treatment alone is more complicated and will be discussed later.

Spectra of the 6H–SiC sample after 1 h [Fig. 3(a)], 4 h [Fig. 3(b)], and 16 h [Fig. 3(c)] oxidations are illustrated in Fig. 3. The solid curves are spectra of the sample measured after oxidation and the dotted curves are measured after subsequent dry heat treatment. After each of the oxidations, no signal is visible above the noise level of the EPR spectrum. After dry heat treatment we observe the activation of a center at approximately $g=2.0026\pm0.0001$ (3494 G). The signal-to-noise ratio of the spectra for the 6H–SiC sample is lower than for the 3C–SiC samples because the high conductivity of the 6H–SiC substrate decreases the EPR sensitivity.

In Fig. 4, we plot the concentration of the center at $g=2.0026$ in 6H–SiC samples as a function of individual oxidation time. The legend is the same as that used in Fig. 2. The uncertainty in concentration is estimated by visual comparison of the relative noise amplitudes. The dashed line indicates the detection limit of the spectrometer for these samples. As with the 3C–SiC samples, the concentration of centers after each oxidation is increased by the subsequent dry heat treatment, and the concentration of centers decreases after the oxide is removed by HF etching. Here, however, we do not observe the monatomic decrease with oxidation time, which is apparent in the 3C–SiC data. Initially over the 1 and 2 h oxidations, we observe an increase in the concentration of centers activated by the dry heat treatment following which the concentration of centers decreases until it reaches a nearly constant value after the 8 h oxidation. The trends observed for the activation of a center at $g=2.0025$ in the 4H–SiC samples are similar to those observed for the 6H–SiC material.

In Fig. 5, we examine the concentration of centers produced by the dry heat treatment is affected by the successive oxidations of increasing duration. For the 3C–SiC epilayer

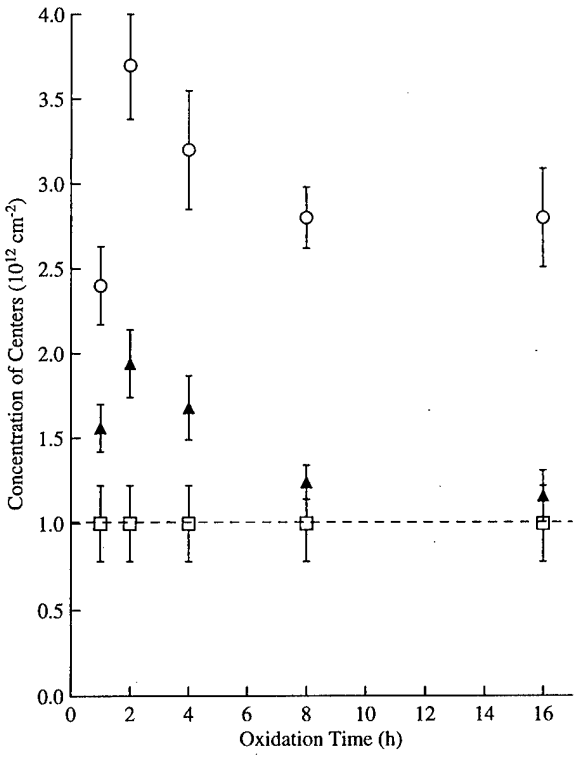


FIG. 4. Concentration of center in the oxidized 6H–SiC sample located at $g=2.0029$ (3492 G) as a function of individual oxidation time. (□) represent the total concentration of centers measured after oxidation, (○) represent the total concentration of center measured after dry heat treatment, and (▲) represent the total concentration of centers measured after the oxide is removed. The error in concentration is due to noise in the EPR spectrum.

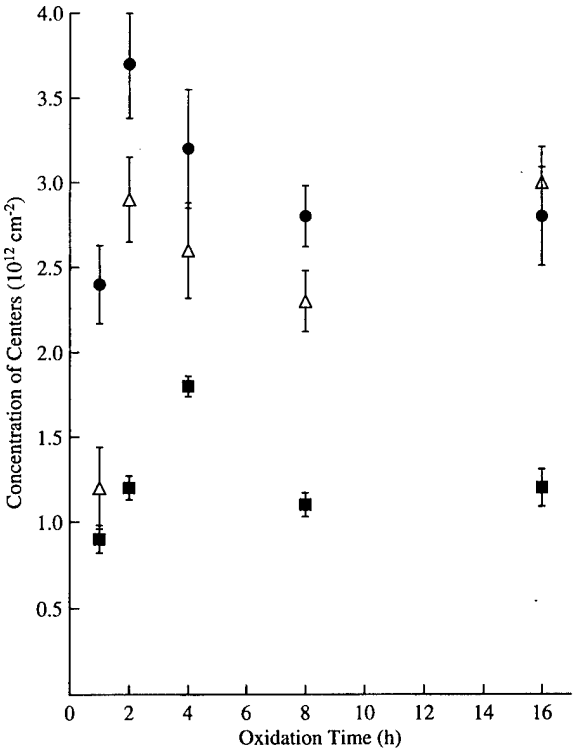


FIG. 5. Concentrations of the center at $g=2.0026$ in the 6H–SiC sample (●), the center at $g=2.0025$ in the 4H–SiC sample (△), and the center at $g=2.0029$ in the 3C–SiC epilayer samples (■) measured after dry heat treatment plotted as a function of individual oxidation time. The error in concentration is due to noise in the EPR spectra.

samples, the concentration of centers produced by the heat treatment is determined by subtracting the concentration of centers measured after oxidation (□ in Fig. 2) from the concentration of centers measured after the subsequent dry heat treatment (○ in Fig. 2) and are indicated by the filled squares in Fig. 5. Because the concentration of the centers after oxidation in both the 4H–SiC and 6H–SiC samples is below the detection level, the concentration of the centers induced by the dry heat treatment is simply the center's concentration measured after dry heat treatment. For the 6H–SiC samples, the concentrations indicated in Fig. 5 by the filled circles are the same data that in Fig. 2 are symbolized by the unfilled squares. The unfilled triangles indicate the concentration of the centers induced by dry heat treatment in the 4H–SiC samples. The concentration of centers in the 4H–SiC (△) and 6H–SiC (●) samples increases until after the 2 h oxidation and then decreases with further oxidation until after 8 h oxidation, when it saturates. The concentration of centers in the 3C–SiC epilayer samples (■) behaves in a similar manner but the increase continues until after the 4 h oxidation. After the 8 h oxidation, the concentration remains approximately the same.

IV. DISCUSSION

There are three kinds of centers that can be identified by their differing responses to the various types of processing. First, there are the signals present on each of the as-prepared polytypes. These are reduced by wet oxidation. Second, there are the centers that can be activated by the dry heat treatment and reduced by wet oxidation. Third, in the 3C–SiC epilayer there is a center that is distributed throughout the epilayer, is reduced during oxidation, and is not affected by the dry heat treatment.

As indicated by the data, the oxidation process reduces the concentration of centers present in the as-prepared material. Before the samples are initially oxidized, the total concentration of surface related centers is approximately 3×10^{13} cm^{-2} for each of the different polytypes. The first 1 h oxidation of the hexagonal polytypes, which removes approximately 12 nm of material from the Si face and 85 nm of material from the C face of the SiC samples, reduces the concentrations of these centers below the EPR detection level (10^{12} cm^{-2}) for these materials. The 1 h oxidation of the 3C–SiC epilayer material, which removes 45 nm of the surface, reduces the center concentration by an order of magnitude. The centers reduced by the initial oxidation are likely created by cutting and polishing. The centers are reminiscent of ones in cut Si samples, which exhibit an EPR signal known as the *D* center. The *D* center arises from Si atoms that are out of register with the Si lattice.[8] In fact, the *D* center is visible in the spectrum of the 3C–SiC epilayers at $g=2.006$ prior to oxidation. For the 3C–SiC material, this center likely arises from the cut edges of the Si substrate upon which the 3C–SiC is deposited. By analogy, we sug-

gest that the centers with g values between 2.0025 and 2.0029 arise from the cut edges of the SiC material. Gerardi and co-workers observed a signal at $g=2.0025$ on 4H–SiC that was abraded with a diamond impregnated file.[9] They were able to remove the centers by etching the material in fused KOH for 10 min at 400 °C. Oxidation would likely have a similar effect where oxygen is used to etch the surface by converting the SiC into the oxidation products.

Each of the three polytypes contains at least one center that can be generated by dry heat treatment. We will focus on the centers at $g=2.0029$ in the 3C–SiC epilayer samples, at $g=2.0026$ in the 6H–SiC samples, and at $g=2.0025$ in the 4H–SiC samples. The g values of each of these centers are the same as those for the centers observed in these materials prior to oxidation indicating that both are likely related to C. However, the surface damage centers and the centers activated by dry heat treatment have different responses to forming gas anneals.[10] Thus, there must exist at least a suitable difference in the structure of these centers.

We suggest two different defect generation models based on the data in Fig. 5. The first is that these centers are present in the as fabricated SiC substrate and are activated by hydrogen effusion during heat treatment. Si vacancies reported by Itoh and co-workers[11] or micropipes well known to exist in the hexagonal substrates could both be candidates for the precursors of these centers. Each dry heat treatment depassivates defects within approximately the same thickness of SiC material located near the SiO_2/SiC interface. As the SiC is removed by successive oxidations, we can probe deeper into the material by depassivation. The data of Fig. 5 could reflect the dependence of defect density depth below the surface. The total material removed by the oxidations is approximately 1.2 and 0.2 μm for the C and Si faces, respectively, of the 4H- and 6H-SiC polytypes and 0.9 μm from the 3C–SiC epilayer. Figure 5 suggests that the defects are distributed in a nonuniform manner through the bulk SiC. A second possible source for the centers is the oxidation itself. During the first few hours of oxidation, the centers increase in concentration. The time dependence of the concentrations could indicate that there is a competition between generation and annealing of the defects during oxidation. Further work is necessary to distinguish between the intrinsic defect model and the oxidation induced defect model.

Because the center located at $g=2.004$ slowly decreases in concentration as the 1 μm 3C–SiC epilayer is removed by oxidation, it appears likely that this center is distributed throughout the 3C–SiC epilayer. Unlike the C-related centers that can be activated by dry heat treatment, this center is not affected by the depassivation procedure. Either this center is not related to hydrogen or the temperatures at which hydrogen depassivation of this defect occurs are much larger than 900 °C. The chemical origin of this signal is unknown.

V. CONCLUSIONS

In order to study the preparedness and stability of the SiC/SiO_2 interface, we have repeatedly oxidized three different polytypes of SiC for increasing lengths of time. After each oxidation, the samples are dry heat treated to depassivate any hydrogen that may be incorporated into the SiO_2/SiC structures. After the first oxidation, we observe at least an order of magnitude reduction in the concentration of centers observed in the as-prepared materials for each of the polytypes. However, the effect produced by the oxidation process is unstable. Dry heat treatments at 900 °C activate centers in each of the polytypes. These centers can be differentiated from the centers in the as-prepared material by forming gas anneals.[10] Although the precursor for the centers activated by the dry heat treatment is unknown, we suggest that the defects are either present in the as-fabricated SiC or are created during oxidation. Finally, we observe a third center, which is present only in the 3C–SiC epilayer. This center appears to be distributed throughout the epilayer, since the repeated oxidations appear to gradually reduce its concentration. This center is not affected by the 900 °C dry heat treatment, indicating that it is different from the centers that are induced by the dry heat treatment.

ACKNOWLEDGMENTS

The 3C–SiC epilayer samples were supplied by HOYA Corp., the 4H–SiC and 6H–SiC wafers were supplied by Cree Research. This work was supported by ONR. P. J. Macfarlane is supported by a fellowship from the Alabama Space Grant Consortium.

[1]L. A. Lipkin and J. W. Palmour, J. Electron. Mater. **25**, 909 (1996).
[2]E. Poindexter, J. Non-Cryst. Solids **187**, 257 (1995).
[3]P. M. Lenahan and P. V. Dressendorfer, J. Appl. Phys. **55**, 3495 (1984).
[4]P. J. Macfarlane and M. E. Zvanut, Appl. Phys. Lett. **71**, 2148 (1997).
[5]P. J. Macfarlane and M. E. Zvanut, Mater. Res. Soc. Symp. Proc. **513**, 433 (1998).
[6]D. J. DiMaria, E. Cartier, and D. Arnold, J. Appl. Phys. **73**, 3367 (1993).
[7]P. J. Macfarlane and M. E. Zvanut (unpublished).
[8]M. Brodsky and R. S. Title, Phys. Rev. Lett. **23**, 581 (1969).
[9]G. Gerardi, E. H. Poindexter, and D. Keeble, Appl. Spectrosc. **50**, 1428 (1996).
[10]P. J. Macfarlane and M. E. Zvanut, J. Electron. Mater. **28**, 144 (1999).
[11]H. Itoh and N. Hayakawa, J. Appl. Phys. **66**, 4529 (1989).

Finite linewidth observed in photoluminescence spectra of individual $In_{0.4}Ga_{0.6}As$ quantum dots

J. L. Spithoven,[a] J. Lorbacher, I. Manke, F. Heinrichsdorff, A. Krost,[b] D. Bimberg, and M. Dähne-Prietsch
Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstr. 36, D-10623 Berlin, Germany

(Received 20 January 1999; accepted 30 March 1999)

The photoluminescence of $In_{0.4}Ga_{0.6}As$ quantum dots was investigated by scanning-near-field optical microscopy at temperatures between 300 and 4 K. Using etched, metal-coated fiber tips, the pure signal of individual quantum dots could be detected. The spectra are found to consist of Lorentzian-shaped emission lines with linewidths ranging from 10 to 20 meV at room temperature, reducing to less than the spectrometer resolution of 1 meV at 4 K. This behavior is in contrast to the expectation of extremely small linewidths amounting to a few μeV, and can be related to lifetime effects mediated by thermal excitation of the carriers in the dots. Furthermore, the spectra show a redshift by about 8 meV upon increasing the excitation intensity, indicating biexcitonic effects.
© *1999 American Vacuum Society.* [S0734-211X(99)04004-4]

I. INTRODUCTION

Quantum dots are an interesting research object because of their discrete electronic energy levels, which can be tuned via their size and shape.[1–3] This behavior can be applied for a variety of photonic devices, e.g., for quantum-dot lasers already working at room temperature.[4] For spectroscopic investigations, extremely narrow photoluminescence (PL) lines with linewidths of a few μeV are expected because of the δ-like density of states and the low radiative recombination rates, which are typically in the order of 10^9/s.[1–3,5] On the other hand, quantum dots are often prepared in a self-organized way by Stranski–Krastanov growth, resulting in an inhomogeneous size distribution and therewith in a corresponding distribution of optical transition energies. In this way, spatially averaging PL experiments regularly yield broadened transition lines.[6,7]

The specific optical characteristics of individual quantum dots, in contrast, can be studied by experiments with high spatial resolution like scanning-nearfield optical microscopy (SNOM), scanning tunneling luminescence (STL), microphotoluminescence (μPL), or cathodoluminescence (CL). In many cases, quantum dots are formed with average lateral distances in the order of 100 nm.[6] STL experiments with their resolution of a few nm allow in principle the study of individual dots,[8,9] but usually require clean surface conditions for tunneling. This is in contrast to the other three methods, which enable the study of *ex situ* prepared samples. The spatial resolution of μPL or CL, on the other hand, is limited to several 100 nm, either by optical diffraction (μPL) or by the extensive diffusion of the hot secondary carriers (CL).[5,10] Therefore, these techniques usually allow to study small groups instead of individual quantum dots. However, CL spectra of a few InAs quantum dots showed transition linewidths below the spectral resolution limit of 150 μeV at temperatures up to 50 K, demonstrating their zero-dimensional electronic properties.[5] In contrast, SNOM with its outstanding spatial resolution in the order of 100 nm or even less[11,12] is ideally suited for the investigation of individual quantum dots.

In this work, we present SNOM-induced PL data taken at different temperatures of $In_{0.4}Ga_{0.6}As$ quantum dots embedded in a GaAs matrix, which were prepared by self-organized growth using metal-organic chemical vapor deposition (MOCVD). Etched and subsequently metal-coated fiber tips allowed to study the PL of individual quantum dots without the disturbing background signal from neighboring dots.[13] In this way, the spectra of single quantum dots were found to consist of Lorentzian-shaped transition lines with finite linewidths exceeding 10 meV at room temperature, decreasing to below the spectral resolution of 1 meV at 4 K. This astonishing observation is in contrast both to the theoretical prediction of extremely narrow lines[1–3] and to experimental results on other quantum-dot samples.[5,9,10] In our case, the observation of broad Lorentzian lines can be related to lifetime effects resulting from thermal excitation of carriers in the dots. At higher excitation intensities, a redshift of the PL signal from single quantum dots is observed, which can be explained by the excitation of biexcitons in the dots.

II. EXPERIMENTAL DETAILS

A. SNOM setup

For the SNOM experiments, two different home-built setups were used. One design is mounted inside of a glass dewar and can be used for measurements down to liquid-nitrogen temperature.[14] A second microscope operates in a continuous-flow He cryostat, allowing measurements at temperatures down to 4 K. A quadrant-sectored tube piezo (EBL 2) with 25 mm length and 3.2 mm diameter allows a scan range of about 100 μm×100 μm at 300 K, which reduces to about 20 μm×20 μm at 4 K. Piezo-driven inertial motors are

[a]On leave from: Faculteit Natuur- and Sterrenkunde, Universiteit Utrecht, P. O. Box 80000, 3508 TA Utrecht, The Netherlands.
[b]Permanent address: Otto-von-Guericke-Universität Magdeburg, D-39016 Magdeburg, Germany.

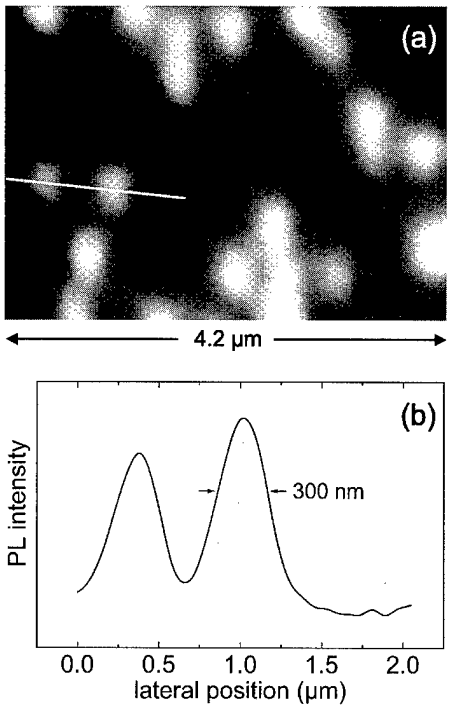


FIG. 1. PL image of $In_{0.4}Ga_{0.6}As$ quantum dots taken at room temperature. (a) Spectrally integrated intensity and (b) cross section along the line in (a).

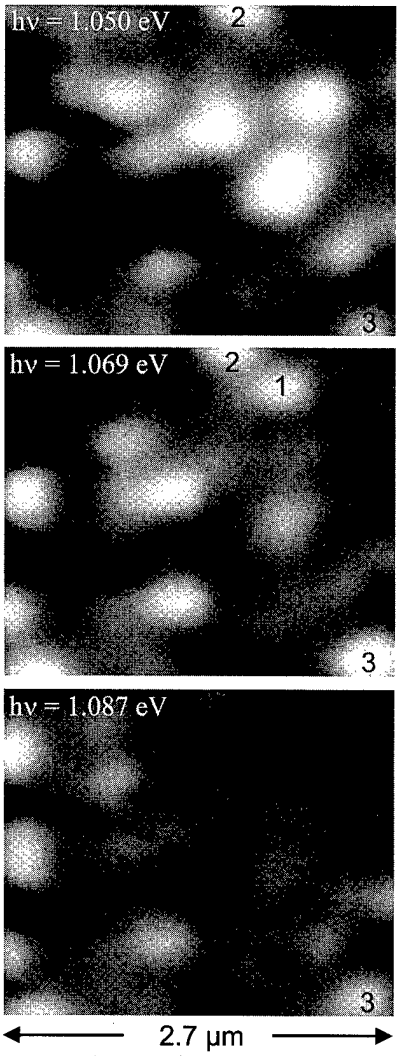


FIG. 2. Spectrally resolved PL images of $In_{0.4}Ga_{0.6}As$ quantum dots taken at room temperature at the same sample area using different detection energies. Three quantum dots are marked with numbers, visible (1) in only one image, (2) in two images, and (3) in all three images.

used for coarse approach. The distance of the SNOM tip from the sample surface is controlled by nonoptical shear-force detection using a quartz tuning fork with the tip glued to one side,[14,15] resulting in a vertical resolution of about 0.5 nm and a lateral resolution of about 15 nm in the topography. The thermal drift can be as low as 0.1 μm per hour at sufficiently stable temperatures.

All SNOM measurements reported in this work were performed in internal-reflection geometry using a 2×1 fiber coupler for feeding the excitation light through the tip onto the sample and collecting the PL light by the tip for detection.[14] For optical excitation, a He–Ne laser with $h\nu = 1.96$ eV was used. The resulting PL signal leaving the 2×1 coupler was analyzed using a monochromator with a spectral resolution around 1 meV and finally detected by a Peltier-cooled InGaAs photodiode with a noise level of about 8×10^{-17} W/$\sqrt{\text{Hz}}$ and a bandwidth of a few Hz. For spectrally integrated PL imaging, the monochromator was replaced by a low-pass filter. This measure prevents the detection both of the PL radiation from recombination outside the dots and of the laser light, which is strongly reflected in the 2×1 coupler or in the tapered fiber tip.

B. Tip preparation

The SNOM tips used here were prepared from multimode fibers with a 50 μm core by chemical etching and subsequent metal coating.[13] Etching was performed in 48% hydrofluoric acid covered with a layer of baby oil.[16] This self-terminating process results in a total opening angle of the tip apex of about 20°. In order to prevent radiation transmission through the side walls of the tip, they were subsequently coated with a metal. This process step was performed in a high-vacuum chamber at a pressure of 10^{-6} mbar by evaporation of 100–200 nm thick Ag or Al films under an angle of about 75° between the evaporator and the fiber axis while rotating the tip. Although we did not observe significant differences between Al- and Ag-coated tips in resolution and transmission, we preferred Ag-coated tips because of their better stability in ambient conditions, allowing to use them for several days. The apex of the tips was often opaque, so that an opening had to be prepared by gently polishing the tip apex at the sample surface using the lateral vibration of the tip with respect to the sample. The opening of the aperture could be monitored easily by the onset of the PL signal.

C. Sample preparation

The $In_{0.4}Ga_{0.6}As$ quantum-dot sample was prepared using MOCVD by growing 0.94 nm $In_{0.4}Ga_{0.6}As$ on an undoped GaAs(001) substrate and subsequent capping with a total of

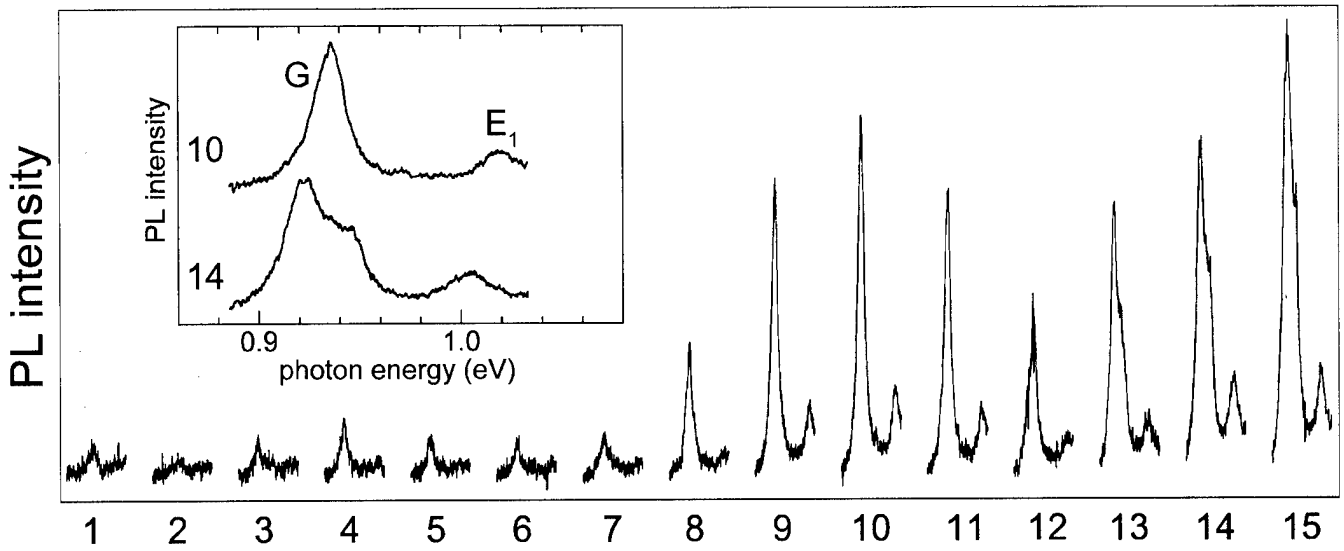


FIG. 3. Series of 15 PL spectra taken at room temperature during one line scan with 180 nm spacing between neighboring spectra. On the left side, up to spectrum 7, almost no PL can be detected since there are no quantum dots close to the tip position. On the right side, the tip passes first one quantum dot (spectra 8–11) and then a group of at least two quantum dots (spectra 13–15). The inset shows spectra 10 and 14 in more detail, with G and E_1 representing the ground-state and first excited-state transition, respectively.

20 nm of GaAs and AlGaAs layers for protection.[6,7] They have already been characterized by transmission electron microscopy (TEM), yielding an area dot density of a few quantum dots per μm^2, and by conventional spatially averaging PL at room temperature.[6] In this latter experiment, the so-called ground-state transition was observed at $h\nu = 0.93$ eV, as well as the first and second excited-state transitions at 1.02 and 1.10 eV. These PL lines were inhomogeneously broadened by at least 40 meV.

III. RESULTS AND DISCUSSION

A. Photoluminescence imaging

In order to demonstrate the ability of our SNOM instrument to investigate the emission from single quantum dots, we first present PL imaging results. Figure 1(a) shows a spatial map of the spectrally integrated PL intensity of the quantum dots taken at room temperature. Most peaks are characterized by an almost circular shape and similar emission intensities, representing the signal from individual dots. In addition, some peaks are broader and brighter and have a more asymmetric shape, so that they can be assigned to small groups of dots. A lateral resolution of about 300 nm is obtained, as revealed from the cross section displayed in Fig. 1(b). The observed areal density of the emission maxima is in the order of $2/\mu m^2$, in reasonable agreement with the results from TEM experiments.[6]

At tip positions between the dots, the signal reduces almost to zero. This behavior demonstrates that there is only negligible background signal from neighboring dots, which is in strong contrast to our previous data obtained using uncoated fiber tips.[14,17] This result clearly demonstrates that etched, metal-coated fiber tips, operated in internal-reflection geometry, allow to resolve the pure PL signal of individual quantum dots with sufficient intensity.

A first analysis of the spectral behavior of the quantum dots is presented in Fig. 2, which shows spectrally resolved PL images taken at room temperature for three different detection energies separated by about 0.018 eV. Some dots are only observed at one single photon energy, for example, the dot marked by 1, but others are found at two or even at all three different energies, like the ones marked by 2 and 3, respectively. This observation already indicates a finite spectral width of the emission from a single dot in the meV range at room temperature.

B. Photoluminescence spectroscopy

In order to analyze this behavior in more detail, PL spectra of single dots were recorded at different positions. Figure 3 shows a series of PL spectra taken at room temperature along one scan line with 180 nm lateral spacing between neighboring spectra. They are characterized by the ground-state transition at 0.935 eV (G) and the first excited-state transition at 1.02 eV (E_1), as shown in the detail in the inset. On the left side of the series, up to spectrum 7, almost no PL signal is observed, which is due to the absence of quantum dots in this area. This behavior again demonstrates that the use of metal-coated fiber tips efficiently eliminates the background signal from neighboring dots by more than an order of magnitude. When going further to the right, the tip passes one single quantum dot, and the PL signal is first increasing to its maximum at spectrum 10, followed by an intensity decrease. On the far right side of the series, around spectrum 14, the PL signal is increasing again because the tip approaches a group of at least two quantum dots, resulting in a shoulder in the peaks, as shown in the inset of Fig. 3.

From an analysis of the lateral variation in the peak intensity observed in Fig. 3, a lateral resolution for the ground-state emission of about 500 nm is obtained, which is due to a

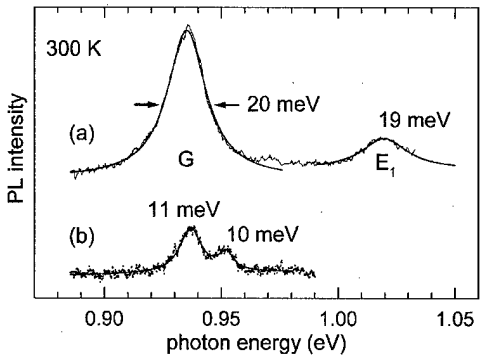


FIG. 4. PL spectra of individual quantum dots, taken at room temperature (data points), together with Lorentzian fit curves (solid lines). (a) Spectrum from one dot, showing the ground-state (G) as well as the excited-state (E_1) transition, (b) ground-state transition from two neighboring dots. The different linewidths (FWHM) are indicated, varying between 10 and 20 meV from dot to dot.

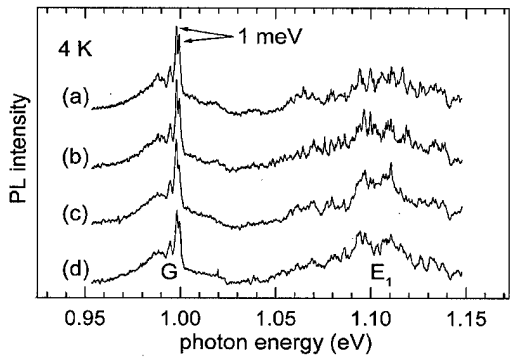


FIG. 5. Four PL spectra taken at a temperature of 4 K at neighboring positions with a lateral spacing of several 10 nm, showing linewidths limited by the experimental resolution of 1 meV. The spectra show an almost identical ground-state emission, but vary considerably in the excited-state emission.

larger opening in the metal layer of the tip used here as compared to the one used in Fig. 1. For the first excited-state transition, in contrast, the resolution is improved to about 400 nm. This behavior is typical for the higher excited-state transitions, which show a quadratic or even higher-power response to the excitation intensity as compared with the linear increase for the ground-state transition.[18] When the tip is too far away from the quantum dot, the excitation intensity is not sufficient to saturate the ground-state transition, resulting in a much weaker emission from the excited-state transition. This behavior is clearly observed when comparing spectra 8 and 10.

The negligible background signal from neighboring dots now enables us to investigate the pure signal of a single quantum dot. In Fig. 4 representative PL spectra taken from selected quantum dots at room temperature are displayed. Figure 4(a) shows the emission from a single dot, including the ground-state and first excited-state transitions. The spectrum shown in Fig. 4(b) displays the ground-state transitions from two neighboring dots.

All spectra are characterized by relatively broad symmetric lines. The spectral shape can well be described by a Lorentzian line, as revealed from a least-squares fit analysis shown in Fig. 4. Fits with a Gaussian line regularly resulted in a worse agreement with the experimental line shape. The analysis of the spectra from several dots results in Lorentzian linewidths [full width at half maximum (FWHM)] scattering between 10 and 20 meV at room temperature, and the ground-state and first excited-state transitions show similar linewidths. The observation of a finite linewidth amounting to several meV is a surprising result, considering the extremely narrow emission lines expected for these zero-dimensional quantum objects on the basis of the optical-transmission rates.[1–3,5,9,10]

Furthermore, the linewidth is found to vary strongly with temperature. Figure 5 shows four spectra taken at 4 K at nominally the same tip position. Because of thermal-drift effects and the long acquisition time for one spectrum of several minutes, the spectra probably stem from slightly different positions with several 10 nm distance. The energy of the ground-state transition is now shifted to about 1 eV because of the thermal variation of the band gaps.[19] It consists mainly of two neighboring lines with widths determined by the monochromator resolution of about 1 meV, so that the actual linewidth is expected to be much lower. They can be assigned to the emission of one or two quantum dots. The origin of the broad background of the ground-state transition is still unknown; however, it can be related to the weak contribution from many dots, probably in the vicinity of the tip apex.

The excited-state transition around 1.10 eV consists of several lines, which are also characterized by linewidths of about 1 meV. They show a much stronger variation with lateral position, again demonstrating the higher lateral resolution at higher-excited transitions. The observation of an extended fine structure indicates a complex term scheme for the excited states in the quantum dots. In addition, neighboring dots may contribute as well.

For a better understanding of the temperature dependence of the spectral properties, we performed an experiment with the tip fixed mechanically on the sample by gently touching the surface. In this way, we could eliminate thermal drift completely and detect the PL emission from the same sample region in the temperature range from 5 to 300 K, as shown in Fig. 6. For this experiment, an uncoated fiber tip was used, so that a stronger background signal is contributing, since the emission of surrounding quantum dots is also collected by the tip.[14,17] In particular for the ground-state transition, a larger group of quantum dots is contributing to a rather unstructured background, while the signal from the excited-state transition is more selective and clearly shows the emission from mainly three quantum dots. Therefore, the variation of the linewidth with temperature can best be monitored using the excited-state transition, but a similar trend is also observed for the ground-state transition.

The linewidth gradually decreases from about 20 meV at room temperature to about 4 meV at 5 K. In this experiment, the spectral resolution was only about 4 meV, indicating a much lower value for the actual linewidth at low temperatures. As has been mentioned before, the actual linewidth of

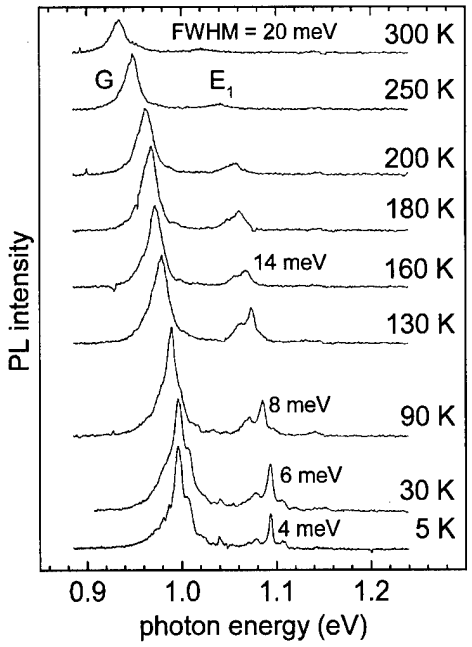


FIG. 6. Temperature dependence of the PL spectra mainly dominated by a group of about three quantum dots, taken with an etched uncoated tip positioned at a fixed location of the surface. The linewidths (FWHM) is decreasing from about 20 meV at 300 K down to the monochromator resolution of about 4 meV at 5 K.

the quantum-dot emission in this temperature region is even less than 1 meV (see Fig. 5). Furthermore, it is observed that the PL intensity is decreasing by less than one order of magnitude when going to room temperature, and the transition energies show the expected redshift by about 60 meV.[19]

C. Discussion of the linewidth

In the following, a first attempt is made to explain the temperature-dependent linewidth and the Lorentzian line shape. Such a behavior is characteristic for a thermal excitation of carriers in the quantum dots, reducing the lifetime in particular at higher temperatures, in this way broadening the spectral lines which are then described by a Lorentzian line shape. In this case, the thermal excitation probability and

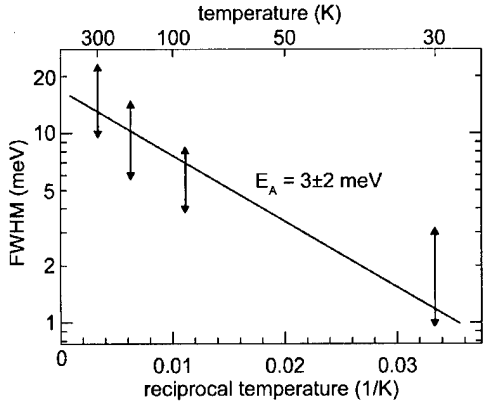


FIG. 7. Arrhenius plot of the linewidth observed in the PL spectra as a function of temperature.

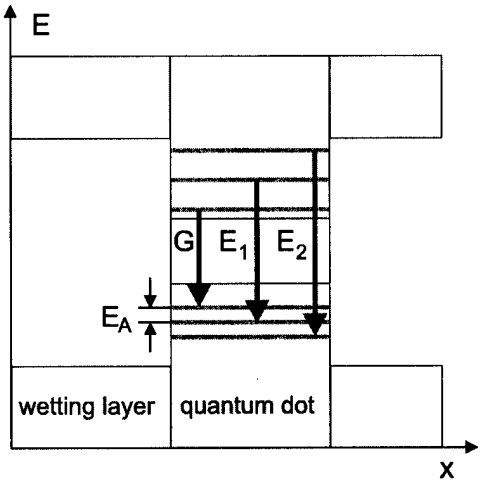


FIG. 8. Schematic model for the possible electronic structure of the $In_{0.4}Ga_{0.6}As$ quantum dots. The ground-state (G) as well as the first and second excited-state (E_1, E_2) transitions are indicated. The binding energies of the hole states are separated by only a few meV, resulting in a thermal activation at higher temperatures.

therewith the linewidth can be described approximately by $\Gamma = \Gamma_0 \exp(-E_A/k_BT)$, where E_A is the activation energy and Γ_0 a prefactor mainly determined by the phonon energies. An Arrhenius plot of the linewidth is shown in Fig. 7. The rather large error bars result from the considerable scatter of the observed linewidths from different dots at a given temperature as well as from a strong influence of the experimental resolution on the accuracy of the linewidth, in particular at lower temperatures. Therefore, the activation energy can only be estimated with a rather large uncertainty to 3 ± 2 meV.

Terms schemes like the one shown in Fig. 8 have been proposed in recent theoretical calculations on the electronic structure of InAs dots.[20,21] In these studies, both electrons and holes were found to have several states in the quantum dot, with a larger energy spacing for the electron states because of their much lower effective mass. In particular larger quantum dots with a rather low ground-state transition energy, like the ones investigated in the present study, are expected to have a variety of states with small energy differences. On the other hand, dipole selection rules or strongly varying oscillator strengths prevent many of the transitions or considerably reduce their intensities. It should be noted here that these more sophisticated calculations are in contrast to previous assumptions of only one electron state in the dot,[22] which we recently stressed to explain our data.[13]

Since the energy spacings between the hole states in sufficiently large dots can be in the order of a few meV,[20,21] the thermal activation of holes can lead to a line broadening, as observed in the present work. The observed scatter in linewidth for different dots can be attributed to different dot sizes and shapes, which in turn have an influence on their electronic structure, and in particular on the spacing between the energy states.

It should be noted that the excited-state transitions were predicted to give rise to a variety of lines,[21] in agreement

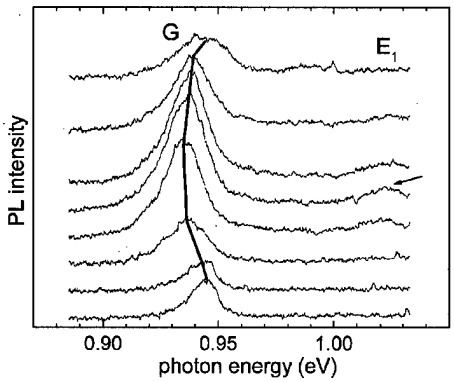


FIG. 9. Spectra taken from a line scan with 200 nm spacing. For large intensities, an energy shift as well as a line broadening is observed, indicating the excitation of biexcitons within the quantum dots.

with the spectra in Fig. 5. The observed variation of these spectra with tip position can well be due to the strong polarization dependence of the transmission probabilities, which was also derived in the calculations.[20,21] Furthermore, the thermal variation of the spectral intensity observed in Fig. 6 can be explained within this model: At thermal energies exceeding the activation energy, a significant fraction of the holes is excited to the other hole states, resulting in a strongly reduced optical transition probability.

Although the present discussion may explain both the observed line shape and the thermal linewidth variation, a conclusive model of the electronic structure of $In_{0.4}Ga_{0.6}As$ quantum dots is not advisable on the basis of the present data. More detailed theoretical investigations, e.g., including the Coulomb interaction of the excitons, are thus required.

D. Intensity-dependent shifts

Figure 9 shows a series of PL spectra taken at room temperature during a line scan with 200 nm spacing between neighboring spectra, while passing a single quantum dot. In addition to the common intensity variation of the PL emission, a shift of the peak emission by about 8 meV to lower transition energies is observed at higher excitation intensities. The maximum shift is obtained when the first excited-state transition becomes visible (see arrow).

This behavior can be explained by the excitation of biexcitons, which consist of two electron-hole pairs in a quantum dot and are therefore mainly excited at a higher excitation intensity. At even higher intensities, triple and other multiple excitons can form, two of them contributing to the ground-state transition, the following exciton pair to the first excited-state transition, and so forth.[18] Biexcitons are expected to show a redshift of a few meV with respect to the single-exciton line in case of an attractive interaction of the two excitons.[2]

In our case of broad Lorentzian lines, the excitonic and the biexcitonic transitions cannot be separated, and their superposition results in an additional broadening as well as in an energy shift with increasing excitation intensities. The assignment of the observed redshift in Fig. 9 to biexcitons is also in agreement with the occurrence of the excited-state transition at comparable excitation energies.

It should be noted that the ground-state emission observed in Fig. 4(a) should also consist of two lines, since the excited-state transition is already observed. Two spectral lines could not be resolved in this spectrum because of the large Lorentzian linewidth, but fits with two Lorentzians separated by 8 meV also resulted in reasonable agreement with the experimental line shape. Also the complex spectra observed at low temperatures in Fig. 5 can now be discussed in terms of multiple excitons. Both the ground-state emission consisting mainly of two lines and the multiline excited-state emission can, at least partially, be caused by excitonic interactions.

IV. CONCLUSION

In this work, we have investigated the PL of $In_{0.4}Ga_{0.6}As$ quantum dots with SNOM. Using etched metal-coated fiber tips we were able to investigate the pure PL of single quantum dots, even for optical transition energies as low as 0.9 eV. In this way, the spectral shape could be studied in detail, resulting in Lorentzian lines with different linewidths between 10 and 20 meV at room temperature, reducing to less than 1 meV at 4 K. The expected sharp emission lines with widths in the μeV region, which are independent of temperature, could not be observed. It is assumed that the thermal broadening is due to a lifetime effect, related to low excitation energies of the carriers in the dots. The spectra also show a redshift of about 8 meV with increasing excitation intensities, which can be explained by the excitation of biexcitons in the quantum dots.

The $In_{0.4}Ga_{0.6}As$ quantum dots investigated in this work show a remarkable linewidth behavior, which has not been observed up to now. In our initial attempts to explain the data, thermal excitation of holes in the dot is assumed. For a conclusive model, however, more information about the system is required, so that both a more detailed theoretical treatment and more systematic experimental studies have still to be performed. In order to investigate the line shape variation and the possible biexcitonic shifts in more detail, further SNOM experiments at low temperatures are currently in preparation.

ACKNOWLEDGMENTS

W. Busse, C. K. Kim, T. Kalka, and A. Bauer are acknowledged for assistance during the experiment, and M. Grundmann and R. Heitz for discussion of the results. Part of this work was supported by the Sonderforschungsbereich 296 of the Deutsche Forschungsgemeinschaft.

[1]M. J. Kelly, *Low-Dimensional Semiconductors* (Clarendon Press, Oxford, 1995).
[2]U. Woggon, *Optical Properties of Semiconductor Quantum Dots* (Springer, Berlin, 1996).
[3]D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum Dot Heterostructures* (Wiley, Chichester, 1999).

[4]F. Heinrichsdorff, A. Krost, M. Grundmann, D. Bimberg, A. Kosogov, and P. Werner, Appl. Phys. Lett. **68**, 3284 (1996).
[5]M. Grundmann *et al.*, Phys. Rev. Lett. **74**, 4043 (1995).
[6]F. Heinrichsdorff, A. Krost, M. Grundmann, D. Bimberg, A. Kosogov, P. Werner, F. Bertram, and J. Christen, in *The Physics of Semiconductors*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), p. 1321.
[7]F. Heinrichsdorff, M.-H. Mao, N. Kirstaedter, A. Krost, D. Bimberg, A. O. Kosogov, and P. Werner, Appl. Phys. Lett. **71**, 22 (1997).
[8]S. F. Alvarado, Ph. Renaud, D. L. Abraham, Ch. Schönenberger, D. J. Arendt, and H. P. Meier, J. Vac. Sci. Technol. B **9**, 409 (1991).
[9]M. Markmann, A. Zrenner, G. Böhm, and G. Abstreiter, Phys. Status Solidi A **164**, 301 (1997).
[10]Y. Nagamune, H. Watabe, M. Nishioka, and Y. Arakawa, Appl. Phys. Lett. **67**, 3257 (1995).
[11]M. A. Paesler and P. J. Moyer, *Near-Field Optics: Theory, Instrumentation, and Applications* (Wiley–Interscience, New York, 1996).
[12]R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy* (Cambridge University Press, Cambridge, 1994).
[13]I. Manke, J. Lorbacher, J. L. Spithoven, F. Heinrichsdorff, and M. Dähne-Prietsch, Surface and Interface Analysis **27**, 491 (1999).
[14]I. Manke, D. Pahlke, J. Lorbacher, W. Busse, T. Kalka, W. Richter, and M. Dähne-Prietsch, Appl. Phys. A: Mater. Sci. Process. **66**, S381 (1998).
[15]K. Karrai and R. D. Grober, Appl. Phys. Lett. **66**, 1842 (1995).
[16]D. R. Turner, U.S. Patent 4 469 554 (1984).
[17]D. Pahlke, I. Manke, F. Heinrichsdorff, M. Dähne-Prietsch, and W. Richter, Appl. Surf. Sci. **123/124**, 400 (1998).
[18]M. Grundmann and D. Bimberg, Phys. Rev. B **55**, 9740 (1997).
[19]Y. P. Varshni, Physica (Utrecht) **34**, 149 (1967).
[20]H. Jiang and J. Singh, Appl. Phys. Lett. **71**, 3239 (1997).
[21]O. Stier, M. Grundmann, and D. Bimberg, Phys. Rev. B **59**, 5688 (1999).
[22]M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **52**, 11969 (1995).

Atomic structure of stacked InAs quantum dots grown by metal-organic chemical vapor deposition

O. Flebbe, H. Eisele,[a] T. Kalka, F. Heinrichsdorff, A. Krost,[b] D. Bimberg, and M. Dähne-Prietsch
Institut für Festkörperphysik, Technische Universität Berlin, D-10623 Berlin, Germany

(Received 19 January 1999; accepted 26 May 1999)

We present a detailed cross-sectional scanning-tunneling microscopy investigation of threefold stacked InAs quantum dots in a GaAs matrix grown by metal-organic chemical vapor deposition. The dots are vertically aligned and show a layer-dependent size. Detailed images with atomic resolution indicate that the dots consist of pure InAs with a shape well described by prisms with {110} and additional {111} side faces as well as a (001) top face. The wetting layer is found to be inhomogeneous, while no In diffusion into the overgrowing GaAs layer is observed. These findings demonstrate that growth in the present sample is more related to kinetics than to equilibrium conditions. Furthermore it is demonstrated that the image contrast is strongly influenced by a cleavage-induced outward relaxation of the strained quantum dots. © *1999 American Vacuum Society.* [S0734-211X(99)08704-1]

I. INTRODUCTION

During the last few years quantum dots have been investigated intensively, mainly because of their peculiar electronic properties characterized by a δ-function-like electronic density of states corresponding to atom-like energy levels.[1] Of particular interest are InAs or InGaAs quantum dots embedded in a GaAs matrix, since such dots can be grown easily in a self-organized way using either molecular-beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD).[2–4] Recently, room-temperature lasing has been observed for such systems,[4,5] with wavelengths between 1.1 and 1.3 μm, which are of high interest for applications in telecommunications.

For InAs growth on GaAs(001) surfaces, the large lattice mismatch of about 7% leads to a Stranski–Krastanov growth mode, characterized by a wetting layer of about 1.5 monolayers thickness, followed by the self-organized formation of strained quantum dots.[6] Subsequent overgrowth of the first dot layer with a thin GaAs layer results in a lateral strain variation on the newly formed GaAs surface, a property that can be used for the growth of vertically stacked dots: Further deposition of InAs will now take place predominantly on top of the buried first-layer dots, since the local lattice constant of the GaAs surface is expanded at these positions, being energetically more favorable for InAs nucleation. In this way, self-assembled stacks of vertically aligned quantum dots can be grown.[1,7,8] Such structures are of high interest for laser devices because of their higher dot density compared with the single-layer case.[4,5]

The dot structure, notably their shapes, sizes, and area densities, depends on a variety of growth parameters like temperature, growth velocity, partial pressure of the different precursor materials, growth-interruption timing, and finally on the amount of the deposited material. Furthermore the growth techniques like MBE or MOCVD will have an influence.

The size and the shape of the quantum dots are important factors for their electronic and optical properties.[9] In recent years, mainly three experimental techniques were used for structural investigations of quantum dots: (1) transmission electron microscopy (TEM) of buried dots[10–12] only provides integrated information on columns consisting of about 100 atoms; it is furthermore strongly sensitive to the strain field of the sample in the dot region. (2) X-ray diffraction was recently applied for the first time to quantum dots,[13] but has to be further developed since it suffers from small cross sections and thus extended measuring times. (3) Top-view atomic force microscopy or scanning-tunneling microscopy (STM) can be used to determine the shape and size of unburied dots directly.[14–18] In order to conserve the surface structure, an *in situ* sample transfer to the microscope has to be managed, requiring further technical development, in particular in the case of MOCVD growth. In addition, structural changes may occur upon overgrowth of the dot layer; in particular stacked dots are not accessible.

A different and more promising method for the investigation of buried quantum dots is cross-sectional scanning-tunneling microscopy (XSTM). In such an experiment the sample, which has been grown, e.g., along the [001] direction, is cleaved at the perpendicular (110) surface. In this way, a cross-sectional (110) surface through the dot layers is obtained, which can be imaged by STM with atomic resolution.[19–21] It should be noted that such an experiment requires an additional lateral-positioning mechanism for matching the scan range of the tip with the layer of interest, which is furthermore located very close to the edge of the cleavage surface.

Recently, XSTM studies of stacked InAs dot structures in a GaAs matrix grown by MBE were presented.[22–25] In these studies, rather uniform dots with disk-like shapes were ob-

[a]Electronic mail: ak@physik.tu-berlin.de
[b]Permanent address: Otto-von-Guericke-Universität Magdeburg, D-39016 Magdeburg, Germany.

served, showing an almost perfect vertical stacking. In addition, In atoms were found to diffuse out of the wetting layer during GaAs overgrowth.[22,24] Such a segregation also resulted in a roughening of the InAs–GaAs interface.[23]

In a recent letter, we presented the first XSTM results of stacked InAs dots grown by MOCVD with atomic resolution.[26] In this article, the results are presented in more detail. Imaging a variety of threefold dot stacks enables us to determine the atomic structure of the dots as well as the structure of the wetting layer. In particular it is found that stoichiometrically pure InAs dots were grown with prismatic shapes, which are different from those observed for the MBE-grown samples.[22–25] Inhomogeneous wetting layers are formed without In diffusion into the GaAs overlayer, the latter again in contrast to the findings for the MBE-grown dots.[22,24] The buckled GaAs overgrowth indicates that the growth conditions are far from thermal equilibrium, also resulting in a less perfect vertical stacking. As already observed in our initial XSTM experiments of fivefold stacked InAs dots also grown by MOCVD,[27] the image contrast mainly results from a cleavage-induced outward relaxation of the strained quantum dots. For a quantitative analysis of the XSTM images, simulations of this relaxation by elasticity theory[27] were performed.

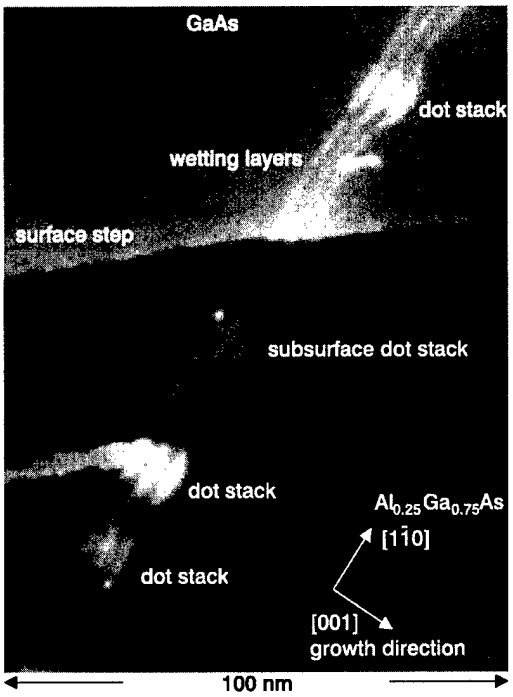


FIG. 1. Overview XSTM image of the dot-stack layer and the surrounding buffer layers, taken at $V_S = +2.6$ V and $I_T = 100$ pA.

II. EXPERIMENT

A. MOCVD growth conditions

The sample was grown by MOCVD at a deposition rate of about 0.18 nm/s. Similar growth parameters were used for this sample as for the one for which room-temperature lasing was observed.[4,5] The dot region studied in this work consists of three InAs dot layers in a GaAs matrix, surrounded by GaAs buffer layers and $Al_{0.25}Ga_{0.75}As$ waveguide layers.

The InAs dot layers were grown at a temperature of 485 °C, using the following procedure: In order to obtain a smooth GaAs surface, the growth was first interrupted under AsH_3 pressure for 10 s. Then a 0.47 nm (corresponding to 1.7 monolayers) thick InAs layer was grown. For dot formation, the growth was then interrupted for 4 s, in this case without AsH_3 pressure, so that only the H_2 atmosphere was present. Subsequently the dot layer was covered by a 3.5 nm thick GaAs overlayer. Then the growth was interrupted again for 2 s under AsH_3 pressure, followed by the next InAs deposition cycle as above. In top-view TEM images taken of similar samples, a density of the dot stacks of $(4 \pm 2) \times 10^{10}/\text{cm}^2$ was observed.

B. XSTM experiment

For the XSTM experiments a home-built system was used, operating at room temperature. The tips were prepared from an annealed tungsten wire by electrochemical etching. Subsequently they were sharpened by Ar^+ ion etching inside a separate high-vacuum chamber.[28] These tips were transferred through air to the STM chamber and directly used for imaging without further treatment.

For sample preparation, a small rectangular piece was cut out of the wafer. In order to define the cleavage edge, a small part of the (001) growth surface was marked by a short notch. After mounting the sample to the sample holder, it was electrically connected with indium. Then the sample was transferred into the ultrahigh vacuum (UHV) chamber, which operates at a base pressure below 1×10^{-8} Pa. Sample cleavage was performed inside the UHV. It should be noted that the rotational orientation of the sample is unknown in the present study, so that we cannot determine whether the (110) or the ($\bar{1}$10) surface was formed by cleavage. In the following, the surface imaged by STM is thus always labeled (110).

III. RESULTS AND DISCUSSION

A. Growth behavior of stacked-dot structures

In order to give an overview on the growth behavior of stacked quantum dots, an XSTM image of a larger area around the dot layers is presented in Fig. 1. The [001] growth direction is indicated, which will be the same for all XSTM images shown in this work.

From the left bottom to the right top side of Fig. 1, the cross sections of the three wetting layers are visible in the form of noninterrupted bright lines. In the case of XSTM images of quantum dots, it has to be taken into account that the cleavage occurs at arbitrary cross sections of the dot layer. In this way, different cross-sectional planes of the dots are obtained, with the remaining parts of the dots having different depths below the cleavage surface. Figure 1 shows three dot stacks in the form of bright triple protrusions, which were cut by the cleavage. Furthermore, one subsurface

dot stack is identified, located underneath the cleavage surface, which can be identified by the surrounding strain field (as will be explained in Sec. III B 2).

Figure 1 clearly demonstrates the Stranski–Krastanov growth mode, consisting of a wetting layer (leftmost line) and the formation of quantum dots on top (towards the right in the image). In the following layers, the stacking of the dots is observed, due to a preferential growth of the next-layer dots on top of dots already buried by the GaAs overlayer. In addition, the size of the dots is increasing from one layer to the next one, as can be seen, e.g., in the case of the stack at the top of the image.

Furthermore, the second and third wetting layers are not plane any more, but have a wavy appearance, in particular close to the dots. While the separation far away from the dot stacks amounts to about 4.0 nm, in agreement with the nominal distance outlined from the growth parameters, it is reduced in close vicinity of the dot stacks to values around 3.3 nm. On the other hand, the base lines of the dots within a dot stack are separated by a larger distance of about 4.5 nm. This observation indicates an inhomogeneous overgrowth by the GaAs layer, which is related to incomplete lateral segregation during MOCVD growth.

The dot layers are surrounded on both sides by the GaAs buffer layers and the $Al_{0.25}Ga_{0.75}As$ waveguide layers. In the center of the image, a onefold surface step is crossing the dot layer, and on the left side a finger-like terrace is located. These features are caused by cleavage imperfections. In addition, some bright spots are found on the cleavage surface, in particular at the step edges, which occasionally move during scanning and can thus be attributed to adsorbate molecules.

About 98% of the observed dot stacks are well aligned. Very few images show an unusual stacking, as shown in Fig. 2. The dot stack in the upper center of Fig. 2(a) is mostly cut away by the cleavage, so that only a weak contrast is observed. In addition, it is vertically not well aligned and the third dot is cut away almost completely, so that only one edge and the wetting layer remain at the cleavage surface. In Fig. 2(b) two dot stacks are imaged, which are located very close to each other. Here, the third dot of the upper stack is missing completely, while the third one of the lower stack is much larger. Figure 2(c) again shows a dot stack where the vertical alignment is not good.

Both the poor vertical alignment and the missing dot in the third layer are probably due to a neighboring dot stack, modifying the strain field in such a way that it is not symmetric any more around the center of the first-layer dot. This asymmetry then leads to a lateral offset of the energetically optimum position for the next dot and therefore to nonvertical alignment.[8] In the extreme case of a closely neighboring stack, this effect can be so strong that only one site is preferred for dot formation, as shown for the third layer in Fig. 2(b). Such a coalescence of neighboring dot stacks leads to a decreasing area density in subsequent dot layers.

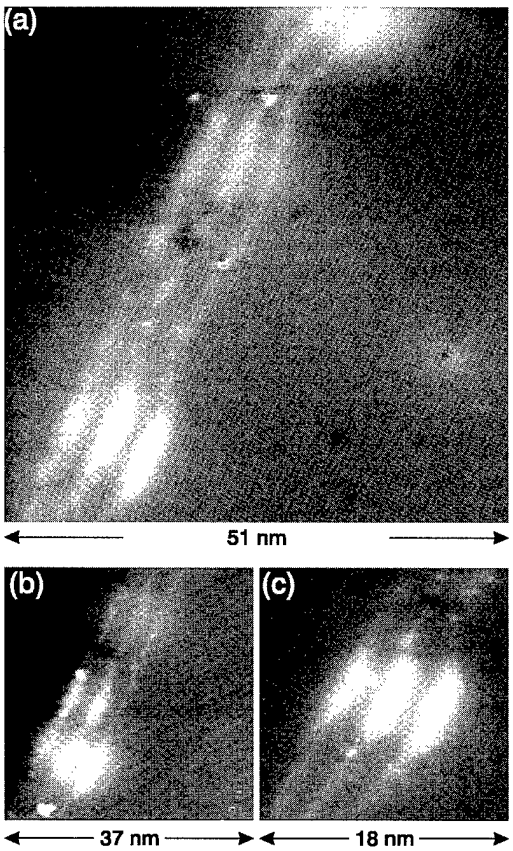


FIG. 2. XSTM images of stacking imperfections: (a) taken at $V_S = +1.5$ V and $I_T = 70$ pA, (b) at $V_S = -3.0$ V and $I_T = 85$ pA, and (c) at $V_S = +1.8$ V and $I_T = 65$ pA.

B. Contrast mechanisms

Before we discuss the atomic structure of the quantum dots and the wetting layer, the possible contrast mechanisms active in XSTM experiments at strained layers have to be reviewed. STM only monitors the height of the tip, depending on both electronic and structural properties of the sample, superposing each other in XSTM images.

1. Electronic contrast

The different electronic structure, e.g., of InAs and GaAs leads to a chemical contrast. This is due to the number of states available for tunneling, which depends on the tunneling voltage and the energetic position of the electronic states, e.g., the band edges in the case of infinite crystals. In particular, this effect is pronounced at low tunneling voltages. In this way, the tip usually retracts above surface regions with a smaller band gap in order to keep a constant tunneling current, in this way yielding brighter regions in the image. Such a behavior was observed for broad quantum wells,[19] as well as for single atoms[29–31] and for depletion layers where band bending contributes.[32] Furthermore, in samples containing potential wells, like the wetting layer or the quantum dots in the present case, the confined electronic states with their wave functions tailing into the matrix material will play a role.

In addition, the image depends on the sample voltage and, in particular, on its polarity.[33,34] In this way atom-selective imaging can be performed with the STM: At positive sample bias, where the tip electrons tunnel into the empty sample states, the dangling bonds of the cation sublattice are imaged, i.e., in our case those of the In or Ga surface atoms. Images taken at negative sample bias, in contrast, are sensitive to the filled dangling bonds from the As atoms at the surface.

2. Structural contrast

Several properties contribute to the surface topography. Inside the relaxed (110) surface unit cell of GaAs or InAs, only every second monolayer along the [001] direction is visible in the form of a zigzag chain.[34] Additional larger-scale topographic features of the cleavage surface, such as surface steps, terraces, or adsorbate molecules, will be more or less imaged in a realistic way, but these effects are not interesting for the present work, at least for ideally cleaved clean (110) surfaces.

An additional topographic effect is due to strain in the structure. On the cleavage surface the strained dots relax outwards, forming smooth topographic protrusions on the cleavage surface, as we demonstrated recently.[27] The magnitude of this relaxation depends on the amount and the atomic arrangement of InAs inside GaAs underneath the cleavage surface, and values of about 0.3 nm can be obtained for InAs quantum dots. Since this effect is of pure topographic nature, it is independent of the tunneling conditions. Even single impurity atoms may relax outwards or inwards from the cleavage surface, caused by the bond lengths to the neighboring atoms.[29,31]

In order to get an impression for the size and the shape of the dots, we have simulated the strain relaxation at the cleavage surface for several different dot shapes by a finite element method based on continuum mechanical theory, as described in Ref. 27. A detailed report on these simulations will be given elsewhere.[35] Using these simulations, we are able to separate topographic effects due to structural strain relaxation from electronic effects in the image contrast.

C. Structural properties of the dots

In this section, we will analyze the atomic structure of single dots. First of all, it should be noted that we never observed any dislocations or other growth-related imperfections of the zincblende lattice in our XSTM images. Such a perfect growth is also expected, since room-temperature lasing was observed at similarly grown samples.[4,5]

1. Stoichiometry of the dots

Figure 3(a) gives a close view of a typical dot stack. The superposed linear-grating structure is due to the III–V zigzag chains, where the filled states and therefore the dangling bonds of the As atoms are imaged because of the negative sample bias V_S. The five large white bumps are assigned to adsorbate molecules. Some minor imperfections are observed in certain scan lines due to an unstable tip, in particular in the upper part of the image.

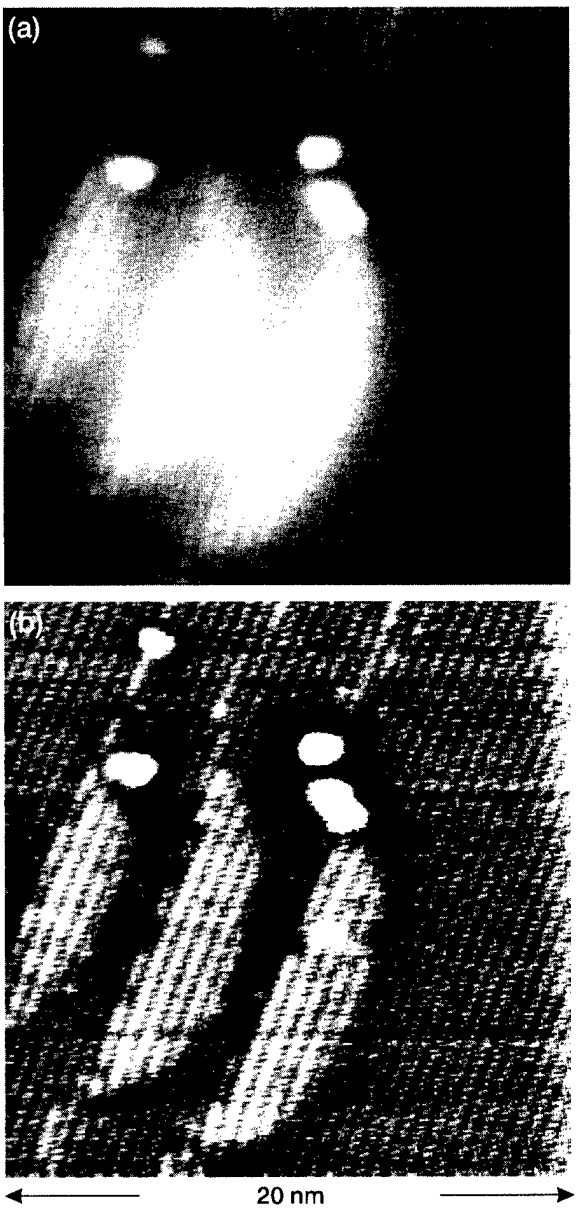


FIG. 3. (a) Filled-state image of a threefold dot stack, taken at $V_S = -1.7$ V and $I_T = 100$ pA. (b) The same image after high-pass filtering in order to outline the chemical contrast.

Usually it is not easy to decide from such a STM image where the interface line separating the dot from the surrounding matrix is located. This is caused by the superposition of the chemically induced electronic contrast with the smooth structural height variation due to strain relaxation at the cleavage surface. In order to separate both effects, a high-pass filtered image is shown in Fig. 3(b). In this way the sudden tip-height variation resulting from the chemical difference is much better visible, and the interface line can be outlined clearly.

The base lines of these dots do not consist of homogeneous InAs, which is probably due to a certain roughness of the GaAs starting surface. On the other hand, no indication is observed for GaAs inside the dots, since there are no sudden height variations along the zigzag chains. Thus the formation

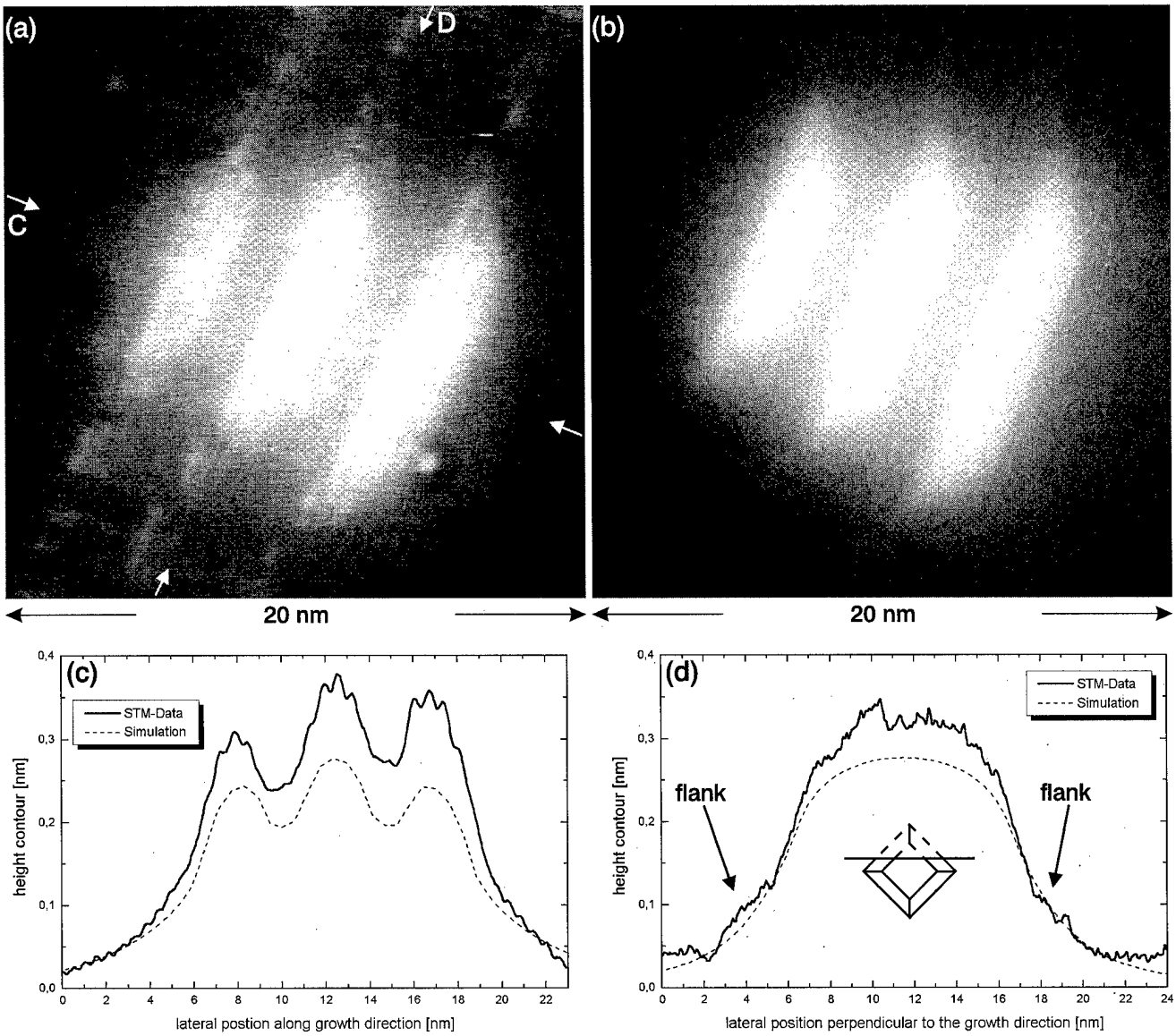


FIG. 4. (a) Filled-state XSTM image of a typical dot stack, taken at $V_S = -2.4$ V and $I_T = 85$ pA, (b) simulated image of the strain relaxation for a dot structure determined from (a). Height contours along the growth direction through the center of the dot stack, as indicated by C in (a), are plotted in (c), while (d) shows corresponding data taken perpendicular to the growth direction at the second zigzag chain of the second dot, as indicated by D in (a). The simulated height contours from (b) are also plotted in (c) and (d). In (d) the flanks induced by the buried material are marked, resulting from a cleavage position as shown schematically in the inset.

of stoichiometrically pure InAs dots can be derived from our data. It should be noted, that even the strain relaxation at the $Ga_{0.02}In_{0.98}As(110)$ surface would lead to a visible inhomogeneity in the images, as we demonstrate elsewhere.[36] In addition, no In atoms are observed in the GaAs overlayers. Thus only negligible intermixing of both materials seems to occur. This observation is in contrast to findings from top-view STM experiments on MBE-grown InAs dots.[15]

2. Size of the dots

Figure 4(a) shows a filled-state XSTM image of another representative dot stack. The arrows mark the positions of the height-contour curves through the center of the stack along the growth direction and perpendicular to the growth direction along the second zigzag chain of the second dot, as displayed in Figs. 4(c) and 4(d), respectively.

In the height-contour curve in Fig. 4(d), flanks are visible besides the dots, indicating that InAs is buried underneath the surface, leading to a weak strain relaxation in these areas. Thus it can be concluded that the cleavage of this stack occurred above the dot center, leaving more than one half of the dot underneath the cleavage surface, as indicated in the inset of Fig. 4(d). Using the onsets of these flanks, it is found that the subsurface dot extends from about 3 to 19 nm on the lateral-position scale, yielding an actual diameter of about 16 nm. The kink-like transitions from the flanks to the large

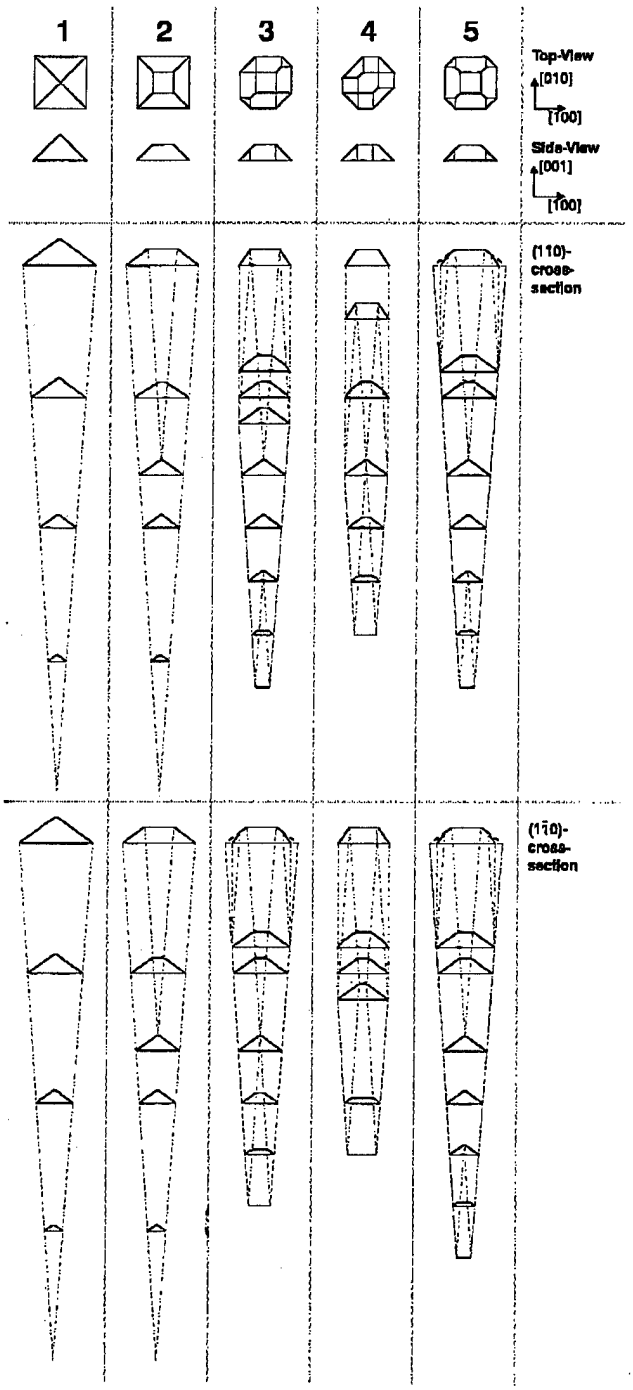


FIG. 5. Overview of the most relevant dot shapes an their possible cross sections with the (110) and ($\bar{1}$10) surface.

bump observed at lateral positions of about 5 and 17 nm, in contrast, represent the position of the InAs-GaAs interface directly at the cleavage surface.

From an inspection of about 30 dot stacks, average dot sizes can be determined, using the lateral extensions along the base line. In this way, (13.6±1.0) nm is obtained for the base diameter of the first dot layer, (15.2±1.0) nm for the second one, and (16.8±1.0) nm for the third one. The heights of the dots show a similar trend: In the first dot layer, most often only four zigzag chains can be identified as InAs like, while the second and third one usually consist of five zigzag chains (see Figs. 3 and 4). For the evaluation of the size of the dots, their shape was assumed to be described by truncated pyramids with a quadratic base, as shown in the inset of Fig. 4(d). In this way it is derived that the dots in the first, second, and third layer consist of 5600±1000, 7400 ±1200, and 9400±1400 atoms, respectively.

Two mechanisms are discussed in order to explain the observed increase in dot size, which is observed despite the nominally identical InAs layer thickness: First it can be assumed that some InAs diffuses during the GaAs-overgrowth process towards the surface, increasing the average In content of the upper layers.[37] Since no In atoms are observed inside the GaAs overlayers including the top buffer layer, this mechanism seems to be less probable.

The second mechanism is based on the lateral strain of the GaAs overlayer, which is optimum for InAs nucleation directly on top of the buried dot from the layer underneath. In this way, more InAs material is attracted from the surrounding wetting layer to accumulate in the dot region, resulting in the formation of larger dots at the expense of the wetting layer. In subsequent layers, the conditions for selective nucleation are further improved because of the increasing lateral strain of the overgrowth layer, leading to even larger dots. Since about 80% of the total InAs deposition is forming the wetting layer, different dot sizes will only lead to negligible layer-dependent variations of the wetting-layer thickness.

3. Shape of the dots

In order to determine the shape of the dots, we have to imagine the possible dot structures and the shape of their cross sections with the (110) or ($\bar{1}$10) cleavage planes. In theoretical studies based on surface energies, shapes with {110} or {111} faces were proposed.[38] Among the 19 possible dot shapes with these low-indexed side faces, Fig. 5 displays the five most relevant ones. The interface lines derived from the material contrast in the XSTM images, which describe the transition from InAs to GaAs at the surface, are forming various polygons, depending on the position where the cleavage plane cuts the dot. In the lower part of Fig. 5, the corresponding series of contrast polygons is shown, both for the (110) and the ($\bar{1}$10) cleavage plane. Since only {110} or {111} side faces are assumed to occur, up to three different angles between the interface lines and the base line of the dot are expected to be observed in the images: 0° for the (001) top surface, and 35.3° for the {101} as well as 54.7° for the {111} side faces.

A variety of images have been analyzed carefully with respect to the shape. A detailed shape analysis is possible on the basis of Fig. 6. In Fig. 6(a) an empty-state image of a dot stack is presented, where the interface line appears rather curved. The high-pass filtered image in Fig. 6(b), on the other hand, yields more detailed information on the shape. In particular, a kink is visible in the interfaces lines at the sides of the dot. This is demonstrated in Fig. 6(c), where the possible structure of the cross-sectional plane is outlined sche-

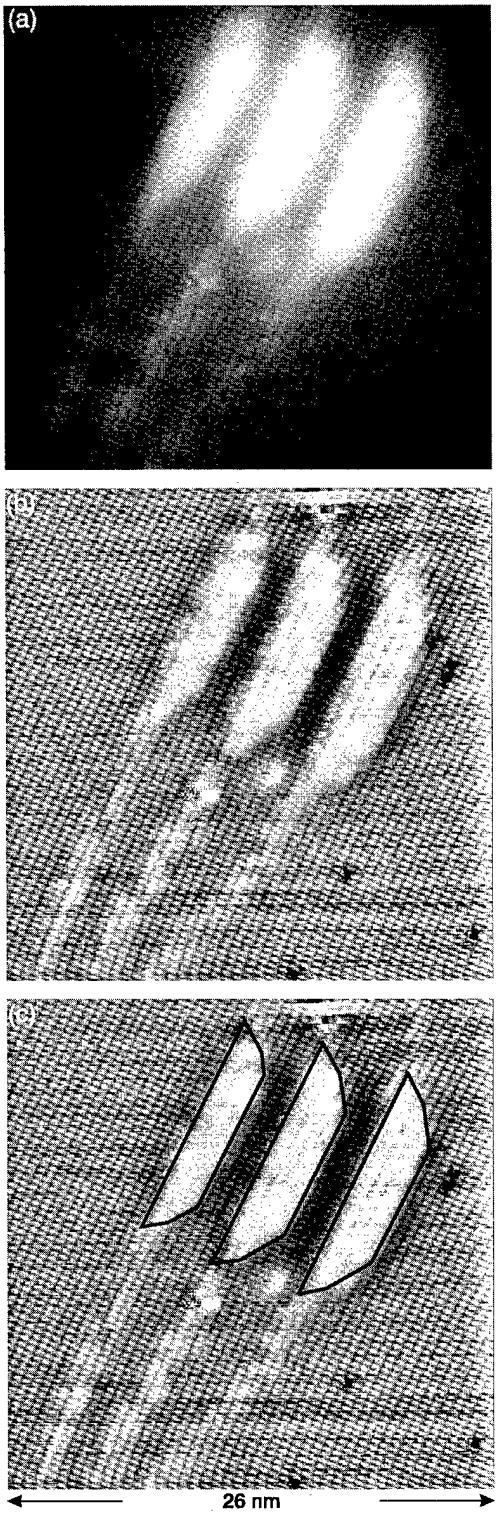


FIG. 6. (a) Empty-state XSTM image of a threefold dot stack, taken at $V_S = +1.4$ V and $I_T = 65$ pA. (b) The same image after high-pass filtering. (c) Schematic outline of the dot contour at the cleavage surface.

matically. The steeper section between the base line and the third zigzag chain of the dots confines an angle of 55°±5° with the base line, while the flatter section between the third zigzag chain and the top of the dot is characterized by an angle of 35°±5°. Such a cross-sectional structure can only be induced from dot shapes 3, 4, or 5 in Fig. 5, but does not correspond to shape 1 or 2.

For these shapes, cross-sectional cuts exhibiting only side lines with the angle of 35.3° are also expected, while the angle of 54.7° is absent. Indeed we occasionally observed such XSTM images, like the ones in Figs. 4 and 7. In this way it can be assumed that all dots are characterized by shapes 3, 4, or 5. This view is supported by an additional statistical observation: We never found cross sections with a short base line along the $[\bar{1}10]$ direction, even in cases when the dots appeared very flat as, e.g., the upper dot stack in Fig. 2(a) or the center one in Fig. 1. Such a cross-sectional image should be observed rather frequently in the case of shape 2. Furthermore it should be noted that triangular cross sections, as expected from the nontruncated pyramidal shape 1, were never observed.

Nevertheless the actual shape of the dots could not be determined securely on the basis of the present data. The observed interface lines indicate the existence of a (001) top face as well as {101} and {111} side faces, corresponding to shapes 3, 4, or 5 in Fig. 5. By comparing these dot shapes to those proposed from theoretical calculations,[38] the experimentally obtained dot heights are much smaller than the theoretical ones. This indicates an incomplete segregation of InAs during dot formation, so that their shape and size is not in thermal equilibrium when the overgrowth process has already started. A further indication for this kinetic limitation is the buckled shape of the wetting layer on top of the dots, which is due to incomplete GaAs segregation during overgrowth.

4. Strain-relaxation effects

For a simulation of the cleavage-induced strain relaxation of the stack shown in Fig. 4, we used dot shape 2, with the dot located mainly underneath the cleavage surface. The simulated outward relaxation is shown in Fig. 4(b). It is noted that this image does not contain the structural information within the III–V unit cell, resulting in missing zigzag chains. The image looks similar to the measured one, with the exception that there is a much higher contrast at the interface lines. Since the actual dot shape is assumed to play only a minor role, this difference can be attributed to the continuum-mechanical simulation that was performed here,[39] which gives a better description of the strain relaxation inside of one material than at material boundaries. A valence-force field simulation,[40] on the other hand, would result in a better approximation of the material interfaces. At present, such a time-consuming simulation could not be performed.

The height-contour curves along the lines indicated in Fig. 4(a), resulting from the simulation, are also presented in Figs. 4(c) and 4(d). By comparing the experiment with simulation in Fig. 4(d), the difficulties in simulating the interface are obvious, as mentioned above. In particular the flanks, induced by the buried material, are not described appropriately. Furthermore, the total elevation derived from the simulation is smaller than the measured protrusion, which is caused by an additional electronic contrast.

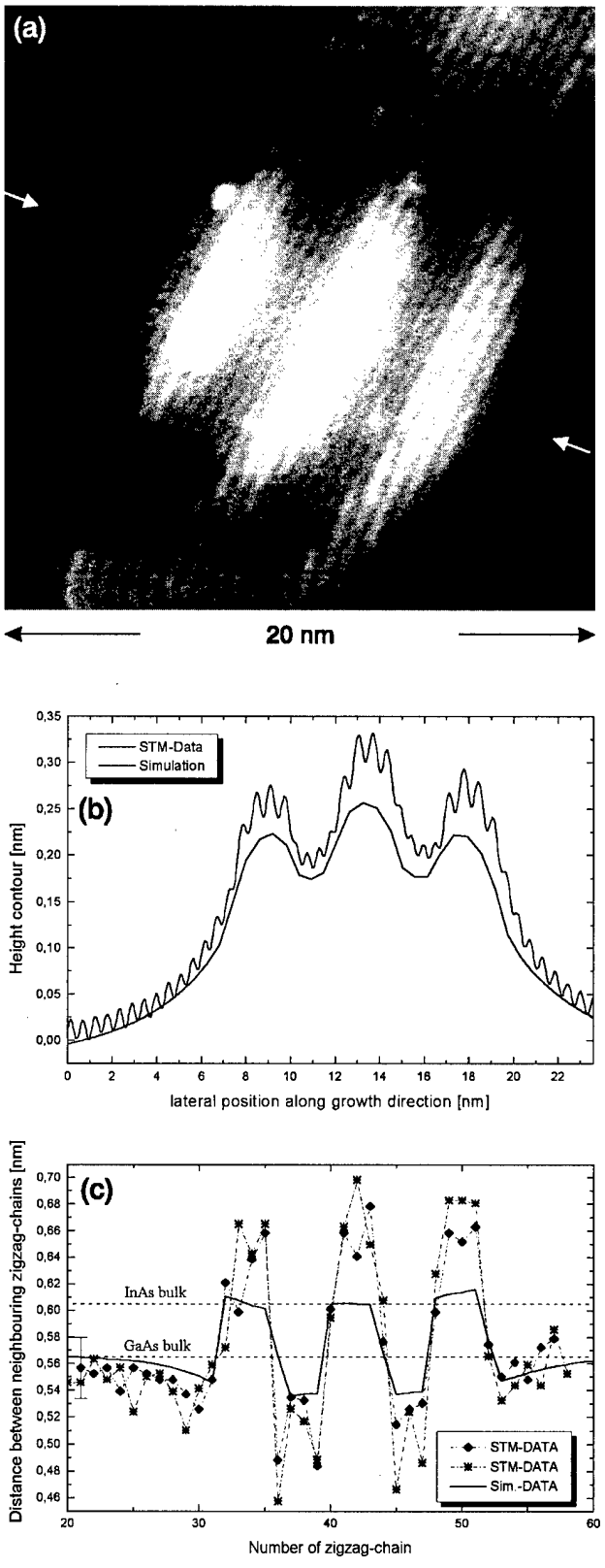


FIG. 7. (a) Filled-state XSTM image of a dot stack with high resolution, taken at $V_S = -2.5$ V and $I_T = 110$ pA. (b) Height contour through the center of the stack along growth direction, as indicated in (a), together with simulated data. (c) Distances between neighboring zigzag chains across the dot stack, together with the simulated data as well as the bulk lattice constants of GaAs and InAs for comparison.

The strain simulation can also provide data on the strain-induced lateral relaxation in the [001] direction. A few XSTM images allow us to determine the distance of the zigzag chains. Figure 7(a) shows a filled-state image of a dot stack. For the simulation, shape 4 in Fig. 5 was used, with a height of 10 monolayers, equivalent to five zigzag chains, and a separation of the base lines amounting to 16 monolayers. In Fig. 7(b), the height contour curve in the growth direction through the center of the dot stack is drawn, together with the result from the simulation, which show remarkable agreement.

In Fig. 7(c) the distance between neighboring zigzag chains is plotted, together with the results of the simulation. In order to demonstrate the reproducibility, a second data set from a similar image of the same dot stack is shown. The nominal lattice constants of InAs and GaAs are also indicated. It is obvious that experiment and simulation agree well with respect to the sudden distance variations at the InAs–GaAs interfaces. This behavior again supports our assignment of the bright zigzag chains to InAs, while the darker ones correspond to GaAs.

In the GaAs buffer layers, the experimental zigzag-chain distances are in good agreement with the calculated ones. At the dots, however, the experimental values are considerably larger than the calculated ones, while they are smaller at the intermediate GaAs layers. On the other hand, such a difference is also expected, since the STM imaged the filled dangling bonds, which may show a different distance than the corresponding As atoms because of the strain-induced surface curvature. In this way, the measured distance between neighboring dangling bonds is expanded at the dots and compressed at the surrounding GaAs layers. However, this geometrical effect could not be simulated within this work, and no measurement of the buckling behavior of strained III–V surfaces exists until now.

5. Effects of the electronic structure

Up to now, we did not perform tunneling spectroscopy for studying the electronic structure of the dots. However, one effect visible in Fig. 6(a) can be assigned to the electronic properties of the dots: At low positive sample voltages, it is not possible to image the zigzag chains at the single dots, while the surrounding GaAs region shows atomic resolution. Although the origin of this effect is not understood up to now, it can be speculated that the localized electron quantum states in the dot act as charge traps for tunneling electrons. In this case, charge fluctuations will result in strong potential fluctuations, which could prevent atomic resolution. However, more detailed experiments have to be performed in order to study this effect.

D. Properties of the wetting layers

Considering the nominal InAs exposure of 1.7 monolayers and the amount of about 0.2–0.3 monolayers of InAs used for dot formation at the different dot layers, an average remaining thickness of the wetting layer of 1.4–1.5 monolayers is determined. Therefore, lateral-thickness fluctuations

of 1–2 monolayers are expected. In addition, only every second monolayer is imaged by XSTM at the (110) surface.

Figure 8(a) shows an empty-state image with a close view of the wetting layer, where one dot stack is located directly above the image. The white bump in the center and the six dark depressions are assigned to adsorbate molecules. As expected, fluctuations in the thickness of the wetting layers are observed. The typical wavelength of these fluctuations is much larger than expected for a statistical distribution of InAs in the second monolayer. Furthermore the distances between two wetting layers vary from five to six zigzag chains in the upper region to seven to eight in the lower region of the image, as already observed in the overview images in Figs. 1 and 2. This behavior demonstrates that there is already a considerable roughness at the GaAs overgrowth surfaces, which again can be related to an incomplete smoothing of the surface during growth.

In Fig. 8(b) two height contours along lines L and U through the wetting layers are plotted. These data were averaged by a few nm along the $[\bar{1}10]$ direction in order to suppress the corrugation within one zigzag chain. From the simulation of strain relaxation of a one monolayer thick InAs layer, an elevation of the cleavage surface at the wetting-layer position of less than 0.03 nm is calculated. On the other hand, the corrugation in Fig. 8(b) shows a maximum height difference of about 0.2 nm. Thus this height contour cannot be explained by structural effects alone and should be mostly of electronic origin. In calculations of the wave function of the bound electron state of one or two InAs monolayers within GaAs,[41–43] a tailing of the wave function into the GaAs layers similar to the observed height contour in Fig. 8(b) was determined. The calculations yield a penetration depth of about three zigzag chains, in good agreement with our observations of 2–3 zigzag chains. Thus we can explain the observed corrugation by imaging the confined states of the wetting layer tailing into GaAs.

For line L the distances between neighboring zigzag chains are analyzed, as described in Sec. III C 4, and drawn in Fig. 8(c) together with the GaAs lattice constant and the nominal lattice constant of 1 monolayer InAs inside GaAs.[44] A considerably larger distance between the two neighboring zigzag chains is found at the wetting layer, which can even be observed directly in the XSTM image in Fig. 8(a) in the form of a deeper depression. This behavior results from the larger lattice constant of InAs as well as from the strain-induced expansion mainly in the growth direction, which is induced by the compression perpendicular to the [001] direction.

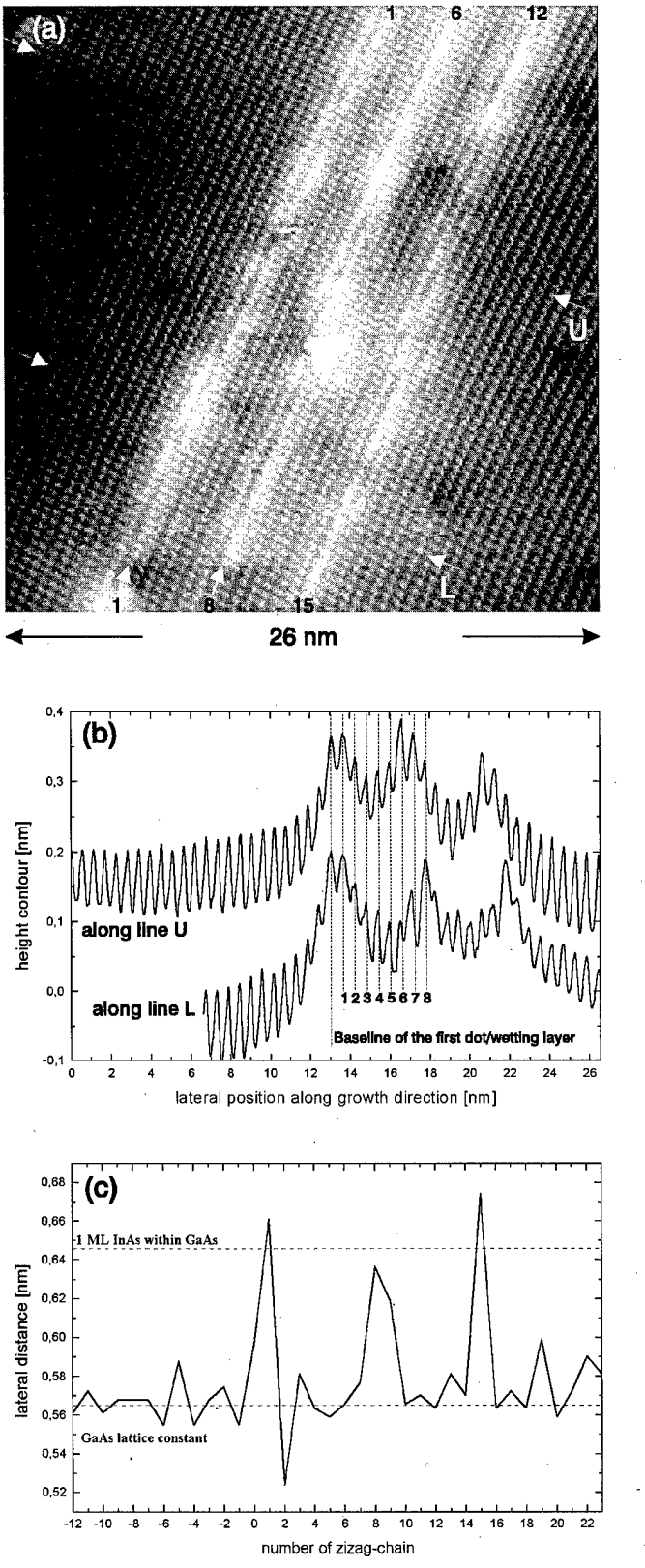


FIG. 8. (a) Empty-state XSTM image of a wetting-layer region, taken at $V_S = +1.5$ V and $I_T = 65$ pA, with a numbering of the zigzag chains. (b) Height-contour curves taken along the growth direction, as indicated in (a) by U and L. (c) Distances between neighboring zigzag chains across the wetting layers, together with the lattice constants of GaAs bulk and of one monolayer of InAs within GaAs (Ref. 44).

IV. CONCLUSION AND OUTLOOK

In this work, we presented the first XSTM results of stacked InAs quantum dots grown by metal-organic chemical vapor deposition, enabling us to study the detailed structure with atomic resolution. The buffer and dot layers were found to be grown in a perfect way without any dislocation. No

intermixing of the InAs and GaAs was observed in the dot region, so that the formation of stoichiometrically pure InAs quantum dots can be concluded.

While the vertical alignment within a stack is quite good, the overgrowing GaAs layer does not smooth out the surface completely, in particular above the dots and within their close environment. This behavior is related mainly to an incomplete surface segregation, leading to kinetically limited structures rather than to a situation expected for thermal equilibrium.

In a recent XSTM work on stacked InAs dots grown by MBE at $T=485\,^{\circ}C$ with a much lower growth rate of 0.028 nm/s, a more symmetric disk-like shape and a more uniform size was observed, together with In diffusion into the surrounding GaAs matrix.[22,24] In contrast, InAs dots grown at $T=520\,^{\circ}C$ at a higher rate of 0.55 nm/s show a weak In diffusion only at the material boundary.[23] Thus different growth methods and parameters yield rather different dot structures, which can be related to a strong influence of growth kinetics.

While the dot sizes are rather uniform within one layer, the size of the dots increases for subsequent layers within a stack. This effect is related to lateral strain of the GaAs surface above the buried dot, increasing for subsequent layers. Thus the nucleation of InAs at the strained GaAs surface above a buried dot is increasingly preferred from layer to layer. The dot shape can be described by prisms with {101} and additional {111} side faces and a (001) top face.

Cleavage-induced strain relaxation plays a major role for explaining the large contrast observed when imaging the dots. At the wetting layers, in contrast, electronic effects dominate, and the wave function of the confined electron state could be imaged.

Finally it should be noted that the present experiment provides detailed structural information on a specific sample grown far from thermal equilibrium. For an analysis of the equilibrium shape of InAs quantum dots, XSTM experiments on samples grown at higher temperatures, slower growth rates, and/or with longer growth interrupts are in preparation.

ACKNOWLEDGMENTS

The authors would like to thank M. Scheffler, P. Kratzer, Q. Liu, and L. Wang for the discussion about the shape of the dots, and C. Preinesberger for the improvement of the STM. This work was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 296, TP A4 and A7.

[1]D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum Dot Heterostructures* (Wiley, Chichester, 1999).
[2]D. Leonard, M. Krishnamurti, C. M. Reeves, S. P. Denbaars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993).
[3]D. Bimberg *et al.*, Thin Solid Films **267**, 32 (1995).
[4]F. Heinrichsdorff, A. Krost, M. Grundmann, D. Bimberg, A. Kosogov, and P. Werner, Appl. Phys. Lett. **68**, 3284 (1996).
[5]F. Heinrichsdorff, Ph.D. thesis, Technische Universität Berlin, 1998.
[6]I. N. Stranski and L. von Krastanov, K. Preuss. Akad. Wiss., Phys. Math. Kl. **146**, 797 (1937).
[7]N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin, P. S. Kop'ev, and Zh. I. Alferov, Semiconductors **32**, 343 (1998).
[8]J. Tersoff, C. Teichert, and M. G. Lagally, Phys. Rev. Lett. **76**, 1675 (1996).
[9]M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **59**, 5688 (1999).
[10]R. Ruvimov, Z. Liliental-Weber, J. Washburn, N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin, P. S. Kop'ev, Zh. I. Alferov, and D. Bimberg, Phys. Solid State **40**, 781 (1998).
[11]Q. Xie, P. Chen, and A. Madhukar, Appl. Phys. Lett. **65**, 2051 (1994).
[12]F. Heinrichsdorff, A. Krost, M. Grundmann, D. Bimberg, A. Kosogov, P. Werner, F. Bertram, and J. Christen, in *The Physics of Semiconductors*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), p. 1321.
[13]A. A. Darhuber *et al.*, Appl. Phys. Lett. **70**, 955 (1997).
[14]J. M. García, G. Medeiros-Riberio, K. Schmidt, T. Ngo, J. L. Feng, A. Lorke, J. Kotthaus, and P. M. Petroff, Appl. Phys. Lett. **71**, 2014 (1997).
[15]B. A. Joyce, J. L. Sudijono, J. G. Belk, H. Yamaguchi, X. M. Zhang, H. T. Dobbs, A. Zangwill, D. D. Vvedesky, and T. S. Jones, Jpn. J. Appl. Phys. **36**, 4111 (1997).
[16]Y. Hasegawa, H. Kiyama, Q. K. Xue, and T. Sakurai, Appl. Phys. Lett. **72**, 2265 (1998).
[17]L. N. Bolotov, A. Nakamura, V. P. Evtikhiev, V. E. Tokranov, and A. N. Titkov, Surf. Interface Anal. **27**, 533 (1999).
[18]N. P. Kobayashi, T. R. Ramachandran, P. Chen, and A. Madhukar, Appl. Phys. Lett. **68**, 3299 (1996).
[19]R. M. Feenstra, Semicond. Sci. Technol. **9**, 2157 (1994).
[20]A. R. Smith, K.-J. Chao, C. K. Shih, Y. C. Shih, and B. G. Streetman, Appl. Phys. Lett. **66**, 478 (1995).
[21]J. Harper, M. Weimer, D. Zhang, C.-H. Lin, and S. S. Pei, J. Vac. Sci. Technol. B **16**, 1389 (1998).
[22]W. Wu, J. R. Tucker, G. S. Solomon, and J. S. Harris, Appl. Phys. Lett. **71**, 1083 (1997).
[23]B. Legrand, B. Grandidier, J. P. Nys, D. Stiévenard, J. M. Gérard, and V. Thierry-Mieg, Appl. Phys. Lett. **73**, 96 (1998).
[24]G. S. Solomon, W. Wu, J. R. Tucker, and J. S. Harris, Physica E **2**, 709 (1998).
[25]B. Lita, R. S. Goldman, J. D. Phillips, and P. K. Bhattacharya, Appl. Phys. Lett. **74**, 2824 (1999).
[26]H. Eisele, O. Flebbe, T. Kalka, C. Preinesberger, F. Heinrichsdorff, A. Krost, D. Bimberg, and M. Dähne-Prietsch, Appl. Phys. Lett. **75**, 106 (1999).
[27]H. Eisele, O. Flebbe, T. Kalka, and M. Dähne-Prietsch, Surf. Interface Anal. **27**, 537 (1999).
[28]T. Kalka, C. Preinesberger, S. Vandré, and M. Dähne-Prietsch, Appl. Phys. A: Mater. Sci. Process. **66**, S1073 (1998).
[29]M. Pfister, M. B. Johnson, S. F. Alvarado, H. W. M. Salemink, U. Marti, D. Martin, F. Morier-Genoud, and F. K. Reinhart, Appl. Phys. Lett. **67**, 1459 (1995).
[30]A. J. Heinrich, M. Wenderoth, M. A. Rosentreter, K. Engel, M. A. Schneider, R. G. Ulbricht, E. R. Weber, and K. Uchida, Appl. Phys. A: Mater. Sci. Process. **66**, S959 (1998).
[31]K.-J. Chao, C.-K. Shih, D. W. Gotthold, and B. G. Streetman, Phys. Rev. Lett. **79**, 4822 (1997).
[32]S. L. Zuo, W. G. Bi, C. W. Tu, and E. T. Yu, J. Vac. Sci. Technol. B **16**, 2395 (1998).
[33]Ph. Ebert, B. Engels, P. Richard, K. Schröder, S. Blügel, C. Domke, M. Heinrich, and K. Urban, Phys. Rev. Lett. **77**, 2997 (1996).
[34]R. M. Feenstra, J. A. Strocio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. **58**, 1192 (1987).
[35]H. Eisele, O. Flebbe, and M. Dähne-Prietsch (unpublished).
[36]H. Eisele, O. Flebbe, T. Kalka, F. Heinrichsdorff, A. Krost, D. Bimberg, and M. Dähne-Prietsch, Phys. Status Solidi B (in press).
[37]N. N. Ledentsov *et al.*, Phys. Rev. B **54**, 8743 (1996).
[38]N. Moll, M. Scheffler, and E. Pehlke, Phys. Rev. B **58**, 4566 (1998).
[39]J. F. Nye, *Physical Properties of Crystals* (Oxford University Press, Oxford, 1972), p. 131.
[40]P. N. Keating, Phys. Rev. **145**, 637 (1966).
[41]G. H. Li, A. R. Goñi, C. Abraham, K. Syassen, P. V. Santos, A. Cantarero, O. Brandt, and K. Ploog, Phys. Rev. B **50**, 1575 (1994).
[42]K. Shiraishi and E. Yamaguchi, Phys. Rev. B **42**, 3064 (1990).
[43]A. R. Goñi (unpublished).
[44]O. Brandt, L. Tapfer, R. Cingolani, K. Ploog, M. Hohenstein, and F. Phillipp, Phys. Rev. B **41**, 12599 (1990).

Quantum dot-like behavior of GaInNAs in GaInNAs/GaAs quantum wells grown by gas-source molecular-beam epitaxy

H. P. Xin,[a] K. L. Kavanagh, Z. Q. Zhu,[b] and C. W. Tu
Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, California 92093-0407

(Received 20 March 1999; accepted 22 April 1999)

A quantum dot-like behavior of GaInNAs in $Ga_{0.7}In_{0.3}N_xAs_{1-x}$/GaAs quantum wells (QWs) grown by gas-source molecular-beam epitaxy has been studied using high-resolution x-ray rocking curves (XRC), cross-sectional transmission electron microscopy (XTEM), and photoluminescence (PL) spectroscopy. XRC determines the average alloy composition and thickness of the QWs to be In 0.30 and N 0–0.030, and 6.2 nm, respectively. XTEM images show that the wells of both $Ga_{0.7}In_{0.3}As$/GaAs and $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$/GaAs are undulated with lateral variations in strain, but the latter is much rougher. For $Ga_{0.7}In_{0.3}As$/GaAs QWs, rapid thermal annealing results in the splitting of a broad excitonic emission into two peaks presumably due to In composition fluctuation. The separation between these two peaks increases with increasing N concentration. Increasing excitation intensity results in a significant blueshift for the low-energy peak, while little blueshift for the high-energy peak. The earlier PL behavior can be interpreted by a simple model where the low-energy peak originates from In and N-rich regions in the wells acting as quantum dots. The high-energy peak is likely due to the excitons of a more two-dimensional QW. © *1999 American Vacuum Society.* [S0734-211X(99)03204-7]

I. INTRODUCTION

In optical fiber communication field, long wavelength (at 1.3 and 1.55 μm) semiconductor lasers are often used as light sources to minimize the transmission loss at the optical fiber windows. However, the lasing properties of currently used 1.3 and 1.55 μm GaInAsP/InP laser diodes are still poor at higher operating temperature due to poor electron confinement (a result of the small conduction band offset). This conventional material system is also not suitable for vertical-cavity surface-emitting lasers (VCSELs) since its poor refractive index contrast as well as low-thermal conductivity have made epitaxially grown distributed Bragg reflectors (DBRs) impractical. The GaInNAs/GaAs system was proposed to overcome these two problems.[1] On the one hand, it has a large band gap bowing, resulting in light emission at 1.3 μm and a large conduction band offset due to the large electronegativity of N atoms. On the other hand, GaAs–AlAs DBRs can be easily grown on a GaAs substrate for VCSELs operating at 1.3 μm.

The first current-injection $Ga_{0.7}In_{0.3}N_{0.004}As_{0.996}$/GaAs quantum well (QW) laser grown by gas-source molecular-beam epitaxy (GSMBE) was obtained by Kondow *et al.*, where a room-temperature (RT) continuous-wave lasing near 1.18 μm was achieved.[2] A laser operating at near 1.3 μm under RT pulse operation using a GaInNAs active layer grown by metalorganic chemical vapor deposition was realized by Sato *et al.*[3] Recently, a GaInNAs/GaAs VCSEL diode near 1.18 μm[4] and a GaInNAs/GaAs QW laser with continuous-wave lasing operation at 1.3 μm have been achieved.[5]

Incorporation of nitrogen in GaInAs/GaAs, however, results in lower photoluminescence (PL) intensities and wider linewidths.[2–6] This could be due to alloy composition fluctuation and/or associated nonradiative centers, and would increase the laser threshold. Therefore, understanding the distribution of N and In in GaInNAs is especially valuable. In this article, a quantum dot-like behavior of GaInNAs due to a correlated nonuniform distribution of N and In in GaInNAs/GaAs QWs is observed. High resolution x-ray rocking curves (XRC), cross-sectional transmission electron microscopy (XTEM), and low-temperature PL techniques are used to characterize the samples before and after rapid thermal annealing (RTA).

II. EXPERIMENTAL DETAILS

7-period GaInNAs/GaAs multiple QWs (MQWs) were grown on semi-insulating (100) GaAs substrates in a modified Varian Gen-II system equipped with two 2200 l/s cryopumps. 7N elemental Ga, In, and thermally cracked AsH_3 were used. High-purity N_2 was injected through a N radical beam source (Oxford Applied Research Model MPD21) operated at a radio frequency (rf) of 13.56 MHz with 200 W power to generate active N species. To determine the In and N composition in GaInNAs, a reference sample GaInAs/GaAs QWs was grown at the same growth conditions without striking the rf source. For the GaAs buffer and cap layers, the growth temperature was 600 °C. For GaInNAs/GaAs MQW, the growth temperature was decreased to 480 °C in order to incorporate In and N into the well layers.

For XTEM observation, a sample consisting of a 2-period $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$ (62 Å)/GaAs (169 Å) QWs and a

[a]Electronic mail: hxin@sdcc10.ucsd.edu
[b]Permanent address: National Laboratory of Functional Materials for Informatics, Shanghai Institute of Metallurgy, Chinese Academy of Sciences, People's Republic of China.

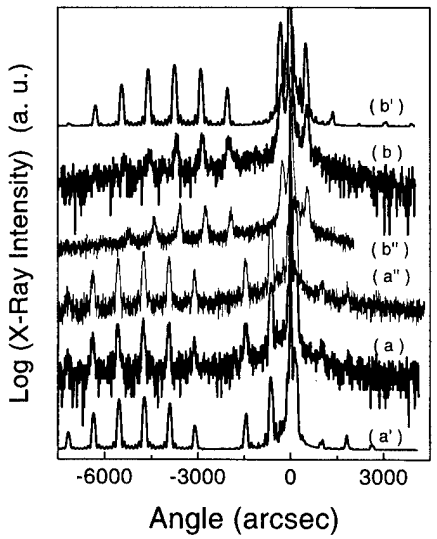


FIG. 1. (a) (400) XRC of the $Ga_{0.7}In_{0.3}As$/GaAs MQW, (a) dynamical theory simulation of (a), (a″) (400) XRC of the 850 °C annealed $Ga_{0.7}In_{0.3}As$/GaAs MQW; (b) XRC of the $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$/GaAs MQW with 1.5 sccm N_2 flow rate, (b′) dynamical theory simulation of XRC of (b), (b″) (400) XRC of the 850 °C annealed $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$/GaAs MQW.

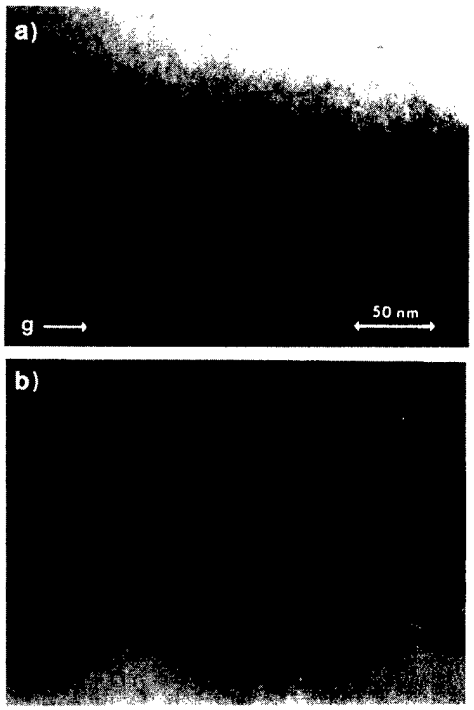


FIG. 2. Bright field XTEM micrographs of an as-grown sample consisting of a 2-period $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$/GaAs QW and a 2-period $Ga_{0.7}In_{0.3}As$/GaAs QW, $g=(220)$, (a) as-grown sample, (b) 850 °C annealed sample.

2-period $Ga_{0.7}In_{0.3}As$ (62 Å)/GaAs (169 Å) QWs was also grown on an n^+-GaAs (100) substrate using the same growth conditions.

XRC measurement was performed using a Phillip x-ray diffractometer. RTA was performed on these samples for 10 s by using halogen lamps and flowing N_2 ambient. During RTA, the samples were put on a GaAs wafer face-to-face to prevent the loss of arsenic at elevated temperatures. Transmission electron microscopy (TEM) was carried out using a Philips CM30 transmission electron microscope operating at a source voltage of 300 KV. Low-temperature PL measurement was carried out by mounting the samples in a liquid He cryostat and using the 514.5 nm line of an Ar^+ laser as the excitation source. A thermo-electrically cooled Ge photodiode was used to detect the signal at the exit of a 50 cm monochromator through an amplifier.

III. RESULTS AND DISCUSSIONS

The layer composition and thickness are determined from high-resolution XRC and simulations based on the dynamical theory. Figure 1 shows (400) XRCs of (a) GaInAs/GaAs MQW and (b) GaInNAs/GaAs MQW with a 1.5 sccm N_2 flow rate. Compared with that of the GaInAs/GaAs QWs, the zeroth order peak of the GaInNAs/GaAs MQW shifts closer toward the GaAs substrate peak, demonstrating that adding N into the GaInAs layer reduces the net compressive strain of the system. The satellite peaks, however, are broader than that of N-free samples, presumably due to composition fluctuations and worse interfaces of GaInNAs/GaAs. Also shown in Fig. 1 are the dynamical-theory simulation results, curves (a′) and (b′) corresponding to curves (a) and (b), respectively. The best fitting shows that curves (a) and (b) correspond to $Ga_{0.7}In_{0.3}As$ (62 Å)/GaAs (169 Å) MQW and $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$ (62 Å)/GaAs (169 Å) MQW, respectively. After RTA at 850 °C, no obvious change of (400) XRCs is observed for both MQWs, shown in curves (a″) and (b″), respectively. It indicates no obvious loss of N and arsenic during RTA.

For the sample consisting of a 2-period $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$/GaAs QW and a 2-period $Ga_{0.7}In_{0.3}As$/GaAs QW, TEM bright field image of an as-grown sample and 850 °C annealed sample were taken with a 2-beam, $g=(220)$ diffraction condition, shown in Figs. 2(a) and 2(b), respectively. The thickness of QWs from XTEM images agrees well with XRC results. The wells of both $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$/GaAs and $Ga_{0.7}In_{0.3}As$/GaAs are undulated, showing a large lateral variation in strain. This is due to lateral variations in composition induced by large lattice mismatch (2.2% lattice mismatch for $Ga_{0.7}In_{0.3}As$/GaAs and 1.7% lattice mismatch for $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$/GaAs), but it is interesting that the $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$ layers (the bottom two layers in each case) show significantly more lateral undulation in spite of the lower average strain. This lateral variations in strain of the well layers are presumably due to nonuniform In and N concentrations, which agrees well with XRC results that the satellite peaks of GaInNAs/GaAs MQW are broader than that of N-free MQW. Figure 2(b) shows that the degree of undulation in lateral strain for both materials decreases after annealing, likely caused by interdiffusion of Ga and In out of and within the same layer and the compositions of N and In become predominantly bimodal.

Figure 3 shows 10 K PL spectra of the 7-period $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs MQW treated by RTA at a different temperature. With increasing annealing temperature,

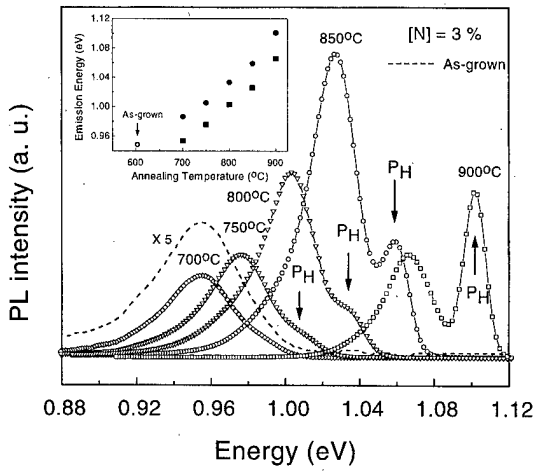


FIG. 3. 10 K PL spectra of a $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs 7-period MQW treated by RTA at different temperatures. The inset is the relationship of peak P_L and P_H emission energy as a function of RTA temperature.

RTA further reduces the lateral strain variation and improves the composition uniformity, so the PL linewidth decreases. The integrated intensity of P_L also is decreased, and peak P_H emission becomes stronger compared with peak P_L. P_L and P_H emission energies are plotted as a function of the annealing temperature in the inset of Fig. 3. Since Rao *et al.* reported that annealing caused only a negligibly small blueshift in the GaNAs/GaAs sample even with a higher N content (~4%),[7] the blueshift in our GaInNAs/GaAs MQWs comes mainly from the interdiffusion of Ga and In. Therefore, the N distribution remains unchanged, and both P_L and P_H have the same blue shift with increasing RTA temperature.

10 K PL spectra of a 7-period $Ga_{0.7}In_{0.3}As$/GaAs MQW are shown in Fig. 4(a), where the dashed line and solid line correspond to as-grown and 900 °C annealed samples, respectively. RTA causes a blueshift of the PL peak due to the interdiffusion of In and Ga between well and barrier layers. For the as-grown $Ga_{0.7}In_{0.3}As$/GaAs QWs, the PL linewidth is broad, presumably due to larger interface roughness and In composition nonuniformity. RTA decreases the PL linewidth and results in a splitting of the broad as-grown PL emission into two peaks, a lower energy peak P_L and a higher energy P_H. The In composition fluctuation is more evident in the $Ga_{0.6}In_{0.4}As$/GaAs MQW, shown in Fig. 4(b). Due to too large a strain (about 2.8% lattice mismatch), the $Ga_{0.6}In_{0.4}As$ phase separates and forms two broad PL peaks. The In composition in the $Ga_{0.7}In_{0.3}As$ layers after annealing becomes predominantly bimodal due to further phase segregation into two major compositions. The presence of two peaks is presumably due to higher and lower In composition regions, consistent with the TEM data in Fig. 2(b). This is similar to the PL properties of an $In_{0.20}Ga_{0.80}N$ (3 nm)/$In_{0.05}Ga_{0.95}N$ (6 nm) MQW,[8,9] which also shows clustering and two PL peaks. The main PL peak at 2.920 eV is attributed to the excitons localized at trap centers within the well, which originates from In-rich regions in the wells acting as quantum dots (QDs). Another weak peak at 3.155 eV is due to the

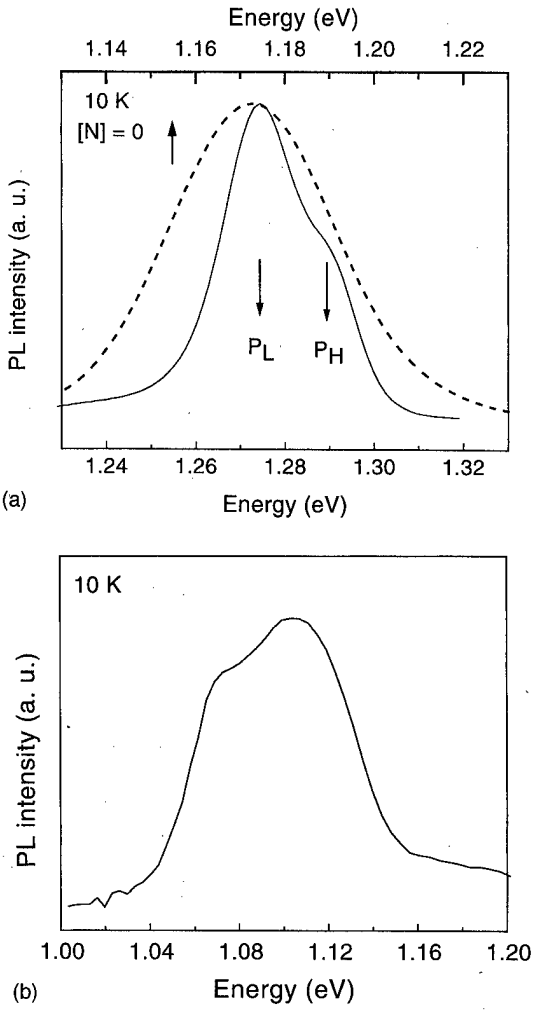


FIG. 4. (a) 10 K PL of a $Ga_{0.7}In_{0.3}As$/GaAs 7-period MQW, where the dashed line and solid line correspond to as-grown and 900 °C annealed samples, respectively. (b) 10 K PL of an as-grown $Ga_{0.6}In_{0.4}As$/GaAs 5-period MQW.

excitons at the $n=1$ quantized level between the conduction and valence band. With a higher power excitation, a 30 meV blueshift is observed for the main low-energy emission band and little blueshift for the 3.155 eV high-energy peak, due to the more easily band filling of the localized tail states than that of quantized energy levels in two-dimensional (2D) layers. Therefore, we attribute P_L to originate from In- and N-rich regions in the wells acting as QDs and P_H to excitons from QWs. Such behavior has not been observed previously for GaInAs/GaAs MQWs, however, probably because they have either a lower In composition (less strain) or a higher In composition but thinner well or because the annealing temperature is lower.[10]

For $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs MQW, RTA results in a splitting of the broad as-grown peak into two peaks P_L and P_H more clearly, as shown in Fig. 5. The In and N composition fluctuation after annealing becomes predominantly bimodal. The separation becomes larger and the relative intensity of P_H over P_L increases with higher N concentration. Since there is a big miscibility gap in mixed group V

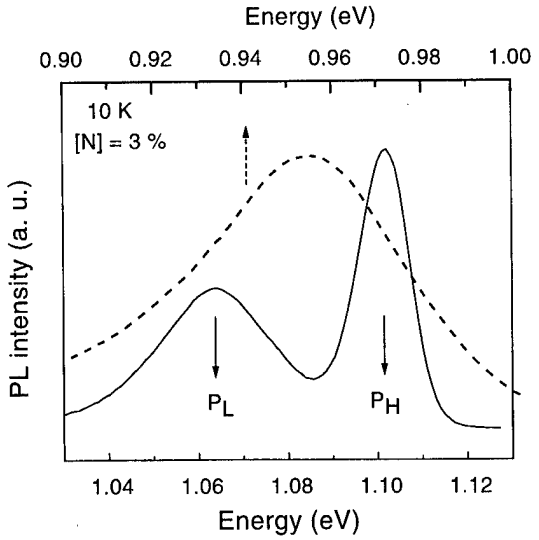


FIG. 5. 10 K PL of a $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs 7-period MQW, where the dashed line and solid line correspond to as-grown and 900 °C annealed samples, respectively.

nitride–arsenide,[11] incorporation of N in GaInAs tends to phase separate and results in the fluctuation of N distribution. This is consistent with the XTEM images (Fig. 2), where the lateral strain variation in the $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}$ layer is significantly larger than the $Ga_{0.7}In_{0.3}As$ layer. Due to the large strain, there is also a fluctuation of In distribution. As a result, N with smaller atomic radius could be preferentially localized in In-rich region to reduce local strain, further shifting P_L peak to low energy, so the separation between P_L and P_H increases. Since the composition fluctuation of clustering region is larger, the corresponding emission peak P_L is much broader than P_H.

Figure 6 shows the temperature-dependent PL of a 900 °C annealed $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs MQW. With increasing temperature, both of the two peaks become weak due to more nonradiative centers. The P_H intensity decreases much faster than that of P_L since the carriers can transport to N- and In-rich regions from 2D regions easily at elevated temperatures. As a result, P_H is not observable at 160 K.

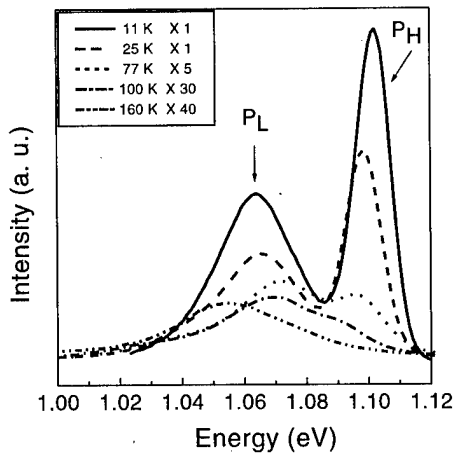
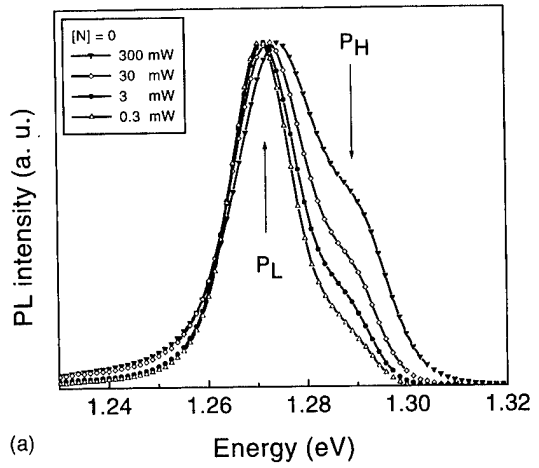


FIG. 6. Temperature-dependent PL of 900 °C annealed $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs MQW.



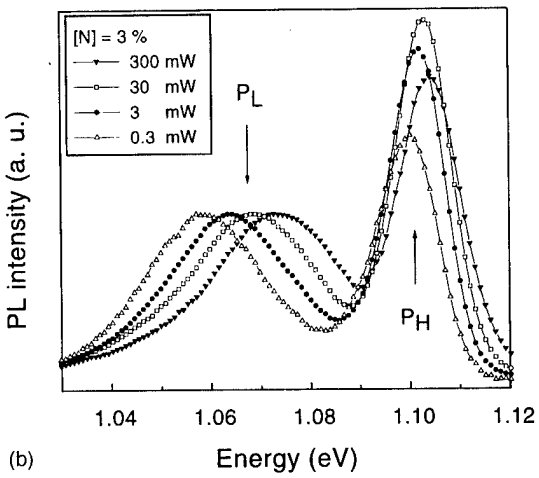


FIG. 7. 10 K PL as a function of laser power for a 7-period (a) 900 °C annealed $Ga_{0.7}In_{0.3}As$/GaAs MQW, (b) 900 °C annealed $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs MQW.

To provide more support for the earlier model, PL measurements were carried out using different excitation laser power. Figures 7(a) and 7(b) show 10 K PL spectra as a function of laser power for the 7-period $Ga_{0.7}In_{0.3}As$/GaAs MQW and $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs MQW after annealing at 900 °C, respectively. For the 7-period $Ga_{0.7}In_{0.3}As$/GaAs MQW in Fig. 7(a), peak P_H from 2D regions is too weak to be observed with 0.3 mW laser power since most free carriers in the 2D region transport to the In-rich regions and results in stronger intensity of P_L. With increasing laser power, the relative intensity of peak P_H to P_L increases due to carrier saturation in the In-rich region and more carriers in the 2D region.

For the $Ga_{0.7}In_{0.3}N_{0.03}As_{0.97}$/GaAs MQW in Fig. 7(b), peak P_H is relatively stronger compared to P_L even with 0.3 mW laser power. Incorporation of N results in more nonradiative centers in QWs.[12] The intensity of P_L decreases faster than that of P_H due to two reasons. On the one hand, few carriers in 2D regions can transport to the N- and In-rich region. On the other hand, there are more nonradiative cen-

ters in the N- and In-rich regions. As a result, a relatively strong P_H is observed. Since the peak P_L originates from the In and N-rich regions in the well acting as QDs, the corresponding localized density of states is smaller than that of P_H. Therefore, the blueshift is larger for P_L than that for P_H with increasing excitation power, mainly due to the easier band-filling effects.

IV. CONCLUSIONS

In summary, a quantum dot-like behavior of GaInNAs due to a nonuniform distribution of N and In is observed in GaInNAs/GaAs QWs. XTEM images show that the wells of both $Ga_{0.7}In_{0.3}As/GaAs$ and $Ga_{0.7}In_{0.3}N_{0.02}As_{0.98}/GaAs$ are undulated with lateral variations in strain, but the latter is more pronounced. RTA results in a splitting of the as-grown broad PL emission into two peaks P_L and P_H, where the In and N composition fluctuation after annealing became predominantly bimodal. To reduce strain, N is preferentially localized in the In-rich regions, which results in larger separation between these two peaks. P_L peak is originated from In and N-rich regions in the wells acting as QDs, whereas P_H is likely due to the excitons of a more 2D QW. Increasing excitation intensity resulted in a significant blueshift for P_L and little blueshift for P_H, due to the easier band filling of QDs than that of 2D layers.

ACKNOWLEDGMENTS

The authors gratefully acknowledge D. J. Qiao and Professor S. S. Lau for their help on RTA. The authors wish to acknowledge the support of an UC MICRO program with Rockwell International, the DARPA Heterogeneous Optoelectronics Technology Center, NSF (DMR-9704159) and ARO MURI.

[1]M. Kondow, K. Uomi, A. Niwa, T. Kitatani, S. Watahiki, and Y. Yazawa, Jpn. J. Appl. Phys., Part 1 **35**, 1273 (1996).
[2]M. Kondow, K. Nakahara, T. Kitatani, Y. Yazawa, and K. Uomi, Electron. Lett. **32**, 2244 (1996).
[3]S. Sato, Y. Osawa, T. Saitoh, and I. Fujimura, Electron. Lett. **33**, 1386 (1997).
[4]M. C. Larson, M. Kondow, T. Kitatani, K. Nakahara, K. Tamura, H. Inoue, and K. Uomi, IEEE Photonics Technol. Lett. **10**, 188 (1998).
[5]K. Nakahara, M. Kondow, T. Kitatani, M. Larson, and K. Uomi, IEEE Photonics Technol. Lett. **10**, 487 (1998).
[6]H. P. Xin and C. W. Tu, Appl. Phys. Lett. **72**, 2442 (1998).
[7]E. V. K. Rao, A. Ougazzaden, Y. Le Bellego, and M. Juhel, Appl. Phys. Lett. **72**, 1409 (1998).
[8]Y. Narukawa, Y. Kawakami, M. Funato, S. Fujita, S. Fujita, and S. Nakamura, Appl. Phys. Lett. **70**, 981 (1997).
[9]Y. Narukawa, Y. Kawakami, S. Fujita, S. Fujita, and S. Nakamura, Phys. Rev. B **55**, R1938 (1997).
[10]G. Zhang, J. Nappi, A. Ovtchinnikov, H. Asonen, and M. Pessa, J. Appl. Phys. **72**, 3788 (1992).
[11]L. G. Ferreita, S. W. Wei, and A. Zunger, Phys. Rev. B **40**, 3197 (1989).
[12]H. P. Xin, K. L. Kavanagh, M. Kondow, and C. W. Tu, J. Cryst. Growth **202**, 419 (1999).

Influence of active nitrogen species on high temperature limitations for (000$\bar{1}$) GaN growth by rf plasma-assisted molecular beam epitaxy

T. H. Myers,[a] M. R. Millecchia, and A. J. Ptak
Department of Physics, West Virginia University, Morgantown, West Virginia 26506

K. S. Ziemer and C. D. Stinespring
Department of Chemical Engineering, West Virginia University, Morgantown, West Virginia 26506

(Received 2 March 1999; accepted 20 April 1999)

A reduced growth rate for plasma-assisted molecular beam epitaxy GaN growth often limits growth to temperatures less than 750 °C. The growth rate reduction is significantly larger than expected based on thermal decomposition. Characterization of various rf plasma source configurations indicated that a flux consisting predominantly of either atomic nitrogen or nitrogen metastables can be produced. The use of atomic nitrogen, possibly coupled with the presence of low energy ions, is associated with the premature decrease in growth rate. When the active nitrogen flux consists primarily of nitrogen metastables, the temperature dependence of the decrease is more consistent with decomposition rates. A significant improvement in electrical properties is observed for growth with molecular nitrogen metastables. In addition, atomic hydrogen stabilizes the growing surface of (000$\bar{1}$) GaN. © *1999 American Vacuum Society.* [S0734-211X(99)03304-1]

I. INTRODUCTION

Applications of blue and ultraviolet optoelectronic devices based on GaN have been recognized for many years. Recent advances in epitaxial GaN growth by metalorganic chemical vapor deposition (MOCVD) are leading to the rapid commercialization of this material system. Significant progress is also being accomplished by molecular beam epitaxy (MBE) growth using active nitrogen species.[1–5] Several issues remain to be resolved for MBE growth. While overall growth rate is no longer an issue, growth of GaN by plasma-assisted molecular beam epitaxy is typically limited to temperatures less than 750 °C due to a greatly reduced growth rate.[6–11] The temperature for the onset of decreased growth rate varies from group to group, but is typically lower than expected based on thermal decomposition rates for GaN. Recent results[12] also indicate that a significant increase in Ga flux can be required in order to obtain Ga-stabilized growth as the temperature is increased above 700 °C, suggesting a significant increase in Ga desorption. We present evidence that the observed decrease in growth rate and increase in Ga desorption are linked to the large reactivity of atomic nitrogen. Our results indicate that metastable molecular nitrogen may be the preferred active nitrogen specie for both growth at higher temperature and for improved electrical properties.

II. EXPERIMENTAL DETAILS

The GaN layers for this study were grown at West Virginia University by MBE in a custom system. A standard MBE source provided the Ga flux. Two rf plasma sources were used to produce active nitrogen, an Oxford Applied Research CARS-25 and an EPI Vacuum Products Unibulb source. The Oxford source featured a removable aperture plate allowing investigation of various hole configurations while maintaining the same overall conductance. The EPI source had a 400-hole aperture with an approximately 50% increase in conductance over the Oxford configuration. Characterization of these sources has been extensively reported elsewhere.[13] In brief, the sources were in direct line of sight to an Extrel quadrupole mass spectrometer whose repeller grid was biased separately to determine ion energies. An electrostatic plate could also be used to completely deflect ions out of the flux. The ionizer energy was set high enough to ionize both molecular and atomic nitrogen (typically >15 eV), but below the dissociation threshold of molecular nitrogen (<28 eV). During the EPI source characterization, considerable molecular nitrogen ions were produced with ionization energies approximately 6 eV lower than normally necessary to ionize molecular nitrogen. This energy corresponds to the $A\,^3\Sigma_u^+$ metastable state of the nitrogen molecule,[14] indicating the EPI source produces a significant flux of molecular nitrogen metastables. The observation of metastables in the EPI source flux is consistent with the previous spectroscopic study indicating that excited molecular nitrogen is generated in the plasma inside the source.[15]

Atomic hydrogen was produced using a thermal cracker (EPI-AHS). Typically, 1×10^{-6} Torr beam equivalent pressure (BEP) of hydrogen was passed through the thermal source operating at 9.5 A, although the hydrogen flux was varied for several samples. Literature supplied with the source indicated that dissociation efficiency was between 5% and 10% for this operating condition. Therefore, the sample was exposed to both atomic and molecular hydrogen during growth. Considering system pumping speed for hydrogen source-to-substrate distance, this gave an atomic hydrogen flux of one monolayer equivalent every 1–3 s.

Determination of relative sample growth rates were performed *in situ* by analyzing interference effects in reflectance measured from the growing layer using 680 nm light from a semiconductor laser. These measurements were converted to

[a]Electronic mail: tmyers@wvu.edu

TABLE I. Typical flux of ions and atomic nitrogen, and the actual incorporation rate into GaN.

	Aperture				
Source	Hole diameter (mm)	Number of holes	Ion flux 10^{13} ions cm^{-2} s^{-1}	N atom flux 10^{15} atoms cm^{-2} s^{-1}	N incorporation in GaN 10^{15} atoms cm^{-2} s^{-1}
Oxford	1.0	9	7.6	4.5	0.58
CARS-25[a]	0.5	37	3.8	3.0	0.26
	0.2	255	2.3	2.3	0.19
EPI 600 W[b]	0.2	400	0.003	0.63	1.9
EPI 300 W[b]	0.2	400	0.001	0.28	1.1

[a]600 W, 6 sccm.
[b]2 sccm.

an absolute growth rate by using total sample thickness and the growth time. The total thickness was determined from interference fringes in *ex situ* optical transmittance measurements using a Cary-14 spectrophotometer. This method of determination of total thickness was found to agree with values determined by transmission electron microscopy.

Desorption mass spectroscopy (DMS) was also performed during growth, primarily to observe the desorbed Ga flux. A differentially pumped UTI Model 100C quadrupole mass spectrometer was placed in direct line of sight with the growing layer at normal incidence. The field of view was limited to the center of the samples using a series of apertures. All samples were grown on *c*-plane sapphire substrates (Union Carbide Crystal Products). Prior to growth, the substrates were degreased and etched in a phosphoric/sulfuric (1:3) acid mixture heated to 140 °C. Based on our earlier study,[6] buffer layers were grown by heating the substrate to 730 °C under an atomic hydrogen flux for 20 min and then cooling to 630 °C for the growth of a 200 Å thick GaN buffer layer under highly Ga-stable conditions. This procedure led exclusively to the nucleation and growth of (000$\bar{1}$1)-oriented (or N-polarity) GaN as determined using the polarity-indicating etch described by Seelmann-Eggebert *et al.*[16]

III. DEPENDENCE OF GROWTH RATE ON ACTIVE NITROGEN SPECIES

Determining the effect of various active nitrogen species on layer growth is complicated when using rf plasma sources. These sources typically produce a complex mixture of active nitrogen superimposed on a background of presumed inert molecular nitrogen. This is illustrated by the source characterization data contained in Table I. The conditions shown are for those resulting in a maximum growth rate, plus one intermediate condition for the EPI source. The ion flux given is for the total ion flux (atomic plus molecular). The atomic ion flux was typically two to three times larger than that of the molecular ions. As reported previously,[17] the maximum ion energy for the Oxford source ranged from ~45 eV for the 9-hole aperture to 10 eV for the 255-hole aperture. The relatively insignificant ion flux from the EPI source had a maximum energy of about 3 eV.

Also included in Table I is the measured incorporation rate of nitrogen into the growing GaN (based on growth rates). Several observations can be made. Neither source produces enough ionic nitrogen to account for the observed growth rate, indicating that growth is due to other nitrogen species. The Oxford source configurations studied produced primarily atomic nitrogen, with little indication of the presence of molecular metastables based on the electron energies required for molecular ionization. Our study with this source indicates that atomic nitrogen is relatively inefficient for growing GaN, requiring about ten atoms in the flux for each one incorporated into the growing layer. In contrast, the EPI source configuration used produced significantly less atomic nitrogen and yet gave a factor of 3–5 increase in growth rate. These results, coupled with the ionization evidence for molecular metastables during source characterization, indicate that metastable molecular nitrogen is the dominant active nitrogen specie for our EPI source configuration.

Figure 1 illustrates the relative growth rate of GaN as a function of temperature for various configurations of the EPI and Oxford sources. Growth using a predominantly atomic nitrogen flux (the Oxford source) led to the early onset of decreased growth rate as reported previously.[18] Reported rates for Ga desorption from GaN surfaces indicate that Ga

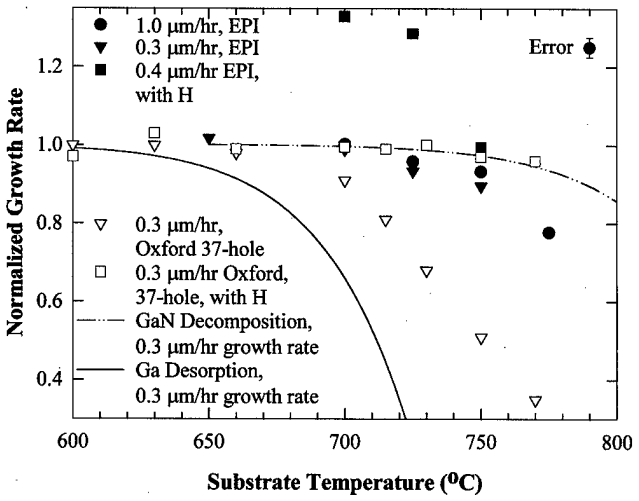


FIG. 1. Relative growth rates of GaN for various conditions. The growth rates were normalized to the low temperature rates. The rates for samples grown with atomic hydrogen were normalized to the low temperature rates of samples grown without atomic hydrogen but identical in other respects.

TABLE II. Mechanisms occurring during the growth of GaN.

Growth	Competition to growth
$Ga+\frac{1}{2}N_2^+ + e^- \rightarrow GaN$	$GaN \rightarrow Ga\uparrow + \frac{1}{2}N_2\uparrow$ (Decomposition)
$Ga+\frac{1}{2}N_2^* \rightarrow GaN$	$Ga_{adsorbed} \rightarrow Ga\uparrow$
$Ga+N \rightarrow GaN$	$GaN+N \rightarrow Ga+N_2\uparrow$
$Ga+N^+ + e^- \rightarrow GaN$	$GaN+N^+ + e^- \rightarrow Ga+N_2\uparrow$

desorption may play significant role over this temperature range.[19] To illustrate this, the temperature dependent growth rate decrease expected if the only contributing factor was the increased Ga-desorption rate is also shown in Fig. 1. This trend is also consistent with a recent study[12] using reflection high energy electron diffraction (RHEED) which indicated a rapidly increasing Ga flux is required to maintain Ga-stabilized conditions for rf plasma MBE growth above 700 °C. However, Ga-desorption alone is not the origin of the decreasing growth rate as increasing the Ga overpressure does not overcome the decreased growth rate for a given temperature.[6,18]

The relative growth rate versus temperature is shown for two operating conditions of the EPI source. For Ga-stable conditions and growth rates comparable to the Oxford source, the decrease in growth rate is now shifted to a higher temperature. A similar trend is observed for growth at 1 μm/h. Shown for comparison is the growth rate dependence on temperature expected if the only contributing factor was GaN decomposition.[19] The decrease in growth rate is now more comparable to the decomposition rate and is similar to that reported for ammonia-based MBE.[20,21] The results with the EPI source indicate that Ga desorption is not the dominant limiting factor, and that the role of active nitrogen species must be analyzed.

Table II contains a subset of possible reactions occurring during Ga-stable growth. The first column lists reactions leading to growth, while the second column details reactions in competition with growth. Of particular relevance is that ionic and neutral atomic nitrogen can participate both in the growth and in the decomposition of GaN. While rate constants for these reactions are not known there is a significant driving force based on free energy considerations.[14] This may explain the relatively poor efficiency for growth with atomic nitrogen indicated in Table I. Competition between growth, surface decomposition, and adsorbed nitrogen capture may limit the efficacy of atomic nitrogen. Such a situation would promote point defect formation, supported by the poor electrical properties discussed later. The decrease in growth rate observed at 775 °C for the EPI source may also be related to the residual atomic nitrogen flux.

Another interesting scenario[22] has been proposed suggesting a mechanism allowing improved growth using excited molecular nitrogen. Both atomic and metastable molecular nitrogen contain significantly more energy than required for GaN formation.[14] Incorporation of atomic nitrogen releases this energy into the lattice where it can drive unfavorable reactions. In contrast, the excited molecule can incorporate one atom into growing GaN while the other desorbs, carrying away the excess energy.

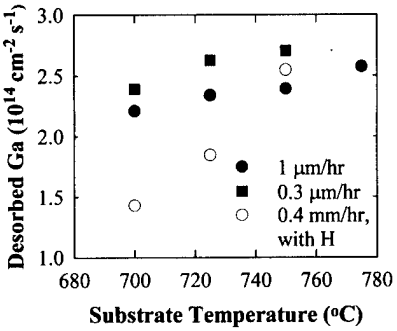


FIG. 2. Desorbed Ga flux vs temperature for two different growth rates and for growth under an atomic hydrogen flux.

DMS was used to monitor reflected Ga flux during growth with the EPI source. Reduction of the desorbed Ga flux to values less than $0.3\times10^{14}\,cm^{-2}\,s^{-1}$ led to N-stable growth, as indicated by the RHEED pattern switching from a streaky, two-dimensional (2D) pattern to a spotty 3D pattern. In general, we maintained Ga-stable conditions during growth with a desorbed Ga flux between 0.5 and $1\times10^{14}\,cm^{-2}\,s^{-1}$. Figure 2 shows the temperature dependence of the desorbed Ga flux for different growth conditions measured with DMS. In obtaining the data shown in Fig. 2, the Ga- and nitrogen-source operating conditions were held constant. An approximate 20% increase in desorbing Ga flux was observed between 700 and 780 °C. While this correlates well with the observed decrease in growth rate, the observed increase in Ga desorption is significantly less than the rate indicated in Fig. 1. This gives further evidence that while Ga desorption plays a role in GaN growth, it is not a significant contributor for Ga-stable growth at these temperatures.

Figure 3 compares the electrical properties of samples grown with the various source configurations. As shown in Fig. 3(a), a significant increase in mobility occurred when using the EPI source, accompanied by a significant decrease in carrier concentration as indicated in Fig. 3(b). The results for the Oxford source are comparable to most values reported for rf plasma MBE, while the EPI results are consistent with improved electrical properties also observed by other groups[4,5,23] using a similar source configuration. Further indication of improvement in material quality is the observation of free excitonic transitions in preliminary photoluminescence measurements made on our GaN grown with the EPI source. Our current study indicates that growth with predominantly atomic nitrogen may result in significant point defects limiting layer quality. Indeed, our highest mobility values with the Oxford source were obtained for growth under a hydrogen flux, which may stabilize the growing surface as discussed later. Although lower carrier concentrations could be obtained with the Oxford source for growth without hydrogen, the accompanying mobilities were also significantly smaller.

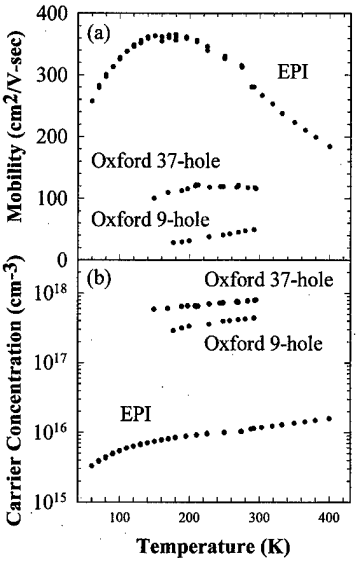


FIG. 3. (a) Mobility and (b) *n*-type carrier concentration for GaN grown using an EPI Unibulb source and an Oxford CARS-21 source. The results for the Oxford source are for samples grown under an atomic hydrogen flux.

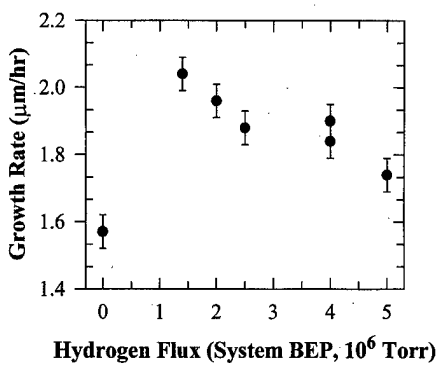


FIG. 4. Effect of atomic hydrogen on the growth rate of GaN when using the EPI source.

IV. EFFECTS OF ATOMIC HYDROGEN

The dramatic effect of atomic hydrogen on the growth of GaN by rf plasma MBE reported in previous studies[17,18] can now be better understood, at least for N-polarity growth. A bulk terminated surface would have a nitrogen dangling bond at the surface. This is normally accommodated by the formation of a surface Ga adlayer, as described by Smith *et al.*[24] The dangling bond, however, would make this surface more prone to decomposition through attack by atomic nitrogen. The rapid decrease in growth rate with increasing temperature may be related to the thermal activation of this reaction. If atomic hydrogen is available, it can also attach to the dangling bond thereby passivating the surface. The presence of a N–H bond would prevent the enhanced decomposition due to attack by atomic nitrogen. This scenario is supported by the recovery of the growth rate for the Oxford source when using atomic hydrogen as shown in Fig. 1. At the highest temperature investigated, growth with atomic hydrogen results in a growth rate about a factor of 3 larger than growth without atomic hydrogen. Growth under atomic hydrogen using the Oxford source appeared to be N stable as indicated by a spotty 3D RHEED pattern, the growth characteristics of inversion domains, and Ga-limited growth rates.[6,18] The recovery to the 0.3 μm growth rate indicated in Fig. 1 is reflective of the Ga-limited nature coupled with the fact that the Ga flux was kept constant for this comparison. Increasing the Ga flux led to steadily increasing growth rates until Ga condensation occurred at the temperatures investigated, to a maximum growth rate of 0.5 μm/h in our study. However, the relatively poor electrical properties of these growths resulted in our efforts being focused elsewhere.

The effect of atomic hydrogen on growth with the EPI source was significantly different, possibly reflecting the difference in the predominant active nitrogen specie. Rather than recovering the growth rate for a fixed Ga flux, an increase in the growth rate was observed for the same growth conditions as the 0.3 μm case, as indicated by relative growth rates larger than unity in Fig. 1. Interestingly, the increase at the lower temperature was approximately the same magnitude as the measured atomic nitrogen flux. The effect of the atomic hydrogen on growth rate became less for higher temperatures, with the growth rates being comparable with and without atomic hydrogen at the highest temperature investigated. Figure 2 also contains the desorbed Ga flux observed for growth under atomic hydrogen. Since the same source conditions were used as the 0.3 μm case, the reduced Ga desorption reflects the increase in growth rate. Increasing the substrate temperature led to a larger increase in desorption rate than for growth without atomic H.

The effect of increasing the atomic hydrogen flux for fixed growth conditions is shown in Fig. 4. Here, the base conditions lead to a GaN growth rate of about 1.6 μm/h. Addition of an atomic hydrogen flux increased the rate to more than 2 μm/h. Further increase in the atomic hydrogen flux then led to a steadily decreasing growth rate. The BEP shown reflect the total hydrogen flux. Based on published cracking efficiencies for this type of source along with relative ion gauge sensitivities, we estimate that 1×10^{-6} Torr BEP corresponds to approximately 0.5–1 monolayer/s equivalent of atomic H. The temperature is large enough for N–H bond breaking coupled with hydrogen desorption, and so any "passivation" effect must be dynamic. It is conceivable that the interaction of atomic hydrogen with the growing GaN surface is a complex situation. A relatively low concentration may protect the growing surface from attack by atomic nitrogen, while at higher concentrations the atomic hydrogen begins to compete with Ga for nitrogen bonds, resulting in a higher probability for Ga desorption. In addition, there is some evidence that hydrogen may even enhance the surface decomposition of GaN at higher temperatures, such as those used for MOCVD growth.[25,26] Since hydrogen is used in most successful growth approaches for GaN, further study is warranted.

V. CONCLUSIONS

The flux from rf plasma sources is a complex mixture of ionic and neutral atomic and molecular nitrogen. Studies of

growth rate as a function of temperature suggest the GaN surface is prone to "attack" by neutral and ionic atomic nitrogen above 700 °C, promoting decomposition. Growth using neutral metastable molecular nitrogen results in a temperature-dependent growth rate similar to that of growth with ammonia. Hydrogen can be used to stabilize the growing surface, at least for N-polarity growth. Too much hydrogen may promote Ga desorption, possibly due to competition for N bonds. Growth with predominantly metastable nitrogen also resulted in improved electrical quality. Metastable or low energy ionic molecular nitrogen may be preferable to neutral or ionic atomic nitrogen for MBE growth.

ACKNOWLEDGMENTS

The authors would like to thank L. S. Hirsch for assisting with GaN growth, and M. Moldovan for performing preliminary photoluminescence measurements. They would also like to acknowledge continuing useful conversations with P. I. Cohen, R. M. Feenstra, N. Newman, J. S. Speck, and R. P. Muller. This work was supported by ONR Grant No. N00014-96-1-1008 and monitored by Colin E. C. Wood.

[1]T. D. Moustakas, T. Lei, and R. J. Molnar, Physica B **185**, 36 (1993).
[2]M. Smith, G. D. Chen, J. Z. Li, J. Y. Lin, H. X. Jiang, A. Salvador, W. K. Kim, O. Aktas, A. Botchkarev, and H. Morkoç, Appl. Phys. Lett. **67**, 3387 (1995).
[3]E. J. Tarsa, B. Heying, X. H. Wu, P. Fini, S. P. DenBaars, and J. S. Speck, J. Appl. Phys. **82**, 5472 (1997).
[4]J. M. Reifsnider, D. W. Gotthold, A. L. Holmes, and B. G. Streetman, J. Vac. Sci. Technol. B **16**, 1278 (1998).
[5]M. A. L. Johnson, J. D. Brown, N. A. El-Masry, J. W. Cook, Jr., J. F. Schetzina, H. S. Kong, and J. A. Redmond, J. Vac. Sci. Technol. B **16**, 1282 (1998).
[6]Z. Yu, S. L. Buczkowski, N. C. Giles, T. H. Myers, and M. R. Richards-Babb, Appl. Phys. Lett. **69**, 2731 (1996); S. L. Buczkowski, Z. Yu, M. R. Richards-Babb, N. C. Giles, T. H. Myers, and L. T. Romano, Mater. Res. Soc. Symp. Proc. **449**, 197 (1997).
[7]H. Morkoc, A. Botchkarev, A. Salvador, and B. Sverdlov, J. Cryst. Growth **150**, 887 (1995).
[8]H. Reichert, R. Averbeck, A. Graber, M. Schienle, U. Straub, and H. Tews, Mater. Res. Soc. Symp. Proc. **449**, 149 (1997).
[9]R. C. Powell, N.-E. Green, Y.-W. Kim, and J. E. Greene, J. Appl. Phys. **73**, 189 (1993).
[10]S. Einfeldt, U. Birkle, C. Thomas, M. Fehre, H. Heinke, and D. Hommel, Mater. Sci. Eng., B **50**, 12 (1997).
[11]T. G. Anderson, K. Nozawa, and Y. Horikoshi, J. Cryst. Growth **175-176**, 117 (1997).
[12]A. R. Smith, V. Ramachandra, R. M. Feenstra, D. W. Greve, A. Ptak, T. Myers, W. Sarney, L. Salamanca-Riba, M. Shin, and M. Skowronski, MRS Internet J. Nitride Semicond. Res. **3**, 12 (1998).
[13]"Characterization and modeling of an rf plasma source of active nitrogen for GaN growth," M.S. thesis in Chemical Engineering, Sumeet Kumar.
[14]N. Newman, J. Cryst. Growth **178**, 102 (1997); N. Newman, *Semiconductors and Semimetals 50* (Academic, New York, 1998), pp. 55–101.
[15]M. A. L. Johnson, Zhonghai Yu, C. Boney, W. C. Hughes, J. W. Cook, Jr., J. F. Schetzina, H. Zho, B. J. Skromme, and J. A. Redmond, Mater. Res. Soc. Symp. Proc. **449**, 215 (1997).
[16]M. Seelmann-Eggebert, J. L. Weyher, H. Obloh, H. Zimmermann, A. Rar, and S. Porowski, Appl. Phys. Lett. **71**, 2635 (1997).
[17]A. J. Ptak, K. S. Ziemer, M. R. Millecchia, C. D. Stinespring, and T. H. Myers, MRS Internet J. Nitride Semicond. Res. **4S1**, G3.10 (1999).
[18]T. H. Myers, L. S. Hirsch, L. T. Romano, and M. R. Richards-Babb, J. Vac. Sci. Technol. B **16**, 2261 (1998).
[19]D. D. Koleske, A. E. Wickenden, R. L. Henry, W. J. DeSisto, and R. J. Gorman, J. Appl. Phys. **84**, 1998 (1998).
[20]N.-E. Lee, R. C. Powell, Y.-W. Kim, and J. E. Greene, J. Vac. Sci. Technol. A **13**, 2293 (1995).
[21]R. Held, D. E. Crawford, A. M. Johnson, A. M. Dabiran, and P. I. Cohen, J. Electron. Mater. **26**, 272 (1997).
[22]R. P. Muller, Beckman Institute, Caltech, Pasadena, CA (private communication).
[23]J. S. Speck, Materials Department, University of California—Santa Barbara, private communication.
[24]A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer, and J. E. Northrup, Phys. Rev. Lett. **79**, 3934 (1997).
[25]D. D. Koleske, A. E. Wickenden, R. L. Henry, M. E. Twigg, J. C. Culbertson, and R. J. Gorman, Appl. Phys. Lett. **73**, 2018 (1998).
[26]M. E. Bartram, MRS Internet J. Nitride Semicond. Res. **4S1**, G3.68 (1999).

Nitridation of the GaAs(001) surface: Thermal behavior of the (3×3) reconstruction and its evolution

J. Lu,[a] L. Haworth, P. Hill, D. I. Westwood, and J. E. Macdonald
Department of Physics and Astronomy, University of Wales, Cardiff, P.O. Box 913, Cardiff CF2 3YB, United Kingdom

(Received 19 January 1999; accepted 27 May 1999)

The nitridation of the GaAs(001) surface using a radio frequency atomic nitrogen plasma source in a molecular beam epitaxy growth chamber has been studied. The resulting nitrogen induced GaAs(001) (3×3) reconstruction was investigated by *in situ* reflection high energy electron diffraction and x-ray photoemission spectroscopy (XPS). It was found that this reconstruction is only obtained in the temperature range 400–580 °C with a very low dose of atomic nitrogen. The nitrogen coverage corresponding to the (3×3) reconstruction was determined by quantitative XPS to be 0.30±0.09 ML. Below 400 °C an As–N species of disordered structure was found on the GaAs(001) surface. Subsequent annealing at about 500 °C produced the (3×3) reconstruction. Above 580 °C, nitridation lead to direct formation of β-GaN islands. In addition, the (3×3) reconstruction was found to be unstable to both exposure to atomic hydrogen and annealing. The N desorption activation energy of the (3×3) was estimated to be 2.75±0.55 eV. A surface phase diagram of the (3×3) has thus been deduced. © *1999 American Vacuum Society.*
[S0734-211X(99)08804-6]

I. INTRODUCTION

GaN has become the most interesting wide-band-gap semiconductor material of recent years, mainly due to its rapid development in the optoelectronics industry for the manufacture of blue light emitting diodes[1] and blue laser diodes,[2] but also for its promising future for high frequency, high temperature and high power device applications.[3] GaN has two major polytypes, viz. the stable wurtzite phase (α-GaN) and the metastable zincblende phase (β-GaN). Most of the achievements so far in device applications have been made using α-GaN grown on sapphire substrates. Meanwhile, the zincblende GaN (β-GaN polytype) grown on cubic substrates has drawn attention to its potential advantages over the conventional α-GaN. For example, β-GaN is theoretically predicted to have a higher electron mobility and the device fabrication will be simplified by using substrates such as GaAs,[4–10] which have well-established substrate and device processing procedures and a convenient cleavage plane for laser facet. Nonetheless, due to the large lattice mismatch between β-GaN and GaAs(001) (~20%), as with the heteroepitaxial growth of the wurtzite GaN, the strain relief mechanism and the interface configuration in the initial stages of growth play a key role in the whole growth process.

A (3×3) reconstruction has been observed during the early stage of β-GaN/GaAs(001) molecular beam epitaxy (MBE) growth[5–7] or when clean GaAs(001) is briefly exposed to an atomic N flux.[11–15] The (3×3) reconstructions were obtained at substrate temperatures between 500 and 600 °C in these reports. The (3×3) could persist for about 1 min at 600 °C[11] and under an As_2 flux gradually changes back to the original clean GaAs (2×4) As terminated reconstructions.[12] Gwo *et al.*[13] have proposed an atomic model of the (3×3) which consists of nitrogen dimer rows with a missing row along both ⟨110⟩ directions, implying a N coverage of 4/9 ML. However this model has been questioned[14] due to the large distortions required by the formation of nitrogen dimers, which make the very structure energetically unfavorable. In order to better understand the (3×3) reconstruction, a more comprehensive and systematic investigation has been undertaken to verify its formation temperature depedence and its nitrogen coverage.

II. EXPERIMENT

The GaAs(001) nitridations were performed in a home-built MBE growth chamber with a base pressure of about $\sim 5\times10^{-10}$ mb. Atomic N and H were supplied by an Oxford Applied Research CARS-25 radio frequency plasma source. Reflection high energy electron diffraction (RHEED) with a 15 kV electron gun was used for *in situ* surface structure monitoring. For surface chemical compositional analysis we used a VG ESCALAB Mk II x-ray photoemission spectrometer (XPS) equiped with Al/Mg twin anodes (which give photon energies of 1486.6 and 1253.6 eV for the Al and Mg, respectively). The ESCALAB is connected to the growth chamber via an ultrahigh vacuum (UHV) gate valve. The work function of the XPS spectrometer was calibrated using the Au $4f_{7/2}$ peak at 83.8 eV.

An undoped GaAs (001) wafer was initially cleaved into 1×1 cm square sample pieces. The sample was etched and mounted onto a molybdenum spade with indium prior to being loaded into the growth chamber. Atomic hydrogen was used to clean the GaAs(001) surface at 400 °C for 5 min using a radio frequency (rf) power of 400 W, a chamber pressure of 5×10^{-5} mb, and a plasma intensity measured by an optical emission detector (OED) of about 1.0 V. After this cleaning procedure, a sharp streaky GaAs(001) (2×4)

[a]Electronic mail: luj@cardiff.ac.uk

RHEED pattern was obtained, and surface oxygen and carbon contamination were below the XPS detectable limit, indicating that the surface was atomically flat and clean.

Since the (3×3) reconstruction occurs for very small doses of atomic nitrogen, in order to better control the experiments, the nitridation process was intentionally slowed by operating with the N plasma source shutter closed. Since the shutter is not 100% efficient, a small but definite amount of N still reaches the sample. The rf power for the N plasma was 200 W, and the N_2 flow rate was 1.0 sccm, corresponding to a chamber pressure of 5×10^{-5} mb, giving a plasma intensity measured by the OED of 1.0–1.2 V. With the shutter open these conditions would give an atomic N flux of about 1×10^{14} atoms/cm^2 s, deduced from the thickness of stoichiometric grown β-GaN film measured by a scanning electron microscope. With the shutter closed, it was estimated that approximately 5%–10% of this N flux reaches the sample. The substrate temperature T_s was measured by a thermocouple positioned between the heater and samples, calibrated regularly by an infrared pyrometer with estimated errors of ±5 °C.

III. RESULTS AND DISCUSSIONS

A. RHEED observation

1. Temperature dependence of the (3×3)

Under UHV, the atomic hydrogen cleaned GaAs(001) surface had an As terminated (2×4) reconstruction at 400 °C or lower temperatures [Fig. 1(a)]. When the temperature was increased to 450–580 °C the (2×4) would change to the (4×6) Ga rich reconstruction. At T_s higher than 580 °C a slightly spotty nonreconstructed GaAs (1×1) pattern indicated the surface was roughened.

For nitridation at substrate temperatures T_s higher than 580 °C, the GaAs (1×1) nonreconstructed RHEED pattern first weakened, then after 40–60 s of nitridation new dispersed spots with different spacing from that of GaAs(001) appeared [Fig. 1(b)]. Comparison of the spacings of these spots with those of GaAs(001) leads to the conclusion that these spots correspond to β-GaN islands with the crystallographic relationship GaN[001]‖GaAs[001], GaN[110]‖GaAs[110]. The (3×3) reconstruction was not observed in this temperature range.

For T_s in the range of 400–580 °C, (3×3) patterns were observed for nitridation times in the range of 20–40 s [Fig. 1(c)] the (3×3) first appeared after ~20 s increased to a maximum intensity at ~30 s before decreasing, and finally disappeared at ~40 s. At 500 °C further nitridation for 10 min resulted in a nonreconstructed weak bulk GaAs(001) (1×1) pattern with high background, suggesting a disordered surface structure. Nitridation for 10–20 min at this temperature finally led to a broad weak streaky β-GaN pattern as shown in Fig. 1(d), indicating the formation of a relatively flat β-GaN surface which was still disordered to some extent.

Nitridation at room temperature up to 400 °C led to nonreconstructed GaAs (1×1) streaks with a hazy background, again indicating a disordered surface. Although the (3×3) reconstruction could not be obtained directly, following 40 s of nitridation in this low T_s regime, annealing at 500 °C did result in a nice sharp (3×3) RHEED pattern.

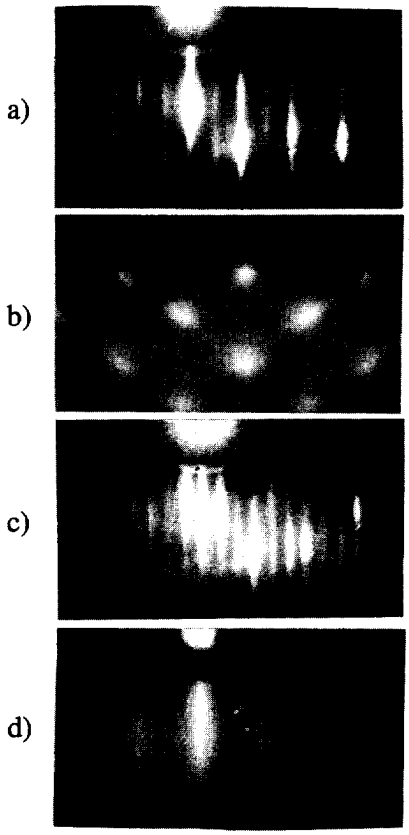


FIG. 1. RHEED pattern during the nitridation of GaAs(001). Only [110] azimuth patterns are shown here. (a) The (2×4) reconstruction of the atomic hydrogen cleaned GaAs(001) surface (b) nitrided at 580 °C, β-GaN spots indicating the formation of β-GaN islands. (c) The (3×3) obtained by slight nitridation at temperature 580 °C. (d) After 20 min nitridation at 500 °C, broad β-GaN streaks.

2. Stability of the (3×3)

The stability of the (3×3) reconstruction was tested by exposing it to atomic hydrogen. Within a few seconds of atomic H exposure at 400 °C, the (3×3) would give way to a (2×4) reconstructed pattern, which was even sharper and brighter than that of the original atomic H cleaned (2×4) pattern, indicating that the surface N species were removed and the clean GaAs(001) surface was recovered. The removal of the surface N atoms was also confirmed by the absence of a N 1*s* peak in XPS spectra. In fact, it was possible to perform multiple cycles of nitridation and atomic H cleaning of the (3×3) on one sample. By contrast, with the same N dose at higher temperature (580 °C) so that the β-GaN islands formed, its spotty RHEED pattern would remain unchanged even after exposing it to atomic H for as long as 20 min at 400 °C, suggesting that Ga–N bonds in those β-GaN islands are stronger.

The removal of surface N in the (3×3) by atomic hydrogen might involve the following surface reaction:

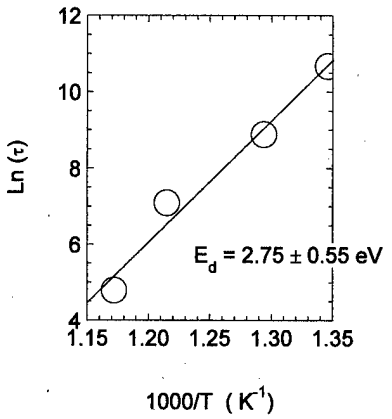


FIG. 2. Thermal stability of the (3×3). τ is the lifetime of the (3×3), and T is the substrate temperature in K. The experimental data points are fitted by a straight line and the gradient of this line gives the N desorption activation energy E_d=2.75±0.55 eV.

$$\text{Ga–N (surface)} + 3\text{H} \Rightarrow \text{Ga} + \text{NH}_3 . \quad (1)$$

Considering the strength of the bulk Ga–N bond of 405 kJ/mol (4.2 eV for an average single bond)[16] and H–N bond strength of 339 kJ/mol (3.5 eV for a single bond);[17] it is energetically unlikely for atomic H to break bulk-like Ga–N bonds and combine with N atoms. However, the surface Ga–N bond in the (3×3) appeared to be weaker than the bulk Ga–N bond, and is probably weaker than the H–N bond, so that it is more likely to be removed by atomic hydrogen.

The stability of the (3×3) was further investigated by annealing the (3×3) at different temperatures. RHEED observation showed that the lifetime of the (3×3) decreased quickly as the annealing temperature increased. Annealing at temperatures of 470 and 500 °C would lead to the return of the clean GaAs(001) (4×6) reconstructions, suggesting that the surface N in the (3×3) was desorbed. However, in the case of annealing at 550 and 580 °C, very weak β-GaN RHEED spots together with (1×1) GaAs(001) bulk streaks could be observed following the decay of the (3×3), indicating the nucleation and growth of small β-GaN islands.

The thermal stability of the (3×3) can be understood as follows. In general, the loss of the (3×3) could be caused by evaporation (desorption) of the N adatoms, or the migration (surface diffusion) of the N adatoms to some energetically favorable nucleation sites to form GaN islands with stronger Ga–N bonds. At medium temperatures (475 and 500 °C), the desorption of N adatoms dominates, whereas at higher temperature (550 or 580 °C), surface diffusion is fast enough to allow nucleation of β-GaN to also occur. Despite the complexity of the process occuring the lifetime of the (3×3) with temperature could be described by a simple Arrhenius equation

$$\tau = \tau_0 \exp(E_d/kT), \quad (2)$$

where τ is the lifetime and E_d is a thermal activation energy for evaporation of N. The experimental data and fitted line is plotted in Fig. 2, where the lifetime was measured by observing the (3×3) RHEED pattern until it returned to clean GaAs(001) patterns. The fitted thermal activation energy E_d is 2.75±0.55 eV, which is larger than the (3×3) activation energy of 2.1 eV obtained by Bandic *et al.*[18] for similar experiments but with an As_2 flux, but still lower than the desorption activation energy of bulk α-GaN of 3.24 eV.[19]

TABLE I. Inelastic mean free paths of some core levels in GaAs.

Core levels	As $2p_{3/2}$	Ga $2p_{3/2}$	Ga $3d$	N $1s$
X-ray source	Al	Al	Mg	Mg
Electron kinetic energy (eV)	163.6	369.6	1234.6	856
IMFP(Å)	6.6	10.7	25.9	19.6

B. XPS study

1. Quantification of the N coverage

XPS was used to analyze the chemical composition of the (3×3) and further nitrided GaAs(001) surfaces. In order to verify the possible XPS surface charging effect on the undoped GaAs(001) sample, an XPS spectrum of a undoped GaAs(001) sample was compared with that of a heavily doped GaAs(001) sample. No detectable spectrum shift was found, indicating that the surface charging effects could be neglected in our XPS measurements.

As $2p_{3/2}$, Ga $2p_{3/2}$, Ga $3d$, and N $1s$ core levels were used for XPS analysis. The XPS surface sensitivity of these core levels can be described by the inelastic mean free paths (IMFP) of photoelectrons from these core levels. We used the Mg anode for Ga $3d$ and N $1s$, and the Al anode for As $2p$ and Ga $2p$. The theoretically calculated IMFPs[20] are listed in Table I.

Quantitative XPS was carried out to determine the N coverage using Ga $3d$ and N $1s$ core level peak areas with an XPS analyzer electron pass energy of 50 eV. This large pass energy was employed due to the low coverage of nitrogen. Even so it was not simple to determine the N $1s$ peak area due to its coincidence with a small Ga Auger *LMM* peak and close proximity to another two Ga Auger *LMM* features labeled $A2$, $A1$, and $A3$, respectively, in Fig. 3. Careful background subtraction was therefore required and two approaches were used and compared. The first was to match the $A1$ and $A3$ profile of a clean GaAs(001) sample with those of nitrided samples by application of a multiplying factor to the spectrum of the clean sample. This modified spectrum of clean GaAs(001) was then subtracted from spectra of nitrided samples. In the second approach, the peak area ratio for the two Ga Auger peaks $A2/A3$ for the clean GaAs(001) was first determined to be 2.5±0.2. Next the peak area for a combined N $1s+A2$ peak and the $A3$ peak area of a nitrided sample were measured. Subsequently, the N $1s$ peak area would be the combined N $1s/A2$ peak area subtracted by ($A3$ area)/2.5. These two methods gave consistent N $1s$ peak areas within an error of less than 10%.

Figure 4 shows the N $1s$ peaks following background subtraction using the first method for nitridation times in the range 0–5 min at T_s=475 °C. Due to the small signal and coinciding with the Ga Auger feature, it is impossible to

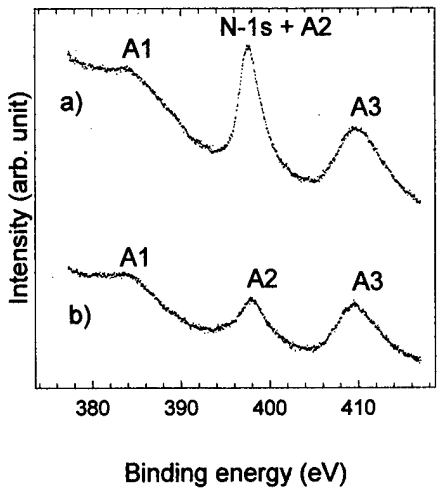


FIG. 3. XPS Ga Auger *LMM* peaks coincident with N 1*s* peak: (a) spectrum of GaAs(001) nitrided at 475 °C for 2 min, (b) clean GaAs spectrum, showing three Ga Auger *LMM* peaks.

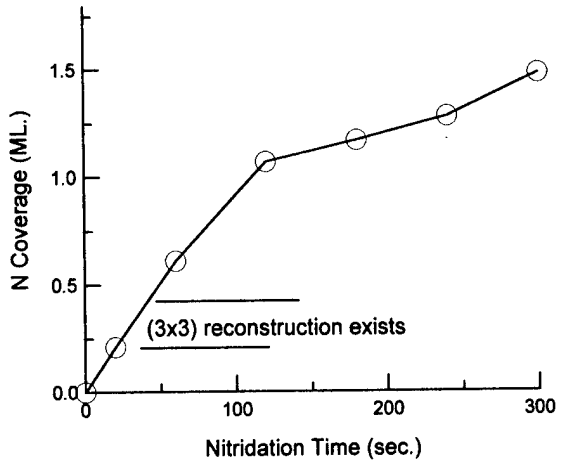


FIG. 5. Nitrogen coverage of nitrided GaAs(001) against nitridation time, derived by quantantive XPS calculation. The nitridation temperature was 475 °C.

precisely determine the peak position. Nevertheless, having compared the N 1*s* peak position of a longer nitrided (>1 min) sample with that of thick β-GaN/GaAs(001) films, no significant chemical shift was found, and the peak full width at half maximum (FWHM) in these two cases are very close.

The N coverages of the GaAs(001) surface were calculated using the approach of Carley and Roberts[21] for submonolayer films. The surface atomic density of the deposited nitrogen can be written as

$$\sigma = I_N \mu_{Ga} T_{Ga} N_A \rho_{GaAs} \lambda_{Ga} \cos\varphi / I_{Ga} \mu_N T_N M_{GaAs}, \quad (3)$$

where I_N and I_{Ga} are the XPS peak areas for N 1*s* and Ga 3*d*, respectively, μ_N and μ_{Ga} are the subshell photonization cross section of the N 1*s* and Ga 3*d*, respectively, for x-ray photon energy of 1253.6 eV (Mg anode) $\mu_N = 0.039 \times 10^6$ b, $\mu_{Ga} = 0.026 \times 10^6$ b,[22] T_N and T_{Ga} are the transmission functions of the XPS analyzer for N 1*s* and Ga 3*d*, respectively (for our XPS analyzer, $T = (E_k)^{-1/2}$,[23] where E_k is kinetic energy of the photoelectrons. So $T_N/T_{Ga} = (E_{k\text{-}N}/E_{k\text{-}Ga})^{-1/2}$), N_A is Avogadro's number, ρ_{GaAs} is the bulk density of the substrate, λ_{Ga} is the IMFP of the Ga 3*d* photoelectron in GaAs ($\lambda_{Ga} = 25.9$ Å from Table I), M_{GaAs} is the molecular weight of GaAs, φ is the angle between detector and sample surface normal, and is zero in our case.

Using the fact that the surface of GaAs (001) corresponds to 6.26×10^{14} atoms/cm^2, the calculated N coverage is plotted in monolayers in Fig. 5. Since the (3×3) reconstruction exists for nitridation times in the range 20–40 s, the (3×3) exists for N coverages in the range from 0.2 to 0.4 ML. The estimated error on these values is ±30%. A nitridation time of 30 s gave the strongest (3×3) and indicates that the N coverage for this reconstruction is most likely 0.30±0.09 ML.

The Carley and Roberts approach used above does not take into account the peak intensity attenuation due to the existence of the overlayer, therefore theoretically it is only valid for submonolayer deposition, and may give erroneous values for over 1 ML films. Therefore calculations considering attenuation caused by inelastic scattering of the photoelectrons were also carried out, all the results were very close to those obtained using Carley and Roberts approach in the N coverage range of 1–1.5 ML, and therefore are not shown in Fig. 5.

A slowing down in nitridation rate can be seen after about 1 ML deposition in Fig. 5. This could be understood as follows. Once the surfaces are bonded to nitrogen, the second layer nitridation must involve both As–N anion exchange reactions. In addition, the 1 ML film already formed on top will hinder the diffusion of incoming N and outgoing As atoms (products of anion exchange), therefore hampering the nitridation reaction.

To summarize this section, XPS N 1*s* and Ga 3*d* peaks were used to quantify the N coverage. Two methods have been used for N 1*s* background subtraction, and they gave consistent results. To calculate the N coverage, an approach

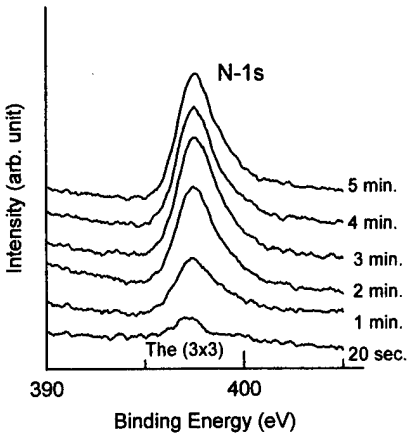


FIG. 4. XPS background subtracted N 1*s* peaks of 475 °C nitrided GaAs(001) for different nitridation times. The (3×3) reconstruction was observed at nitridation time of 20 s.

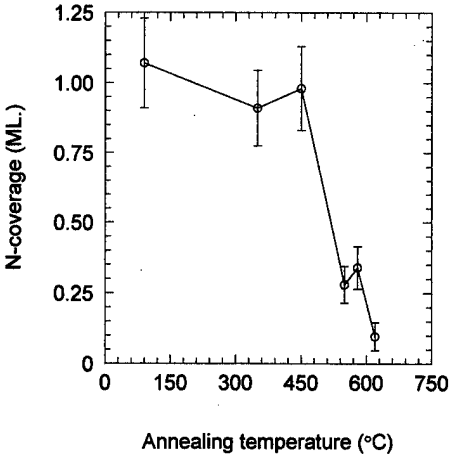


FIG. 6. N coverage against annealing temperatures in experiment A: annealing of 1 ML N covered GaAs(001).

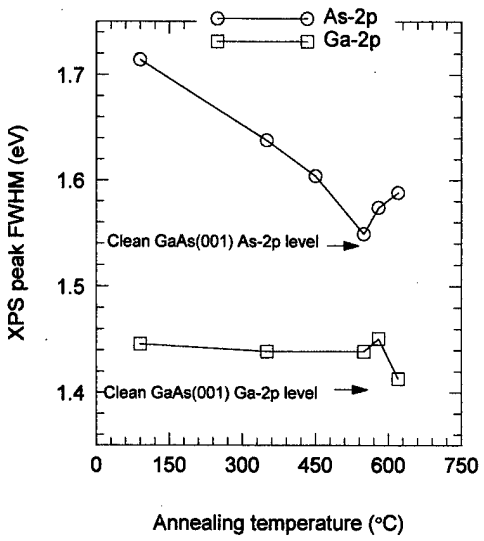


FIG. 7. XPS As $2p_{3/2}$ and Ga $2p_{3/2}$ peak FWHM after annealing in experiment A: annealing of 1 ML N covered GaAs(001).

for the submonolayer and for thicker layers were both used, and are in good agreement within a range of <2 ML. The N coverage for the (3×3) was found to be 0.30±0.09 ML.

2. XPS study of the temperature dependence of nitridation and nitrided films

In order to further investigate the temperature dependence of the nitrided GaAs(001) surface, we carried out two more experiments (due to the change of the geometry of the N source shutter in our system, the N flux was not identical to the previous experiments): (A) Nitridation of GaAs(001) at 90 °C to 1 ML N coverage, followed by annealing for 20 min at increasing temperatures; (B) Nitridation of the GaAs(001) surface for 60 s at different temperatures. XPS scans were performed at room temperature after every anneal or nitridation. XPS peaks were analyzed to determine N coverage and FWHM was measured to analyze the possible change of surface chemical state.

Experiment A: Annealing of 1 ML film. The N coverage measured by XPS against annealing temperature in experiment A is shown in Fig. 6. The N coverage was largely unchanged up to 450 °C but decreased drastically at 500 °C. Since β-GaN RHEED spots were never observed even up to 620 °C, it is clear that most of the N atoms were evaporated at higher temperatures. After the final annealing at 620 °C, a spotty GaAs(001) (1×1) RHEED pattern was obtained. This behavior is consistent with N desorption at 450–550 °C preventing the formation of β-GaN islands and with the RHEED observation of the (3×3) lifetime in this temperature range. Due to the small N coverage, chemically reacted compounds of the substrate As $2p$ and Ga $2p$ peaks cannot be distinguished directly, therefore peak FWHM was used to analyze the possible chemical changes. The As $2p$ and Ga $2p$ peak FWHMs against annealing temperature in experiment A are shown in Fig. 7. After nitridation at room temperature the As $2p$ FWHM was increased, however with annealing at higher and higher temperature it decreased back to that of clean GaAs(001) after annealing at 550 °C. The increased As $2p$ peak FWHM after low temperature nitridation gave evidence of the existence of another chemical state of As, possibly the As–N species which has been reported previously and has ~2 eV higher binding energy.[24] The return to the clean GaAs(001) FWHM after 550 °C annealing is an indicator of the desorption of the N–As species. The relatively small Ga $2p$ peak width change on the other hand, are an inevitable consequence of the smaller chemical shift of Ga $2p$ in the Ga–N species with respect to that in GaAs(001)(0.6 eV).[25] The slight Ga $2p$ FWHM increase at 550 °C might imply a formation of the Ga–N species. It is noticed that there was an increase for As $2p$ FWHM at above 550 °C. We cannot find a satisfactory explanation for this effect to date.

Nitridation to 1 ML at 90 °C also resulted in an energy shift of the whole XPS spectrum, and the value of the shift decreased when the sample was annealed in higher temperatures. This shift may be due to the different band bending effects caused by the surface states of the nitrided sample and by that of clean GaAs surface.

Experiment B: 60 s nitridation at different temperatures. The results of experiment B showed very similar features (Figs. 8 and 9). Again the decrease of N coverage starts at 450 °C, and at ~550 °C there is a N coverage increase, and the As $2p$ and Ga $2p$ peak width have a minimum and maximum, respectively (Fig. 9). The N combination rate (or sticking coefficient) decreased at elevated temperatures (e.g., 500 °C) under this very low N flux. However, it should be noted that under normal growth conditions, the N flux is more than ten times higher, and the Ga–N bond formed during the growth is stronger than that on the GaAs surface, therefore the sticking coefficient is much larger and close to unity up to 680 °C.[26]

In order to clearly show the very existence of the As–N species at different temperatures, a longer time nitridation was carried out and the As $2p_{3/2}$ peak was analyzed. The As $2p_{3/2}$ peaks of samples nitrided at 300 and 500 °C for 20 min are plotted in Fig. 10 with an As $2p_{3/2}$ peak of a clean

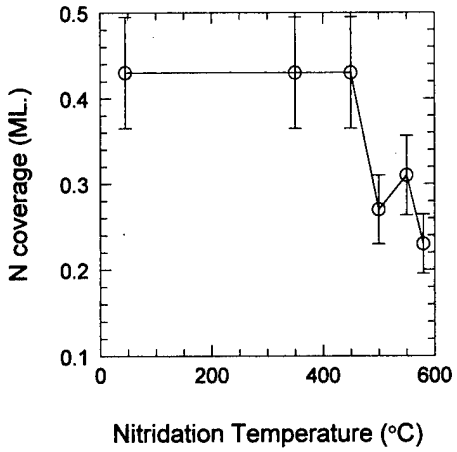


FIG. 8. N coverage in experiment B: GaAs(001) nitride for 60 s at various temperatures.

GaAs(001) sample for comparison. At 300 °C, two peak components can be resolved clearly. Compared with the clean GaAs spectrum, one can see that the first component with lower binding energy is from the Ga–As bonds in the GaAs substrate, the second component of about 2 eV higher binding energy could be assigned to the As–N species in the nitrided layer.[24] By contrast, at 500 °C [well within the (3×3) existing temperature region] the As–N component was small. At the same time the RHEED pattern showed broad β-GaN streaks indicating the formation of β-GaN film.

C. Discussion

From the above RHEED observations and XPS measurements, we can summarize the temperature and N coverage dependence of the (3×3) in a surface phase diagram (Fig. 11). Since there is possibly more than one phase at the same time, and due to the uncertainty of the RHEED observation of development and loss of the (3×3) and the errors in N coverage measurements, some of the boundaries between phases in the diagram are presented as a broad shadowed region. At close to zero N coverage of the diagram, the H cleaned GaAs(001) surface reconstructions under UHV conditions are shown. When slightly nitrided at temperatures in the range of 400–580 °C, these reconstructions would give way to the (3×3). The (3×3) area in the diagram is in a rectangular shape. The strongest (3×3) was found in the middle of the square region. Near both borders of the (3×3) region, this reconstruction was weak as observed by RHEED, and therefore it is possible that the (3×3) reconstruction only occupied a very small proportion of area of the sample surface in these cases. Nitridation beyond the (3×3) regime leads to either spotty β-GaN islands at higher temperature or a hazy high background RHEED pattern at lower temperature.

Gwo's N dimer model[13] is the only published atomic model of the (3×3) to our knowledge. However, the existence of the N dimers have been questioned due to the energetic consideration of the fact that the small bond length between the two N atoms in a N dimer may induce large distortion. The large distortion energy needed makes it ener-

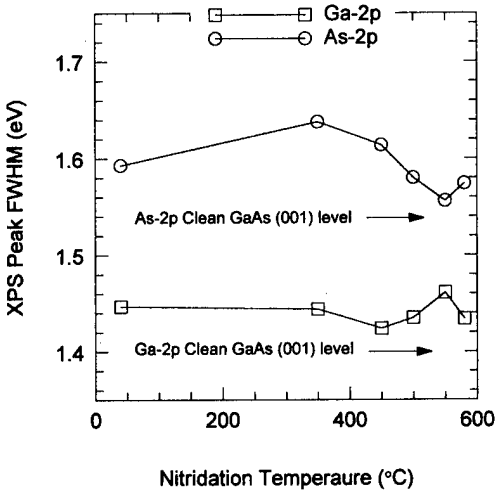


FIG. 9. XPS As $2p_{3/2}$ and Ga $2p_{3/2}$ peak FWHM against nitridation temperatures in experiment B: GaAs(001) nitride for 60 s at various temperatures.

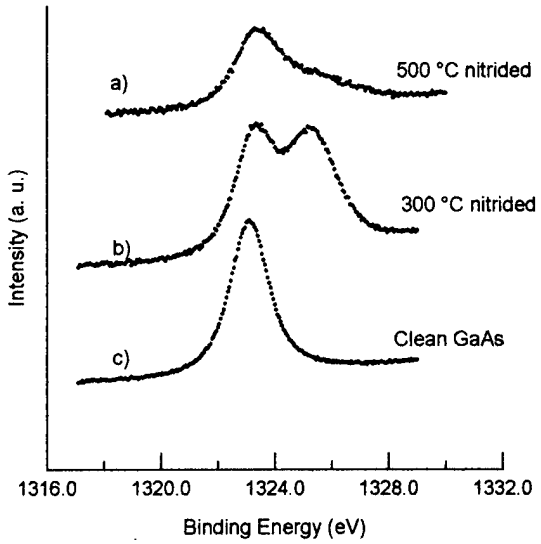


FIG. 10. XPS As $2p_{3/2}$ peaks of: (a) 20 min nitrided at 500 °C, (b) 20 min nitrided at 300 °C, (c) clean GaAs(001) (2×4) surface.

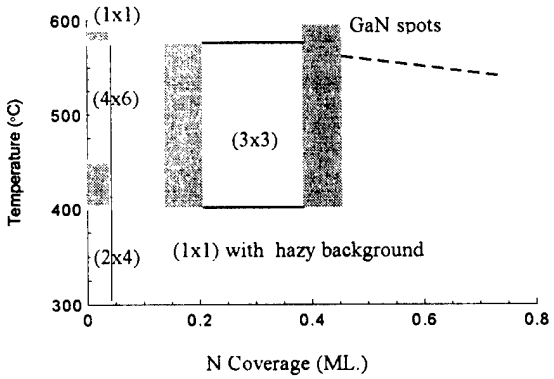


FIG. 11. Phase map of nitrided GaAs(001) surface. The shadow regions in boundary between two phases are due to the uncertain nature of the RHEED observation and the errors in N coverage measurements.

getically unfavorable. Our quantitative XPS showed that when the (3×3) RHEED pattern was strongest, i.e., when maximum proportion of the surface was covered by the (3×3), the N coverage was 0.30±0.09 ML. This N coverage is in the vicinity of, but still significantly smaller than, that derived by Gwo's model (0.44 ML). However, taking into account the possibly unforseen systematic errors the measured N coverage may be taken to moderately support Gwo's model. Jung *et al.*[12] suggested that the formation of the (3×3) is associated with rearrangement and partial desorption of As atoms. Since we obtained the (3×3) by nitriding either the As stabilized (2×4) or the Ga rich (4×6) GaAs(001) surfaces, it seems that the As–N anion exchange reaction or As rearrangement is not necessary to achieve the (3×3). Although our results support the view that the (3×3) consists of weak Ga–N bonds, the exact atomic structure of the (3×3) is still an open question.

IV. CONCLUSIONS

The nitridation of the GaAs(001) surface, particularly the (3×3) reconstruction in the early stage of the nitridation and its thermal stability, has been studied. RHEED studies showed that slight nitridation of GaAs(001) clean surface at a substrate temperature 400–580 °C leads to a N induced (3×3) reconstruction. After nitridation, at temperatures lower than 400 °C, disordered As–N species were found on the nitrided GaAs(001) surface. The (3×3) reconstruction cannot be obtained directly. However, the (3×3) could be obtained by annealing at 450 °C following low temperature nitridation. At temperatures higher than 580 °C, the β-GaN islands formed quickly and no (3×3) reconstuction was observed. The coverage of N for the (3×3) reconstruction was estimated to be 0.2–0.4 ML by quantitative XPS. The strongest (3×3) RHEED pattern, indicating the maximum proportion of the (3×3) on the surface, was shown to occur at N coverage of 0.30±0.09 ML. The N induced (3×3) was vulnerable to atomic hydrogen, and it is unstable at higher temperature due to the desorption of the N adatoms or the nucleation of β-GaN islands. The characteristic thermal activation energy for removal of the (3×3) is estimated to be 2.75 ±0.55 eV.

ACKNOWLEDGMENTS

The authors would like to thank Chris Dunscombe for very helpful discussions on XPS data analysis. They also gratefully acknowledge the financial support of the UK EPSRC.

[1]S. Nakamura, T. Mukai, and M. Senoh, Jpn. J. Appl. Phys., Part 2 **30**, L1998 (1991).
[2]S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, Jpn. J. Appl. Phys., Part 2 **35**, L217 (1996).
[3]M. A. Khan and M. S. Shur, Mater. Sci. Eng., B **46**, 69 (1997).
[4]S. Strite, J. Ruan, Z. Li, A. Szlvador, H. Chen, D. J. Smith, W. J. Choyke, and H. Mordoc, J. Vac. Sci. Technol. B **9**, 1924 (1991).
[5]O. Brandt, H. Yang, B. Jenichen, Y. Suziki, L. Daweritz, and K. H. Ploog, Phys. Rev. B **52**, R2253 (1995).
[6]O. Zsebok, J. V. Thordson, and T. G. Anderson, MRS Internet J. Nitride Semicond. Res. **3**, 14 (1998).
[7]T. S. Cheng, L. C. Jenkins, S. E. Hooper, C. T. Foxson, J. W. Orton, and D. E. Lacklison, Appl. Phys. Lett. **66**, 1509 (1995).
[8]H. Okumura, S. Misawa, and S. Yoshida, Appl. Phys. Lett. **59**, 1058 (1991).
[9]M. Sato, Inst. Phys. Conf. Ser. **142**, 875 (1996).
[10]S. Fujieda and Y. Matsumoto, Jpn. J. Appl. Phys., Part 2 **30**, L1665 (1991).
[11]R. J. Hauenstein, D. A. Collins, X. P. Cai, M. L. O'Steen, and T. C. McGill, Appl. Phys. Lett. **66**, 2861 (1995).
[12]H. D. Jung, N. Kumagai, T. Hanada, Z. Zhu, T. Yao, T. Yasuda, and K. Kimura, J. Appl. Phys. **82**, 4684 (1997).
[13]S. Gwo and H. Tokumoto, Appl. Phys. Lett. **71**, 362 (1997).
[14]Q. Xue, T. Hashizume, J. M. Zhou, T. Sakata, T. Ohno, and T. Sakurai, Jpn. J. Appl. Phys., Part 2 **36**, L1486 (1997).
[15]J. Lu, L. Haworth, P. Hill, D. I. Westwood, and J. E. Macdonald, Thin Solid Films (in press).
[16]Y. Seki, H. Watanabe, and J. Matsui, J. Appl. Phys. **49**, 822 (1978).
[17]*Handbook of Chemistry and Physics*, 77th ed., edited by D. R. Lide (Chemical Rubber, Boca Rotan, FL, 1996), pp. 9–54.
[18]Z. Z. Bandic, T. C. Mcgill, R. J. Hauenstein, and M. L. O'Steen, J. Vac. Sci. Technol. B **14**, 2948 (1996).
[19]R. Groh, G. Gerey, L. Bartha, and J. I. Pankove, Phys. Status Solidi A **26**, 353 (1974).
[20]S. Tanuma, C. J. Powell, and D. R. Penn, Surf. Interface Anal. **17**, 927 (1991).
[21]A. F. Carley and M. W. Roberts, Proc. R. Soc. London, Ser. A **363**, 403 (1978).
[22]J. J. Yen and I. Lindau, At. Data Nucl. Data Tables **32**, 155 (1985).
[23]VG Technical Document TD8601.
[24]Y. Bu and M. C. Lin, Surf. Sci. **317**, 152 (1994).
[25]P. Hill, D. I. Westwood, L. Haworth, J. Lu, and J. E. Macdonald, J. Vac. Sci. Technol. B **15**, 1133 (1997).
[26]S. E. Hooper *et al.*, J. Cryst. Growth **155**, 157 (1995).

Use of ultrathin ZnSe dipole layers for band offset engineering at Ge and Si homo/heterojunctions

S. P. Wilks,[a] R. H. Williams, M. Pan, and P. R. Dunstan
Department of Electrical and Electronic Engineering, University of Wales Swansea, Swansea SA2 8PP, South Wales, United Kingdom

B. C. C. Cowie
CCLRC Daresbury Laboratory, Warrington, Cheshire, Warrington, Cheshire WA4 4AD, United Kingdom

(Received 19 January 1999; accepted 17 May 1999)

The ability to control semiconductor band discontinuities would allow solid devices to be specifically tailored so that efficiency and performance could be dramatically improved. This article reports the use of an ordered ZnSe monolayer to induce a valence band discontinuity at the Ge homojunction (0.38 eV), at the Ge–Si heterojunction (0.53 eV), and at the Si homojunction (~0.2 eV). Soft x-ray photoemission was used to probe the interfaces as they were formed under ultrahigh vacuum conditions. The effect of overlayer band bending on the interpretation of band offset measurements is discussed. As the interfacial bonding and orientation of the dipole layer are key factors in determining the direction and magnitude of the band modification, x-ray standing wave measurements were performed on the Ge–ZnSe–Ge systems to identify the atomic structure of the junction. Se atoms were always found to bond to the Ge substrate in the *a*-top position, while the Zn atoms adopted the H_3 sites, bonding to the overlayer. The results for these interfaces are interpreted in terms of the charge transfer; other factors such as strain and order are also addressed.
© *1999 American Vacuum Society.* [S0734-211X(99)07504-6]

I. INTRODUCTION

The ability to control or tailor the conduction and valence band profiles across semiconductor junctions has been viewed as somewhat of a *holy grail* in the area of solid state device engineering. Such a technique would allow more efficient devices to be produced, or more novel interface properties to be exploited in material systems that were previously redundant due to the unsuitable magnitude of the band discontinuities formed. This is surprising as the fundamental mechanisms dictating the formation of semiconductor heterojunctions are understood and have been thoroughly investigated both theoretically and experimentally. It is accepted that the model proposed by Tejedor and Flores[1] in 1978 and later developed by Harrison and Tersoff,[2] which took into account the interactions and resulting dipole due to the interface formation, provided the basic platform to understand the microscopic nature of band alignment. This theoretical treatment set the trend for interface calculations to predict the magnitude of the band offsets as a direct result of preferential bonding at the interface. Therefore, due to the thermodynamic nature of the interface, the same characteristic band alignment will always occur for a particular semiconductor heterojunction. Unfortunately, this fixes the band offsets formed at each heterojunction.

This inflexibility means that any device that is based on heterojunction technology [quantum well laser, high electron mobility transistor (HEMT), waveguides] is usually operating under conditions that are far from optimum. Furthermore, this issue is a significant factor that hinders the development of novel devices based on the more conventional semiconductor material systems. Since the natural band offsets are dependent on the charge transfer (intrinsic dipole) across the heterojunction, the interfacial geometry/chemistry and the quantum mechanical nature, the ability to control one or more of these properties allows bands to be selectively engineered. Hence, over the last two decades, much effort has been directed at these issues resulting in several methods to control the effective magnitude and direction of band discontinuities: (a) δ doping,[3,4] (b) multi-quantum barriers (MQBs),[5–7] and (c) interface dipoles.[8–13] All of these methods rely on the production of a sharp potential step at or near the interface region, which adds algebraically to the natural junction profile. This is extremely difficult to generate in practice, normally due to experimental limitations. For example, the use of δ layers to create a potential step requires a sheet of positive donors and negative acceptors to be placed in close proximity (~5 nm). Furthermore, charge neutrality must be achieved, otherwise the resultant profile will be extremely distorted as shown by Wilks and Kestle;[8] an error of 10% can result in a large background potential. Such accuracy is difficult to achieve during the growth of such structures. The use of MQBs also suffers from similar limitations where the production of abrupt superlattices is not a reality for some material systems; a limited success has been demonstrated for the GaAs/AlGaAs system. However, extrinsic interfacial dipoles can circumvent the problems associated with charge neutrality if they can be formed by the congruent deposition of an ionic molecule. This is an area that this article explores in detail.

The concept of band engineering using extrinsic dipoles placed at the heterojunction originated from the work of Niles *et al.*[9] from 1985 onwards. These initial studies involved the introduction of thin metallic interlayers of Al or Au at the interface of a variety of heterojunctions (CdS–Ge,[9]

[a]Electronic mail: s.p.wilks@swansea.ac.uk

Si–CdS,[9] Ge–ZnSe,[10] and Si–GaP[11]). Such studies revealed that the valance band offset increased by 0.1–0.3 eV, depending on the thickness of the metallic interlayer. It was deduced that presence of the intralayer modified the charge transfer across the interface and hence induced a dipole effect. This notion was extended to the use of binary intralayers (GaAs,[12] AlAs,[13] GaP,[13] and AlP[13]) to induce a dipole moment and hence modify the alignment of homojunctions, based on Si or Ge. The results of such investigations yielded discontinuities of between 0.4 and 0.5 eV. In some cases the valence band discontinuity was observed to change polarity when the anion–cation deposition sequence was reversed (e.g., Si–AlP–Si[13]).

However, the chemical nature of the substrate-dipole interface is extremely important as reactions or out-diffusion can destroy the dipole moment. If, for example, the substrate–anion bond is much stronger than the substrate–cation bond, the dynamics of the interface formation and hence the structure, will be dependent on the deposition sequence. This was demonstrated by Marsi *et al.*[13] for the Ge–AlAs–Ge homojunction where the valence band offset possessed the same polarity for anion first and cation first scenarios. This result implied a rearrangement of the interfacial geometry, i.e., clustering, in-diffusion, or even the anion and cation layers interchanging to produce the same dipole orientation. Clearly, the chemical nature and structural geometry at the interface are of vital importance in determining the magnitude and direction of the dipole moment and hence the band offset. In order for the dipole modification mechanism to be understood, idealized systems must be used whereby the heterojunction must retain its integrity on the incorporation of the dipole layer. In essence, the junctions should be prototypical, that is abrupt, ordered, and free from chemical reactions.

The results presented in this article are based on the use of ZnSe as an interfacial dipole layer placed at a Ge homojunction and a Ge/Si heterojunction. Soft x-ray photoelectron spectroscopy (SXPS) measurements were performed during the *in situ* formation of each interface to establish the electronic and chemical nature of the junctions. The effect of the dipole layer on the valence band discontinuities is clearly demonstrated. Furthermore, as the orientation and geometric configuration of the ZnSe molecule at the interface is believed to be a fundamental issue dictating the charge transfer, x-ray standing wave (XSW) was utilized to probe the sited incorporation of the Zn and Se atoms.

II. EXPERIMENT

The photoemission and x-ray standing wave experiments were carried out on the 2 GeV storage ring at the Synchrotron Radiation Source in Daresbury Laboratory, UK. The Ge(111) samples were cut from *n*-type commercial wafers, with a resistivity in the range of 5–20 Ω cm, and inserted into the ultrahigh vacuum (UHV) system. A clean Ge(111) surface was obtained by 12–14 cycles of argon ion bombardment at an energy of 1.2 KeV for 20 min and annealing at about 700 °C for 15 min. After this procedure a sharp $c(2\times8)$ reconstruction was observed by low energy electron diffraction (LEED). The quality of the sample surface was verified from the core level line shapes and the well-known valence band features in the photoemission spectra.[14,15] ZnSe was deposited from a well outgassed W filament onto the clean Ge(111) surface at an elevated temperature of 270 °C to produce an ordered (1×1) surface reconstruction. Silicon or germanium was deposited from a commercial e-beam source with the sample held at room temperature, producing an amorphous overlayer.[15]

As part of this study, two types of systems were investigated, requiring three samples. Sample 1 refers to the Ge–ZnSe–*a*Ge homojunction while samples 2 and 3 related to the Ge–*a*Si and the Ge–ZnSe–*a*Si heterojunction. Sample 2 was required as a control experiment so that the effect of the ordered ZnSe interlayer on the heterojunction in sample 3 could be ascertained. SXPS was used to study each system at all stages during the sequential formation of the interfacial regions. Si and Ge overlayer coverages, ranging from 0.5 up to 20 Å, were considered in this study.

The XSW study was performed during the formation of the Ge–ZnSe–*a*Ge homojunction, with adsorption spectra recorded after the deposition of ZnSe and 2 ML of *a*Ge. In order to triangulate the position of the Ge, Zn, and Se atoms, the (220), (311), and $(3\bar{1}1)$ Bragg reflections were energy scanned, while monitoring the Ge $2p_{3/2}$, Se $2p_{1/2}$, and Zn $2p_{3/2}$ photoelectron core levels. The typical collection time for a set of spectra was around 4 h for each Bragg reflection; no surface contamination was detected during this period.

III. RESULTS AND DISCUSSION

Each system will be considered in turn and the results pertaining to each presented in the following sections. Any similarities and differences will be used to explain the nature of the dipole moment responsible for the modification of the band profile.

A. SXPS study of the Ge/ZnSe/*a*Ge homojunction

ZnSe and Ge have a small lattice mismatch making them ideal candidates for heteroepitaxial growth, previous work has indicated that crystalline ZnSe can be grown on Ge for substrate temperatures around 270 °C. The structural quality of the ZnSe layers used in this study was indicated by a sharp (1×1) LEED pattern, indicative of a highly ordered interface. The Ge $3d$ core level for each stage of the Ge(111)–ZnSe interface formation is depicted in Fig. 1, along with the deconvoluted components. The spectrum obtained from the clean $c(2\times8)$ surface consisted of three components, the largest emanating from the bulk, and the two higher kinetic energy features were associated with the reconstructed surface. These observations are consistent with work reported by others.[14] However, after the deposition of 1 ordered ZnSe ML, the line shape of the Ge $3d$ core level changed dramatically. Again the largest peak was associated with the bulk derived photoelectrons, yet two interface com-

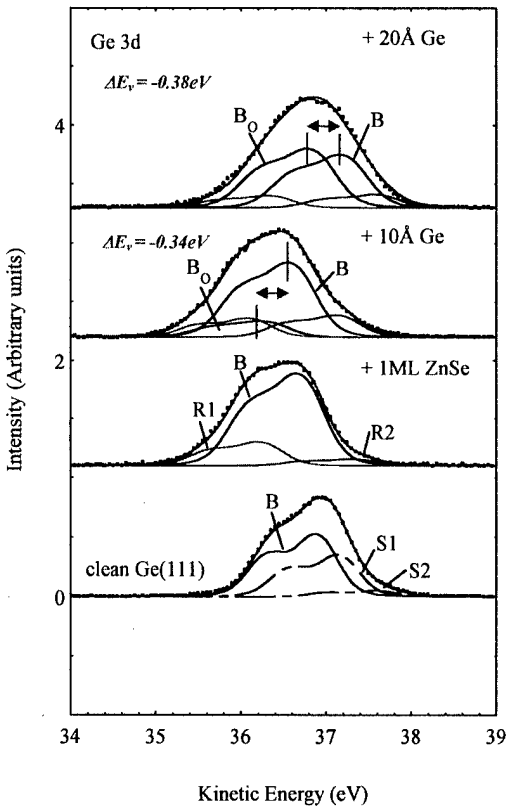


FIG. 1. Evolution of the Ge $3d$ core level spectra; illustrating the clean $c(2\times8)$ surface, 1 ML ZnSe and deposition of 10 and 20 Å of Ge. Dots are data points; the solid curves are fits to the data and deconvoluted components. The relative shift between the overlayer Ge and the substrate Ge is illustrated.

ponents were required to successfully fit the experimental spectrum. The larger feature located at lower kinetic energy was associated with Ge–Se bonds while the higher kinetic energy component was associated with Ge–Zn bonds. Intuitively, on the deposition of a Ge overlayer, one might expect these three features to be attenuated, with the addition of a bulk derived component derived from the deposited Ge. Indeed, this was found to be the case, where the Ge bulk overlayer component grew as the coverage was increased. This can be clearly seen in Fig. 1 where the overlayer related feature is located at a lower kinetic energy (0.30–0.38 eV) with respect to the substrate component. This difference in energy, which stabilized at 0.38 eV for coverages >10 Å, is a direct consequence of the potential difference induced by the dipole moment of the ZnSe interlayer. It is worth noting that the line shape of the bulk overlayer component was slightly broader than the substrate and interface components, allowing for the amorphous nature of the Ge growth at room temperature. (The parameters corresponding to the overlayer line shape were obtained by fitting a thick coverage with a single component.) When this result was compared to a control sample for this interface, i.e., Ge/*a*Ge homojunction, a negligible shift of 0.1 eV between the substrate and overlayer derived components was observed. Clearly, the energy difference of 0.38 eV was due to the presence of the ZnSe dipole layer.

It is also interesting to note the behavior of the Ge–Se and Ge–Zn derived features as a function of the overlayer growth. As expected, the Ge–Se component decreased in intensity as the interface was buried by the increasing Ge coverage. Conversely, the Ge–Zn peak at higher kinetic energy initially increased for the first few depositions of Ge, after which, the intensity of the component decreased as with the Ge–Se feature. This was believed to be a direct consequence of the deposited Ge bonding to the exposed Zn atoms on the surface. These results imply that the ZnSe molecule is bonded at the interface such that Se attaches to the Ge substrate while the Zn atoms are available for bonding with the overlayer. This hypothesis was supported by the variation of Zn:Se ratio as a function of the takeoff angle of the emitted photoelectrons.

B. XSW study of the Ge/ZnSe/*a*Ge homojunction

The adsorption profiles for the (220), (311), and $(3\bar{1}1)$ Bragg reflections are shown in Figs. 2 and 3, for the buried Se and Zn atoms, respectively. The corresponding spectra for the Ge/ZnSe junction are not shown as very little difference was detected after the deposition of the *a*Ge overlayer; clearly this is an indication that the ZnSe layer is stable and fixed at the interface boundary. The adsorption spectra were fitted based on the standard theoretical model[16] to yield the coherent position (P) and the coherent fraction (F). The former parameter provides the distance of the atom from the Bragg reflection plane while the coherent fraction is a measure of the order present in the atomic layer. The results from the deconvolution are presented in Table I. From this data it can be seen that the Se adsorption spectra exhibited high coherent fractions in all three Bragg directions (~0.7–0.8). A simple interpretation of this result would be that 70%–80% of the Se atoms occupy a single adsorption site forming an ordered layer. In contrast, the Zn adsorption spectra yielded much lower coherent fractions (~0.4–0.6) implying that only 40%–60% of the Zn atoms exist in a unique adsorption site. However, low coherent fractions are not always indicative of disorder. It is possible that multiple adsorption sites occur in the Zn layer or even that domain averaging occurs if the atoms are slightly displaced from a high symmetry adsorption site. However, as the coherent fraction is relatively low in all of the directions selected, it is believed that the values are associated with either a degree of disorder in the Zn layer, cluster formation, or that domain averaging does indeed occur. To clarify this point, other Bragg conditions would need to be considered. Furthermore, it is important to note that the coherent fractions reported above for the buried dipole include a 10% reduction as compared to the Ge/ZnSe interface (not shown). This was also true for the Ge substrate adsorption profiles and is indicative of a Ge overlayer that is slightly out of registry with the substrate.

However, meaningful information can still be gleaned from the coherent positions, as this parameter is less sensitive to the disorder or domain averaging effects. In essence, under these conditions, the coherent position will yield the perpendicular distance of the ordered species from the Bragg plane. The adsorption sites of the Zn and Se species were

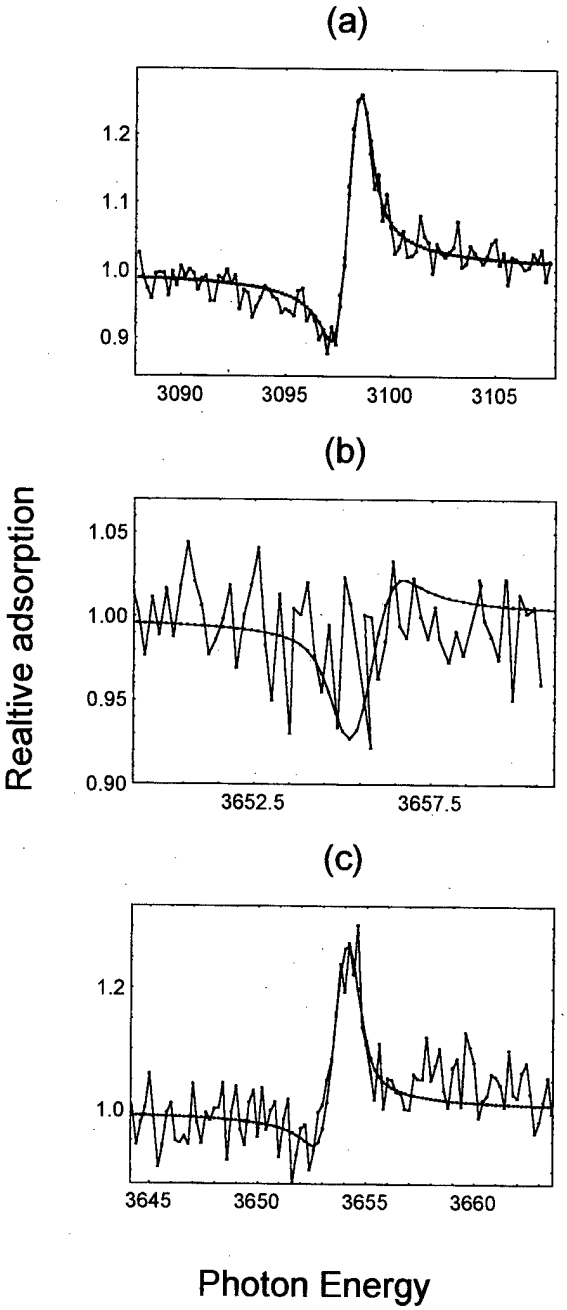


FIG. 2. Se $2p_{1/2}$ XSW profiles corresponding to (a) the (220), (b) the $(3\bar{1}1)$, and (c) the (311) reflections for the Ge–ZnSe–aGe buried interface. The continuous smooth solid line represents the theoretical fit.

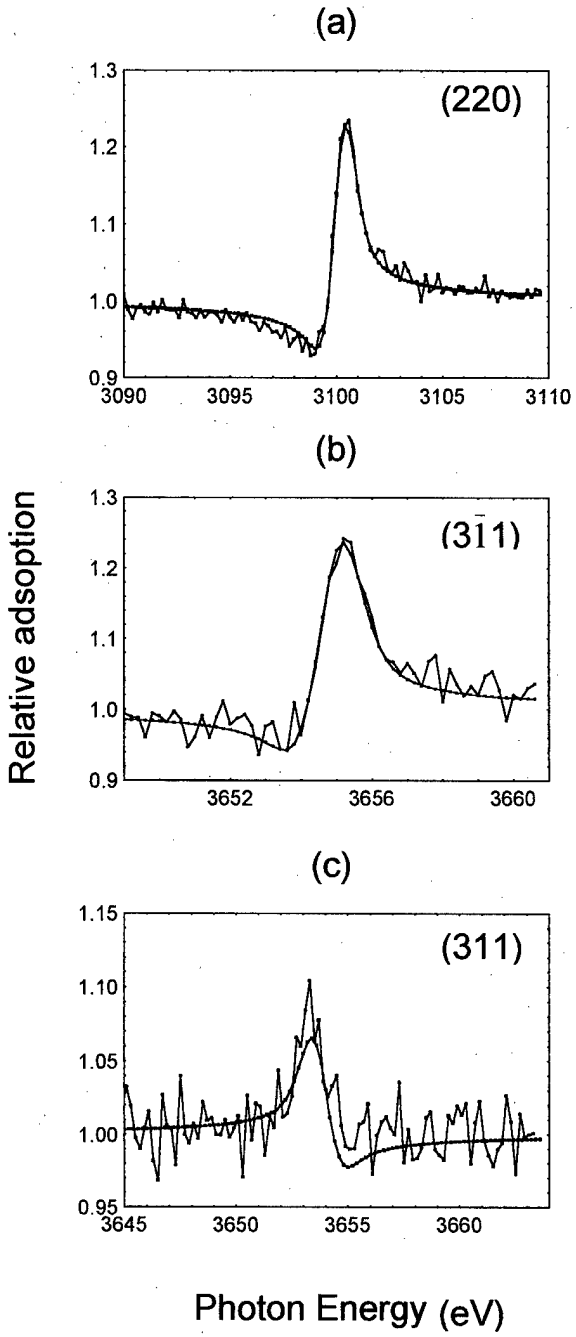


FIG. 3. Zn $2p_{3/2}$ XSW profiles corresponding to (a) the (220), (b) the $(3\bar{1}1)$, and (c) the (311) reflections for the Ge–ZnSe–aGe buried interface. The continuous smooth solid line represents the theoretical fit.

identified by triangulation. This was achieved by calculating the coherent positions expected for the (311) and $(3\bar{1}1)$ Bragg reflections in high symmetry sites, based on the experimental value measured in the (220) direction. The results are displayed in Table II, where the shaded area represents the sites that produced the best fit to the experimental values. Note that higher order coherent positions were also examined due to the invariant nature of the XSW technique to perpendicular translations relative to the Bragg plane (e.g., coherent positions of 0.5 and 1.5 are equivalent). The coherent position associated with the (111) reflection is also calculated as this provides a measure of the displacement in the [111] direction, i.e., perpendicular to the interface. From this comparison it is clear that the Se atoms preferentially locate in the a-top position, while a proportion of the Zn atoms rest in the H_3 site. This implies that Se bonds directly to the Ge substrate while Zn bonds to the aGe overlayer. The signals associated with the overlayer suggest that the growth mode is predominantly epitaxial for the coverages investigated in this study; for thicker layers one might expect a less ordered nature. A cross section through the interface is illustrated in Fig. 4, where the inter-plane distances are labeled. Based on

TABLE I. Coherent positions (P) and the coherent fractions (F) extracted from the theoretical fits of the Se and Zn XSW profiles measured for the (220), ($3\bar{1}1$), and (311) Bragg reflections.

Reflection	Se parameters		Zn parameters	
	P	F	P	F
(220)	0.94	0.70	0.92	0.43
(311)	1.07	0.72	0.72	0.43
($3\bar{1}1$)	0.80	0.72	1.05	0.60

the measurements performed as part of this investigation, the Zn–Se bond length was found to be (2.45±0.06) Å, which is in excellent agreement with the value found in bulk ZnSe. Furthermore, the Ge–Se bond length is (2.71±0.06) Å, 0.2 Å smaller than that based on the sum of the covalent radii, implying a relaxation of the Se towards the Ge. Thus, the vertical separation of the Zn and Se atomic planes was found to be 0.78 Å in the [111] direction.

C. Study of the Ge/ZnSe/*a*Si heterojunction

To determine the valence band discontinuity, ΔE_v, an indirect approach was adopted because the Si and Ge valence band edges could not be resolved in the same spectrum.[15] The energy difference between the clean substrate valence band maximum (VBM) and the overlayer VBM was measured by the linear extrapolation of valence band leading edges. This difference was corrected for the change in substrate band bending, due to the interface formation, by monitoring any shift in the energy of the substrate core level peak

TABLE II. Predicted (111), (311), and ($3\bar{1}1$) layer spacings calculated for symmetry adsorption sites, assuming (a) the experimental Se (220) coherent positions of 0.94 and 1.94, and (b) the experimental Zn (220) coherent positions of 0.92 and 1.92. Gray shading marks site consistent with experimental parameters.

(a)

Measured Se values	Site	Calculated Se layer spacing based on measured $P(220)$		
$P(220)$		$P(111)$	$P(311)$	$P(3\bar{1}1)$
0.94	*a*-top	0.83	1.05	0.83
1.94		1.58	2.29	1.58
0.94	T_4	0.58	1.29	0.58
1.94		1.33	2.54	1.33
0.94	H_3	0.33	0.54	0.33
1.94		1.08	1.79	1.08

(b)

Measured Zn values	Site	Calculated Zn layer spacing based on measured $P(220)$		
$P(220)$		$P(111)$	$P(311)$	$P(3\bar{1}1)$
0.92	*a*-top	0.81	1.02	0.81
1.92		1.56	2.27	1.56
0.92	T_4	0.57	1.28	0.57
1.92		1.32	2.53	1.32
0.92	H_3	0.32	0.53	0.32
1.92		1.07	1.78	1.07

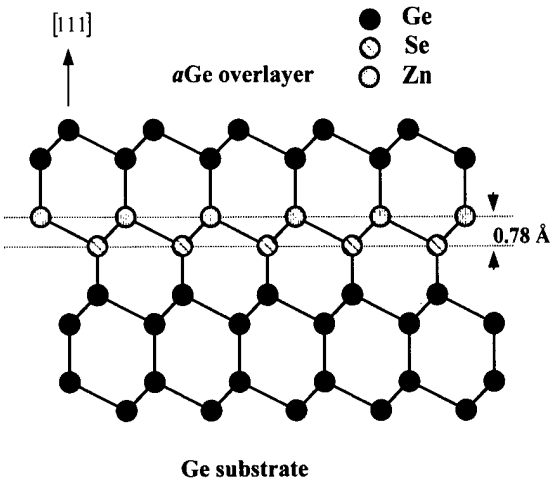


FIG. 4. Cross-sectional view of the Ge(111)–ZnSe–*a*Ge interface. The Se atoms bond to the Ge(111) substrate in the *a*-top position while the Zn atoms bond in the H_3 hollows, adjacent to the *a*Ge overlayer.

and hence the correct discontinuity established. The error in the measurement of band offsets did not exceed 0.07 eV.

For the Ge–*a*Si heterojunction, sample 2, Ge 3*d* core level spectra on the clean Ge(111) surface were deconvoluted into three components, as indicated by the lowest spectrum of Fig. 5. The component labeled B emanated from

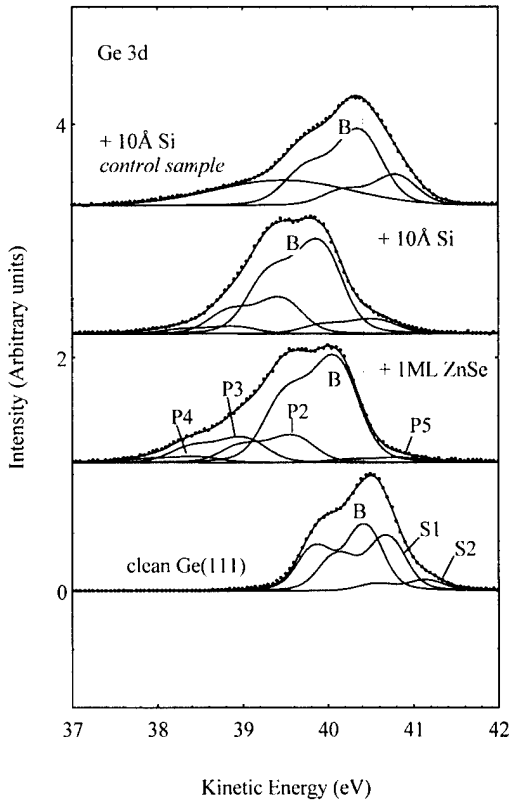


FIG. 5. Ge 3*d* core level spectra taken at photon energy of 75 eV. The lower three spectra are from the Si/ZnSe/Ge(111) sample, the upper one is the spectrum of the Si/Ge(111) sample after 10 Å Si deposition for comparison. Dots are data points; the solid curves are fits to the data and deconvoluted components.

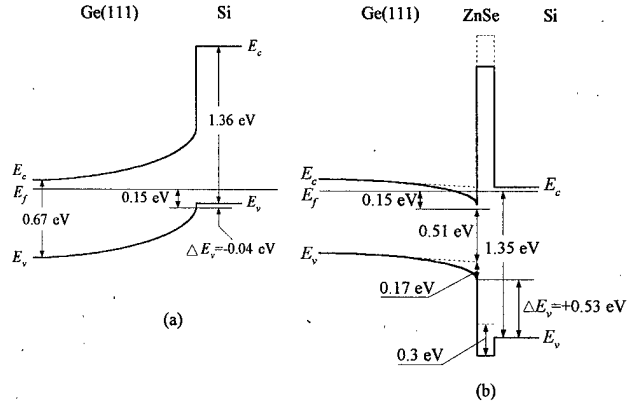


FIG. 6. Schematic diagram of energy bands in (a) a-Si/Ge(111) sample and (b) Si/ZnSe/Ge(111) system, the thin solid lines represent Fermi level E_F, the dashed lines correspond to the band positions after deposition of 1 ML ZnSe, and the thick solid lines correspond to the final positions of bands after 10 Å of Si. The energy band gaps of amorphous Si(a-Si) and ZnSe crystal are from the typical values of 1.36 and 2.58 eV in Refs. 26 and 27.

bulk emission while components S1 and S2 were derived from the reconstructed surface. This is in good agreement with previous studies.[14] The fit parameters for the clean Ge surface were used throughout the deconvolution of the Ge $3d$ signal after the deposition of subsequent layers. As illustrated in the topmost spectrum of Fig. 1 there was no change in the Ge $3d$ line shapes and the peak position of the bulk component after the Si overlayer was deposited. This indicated no change in band bending during the interface formation, which is in agreement with the results of Katnani *et al.*[15] Therefore, the energy difference of -0.04 ± 0.07 eV between the clean Ge(111) VBM and that of the thick Si overlayer established the valence band offset ΔE_v, with the Si valence band maximum at a lower binding energy than the substrate. This is schematically shown in Fig. 6(a). Several theoretical calculations and experimental measurements of the valence band offsets for Si/Ge heterojunctions have been reported, in particular, band offsets of 0.15 ± 0.1 and 0.4 ± 0.1 eV were observed for Si on cleaved Ge(111) by Katnani *et al.*[15] and Mahowald *et al.*,[17] respectively. The result from this study is in reasonable agreement with the previous value bearing in mind that the two starting surfaces were not identical; the relatively poorer agreement with the latter could originate from the different strain configurations noticing the overlayer thicknesses were quite different in the two studies.

The deconvoluted Ge $3d$ core level spectra of sample 3 are also presented in Fig. 5 and correspond to an identical clean surface, the subsequent deposition of 1 ML of ZnSe and successive overlayers of Si. The deposition of 1 ML of ZnSe produced dramatic changes on the core level line shapes where an obvious shoulder on the low kinetic energy side appeared. This shoulder is indicative of Ge–Se bonding as demonstrated by the previous photoemission studies detailed earlier in this article.[18,19] In the best fit, three components (P2, P3, and P4) were required to describe the Ge–Se bonding, a large peak (B) for the bulk derived feature, and an additional small component (P5) at higher kinetic energy. For this experiment the resolution of the system was improved and thus more peaks were identifiable. From the previous study detailed earlier, P5 was attributed to the Ge–Zn bonds and the uncovered bare areas of Ge substrate. Another important feature of the Ge $3d$ bulk component (B) was its movement by 0.5 ± 0.03 eV to lower kinetic energy after the deposition of ZnSe, implying a flattening of the bands as compared to the clean surface and a passivation of the surface.

After the initial shift in the Ge $3d$ bulk signal on ZnSe deposition, it continued to shift to lower kinetic energy for increasing Si coverages; after 10 Å of Si, an additional $\sim0.2\pm0.03$ eV movement was observed. Energy shifts in intralayer core levels were also observed during the interface formation. The Zn $3d$ and the Se $3d$ core levels both shifted to lower kinetic energy by $\sim0.3\pm0.03$ eV, yet the energy position of Si $2p$ core level remained relatively constant.

Photoemission spectra were taken near the valence band maximum to monitor the evolution of the band edges. After the deposition of ZnSe, emissions from both the ZnSe layer and Ge substrate were clearly visible and thus resolved within the same spectrum. A direct measurement yielded a "valence band discontinuity" of 1.12 ± 0.07 eV between ZnSe and Ge. Similar results of 0.95 ± 0.05 and 1.29 ± 0.03 eV were reported for ~20 Å of ZnSe grown on a Ge(001) substrate and a Ge(110) substrate, respectively.[20,21] A well characterized valence band of ZnSe might not be fully developed since only 1 ML of ZnSe was used in this study. Successive deposition of the Si overlayer gradually masked the Ge related features in the valence band spectra. At 10 Å Si coverage, the Si-induced leading edge in valence band had well developed.

Based on the above discussion of core level and band edge movements, an energy band diagram at the Si coverage of 10 Å is shown in Fig. 6(b) for sample 3. All VBM were corrected for band bending extracted from core level movements. A valence band offset of $+0.53\pm0.07$ eV was deduced from the final positions of valence band edges, with higher binding energy on the Si side. Furthermore, the change in the energy separation between Ge $3d$ and Si $2p$ core levels, when comparing sample 2 and sample 3, also supported the above observation. From Figs. 6(a) and 6(b), the ZnSe intralayer induced modification of the valence band offset is evident. Specifically, the offset increased significantly in magnitude from -0.04 ± 0.07 to $+0.53\pm0.07$ eV upon the introduction of the ZnSe intralayer at the interface. It is believed that the change in ΔE_v is attributable to the modifications of the intrinsic interface dipole.

Above 10 Å, the interface becomes unstable and results in the out-diffusion of Se into the aSi overlayer; this was evident from the attenuation plots of the various core levels. When considering the thermodynamic aspects, this is not surprising, as the Si–Se bond is the most stable. However, the mechanism for the breakdown of the interface is not clear yet could possibly be related to strain as the critical thickness of Si on Ge is approximately 11 Å.

D. Comparison with theory

The electrostatic dipole due to the ZnSe was estimated based on the simple microscopic capacitor model[22] and electronegativity considerations,[23] respectively. In the microscopic capacitor picture, the effect of the ZnSe intralayer was viewed as a two-plate capacitor; whereas charge transfer was assumed to only occur between Zn and Se atomic planes in the electronegativity estimation. The resulting potential across the intralayer can be expressed in a common form for both the approaches as, $V=qd/\epsilon\epsilon_0$, where q is the charge per unit area, ϵ is the relative dielectric constant of the material, and ϵ_0 is the permittivity of vacuum. The term d is either the distance of the capacitor which is of the value of 0.78 ± 0.08 Å[19] or the Zn–Se bond length (2.45 Å). Since the intralayer was very thin, the dielectric constant ϵ would fall in between the two extreme values of 16.0 for Ge and 9.12 for ZnSe.[24] Thus V was found to be in the range of 0.64–1.12 and 0.31–0.55 eV in the capacitor and the electronegativity models, respectively. Considering the realistic interface configurations,[19] V had a value of 0.1–0.17 eV or less for the corrected Zn to Se plane separation in the electronegativity model. It seems therefore that the estimation from the capacitor model is closer to the observed value. The two calculations were very approximate and have their limitations,[24] so it is not clear which model better describes the physical situation. In addition, these models neglect the nature of the realistic interface, for example, the complex bonding of the ZnSe layer and indeed its geometric configuration. Furthermore, from the XSW results, it is known that a small degree of disorder is present at the junction where the dipole effect could possibly be reversed producing a lateral inhomogeneity in the modification of the band offset. Clearly all of these effects must be included in any realistic calculations of the dipole effect for true comparison with experimental data. However, it is certain that the presence of ZnSe intralayer set up an electric field and hence an additional dipole at the interface, which modified the interface charge distribution. The induced change in valence band offset had the same sign and magnitude as the estimation of the dipole, indicating that the ZnSe dipole is one of the major contributions responsible for the changes in the band offsets.

E. Effect of band bending on the measured ΔE_v

The extraction of the valence band discontinuity from SXPS measurements assumes that there is no band bending in the overlayer materials system, i.e., the bands are flat from the interface to the surface as shown in Fig. 6(b) for the Ge/ZnSe/aSi heterojunction. However, there has been much discussion in the literature recently regarding apparent discontinuities as a result of heavy band bending in the overlayer, resulting in distorted and shifted photoelectron spectra.[25] This is an important issue that should be considered when extracting band offsets from SXPS data.

A model was developed to calculate the core level line shape for the Si overlayer in the presence and absence of band bending. The conduction and valence band profiles across the Ge/ZnSe/aSi interfaces were calculated by solving Poisson's equation, assuming known parameters extracted from the results presented in Sec. III C. In particular, the movement of the bulk derived component from the Ge substrate and the valence band discontinuity between the Ge/ZnSe systems. The latter assumption may appear contradictory, however the Ge and ZnSe valance bands were resolved within the same spectra, providing a direct measurement of this offset, which was in good agreement with the work of others. In addition, the Fermi level at the surface of the aSi overlayer was presumed to be pinned at the midgap position, causing the bands to bend upwards in the overlayer. The calculation proceeded by summing the contributions to the Si $2p$ spectra from each atomic layer, taking into account the reduction of the core level intensity due to the finite mean free path and the variation in kinetic energy of each core level as a result of the band bending. As can be seen in Fig. 7, the effective centroid of the peak will be distorted to higher kinetic energy as compared to the flatband case. It is clear that as the band bending increases, the core level line shape will be distorted, and asymmetric towards lower kinetic energy. To illustrate the effect of band bending, a band discontinuity of -0.53 eV was used in the calculation and the movement of the peak centroid was compared to that of the flatband case. The simulation revealed that the Si $2p$ core level would appear 0.3 eV lower in kinetic energy, due to the presence of band bending. This would result in a measured discontinuity (ΔE_{vm}) of 0.23 eV. Therefore, in order to achieve a measured value of 0.53 eV, the valence band discontinuity, due to the presence of the dipole layer, would have to be much larger ($\sim$1 eV) which would result in severe deformation of the Si $2p$ core level line shape. Experimentally, this was not observed. This would suggest that the band bending in the aSi overlayer was negligible, or smaller than the resolution of the SXPS technique ($\sim$0.1 eV).

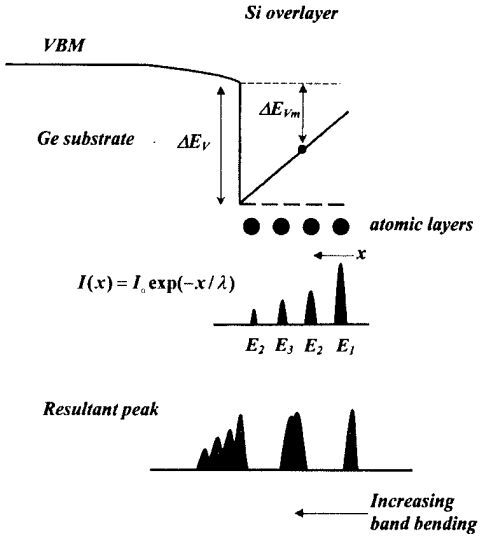


FIG. 7. Schematic representation of the effect of band bending on the measurement of the valence band discontinuity and core levels in comparison to the flatband scenario.

IV. CONCLUSIONS

Photoemission spectroscopy was used to study the effect of placing a ZnSe interlayer at a Ge homojunction and a Ge/*a*Si heterojunction. Effective valence band discontinuities of −0.38 and −0.53 eV, respectively, were measured due to the presence of the ZnSe dipole as compared to sample with no interlayer. It is evident that the ZnSe layer modifies the charge transfer across the interface, resulting in an additional potential step. Clearly the structure and orientation of the interlayer within the interface are important parameters that determine the magnitude and direction of the dipole moment associated with the ZnSe layer. Hence, XSW was applied to identify the incorporation sites of the Zn and Se atoms at the Ge homojunction. Results implied that Se bonded to the Ge substrate in the *a*-top position while the Zn atom bonded in the H_3 site within the next atomic layer, closest to the *a*Ge overlayer. Furthermore, the XSW technique identified an element of disorder within the Zn layer, possibly due to multiple incorporation sites. These measurements yielded a Se–Zn bond length of 0.78 Å in the [111] direction, normal to the interface, enabling the potential due to the presence of the dipole to be estimated based on theoretical models. The best agreement was obtained using the capacitor-based model, predicting a discontinuity of (0.64–1.12 eV). Finally, the effect of band bending on the interpretation of the valence band discontinuity was found to be negligible for the interfaces studied in this article.

ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support of EPSRC for funding and beamtime at the SRS facility. The technical supports offered by the Electrical and Electronic Engineering Department at Swansea and at the SRS at Daresbury are gratefully noted. In addition, the authors would wish to thank the contributions of Dr. S. Clark and Dr. S. R. Burgess.

[1]C. Tejedor and F. Flores, J. Phys. C **11**, L19 (1978).
[2]W. A. Harrison and J. Tersoff, J. Vac. Sci. Technol. B **4**, 1068 (1986).
[3]F. Capasso, A. Y. Cho, K. Mohammed, and P. W. Foy, Appl. Phys. Lett. **46**, 664 (1985).
[4]T.-H. Shen, M. Elliott, R. H. Williams, and D. I. Westwood, Appl. Phys. Lett. **58**, 842 (1991).
[5]K. Iga, H. Uenohara, and F. Koyama, Electron. Lett. **22**, 1008 (1989).
[6]A. Kestle, S. P. Wilks, D. I. Westwood, M. Ke, M. Elliott, and R. H. Williams, Mater. Sci. Eng., B **35**, 145 (1995).
[7]S. P. Wilks and A. Kestle, Appl. Surf. Sci. **117**, 334 (1997).
[8]S. P. Wilks and R. H. Williams, J. Electron Spectrosc. Relat. Phenom. **72**, 49 (1995).
[9]D. W. Niles, G. Margaritondo, P. Perfetti, C. Quaresima, and M. Capozi, Appl. Phys. Lett. **47**, 1092 (1985).
[10]D. W. Niles, E. Colavita, G. Margaritondo, P. Perfetti, C. Quaresima, and M. Capozi, J. Vac. Sci. Technol. A **6**, 962 (1986).
[11]J. T. McKinley, Y. Hwu, D. Rioux, A. Terrasi, F. Zanini, G. Margaritondo, U. Debska, and J. K. Furdyna, J. Vac. Sci. Technol. A **8**, 1917 (1990).
[12]J. T. McKinley, Y. Hwu, B. E. C. Koltenbah, G. Margaritondo, S. Baroni, and R. Resta, Appl. Surf. Sci. **56–58**, 762 (1992).
[13]M. Marsi, S. L. Rosa, Y. Hwu, F. Gozzo, C. Coluzza, A. Baldereschi, G. Margaritondo, J. T. McKinley, S. Baroni, and R. Resta, J. Appl. Phys. **71**, 2048 (1992).
[14]A. L. Wachs, T. Miller, A. P. Shapiro, and T.-C. Chiang, Phys. Rev. B **35**, 5514 (1987).
[15]T. Miller, T.-C. Hsieh, and T.-C. Chiang, Phys. Rev. B **33**, 6983 (1986); P. M. J. Marée, K. Nakagawa, F. M. Mulders, and J. F. Van der Veen, Surf. Sci. **191**, 305 (1987); A. D. Katnani, N. G. Stoffel, R. R. Daniels, T.-X. Zhao, and G. Margaritondo, J. Vac. Sci. Technol. A **1**, 692 (1983).
[16]J. Zegenhagen, Surf. Sci. Rep. **18**, 199 (1993).
[17]P. H. Mahowald, R. S. List, W. E. Spicer, J. Woicik, and P. Pianetta, J. Vac. Sci. Technol. B **3**, 1252 (1995).
[18]K. A. Saynor, Ph.D. thesis, University of Wales, Cardiff, 1995.
[19]S. R. Burgess, Ph.D. thesis, University of Wales, Cardiff, 1997.
[20]G. Bratina, L. Vanzetti, L. Sorba, G. Biasiol, A. Franciosi, M. Peressi, and S. Baroni, Phys. Rev. B **50**, 11723 (1994).
[21]S. P. Kowalczyk, E. A. Kraut, J. R. Waldrop, and R. W. Grant, J. Vac. Sci. Technol. **21**, 482 (1982).
[22]A. Muñoz, N. Chetty, and R. M. Martin, Phys. Rev. B **41**, 2976 (1990).
[23]P. Perfetti, C. Quaresima, C. Coluzza, C. Fortunato, and G. Margaritondo, Phys. Rev. Lett. **57**, 2065 (1986).
[24]A. Franciosi and C. G. Van de Walle, Surf. Sci. Rep. **25**, 1 (1996).
[25]M. Moreno, H. Yang, M. Höricke, M. Alonso, J. A. Martin-Gago, R. Hey, K. Horn, J. L. Sacedón, and K. H. Ploog, Phys. Rev. B **57**, 12314 (1998).
[26]E. C. Freeman and W. Paul, Phys. Rev. B **20**, 716 (1979).
[27]G. Margaritondo, Surf. Sci. **168**, 439 (1986).

AlN and GaN epitaxial heterojunctions on 6H–SiC(0001): Valence band offsets and polarization fields

Angela Rizzi[a)]
Institut für Schicht- und Ionentechnik (ISI), Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany and Istituto Nazionale di Fisica della Materia, Dipartimento di Fisica, Università di Modena, I-41100 Modena, Italy

Roberta Lantier
Institut für Schicht- und Ionentechnik (ISI), Forschungszentrum Julich GmbH, D-52425 Julich, Germany

Fulvia Monti
Institut für Schicht- und Ionentechnik (ISI), Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany and Istituto Nazionale di Fisica della Materia, Dipartimento di Fisica, Università di Modena, I-41100 Modena, Italy

Hans Lüth
Institut für Schicht- und Ionentechnik (ISI), Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany

Fabio Della Sala, Aldo Di Carlo, and Paolo Lugli
Instituto Nazionale di Fisica della Materia, Dipartimento di Ingegneria Elettronica, Università di Roma, "Tor Vergata," Italy

(Received 24 February 1999; accepted 20 April 1999)

From a series of *in situ* photoemission experiments macroscopic electric fields are clearly demonstrated in SiC/AlN, SiC/GaN, and GaN/AlN heterostructures grown by molecular beam epitaxy on 6H–SiC(0001). A significant contribution is due to the *spontaneous* polarization; the *piezoelectric* term alone would not explain the sign of the field measured in SiC/AlN. The experimental field has lower intensity as compared to theory: the role of electronic gap states at the surface is pointed out. A self-consistent tight-binding approach which is able to describe polarization fields, dielectric screening, and free carrier screening is applied for a more consistent theoretical discussion of the experimental data. The valence band offset (VBO) has been determined for all heterojunctions under study and the apparent dependence on the overlayer thickness, due to the presence of the strong polarization fields, has been pointed out in view of a correct determination of the VBO. The VBOs at the heterojunctions obtained by extrapolation to zero overlayer thickness are (substrate/overlayer): $[(1.5–1.7)\pm0.1]$ eV for SiC/AlN, $[(0.7–0.9)\pm0.1]$ eV for SiC/GaN, (-0.3 ± 0.1) eV for AlN/GaN, and $[(0.15–0.4)\pm0.1]$ eV for GaN/AlN. © *1999 American Vacuum Society.* [S0734-211X(99)02704-3]

I. INTRODUCTION

Group III nitride heterostructures are meanwhile considered in many applications, optoelectronic as well as power microelectronic devices.[1] When designing and simulating device heterostructures the energy band scheme has to be known, in particular band offsets and macroscopic electrostatic fields play a central role in the device performance. Beside space charge regions, which we are very familiar with in common semiconductor heterostructures, polarization charges resulting from polarization discontinuities at the heterointerfaces characterize nitride heterostructures.

The nitride equilibrium crystal structure is wurtzite, in which the uniaxial symmetry causes the presence of a *spontaneous polarization* directed along the principal *c* axis. Such materials are commonly called pyroelectrics, but not necessarily a temperature variation has to be considered to tackle these polarization effects. The polarization discontinuity at a heterojunction, for instance, manifests itself with charges at the interface and with electrostatic fields, which in the nitrides can be of the same order of magnitude as in ferroelectrics (~ MV/cm). Furthermore, in the case of thin epitaxial strained layers the *piezoelectric polarization* also has to be taken into account. Since this latter one is more common for semiconductor heterostructures[2] most of the experiments concerned with macroscopic electric fields in III–N heterostructures so far have emphasized just the piezoeffect, overlooking the important contribution of the spontaneous polarization.[3–7] On the other hand Bernardini *et al.* have very clearly pointed out the magnitude and the practical importance of the spontaneous polarization by *ab initio* calculations, which also provide the theoretical piezoelectric coupling constants of the nitride binary compounds.[8,9]

In this article, after recalling the main concepts on polarization effects in wurtzite nitride heterostructures, we present a series of photoemission experiments on the 6H–SiC/AlN, 6H–SiC/GaN, and GaN/AlN heterostructures. A systematic *in situ* analysis considering, in particular, the dependence of the observed effects on the overlayer thickness reveals the presence of strong built-in electrostatic fields, which are induced by the macroscopic polarization, but lowered by the

a)Electronic mail: rizzi.angela@unimo.it

presence of surface electronic states. The valence band offset at the heterojunctions are determined by XPS. A self-consistent tight-binding (TB) approach which is able to describe polarization fields, dielectric screening, and free carrier screening is applied for a more consistent theoretical discussion of the experimental data.

II. POLARIZATION FIELDS IN THE WURTZITE HETEROSTRUCTURES

In this section some basic concepts on the macroscopic polarization in wurtzite heterostructures shall be pointed out, that will be helpful for the discussion of the experimental results. For a deeper insight into this field, the reader is referred to the original theoretical work.[8–12]

In an anisotropic dielectric crystal the linear relationship between electrical induction $\mathbf{D}$ and electrical field $\mathbf{E}$ is more generally characterized by the presence of a constant term $\mathbf{D}_0$ such that $D_i = D_{0i} + \varepsilon_{ik} E_k$, where ε_{ik} is the second-rank dielectric tensor.[13] For most crystallographic symmetries $\mathbf{D}_0$ vanishes, whereas in the remaining cases the dielectric medium shows a *spontaneous polarization* even in the absence of an external field. Since the physical properties of a crystal remain unchanged by any symmetry transformation, crystals that have a direction invariant for such transformations (e.g., uniaxial symmetry) have nonvanishing $\mathbf{D}_0$, which then lies in this direction. The wurtzite structure belongs to this crystal class. The spontaneous polarization can induce very strong electric fields, which cannot exist in large crystals because this would lead to a nonequilibrium situation: even if small, a finite conductivity always induces a current in the presence of a polarization field, which flows until the charges built up at the surface compensate the internal field. This effect does not hold for very thin heterostructure layers, where the electrostatic field can well be sustained.

Theoretically the polarization calculated for a periodic system is a *transverse* polarization $\mathbf{P}^T$, i.e., it is derived in the absence of a depolarization field. In reality polarization charges present at surfaces and interfaces under the assumption of a uniform polarization give rise to the *longitudinal* polarization $\mathbf{P}^L$, the one measured in the presence of a depolarization field. In the regime of linear response and in presence of ion relaxation $P^T = \varepsilon(0) P^L$, where $\varepsilon(0)$ is the static dielectric constant.[14,15]

In the following we will restrict ourselves to contributions along the wurtzite c axis and for the general case of strained layers we get for the macroscopic *transverse* polarization $\mathbf{P}^T = \mathbf{P}^{sp} + \delta\mathbf{P}_3$. The latter term is the piezoelectric polarization obtained via the piezoelectric stress coefficients $e_{31} = e_{32}$ and e_{33} as

$$\delta P_3 = e_{31}(\epsilon_1 + \epsilon_2) + e_{33}\epsilon_3, \tag{1}$$

where $\epsilon_1 = \epsilon_2$ and ϵ_3 are the strains in the basal plane and along the c axis, respectively. The positive direction of the c axis points from the cation to the anion, along the longitudinal bond. Under the assumption of biaxial stress for thin unrelaxed epitaxial layers the strain components along the c axis and in the basal plane are related to the elastic constants through $\epsilon_3 = -2(C_{13}/C_{33})\epsilon_1$.[16]

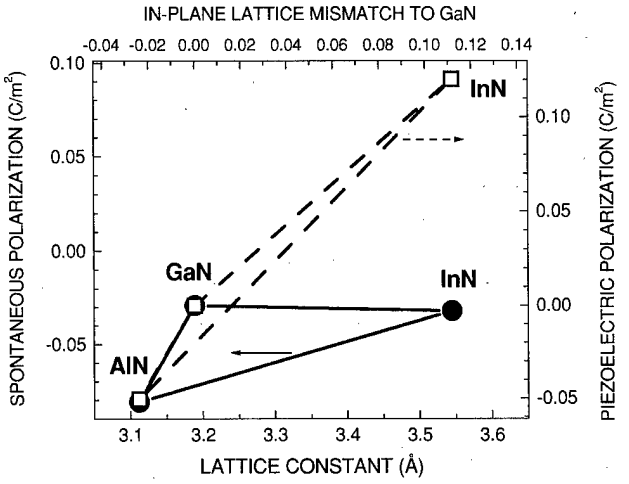


FIG. 1. Spontaneous (filled circles) and piezoelectric (open squares) polarizations for the binary III–N compounds. The piezoelectric polarization is calculated for a pseudomorphic epitaxial layer on a GaN thick substrate, with the assumption of biaxial strain in the basal plane. The spontaneous polarizations are taken from Refs. 8 and 9.

In Fig. 1 theoretical spontaneous and piezoelectric polarizations are shown for the AlN, GaN, and InN system. The connecting lines represent the corresponding quantities for the alloys under the assumption of linear interpolation between the binary compounds. The piezoelectric terms are calculated for pseudomorphic epitaxial layers on GaN ($\delta P_3^{GaN} = 0$), which is the most common case in device heterostructures. Relevant for the built-in electrostatic fields in multiquantum well heterostructures (MQWs) are the polarization differences between two interfacing layers.[9] It can clearly be seen that while in the case of InGaN/GaN the piezoelectric contribution is dominant this is not at all the case for AlGaN/GaN heterostructures. Here the polarization differences are quite exactly the same and therefore, both contributions must be considered equally for a correct determination of the electrostatic fields.

In the case of interest here, a single heterojunction with a thin overlayer on top of a thick field free substrate, the polarization fields in each layer i are simply given by[9]

$$\mathbf{E}_i = -\frac{\mathbf{P}_i^T}{\varepsilon_i \varepsilon_0}. \tag{2}$$

In Fig. 2 the band scheme is plotted for some selected single heterojunctions. With the assumption that the polarization fields can be sustained in very thin overlayers (<10 nm of thickness) such fields are expected to be very high, of the order of 1–10 MV/cm. This is true even if only the *spontaneous polarization* is considered, in absence of strain (continuous lines). In the presence of strain, only the valence band is indicated under the assumption that the valence band discontinuity does not change significantly with the strain. There are some indications in the literature that at least for wurtzite GaN only small energy changes with strain are seen

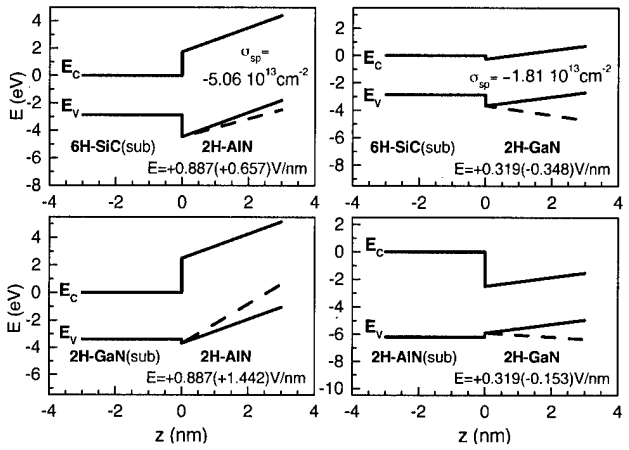


FIG. 2. Calculated band scheme at some nitride heterojunctions by considering the contribution of only the spontaneous polarization for *relaxed* heterostructures according to Eq. (2) (continuous line). For each heterojunction as well the valence band for the case of a *pseudomorphic* overlayer, spontaneous plus piezoelectric contributions, is shown (dashed line). Space and surface state charges are not considered. Polarization charge densities at the surface of the relaxed GaN and AlN overlayers as well as the electrostatic fields are shown (in parentheses for a pseudomorphic overlayer).

in the highest valence band.[17] Of course the gap energy and the conduction band minimum of the strained layer will be more affected by strain but these effects are not indicated here.

In the following we will show how photoemission experiments can give insight into these polarization properties at single heterojunctions. The band offsets shown in Fig. 2 are those determined from our experiments (Sec. IV C).

III. EXPERIMENT

AlN and GaN epitaxial layers are grown by MBE with a rf plasma source for the activated nitrogen supply. The GaN/SiC and AlN/SiC heterostructures are grown on ultrahigh vacuum clean 6H–SiC(0001)–(1×1) Si terminated substrates at a temperature of 620–640 °C. The AlN on GaN heterostructure is grown on a thick (~1 μm) GaN epitaxial layer and the GaN on AlN on a thick AlN layer. In both cases the thick layers are deposited on a LT-AlN (20 nm)/6H–SiC substrate. The growth temperatures are 790 °C for GaN, 790–870 °C for AlN, and 620 °C for the LT-AlN buffer. For all heterostructures under study successive thin overlayers are deposited with a total thickness between 0.5 and 6 nm. After each deposition the sample is transferred *in situ* into the XPS unit, where photoemission spectra are measured with monochromatized Al $K\alpha$ radiation ($h\nu$ = 1486.6 eV) and with an overall energy resolution of 0.43 eV. After each deposition step both AlN and GaN surfaces show a clear (1×1) electron diffraction pattern (LEED), with the same orientation as on the substrate.

Convergent beam electron diffraction (CBED) pattern analysis was applied to determine the polarity of nitrides grown by MBE on 6H–SiC_{Si}. In this experiment the CBED pattern on a thick epitaxial GaN layer grown on an AlN (buffer)/6H–SiC_{Si} substrate was measured and the polarity

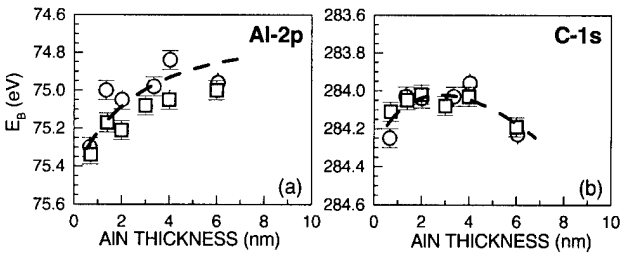


FIG. 3. XPS core level energy positions measured after successive AlN deposition steps. (a) Al 2p from the AlN overlayer; (b) C 1s from the SiC substrate. Two independent sets of experiments (squares and circles) are shown.

was evaluated by comparison with simulations.[18] The obtained result is: GaN growth along the [0001] or Ga-face growth, which confirms analogous recent results.[19] Here the same polarity (Al-face growth or correspondingly Si–N interface bond) is reasonably assumed as well for the thin AlN interface layer.

The core level shifts observed in the following is explained by the presence of an electric field in the overlayer, which bends the energy band diagram. In fact, due to the surface sensitivity of XPS the main contribution to the photoemitted signal arises from the topmost layers whereas the intensity from underlying regions is exponentially reduced according to the escape depth of the photoelectrons. The given data representation in Figs. 3–5 with decreasing binding energy scales directly reflects the band bending direction.

The valence band discontinuity at a given heterojunction is determined by three independent core level XPS measurements.[20] In this study the Al 2p, Ga 3p, and C 1s core levels are considered for the overlayers and the substrate, respectively. First the energetic distance of the core level relative to the valence band maximum E_V is determined on bulk standards. The energy difference $\Delta E_{cl}=[E_{cl}^{\text{sub}}-E_{cl}^{\text{ov}}]$ of the two core levels is then measured at the interface in the actual heterojunction and the valence band offset is given by

$$\Delta E_V=E_V^{\text{sub}}-E_V^{\text{ov}}=\Delta E_{cl}+[E_{cl}-E_V]^{\text{ov}}-[E_{cl}-E_V]^{\text{sub}}. \tag{3}$$

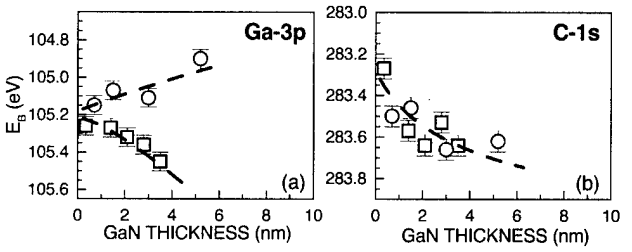


FIG. 4. XPS core level energy positions measured after successive GaN deposition steps. (a) Ga 3p from the AlN overlayer; (b) C 1s from the SiC substrate. Two independent sets of experiments (squares and circles) are shown.

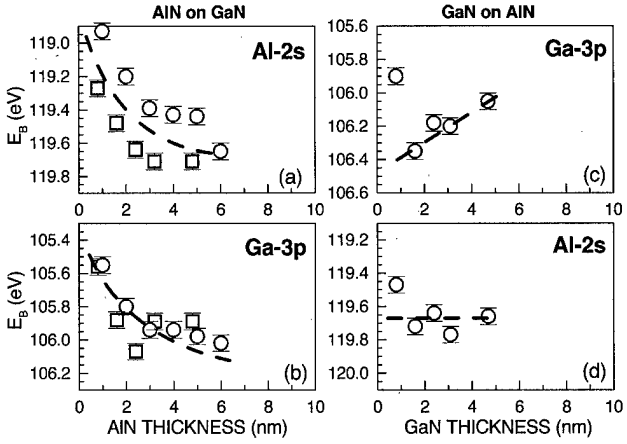


FIG. 5. XPS core level energy positions measured after successive deposition steps. (a) Al $2s$ from the AlN overlayer; (b) Ga $3p$ from the GaN "substrate." (c) Ga $3p$ from the GaN overlayer; (d) Al $2s$ for the AlN substrate. Two independent sets of experiments (squares and circles) are shown in (a) and (b).

IV. PHOTOEMISSION RESULTS

A. AlN on 6H–SiC–GaN on 6H–SiC

Figure 3 shows the evolution of the Al $2p$ and C $1s$ core level binding energies measured after deposition of successive AlN overlayers. Two independent sets of data are shown, corresponding to epitaxial layers grown under very similar experimental conditions. The dashed lines are only a guide to the eye for a qualitative recognition of the band bending in the overlayer and in the substrate.

The Al $2p$ shifts to lower binding energies at increasing AlN thickness, up to 6 nm. The slope at lower thicknesses corresponds to an electric field $E^{\mathrm{AlN}}\sim 0.1$ V/nm in the AlN; the shift tends to saturation at higher thicknesses. The C $1s$ after a slight shift to lower E_B shows an increasing binding energy with the overlayer thickness. The value of ΔE_{cl} is determined for each AlN thickness and obviously the shift of the core levels causes an *apparent dependence* of ΔE_V on the AlN thickness [Eq. (3) and Fig. 6].

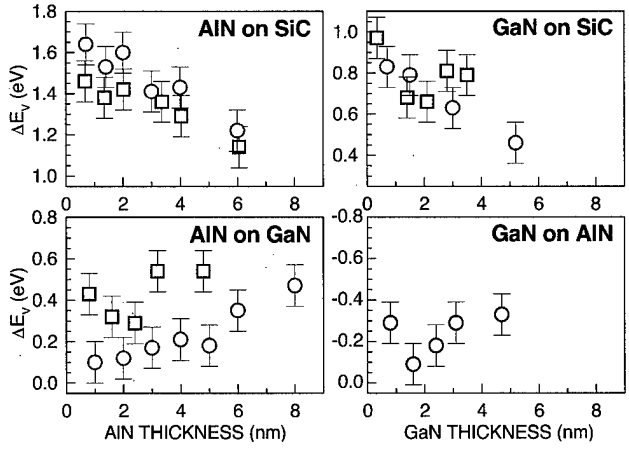


FIG. 6. Valence band offset as determined from the XPS core level measurements through Eq. (3).

TABLE I. Spectroscopic (XPS) energy differences for the bulk materials. For all values an error of 0.1 eV must be considered.

	6H–SiC	AlN	GaN
E_F-E_V (eV)	2.1±0.1	3.59±0.1	2.36±0.1
E_V-E_{cl} (eV)	281.32 C $1s$	71.41 Al $2p$	102.59 Ga $3p_{3/2}$
	150.17 Si $2s$	115.93 Al $2s$	17.46 Ga $3d$

The same analysis has been performed for epitaxial GaN on 6H–SiC and the core level shift in dependence on the GaN thickness is shown in Fig. 4. The two series of data show the same trend for the substrate peak [Fig. 4(b)] whereas a different one is revealed for the overlayer core level [Fig. 4(a)]. The first set of data (open circles) has been obtained on four different SiC/GaN samples, each grown under the same conditions ($T=620\,^{\circ}$C) but with increasing overlayer thicknesses. X-ray absorption spectroscopy at the Ga edge on these samples revealed that the growth of GaN is always relaxed.[21] The electric field extracted here is $E_1^{\mathrm{GaN}}\sim 0.05$ V/nm. The second set of data (open squares) has been obtained from one and the same sample after successive growth steps ($T=620\,^{\circ}$C) and shows an opposite direction of the built-in field, E_2^{GaN} appproximately -0.06 V/nm.

B. AlN on GaN–GaN on AlN

The AlN/GaN and reverse heterojunctions have been realized on a thick GaN (AlN) epilayer, otherwise the experiments were performed within the same scheme as previously. In Fig. 5 the core level shifts for the two heterojunctions are summarized, left AlN on GaN and right GaN on AlN. For the AlN on GaN a clear shift downwards is obtained for both core levels, more pronounced in the overlayer. The field in the AlN extracted by a linear interpolation of the points at lower thickness (<3 nm) is E^{AlN} approximately -0.23 V/nm. It is worth noting, that the thickness and energy scales are such that the slopes in all plots are comparable. More difficult is to identify a clear tendency for the GaN on AlN experimental data, the guide lines shown are drawn neglecting the points at the smallest thickness because of the margin of error.

C. Valence band offsets

The valence band offset has been determined for each thickness by considering the core level energy distance at the heterojunctions, together with their distance from the VBM in their respective bulk materials [Eq. (3)]. The values determined for the bulk standards are summarized in Table I.

As can be seen in Fig. 6, ΔE_V shows a dependence on the thickness for all heterojunctions, which arises from the core level shifts described in the previous sections. To eliminate any apparent dependence arising from the presence of polarization fields, we decided to determine the VBO at the different heterojunctions as extrapolated value at zero thickness. These values are summarized in Table II.

TABLE II. Valence band offsets (eV) as determined by XPS by extrapolating the observed dependence at zero thickness (Fig. 6). A positive value indicates that the compound on the right-hand side has a lower valence band.

Substrate	AlN	GaN
6H–SiC	(1.5–1.7)±0.1	(0.7–0.9)±0.1
AlN		−0.3 ± 0.1
GaN	(0.15–0.4)±0.1	

V. 6H–SiC/AlN AND 6H–SiC/GaN: BAND SCHEME AT THE HETEROJUNCTIONS

A self-consistent sp^3d^5s* tight-binding approach which is able to describe the system in the whole Brillouin zone, thus allowing to consider both direct and indirect semiconductors has been applied.[22,23] Moreover, the microscopic description is able to treat thin overlayers, overcoming all the limitations of the envelope function methods. In this way a comprehensive theoretical discussion of the experimental data is provided. The parameter of our empirical TB model were determined by fitting DFT-LDA band structure as outlined in Ref. 24. Self-consistency is obtain coupling the Schrödinger equation (in the TB basis) with the Poisson equation

$$\frac{d}{dz}D=\frac{d}{dz}\left(-\varepsilon\frac{d}{dz}V+P^T\right)=e(p-n). \tag{4}$$

The (position-dependent) quantities D, ε, and V are, respectively, the displacement field, dielectric constant, and potential. The (position-dependent) transverse polarization P^T is in general the sum of the spontaneous component calculated *ab initio*,[8] and of the piezoelectric component. In this first study, only the band scheme at relaxed overlayers (no piezoelectric component) will be presented and discussed.

The band scheme close to the n-6H–SiC/2H–AlN heterojunction has been calculated [Fig. 7(a)]. The substrate doping is $n=10^{18}\,\mathrm{cm}^{-3}$, the AlN overlayer is undoped. It can be seen that the polarization field bends the band scheme upwards in AlN, with $E=0.75$ V/nm independent of the overlayer thickness. No bending is observed in the substrate.

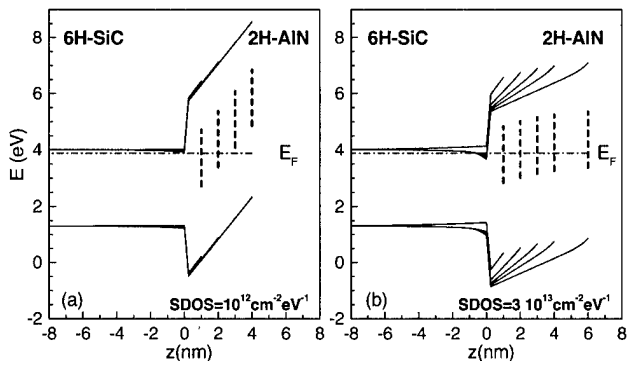


FIG. 7. TB self-consistent calculation of a *relaxed* AlN/SiC heterostructure for growing thickness of the overlayer and taking into account the role of a constant surface density of states SDOS (sketched with vertical dashed lines). (a) and (b) refer to different values of the SDOS at the AlN surface, located from 2.5 to 4.5 eV above the VBM.

Experimentally a similar upwards bending is measured as can be seen from the Al2p core level shifts in Fig. 3(a) however with a field, ~0.1 V/nm, lower than calculated. Taking into account a strain induced field ($E^{pz}=-0.23$ V/nm) in the case of a pseudomorphic AlN overlayer would only slightly reduce the field due to the spontaneous polarization (Fig. 2), not enough to explain the experimental data. It is worth noting here that since the piezoelectric field points in the $(000\bar{1})$ direction, neglecting the contribution of the spontaneous polarization would result in a downwards band bending in AlN, obviously in contradiction with the experiment.

A probable mechanism for the reduction of the polarization field is related to the presence of AlN surface states. These are accounted for in a simplified way, by considering a uniform distribution in the AlN gap. The energy position of the AlN gap states is between 2.5 and 4.5 eV above the AlN VBM; their charge neutrality level is in the middle. This midgap surface density of states (SDOS) is in qualitative agreement with the surface state energies calculated for a (1×1) ideal surface.[25] Furthermore it is consistent with the E_F energy position close to midgap determined experimentally (Table I). Different electronic properties have been inferred for the AlN (0001) surface by Wu and Kahn,[26] which derive a pinning of the Fermi level in the upper part of the gap, 1.4 eV from the CBM. Two main points are noticed here: (i) the AlN in Ref. 26 has been grown on a large mismatched substrate, Si(111), (ii) the surface analysis has been performed in a separate UHV system after repeated nitrogen sputtering and annealing cycles to get a clean surface. We think both are critical issues in achieving a well- defined (intrinsic) surface. The *in situ* analysis of the MBE as grown surface performed in our study guarantees a surface structure closer to the ideal one; furthermore, the chosen growth conditions ensure a two dimensional growth of AlN on 6H–SiC, as confirmed by the exponential attenuation of the substrate core level intensities.

We have performed several calculations for a surface density of states ranging between 10^{12} and $10^{16}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$. Figure 7(b) shows the surface state (SS) effect for an intermediate SDOS ($3\times10^{13}\,\mathrm{cm}^{-2}\mathrm{eV}^{-1}$). At increasing AlN thickness the SS carries an increasing positive charge, which partially compensates the negative surface polarization charge. As a consequence the electric field intensity decreases with the AlN thickness. The energy shift of the valence band edge qualitatively reproduces the experimental Al2p core level shift in Fig. 3(a). Even the substrate C1s energy shift [Fig. 3(b)] can be explained by the downward bending in SiC induced by the positive surface charge in AlN, which attracts electrons from the SiC close to the heterojunction.

The interplay of AlN surface states with the polarization fields is further supported by the adsorption of activated hydrogen on the AlN surface. This experiment has been carried out for several samples and systematically a 0.2 eV shift of the Al2p towards *lower binding energies* is observed. This is consistent with the removal of the surface state DOS in the

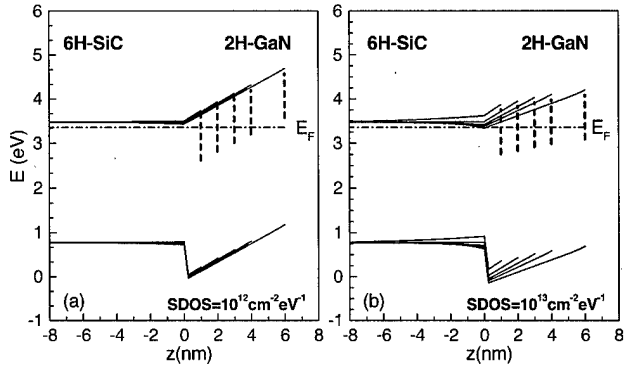


FIG. 8. TB self-consistent calculation of a *relaxed* GaN/SiC heterostructure for growing thickness of the overlayer and taking into account the role of a constant surface density of states SDOS (sketched with vertical dashed lines). (a) and (b) refer to different values of the SDOS at the GaN surface, located from 2.4 to 3.4 eV above the VBM.

gap, i.e., of the compensating positive charge, with consequent increase of the polarization field in the overlayer.

An analogous model has been considered for the *relaxed* 6H–SiC/GaN heterojunction in Fig. 8. A GaN SDOS has been assumed between 2.4 and 3.4 eV, consistent with previous determinations of the Fermi level pinning at such surfaces.[27,28] Even in this case the SDOS induces a reduction of the polarization field in GaN, according to the same charging behavior as for AlN. This scheme would account for the set of data with an upward shift in Fig. 4. A different mechanism must be responsible for the downward shifting squares. A strained SiC/GaN heterostructure would explain the observed results, as seen in Fig. 2. Since the growth conditions are different for the two cases (Sec. IV A) this assumption cannot be excluded *a priori*, but needs experimental confirmation.

We now turn to the AlN on GaN and reverse heterojunctions for which a more qualitative discussion, based on the band scheme in Fig. 2 is presented here. As already pointed out in Sec. IV B the AlN on GaN shows a pronounced shift for thicknesses lower than 3 nm, with a decreasing slope at higher overlayer thicknesses. The critical thickness in a GaN/AlN multilayer structure has been determined to be around 3 nm[6] and a lower field is expected for a relaxed AlN layer as compared with the pseudomorphic structure (Fig. 2). The experimental data are well consistent with this picture, but the band bending has the opposite sign. A straightforward explanation is that the polarity of the AlN grown on GaN has changed to $[000\bar{1}]$ (N face), opposite to that of ultrathin AlN grown directly on 6H–SiC. This could be explained if the growth of ultrathin AlN on GaN at the experimental conditions chosen here is affected by the presence of antiphase domains. A TEM analysis of the kind of samples presented here is in progress.

The last heterojunction is GaN on AlN [Figs. 5(c) and 5(d)]. Since we only have one set of data, these results should be treated as preliminary. If any tendency is observed, this corresponds to the growth of a relaxed GaN overlayer on top of AlN.

VI. VALENCE BAND OFFSETS AND POLARIZATION FIELDS

It is worthwhile to point out the main message here: *if not aware of the role played by the strong polarization fields in nitride heterostructures the VBO determination by XPS might well be erroneous.* In fact, Fig. 6 clearly shows that a band offset value obtained for a single overlayer thickness is meaningless. Since the potential step at a heterojunction is determined by the band structure of the two interfacing solids and extends just over one or two atomic interface planes, in our opinion the experimental value which best represents the band offset is that obtained by extrapolation to zero thickness of the overlayer (Table II).

All heterojunctions but GaN on 6H–SiC are of type I (straddled), whereas the 6H–SiC/GaN is staggered. Our results for 6H–SiC/AlN are in good agreement with the only experimentally determined ones known to us at the same heterojunction,[29] if one considers that there the value of (1.4 $\pm$0.3) has been measured on a 1–2 nm thick AlN. Theoretical calculations give a wide range of values for the polar SiC/AlN and SiC/GaN heterojunctions, depending on the interface compositions,[30,31] in particular valence band offsets for SiC/AlN as high as 2.5 eV are obtained for neutral interfaces with an Al/Si mixed plane and as low as 1.3 eV with a N/C mixed plane.[30] We do not have any experimental evidence of a strong intermixing at the two heterovalent interfaces, which would be evident with reacted core level components at the lowest overlayer thicknesses. Interestingly at the SiC/GaN interface the band discontinuity is completely found in the valence band, which makes the GaN very appealing as a wide band gap emitter in *n*-GaN/*p*-SiC/*n*-SiC HBTs. This band alignment in fact results in a very low base-emitter current, which improves the transistor performances. The value of 0.8 eV confirms the result obtained on device structures[32] and is higher than the valence band barrier of 0.43 eV for holes, determined by temperature dependent current gain measurements by Pankove *et al.*[33] Empirical tight-binding calculation of the branch-point energy of the continuum of interface-induced gap states has been reported to describe the band lineup at semiconductor heterostructures.[34] Values for the VBO of 1.53 and 0.93 eV are derived for SiC/AlN and SiC/GaN, respectively, in excellent agreement with our experimental data.

More puzzling are the data reported in the literature for the AlN/GaN system. A comprehensive table of experimental and theoretical VBO is reported by King *et al.* in Ref. 35. Experimental values reported up to now ranges from 0.5 eV (Refs. 35 and 36) through 0.7–0.8 eV (Refs. 35 and 37) up to 1.36 eV.[38] With the exception of Ref. 36, where the band discontinuity is obtained from the energy position in the gap of the (−/0) acceptor levels of iron, all the other studies are concerned with the VBO determination from XPS experiments. An extensive discussion about the reason of the large differences in the XPS determined VBOs is found in Ref. 35. We in principle agree with the conclusion drawn there, that the discrepancies are most related to differences in the heterostructure material quality, such as surface preparation,

defects, and stoichiometry. It is worth noting that the choice of the Ga $3d$ as the reference core level in GaN, which is the one generally considered in the literature, is in this respect questionable. In fact it is well known that due to the strong hybridization of the d bands with the bottom s-like valence band stemming from N $2s$, the Ga $3d$ levels cannot be treated as inert.[39] This explains the large scattering of the Ga $3d$-VBM energy values observed in the literature (cf. Table II in Ref. 35) with reference to different material quality. For this reason the deeper Ga $3p_{3/2}$ is chosen in our experiment, which also has a binding energy much closer to the Al $2s$ and therefore very similar photoelectron escape depth.

The VBO value of -0.3 ± 0.1 eV for GaN on AlN is smaller than any value reported up to now in the literature.[35] It has already been noticed that the upwards shift of the Ga $3p$ core level with increasing thickness in Fig. 5 is an indication of a nonstrained GaN. We therefore assign this value to a relaxed AlN/GaN heterostructure. More difficult is to give a VBO value for the reverse heterojunction, AlN on GaN (Fig. 6). The large shifts with the overlayer thickness observed in Figs. 5(a) and 5(b) force us to indicate an energy range extracted from the lower thickness values (0.15–0.4) ±0.1. Whether or not this corresponds to a strained heterostructure is unclear even though a trend with the thickness is recognized; furthermore the sign of the polarization fields has to be cleared as discussed in the previous section. We also here determine very small values, as compared with the literature. A tendency towards larger *apparent* VBO values is seen at increasing thicknesses for both AlN/GaN heterostructures (Fig. 6). This is *apparent* in the sense that is masked by the presence of the large polarization fields, however the trend possibly related to the AlN relaxation gives VBO values which increase with strain relaxation. Theoretical studies calculate for wurtzite GaN/AlN(0001) VBOs which range from 0.57 to 0.85 eV.[34,40–42] Even though strain effects are emphasized, the two works which report on strained heterojunctions give values at the opposite sides of the range for the same AlN/GaN system: 0.57 eV[40] and 0.85 eV.[42] The VBOs for nonstrained heterostructures are found in between.[34,41] The value of 0.2 eV calculated by Bernardini and Fiorentini for strained AlN on GaN compares fairly well with our experiment. The large strain induced asymmetry calculated by the same authors, 0.2 eV for AlN/GaN(0001), 0.85 for GaN/AlN(0001), is not reproduced by our experiment. On the other hand there is evidence for a relaxed GaN/AlN(0001) heterostructure in our experiment, which would explain the discrepancy with the theory.

VII. CONCLUSIONS

From a series of *in situ* photoemission experiments macroscopic electric fields are clearly demonstrated in wurtzite nitride heterostructures grown by MBE on 6H–SiC(0001). A significant contribution is due to the *spontaneous* polarization; the *piezoelectric* term alone would not explain the sign of the field measured in SiC/AlN. The experimental field has lower intensity as compared to theory: the role of electronic gap states at the surface is pointed out. For AlN the influence of such band gap states has been evidenced by hydrogen absorption. The valence band offset has been determined for all heterojunctions under study and the apparent dependence on the overlayer thickness, due to the presence of the strong polarization fields, has been pointed out in view of a correct determination of the VBO. The VBOs at the heterojunctions obtained by extrapolation to zero overlayer thickness are (substrate/overlayer): $[(1.5-1.7)\pm0.1]$ eV for SiC/AlN, $[(0.7-0.9)\pm0.1]$ eV for SiC/GaN, (-0.3 ± 0.1) eV for AlN/GaN, and $[(0.15-0.4)\pm0.1]$ eV for GaN/AlN.

ACKNOWLEDGMENTS

The authors wish to thank Dirk Freundt for his earlier contribution to this work, Reinhard Schmidt and Dirk Holz for taking part at some of the experiments; K. Wambach and F. Ringelmann are acknowledged for technical support. We are grateful to F. Bernardini and A. Catellani for many interesting discussions and to S. Frabboni for the CBED analysis. The financial support of the Deutsche Forschungsgemeinschaft (DFG) and the Vigoni Program are kindly acknowledged. Partial support from CNR-MADESS and MURST are also acknowledged (A.R.).

[1]For recent reviews, see, F. A. Ponce and D. P. Bour, Nature (London) **386**, 351 (1997); J. W. Orton and C. T. Foxon, Rep. Prog. Phys. **61**, 1 (1998).
[2]The zinc blende due to the noncentrosymmetrical properties of the ⟨111⟩ polar axes is piezoelectric along this direction.
[3]E. T. Yu, G. J. Sullivan, P. M. Asbeck, C. D. Wang, D. Qiao, and S. S. Lau, Appl. Phys. Lett. **71**, 2794 (1997); E. T. Yu, X. Z. Dang, L. S. Yu, D. Qiao, P. M. Asbeck, S. S. Lau, G. J. Sullivan, K. S. Boutros, and J. M. Redwing, *ibid.* **73**, 1880 (1998).
[4]J. S. Im, H. Kollmer, J. Off, A. Sohmer, F. Scholz, and A. Hangleiter, Phys. Rev. B **57**, R9435 (1998).
[5]L. Hsu and W. Walukiewicz, Appl. Phys. Lett. **73**, 339 (1998).
[6]A. D. Bykhovski, B. L. Gelmont, and M. S. Shur, J. Appl. Phys. **81**, 6332 (1997).
[7]H. S. Kim, J. Y. Lin, H. X. Jiang, W. W. Chow, A. Botchkarev, and H. Morkoç, Appl. Phys. Lett. **73**, 3426 (1998).
[8]F. Bernardini, V. Fiorentini, and D. Vanderbilt, Phys. Rev. B **56**, R10024 (1997).
[9]V. Fiorentini, F. Bernardini, F. Della Sala, A. Di Carlo, and P. Lugli, Phys. Rev. B (to be published).
[10]M. Posternak, A. Baldereschi, A. Catellani, and R. Resta, Phys. Rev. Lett. **64**, 1777 (1990).
[11]R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).
[12]R. Resta, Rev. Mod. Phys. **66**, 899 (1994).
[13]L. D. Landau and E. M. Lifschitz, *Elektrodynamik der Kontinua, Lehrbuch der Theoretische Physik* (Akademie, Berlin, 1980), Vol. VIII.
[14]A. Baldereschi, M. Posternak, and R. Resta, Phys. Rev. Lett. **69**, 390 (1992).
[15]F. Bernardini, V. Fiorentini, and D. Vanderbilt, Phys. Rev. Lett. **79**, 3958 (1997).
[16]A. F. Wright, J. Appl. Phys. **82**, 2833 (1997).
[17]B. Jogai, Phys. Rev. B **57**, 2382 (1998).
[18]P. A. Stadelmann, Ultramicroscopy **21**, 131 (1987).
[19]P. Vermaut, P. Ruterana, and G. Nouet, Philos. Mag. A **76**, 1215 (1997).
[20]J. R. Waldrop and R. W. Grant, Phys. Rev. Lett. **43**, 1686 (1979).
[21]F. Boscherini, R. Lantier, A. Rizzi, F. D'Acapito, and S. Mobilio, Appl. Phys. Lett. **74**, 3308 (1999).
[22]A. Di Carlo, S. Pescetelli, M. Paciotti, P. Lugli, and M. Graf, Solid State Commun. **98**, 803 (1996); A. Di Carlo, Mater. Res. Soc. Symp. Proc. **491**, 389 (1998).
[23]F. Della Sala, A. DiCarlo, P. Lugli, F. Bernardini, V. Fiorentini, R. Scholz, and J.-M. Jancu, Appl. Phys. Lett. **74**, 2002 (1999).

[24]J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. B **57**, 6493 (1998).
[25]J. E. Northrup, R. Di Felice, and J. Neugebauer, Phys. Rev. B **55**, 13878 (1997).
[26]C. I. Wu and A. Kahn, Appl. Phys. Lett. **74**, 546 (1999).
[27]A. Rizzi and H. Lüth, Nuovo Cimento D **20**, 1039 (1998).
[28]V. M. Bermudez, J. Appl. Phys. **80**, 1190 (1996).
[29]S. W. King, M. C. Benjamin, R. J. Nemanich, R. F. Davis, and W. R. L. Lambrecht, Mater. Res. Soc. Symp. Proc. **395**, 375 (1996).
[30]P. Ferrara, N. Binggeli, and A. Baldereschi, Phys. Rev. B **55**, R7418 (1997).
[31]M. Städele, J. A. Majewski, and P. Vogl, Phys. Rev. B **56**, 6911 (1997).
[32]M. Topf, D. Meister, I. I. Dirnstorfere, G. Steude, S. Fisher, B. K. Meyer, A. Kritschil, H. Witte, J. Christen, T. U. Kampen, and W. Mönch, Mater. Sci. Eng., B **50**, 302 (1997).
[33]J. I. Pankove, M. Leksono, S. S. Chang, C. Walker, and B. Van Zeghbroeck, MRS Internet J. Nitride Semicond. Res. **1**, 39 (1996).
[34]W. Mönch, J. Appl. Phys. **80**, 5076 (1996).
[35]S. W. King, C. Ronning, R. F. Davis, M. C. Benjamin, and R. J. Nemanich, J. Appl. Phys. **84**, 2086 (1998).
[36]J. Baur, K. Maier, M. Kunzer, U. Kaufmann, and J. Schneider, Appl. Phys. Lett. **65**, 2211 (1994). Mater. Sci. Eng., B **29**, 61 (1995).
[37]G. Martin, S. Strite, A. Botchkarev, A. Agarwal, A. Rockett, H. Morkoç, W. R. L. Lambrecht, and B. Segall, Appl. Phys. Lett. **65**, 610 (1994); J. Electron. Mater. **24**, 225 (1995); G. Martin, A. Botchkarev, A. Rockett, and H. Morkoç, Appl. Phys. Lett. **68**, 2541 (1996).
[38]J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **68**, 2879 (1996).
[39]V. Fiorentini, M. Methfessel, and M. Scheffler, Phys. Rev. B **47**, 13353 (1993).
[40]M. Buongiorno Nardelli, K. Rapcewicz, and J. Bernholc, Phys. Rev. B **55**, R7323 (1997).
[41]S.-H. Wei and A. Zunger, Appl. Phys. Lett. **69**, 2719 (1996).
[42]F. Bernardini and V. Fiorentini, Phys. Rev. B **57**, R9427 (1998).

Epitaxial growth and electronic structure of lanthanide silicides on *n*-type Si(111)

S. Vandré, T. Kalka, C. Preinesberger, and M. Dähne-Prietsch[a)]
Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstraße 36, D-10623 Berlin, Germany

(Received 20 January 1999; accepted 30 March 1999)

We present an investigation of Gd, Dy, Er, and Lu silicides epitaxially grown on *n*-type Si(111) using scanning-tunneling microscopy, low-energy electron diffraction, and core-level photoemission. With varying silicide film thickness, we observe structurally, electronically, and chemically different silicide phases. In particular we found that the Si-2*p* photoelectron spectra from monolayer silicides are different from those at higher coverages. This observation is attributed to the structure of the layered hexagonal silicide, with a Si vacancy lattice only present in case of multilayer films. Furthermore, we observe peculiar electronic properties: An extremely low band bending is found in the monolayer range, with a Fermi-level position of only 0.08 ± 0.05 eV below the conduction-band minimum of silicon, representing the lowest value ever observed on *n*-type silicon. With increasing coverage, a final Schottky-barrier height of 0.32 ± 0.05 eV is obtained. This behavior is interpreted as a consequence of the developing metallicity of the silicide overlayer and will be discussed in the framework of theoretical models for Schottky-barrier formation. © *1999 American Vacuum Society.* [S0734-211X(99)03404-6]

I. INTRODUCTION

The atomic arrangement and the electronic structure of metal overlayers on semiconductors are of considerable interest, both because of their importance for contacts in microelectronic devices and for a fundamental physical understanding. From a technological point of view, one main goal is to find materials forming a stable interface with low lattice mismatch. Their applicability depends further on their Schottky-barrier height. The barrier heights of common metals such as Al, Ag, Cu, and Au on *n*-type Si fall in the range between 0.7 and 0.8 eV, whereas only a few Schottky contacts are known with lower values, e.g., Hf and Ti silicides with barrier heights scattering between 0.5 and 0.6 eV.[1]

Silicides of the heavy trivalent lanthanides (Ln) on Si(111), in contrast, show even lower Schottky-barrier heights of only 0.3–0.4 eV, currently representing the lowest known values for metal/*n*-Si contacts.[2] This behavior makes them in particular interesting for applications as ohmic contacts for *n*-type Si. Furthermore, the corresponding high barrier height on *p*-type Si substrates[3] is an interesting property for infrared detectors or photovoltaic applications. Moreover, these silicides are characterized by a small lattice mismatch relative to the Si(111) surface, ranging from zero to 2.5%, therewith allowing epitaxial growth with chemically sharp interfaces and a high degree of crystallinity and structural perfection. These interfaces are thus well suited for a microscopic understanding of the structural and electronic properties as well as their influence on Schottky-barrier formation.

Ln silicides can be grown on Si substrates by codeposition of Ln and Si at elevated temperatures or by solid-phase epitaxy, i.e., by Ln deposition on Si and subsequent annealing.[2,4–6] In the latter case, the diffusion processes necessary for silicide formation give rise to a complex interplay of energetics and kinetical processes, requiring a detailed control of the process parameters.

In the present article, we report on a detailed study of the epitaxial growth, the chemical composition, and the Schottky-barrier formation of silicides from the heavy trivalent lanthanides Gd, Dy, Er, and Lu on *n*-type Si(111) by means of scanning-tunneling microscopy (STM), low-energy electron diffraction (LEED), and in particular high-resolution core-level photoemission spectroscopy (PES). The high spectral resolution in the PES experiments allows for the first time to identify chemically different components of the Si-2*p* lines as a function of silicide thickness, while the Ln-4*f* spectra always consist of only one component. This behavior can be interpreted within the AlB_2 structure model for the silicide, developing Si vacancies in case of thicker films.[7] Furthermore, the variation of the Schottky-barrier height with the film thickness of the silicide is monitored using the Si-2*p* substrate emission. In the monolayer regime, we found the lowest barrier heights ever observed for metal/*n*-Si interfaces, increasing towards higher values with growing film thickness.[8] This behavior is assigned to the developing metallicity of the silicide films.

II. EXPERIMENTAL DETAILS

The PES experiments were performed at the PM-2 and PM-5 beamlines of the Berlin storage ring BESSY I, using a hemispherical Leybold EA-11 electron-energy analyzer in normal-emission geometry with an angular acceptance of about $\pm 10°$. The total-system resolution varied between 200 and 300 meV [full width at half maximum (FWHM)], de-

a)Electronic mail: daehne@physik.tu-berlin.de

pending on the photon energy. The reproducibility of the kinetic energies was 0.01 eV. In addition, LEED and STM were used for structural analysis.

The *n*-type Si(111) wafers (10–20 Ω cm) were rinsed with ethanol and destilled water before transfer into the ultrahigh vacuum. For the preparation of clean and well ordered Si(111)7×7surfaces, the samples were thoroughly outgassed at 600 °C for several hours. Then the Si-oxide layer was removed by a series of rapid heating up to 1200 °C, and the sample was subsequently cooled down slowly to room temperature with rates of 50 °C/min between 900 and 700 °C in order to provide sufficient formation time for the 7×7 reconstruction.[9] The sample temperature was controlled by means of an infrared pyrometer.

Silicide films were prepared by vapor-phase deposition of different high-purity Ln metals (Gd, Dy, Er, and Lu) out of W crucibles heated by electron bombardment. The evaporation rates of typically 1 Å/min were controlled using a quartz-crystal oscillator. Subsequent annealing at 500–600 °C yields epitaxially ordered silicide films. Their high degree of crystallinity is evidenced by sharp LEED patterns. The base pressure in the experimental chamber was 5×10^{-11} mbar, rising to 5×10^{-10} mbar during sample preparation.

III. RESULTS AND DISCUSSION

A. Growth of epitaxial silicide films

A variety of techniques has been used to determine the atomic structure of epitaxial Ln silicides on Si(111), with the main emphasis on Dy and Er silicides.[10–17] Since it was found that Dy and Er behave in a similar way, it can be assumed that these findings are representative for most of the members of the chemically similar heavy trivalent Ln.

For submonolayer coverages of 0.5 Å Dy on Si(111) the STM topograph in Fig. 1(a) shows that the Si(111)7×7 surface is inhomogeneously covered by silicide patches with a $2\sqrt{3}\times2\sqrt{3}$R30° superstructure interspersed with domain boundaries, which are related to kinetic limits during layer formation. In a detailed STM study on Er submonolayers it was demonstrated that the corresponding unit cell consists of two inequivalent halves filled with a total of six Er atoms.[16] Therefore this reconstruction will be obtained for coverages around 0.5 monolayers. Most probably this specific structure is not significant for the interface geometry of thicker films, where more Ln is available.

At increasing coverages the growth of flat silicide terraces is observed, as shown in Fig. 1(b). This monolayer silicide, which can be prepared from about 2.6 Å Ln, has a $p(1\times1)$ structure with threefold rotational symmetry. The corresponding atomic structure consists of an ordered hexagonal monolayer of Ln atoms accommodated underneath a buckled Si surface layer, which is similar to the (111) bilayers in bulk Si. This structure is shown in Fig. 2(a) and can be considered a single $LnSi_2$ layer with one Ln and two Si atoms in each (1×1) unit cell.[15]

With further increasing coverage, the growth of epitaxially ordered Ln_3Si_5 layers is observed, as depicted in Fig. 1(c). This bulk-like silicide forms a defected AlB_2-type structure, consisting of stacked hexagonal Ln planes and graphite-like Si planes with an ordered arrangement of vacancies at every sixth Si lattice site,[7] as shown in Fig. 2(b).

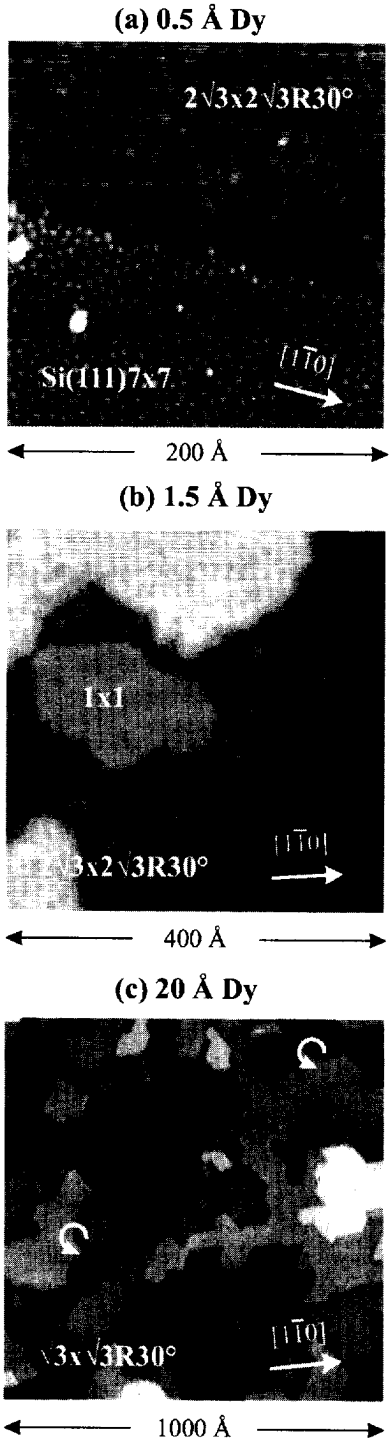


FIG. 1. STM images of Dy silicides on Si(111) prepared by Dy deposition and subsequent annealing. (a) $2\sqrt{3}\times2\sqrt{3}$R30° superstructure of a submonolayer annealed at a temperature of $T_a=500$ °C, (b) coexistence of the $2\sqrt{3}\times2\sqrt{3}$R30° superstructure with (1×1) reconstructed Dy-silicide islands formed at $T_a=500$ °C, and (c) topography of a thick silicide film formed at $T_a=600$ °C. Some screw dislocations are indicated by arrows.

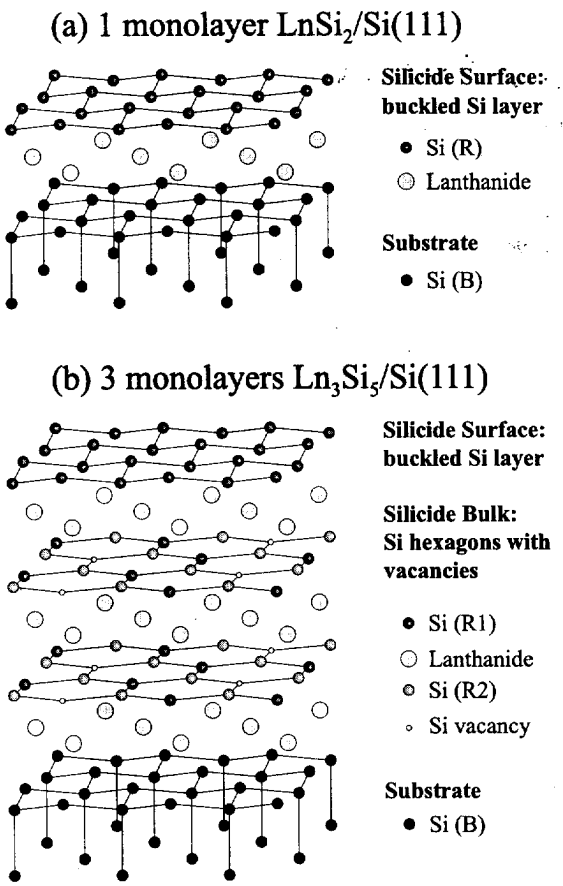


FIG. 2. Sketch of the atomic structure of (a) one monolayer $LnSi_2$ and (b) 3 monolayers Ln_3Si_5 on Si(111) (Refs. 7 and 15). See the text for details.

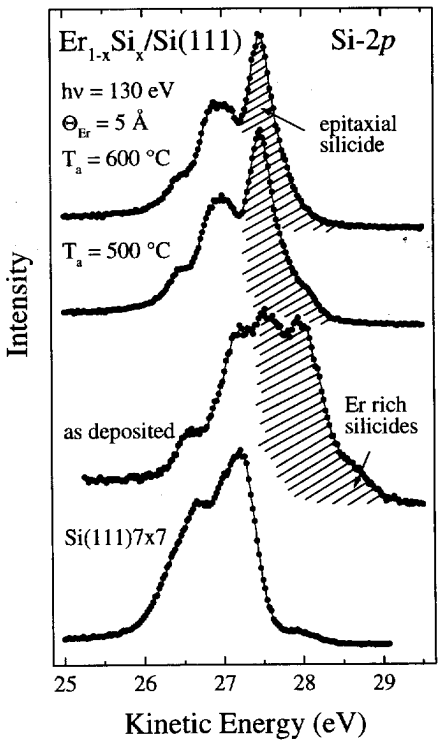


FIG. 3. Photoemission spectra of the Si-2p core level of 5 Å Er on Si(111)7×7 recorded at a photon energy of 130 eV as a function of annealing temperature T_a, demonstrating silicide formation. The spectra are normalized to equal heights. The shaded areas indicate the contribution from Si reacted with Er.

Its surface is again formed by a buckled Si layer without vacancies and was shown to be essentially identical to that of the monolayer silicide.[13–15] In LEED experiments of Gd, Dy, and Er silicide films, we observed the characteristic $\sqrt{3}\times\sqrt{3}$R30° pattern, which is assigned to an ordered superlattice of Si vacancies producing a $\sqrt{3}\times\sqrt{3}$ mesh in the silicide film. In contrast, we found a (1×1) pattern in case of several monolayer thick Lu coverages, indicating a different structure. This may be related to the Ln contraction towards the heavier Ln metals.[18] It should be mentioned that the growth of thick layers is relatively free of defects, but nevertheless accompanied by pinholes, stacking faults, and screw dislocations, as demonstrated in STM images like the one in Fig. 1(c). This growth behavior indicates the strong influence of formation kinetics, resulting in incomplete diffusion of the Ln and Si atoms upon postgrowth annealing.

B. Photoemission results

The electronic structure was studied in detail with PES. In Figs. 3 and 4, spectra of the Si-2p core level are presented for Er silicides.

In order to demonstrate the spectral behavior upon silicide formation, Fig. 3 shows Si-2p spectra of the bare Si(111) 7×7 surface as well as after deposition of 5 Å Er and subsequent stepwise annealing up to 600 °C. After Er deposition the substrate spectrum is superposed by intense features located at the high kinetic-energy side (shaded in Fig. 3). Since Si is the element with higher electronegativity, spectral components from Si reacted with Ln are expected to shift to higher kinetic energies with respect to those from nonreacted Si.[1] Therefore the new features can mostly be assigned to Si diffused into the Er overlayer forming an Er-rich silicide. After annealing at 500 °C a noticeable decrease in intensity of this Er-rich silicide occurs, resulting in the growth of a narrow line at lower kinetic energies dominating the Si-2p spectrum. This behavior indicates the formation of the epitaxial silicide. The disappearance of the Er-rich component is observed after further annealing at 600 °C and is assigned to complete silicide formation.

In the following, the annealing temperatures were adjusted to minimize the contribution from the Ln-rich components in order to achieve the best silicide formation with a large degree of ordering. At these temperatures, the LEED patterns exhibit sharp spots and the STM images show extended flat areas, further demonstrating the high crystallinity.

More details about the silicide formation can be inferred from Fig. 4, where a representative set of Si-2p spectra as a function of Er exposure is shown for optimum annealing temperatures, together with the components of a least-squares fit analysis. The data are taken at different photon energies in order to tune the escape depth of the photoelectrons, resulting in the highest surface sensitivity at $h\nu = 130$ eV. In this way, an unambiguous assignment of the spectral

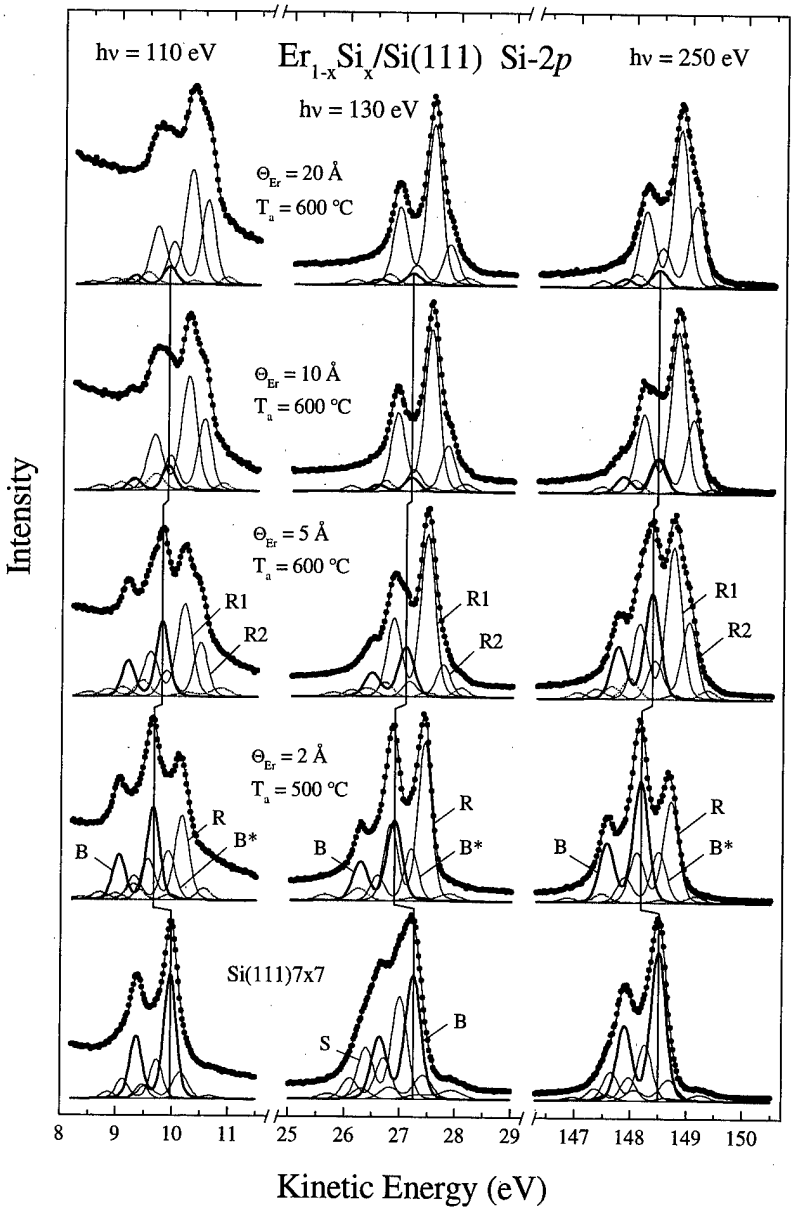


FIG. 4. Photoemission spectra of the Si-2p core level for Er silicides on Si(111) recorded at photon energies of 110, 130, and 250 eV as a function of Er coverage. The spectra (bold dots), normalized to equal heights, are shown together with the corresponding line shape analysis, consisting of substrate bulk (B, B^*, thick solid lines), substrate surface (S, thin solid lines), reacted silicides (R, $R1$, $R2$, thin solid lines), and minor components related to edges, pinholes and screw dislocations (thin dotted lines). The latter components were minimized by selecting an appropriate annealing temperature T_a. The energy variation of the substrate bulk silicon component, which reflects the band bending behavior, is indicated by the solid vertical lines.

components to the silicide or to the underlying substrate is possible on the basis of their relative intensities.

The spectra are well described by a superposition of Lorentzian peaks (FWHM=0.08±0.02 eV) convoluted by a Gaussian for describing both experimental resolution and crystalline order. Asymmetric Doniach–Šunjic shapes were also tried, since the Si(111)7×7 surface as well as the lanthanide silicides exhibit metallic character, but did not result in an improved fit of the line shape. An exponential background was used to simulate the secondary-electron contribution at $h\nu$=110 eV, and an integral background was taken to describe the contribution of inelastic processes at $h\nu$ =130 and 250 eV.

The spectra of the bare Si(111)7×7 surface consist of five identically spin-orbit split doublets, which originate from the substrate bulk (thick solid subspectrum B) and different surface sites (thin solid subspectra S).[19] For Er coverages between 0.5 and 1 Å only weak changes occur (not shown here). According to the manifold lattice sites observed by STM,[16] a variety of weak spectral components contributes to the spectra, making an unambiguous assignment impossible.

At Er coverages of 2 Å, the substrate-bulk signal shifts to lower kinetic energies. This observation is caused by band bending and will be discussed in detail in Sec. III D. In addition, the Si-2p substrate emission is now superposed by one doublet from reacted silicon (thin solid subspectrum R), shifted by 0.54±0.03 eV to higher kinetic energies with respect to the bulk-substrate component. This chemical shift is assigned to the above-mentioned charge transfer from Er to Si. Furthermore, there is still a weak unshifted spectral contribution (B*) from the bulk underneath the initial bare Si(111)7×7 surface or the $2\sqrt{3}\times2\sqrt{3}$R30° reconstructed monolayer, demonstrating that the surface is not completely covered by the silicide monolayer. This behavior is also sup-

ported by the STM image shown in Fig. 1(b) as well as by the nominal Er coverage, which is only sufficient for about 0.8 monolayers of the silicide.

Upon exposure of 5 Å Er and more, i.e., for several monolayer thick films, the substrate component shifts back to higher kinetic energies, and its intensity decreases relative to the emission from the silicide overlayer. The latter is now characterized by two silicide doublets (thin subspectra $R1$ and $R2$) shifted by 0.36 ± 0.04 and 0.65 ± 0.04 eV to higher kinetic energies with respect to the bulk-substrate component. This behavior clearly reflects the structural transition from the $p(1\times1)$ to the $\sqrt{3}\times\sqrt{3}$R30° symmetry observed by STM and LEED,[15] as will be discussed in Sec. III C.

In general, the Ln-silicide components are characterized by very sharp core-level features, indicating the high crystallinity of the silicide. Nevertheless, there are still weak spectral contributions (thin dotted subspectra) at the low- and high-energy sides of the Si-2p emission, which regularly amount to less than 5% of the total spectral intensity. These subspectra are attributed to particular atomic sites at terrace edges, pinholes, or screw dislocations, as observed in Fig. 1(c). It was found that these components have their lowest intensity, when annealing was performed at 600 °C. At even higher annealing temperatures, the substrate emission increases again, already indicating desorption, hole formation, or clustering of the silicide film.

For the other Ln silicides prepared from Gd, Dy, and Lu, a very similar spectral behavior as in Fig. 4 has been observed. This is demonstrated in Fig. 5, where a comparison of Si-2p spectra of samples prepared by deposition of 5-Å-thick Gd, Dy, Er, and Lu films and subsequent annealing is shown. In particular, the silicide components $R1$ and $R2$ appear with similar chemical shifts and relative intensities, as well as component R in case of the monolayer (not shown here). Small variations in intensity will be discussed in Sec. III C.

The respective Ln-4f core-level and valence-band photoemission spectra are presented in Fig. 6. The clean Si(111)7×7 surface displays the typical structure of the density of states of the Si valence band.[20] Minor differences in spectral shape can be attributed to variations in detection geometry. The 4f emission from the trivalent Ln becomes visible after a Ln deposition of 1 Å in all four cases. With increasing Ln coverage, these final-state multiplets strongly increase in intensity relative to the Si valence-band emission and finally dominate. In addition, the 4f emission gradually shifts to higher kinetic energies, i.e., to lower binding energies. In particular, abrupt shifts are observed in the coverage range between 1 and 5 Å, when new structural phases are built. At coverages exceeding 10 Å, the binding energies of the 4f multiplet remain constant and are located close to the characteristic values for the respective elemental Ln metals. A comparison of the 4f final-state multiplets with results from intermediate-coupling calculations[21] clearly shows that the photoemission lines consist of only one component, while the observed energy splittings are increased by about 5%–10% with respect to the theoretical values.

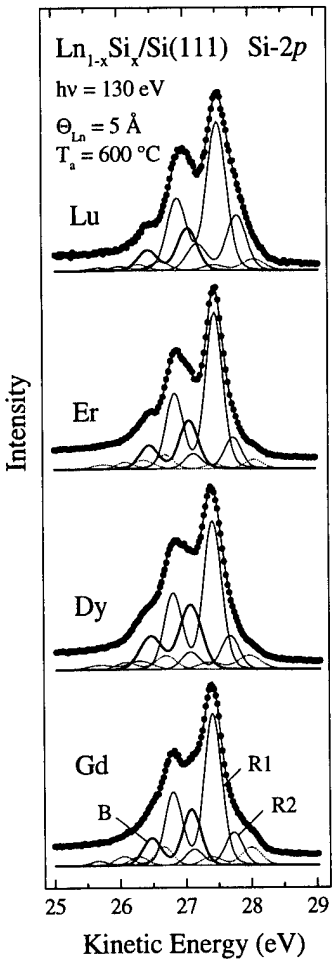


FIG. 5. Photoemission spectra of the Si-2p core level recorded at a photon energy of 130 eV for silicides formed from 5 Å of the various Ln metals. The spectra are normalized to equal heights and are aligned to equal substrate-bulk position. For more details see Fig. 4.

It is further observed in Fig. 6 that the density of states at the Fermi level increases with film thickness. In the monolayer regime, the emission close to the Fermi level is usually much lower than for thick overlayers, indicating a semimetallic behavior. For thicker silicide films, in contrast, a higher density of states with a clear Fermi edge is observed.

C. Origin of the spectral components

In this section the occurrence of the two different spectral components $R1$ and $R2$ will be discussed, which were observed for thick Ln_3Si_5 films, but not in case of the $LnSi_2$ monolayers. We can propose a simple model for an assignment of these components, which is based on the layered crystallographic structure of the epitaxial Ln silicide, where the formation of Si-sp^2 hybrids evokes a strong bonding anisotropy. Along this line, the main chemical interaction leading to different spectral components is occurring within the Si planes, while the interactions between Si planes and the neighboring Ln layers or the vacuum play only a minor role. In this way, surface effects may be neglected as well, despite the buckled surface structure. As shown in Fig. 2(b), there are two inequivalent sites of silicon atoms in the hex-

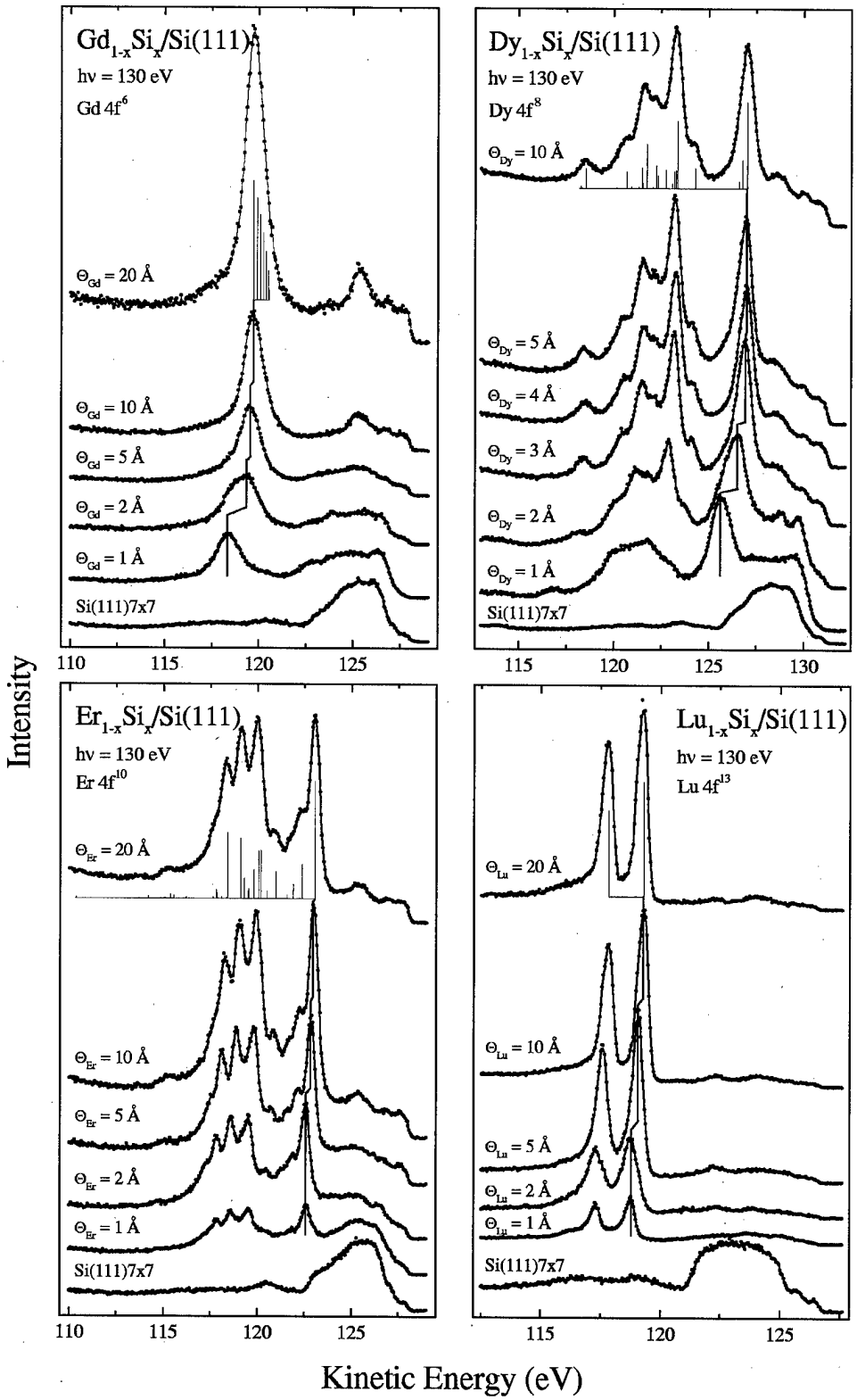


FIG. 6. Photoemission spectra of the Ln-4f final-state multiplets taken at various Ln coverages. For coverages $\Theta_{Ln} \leq 2$ Å the samples were annealed at 500 °C, for higher coverages at 600 °C. The original experimental data (bold dots) are normalized to the photon flux and connected by smoothed solid lines. The results from intermediate-coupling calculations (Ref. 21) with appropriate adjustment of the energy splitting are shown below the respective uppermost spectra.

agonal planes of Ln_3Si_5: those bound to three next silicon atoms and those located next to a silicon vacancy, thus bound only to two silicon atoms and having one dangling-bond orbital. Therefore two spectral components $R1$ and $R2$ are expected in the Si-2p spectra. The observation of only one silicide component R in the monolayer regime is in good agreement with this model, since only one single bonding configuration exists for the Si atoms in the monolayer silicide, as shown in Fig. 2(a). It should be noted here, that the observation of only one spectral component in the Ln-4f spectra is always expected, since only one lattice site exists for all coverages.

The components $R1$ and $R2$ of the Si-2p spectra can now be assigned by using their relative photoemission intensities

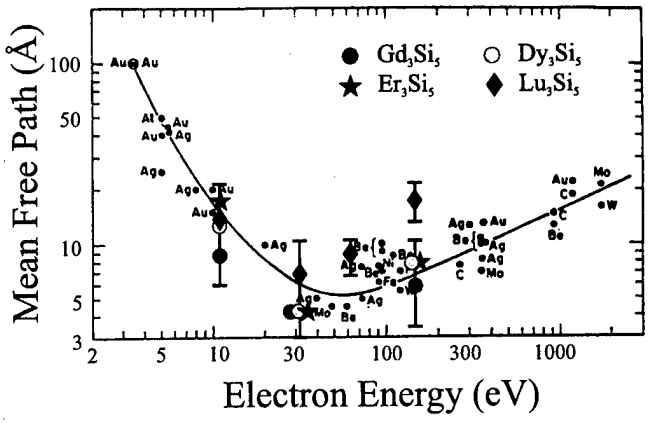


FIG. 7. Mean-free paths of the photoelectrons for the different Ln silicides as a function of kinetic energy relative to E_F shown together with the *universal curve* (Ref. 23). For clarity, symbols lying at the same position are slightly shifted horizontally.

at different photon energies and different coverages. In this way, component $R1$ has to be related to silicon atoms bound to three silicon neighbors and $R2$ to silicon atoms with neighboring vacancies, the latter contributing more at higher bulk sensitivity, as labeled in Fig. 2(b). This assignment is further supported by the observed magnitude of the chemical shifts of $R1$ and $R2$, which reflects theoretical predictions of a larger charge transfer from Ln atoms to the silicon atoms labeled $R2$ than to those labeled $R1$.[22]

Using the relative density of the $R1$ and $R2$ species in the $\sqrt{3}\times\sqrt{3}$R30° structural model, the escape depths at the respective kinetic energies of the photoelectrons can be calculated. The results, shown in Fig. 7, are in fair agreement with the *universal curve*[23] with the exception of the values obtained for the Lu silicide. It is not expected that the escape depth of Lu silicide is significantly different from that of the other silicides. Since thick Lu silicide shows a (1×1) LEED pattern, it is therefore assumed that a silicide structure with a different amount of vacancies is formed.

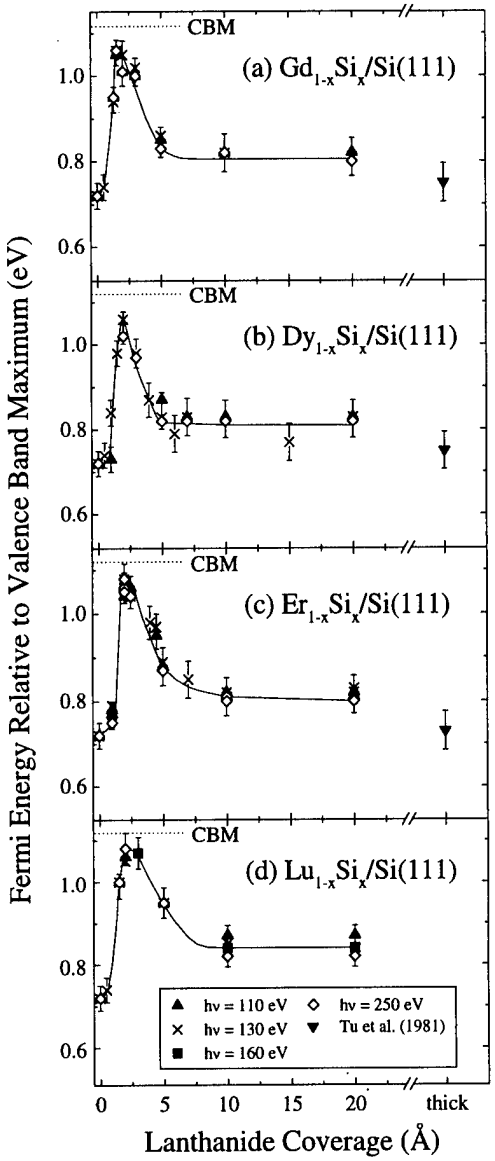


FIG. 8. Development of the Fermi-level position in the band gap of silicon for silicides of Gd, Er, Dy, and Lu as a function of Ln coverage.

D. Band-bending behavior

The variations of the Schottky-barrier height with film thickness for the four different Ln silicides investigated in this study are shown in Fig. 8, as derived from the variation in the kinetic energies of the Si-2p substrate bulk component (labeled B in Fig. 4). All four systems show barrier-height variations quite close to each other, reflecting the remarkable similarity of the heavy trivalent Ln metals.

The initial Fermi-level position of pristine Si(111)7×7 is located at 0.72±0.03 eV above the valence-band maximum (VBM).[24] Only minute variations are observed in the submonolayer range, where the $2\sqrt{3}\times2\sqrt{3}$R30° superstructure is formed, as shown by STM in Fig. 1(a).

At 2 Å coverage the substrate component shifts by about 0.32±0.04 eV to lower kinetic energies, as compared with Si(111)7×7. In this way the position of the Fermi level for one monolayer amounts to only 0.08±0.05 eV below the conduction-band minimum (CBM).[8] This value is in agreement with previous results on $ErSi_2$ monolayers, where a similar Fermi-level position of 0.13±0.05 eV was observed.[25] These values represent the lowest band bending ever observed for metal/n-Si interfaces and indicate almost complete flat-band conditions in the Si substrate. The possible origin of this exceptional result will be discussed in Sec. III E.

With increasing coverage, the Si-2p substrate component decreases in relative intensity, while the silicide emission becomes dominant, so that the band bending information gets gradually weaker. Therefore bulk sensitive measurements had to be performed at $h\nu$=110 and 250 eV in order to monitor the development of the Schottky-barrier height up to higher silicide thicknesses. Nevertheless, the Si-2p spectra at 10 and 20 Å Er are very similar, as observed in Fig. 4. The most striking differences are the decreasing shoulder at kinetic energies around 9.8 eV for the spectra taken at $h\nu$

=110 eV and the increasing dip at kinetic energies around 148.4 eV for the spectra taken at $h\nu=250$ eV. These differences can be considered a fingerprint of the substrate emission, which is expected to contribute more to the spectrum at 10 Å. In this way, the substrate core level is found to shift back to higher kinetic energies with increasing coverage, corresponding to final Schottky-barrier heights of 0.32 ±0.05 eV for all Ln silicides studied here (see Fig. 8). This value is in nice agreement with literature data from current–voltage measurements yielding barrier heights of 0.37 ±0.03 eV.[2]

E. Models for Schottky-barrier formation

In the following, a first attempt is made to discuss the observed band-bending behavior and in particular the remarkable overshoot leading to flat-band conditions in the monolayer range. Similar but not as extreme overshoots have been observed frequently for metal overlayers on semiconductor substrates in the low-coverage regime and have been explained by adsorbate-induced states or defect states at low coverages,[1,26,27] where the maximum Fermi-energy position is observed. At higher coverages different mechanisms take over, which are often related to overlayer metallization. It has been, e.g., discussed that the increasing influence of metal-induced gap states (MIGS), which are metallic states penetrating into the semiconductor at energies of the band gap, leads to a final Fermi-level position close to midgap.[28–30] In general, however, the final distribution of such MIGS and therewith the Fermi-level position does not only depend on the bulk properties of the semiconductor, but also on the particular structure of the interface and the electronic properties of the metal.[31,32]

For the present case, we present a preliminary model based mainly on the efficiency of MIGS for providing sufficient interfacial charge transfer. Mainly three mechanisms are expected to contribute: (1) The simple work-function model[33] neglects interface states completely, so that no interface dipole is present, and the vacuum level is continuous at the interface. In the present case the rather low work function of the Ln and therewith of their silicides should lead to a Fermi-level position close to or even above the CBM. (2) The presence of MIGS, in contrast, will lead to a Fermi-level position slightly above the charge-neutrality level, which is located in the midgap region.[34] The latter mechanism requires a sufficient density of states close to the Fermi energy in the metal layer in order to enable the required interface charge transfer and therewith a discontinuity of the vacuum level. (3) The Fermi-level position depends further on the structure and therewith on the particular chemical bonding at the interface. Variations in this bonding may change the interface dipoles, in this way altering the barrier heights.

In case of the $ErSi_2$ monolayer it was shown by angle-resolved PES[35] that the Fermi surface consists of a small hole pocket around the $\bar{\Gamma}$ point and a small electron pocket around the $\bar{M}$ point of the surface Brillouin zone. These tiny pockets of holes and electrons are characteristic for a semimetal with a rather low density of states close to the Fermi level.[36] This observation is also supported by the valence-band spectra shown in Fig. 6, where only a low emission is observed at the Fermi energy in the monolayer regime. Thus the role of the MIGS is supposed to be strongly reduced, so that the band bending will be determined mainly according to the work-function model.

The final Fermi-level position, in contrast, is obtained for coverages exceeding about 3 monolayers of Ln_3Si_5. At this coverage the overlayer assumes the full metallic character of the bulk silicide, which is characterized by a low resistivity[37] and a high density of states around the Fermi level[38,39] (see also Fig. 6). Therefore we propose that the Fermi-level position observed for thick Ln_3Si_5 is dominantly related to the MIGS.

According to this model, the barrier height varies mainly as a function of the net charge transfer between the silicide film and the semiconductor. In the monolayer regime, the charge transfer is strongly suppressed because of the semimetallic properties of the overlayer, so that MIGS are less efficient than in the multilayer case. Moreover the different chemical compositions of the silicides are supposed to influence the energy distribution of the MIGS and therewith the Fermi-level position. However, it should be noted that the present discussion has to be considered preliminary and is mainly intended to stimulate progress in this rather complex topic. It is in particular obvious that detailed theoretical studies are necessary for a better understanding of the overshoot behavior.

IV. CONCLUSION AND OUTLOOK

In this work the epitaxial growth, the chemical composition, and the Schottky-barrier formation of various trivalent lanthanide silicides on *n*-type Si(111) were investigated by means of STM, LEED, and high-resolution core-level PES.

Highly ordered Ln silicides can be grown on Si(111) with $2\sqrt{3}\times2\sqrt{3}$R30° and (1×1) periodicity in the submonolayer and monolayer range, respectively. For thicker coverages mainly $\sqrt{3}\times\sqrt{3}$R30° structures have been observed, and two different reacted Si-2*p* components were identified in the photoemission spectra, which could be assigned to inequivalent silicon sites within the defected AlB_2 structure.

The Schottky-barrier height was found to vary strongly with the film thickness of the silicide. An extremely low Schottky-barrier height corresponding to flat-band conditions was observed in the monolayer regime. This observation was discussed in the framework of both work-function model and the concept of metal-induced gap states.

An interesting topic would be the exploitation of the flat-band conditions for device applications. It should be noted that our first attempts to conserve the Fermi-level position of the silicide monolayer by Ag overgrowth resulted in strong intermixing. A similar behavior is expected to occur for other metallic overlayers, since both Ln and Si are known to form compounds with most other metals.[40] Furthermore the peculiar electronic properties of the silicide monolayer are probably modified upon metal deposition, therewith destroying the flat-band conditions. On the other hand, the low lat-

tice mismatch between Si and the heavy trivalent Ln silicides should allow an epitaxial overgrowth of the Ln silicide monolayer by further Si layers. In this way, a thin Si layer on top of a monolayer could allow to freeze the flat-band conditions. Furthermore the realization of silicon–silicide superlattices should be possible, which is of high interest for various applications because of their tunable physical properties.

Finally it should be mentioned, that the silicide structure depends strongly on the surface orientation since a variety of structurally different Ln silicides are known to exist.[18,40] On Si(001) the formation of regular Dy silicide wires and clusters was recently observed.[41] Self-organized one-dimensional structures of that kind with their novel physical properties may open up new possibilities for applications in silicon-based integrated circuit technology.

ACKNOWLEDGMENTS

The authors would like to thank G. Kaindl and E. Weschke for the availability of the measurement chamber and to W. Busse, I. Manke, and R. Meier for assistance during the measurements. This work was supported by the Bundesministerium für Bildung und Forschung, Project No. 05 622KTB 2.

[1]W. Mönch, *Semiconductor Surfaces and Interfaces* (Springer, Berlin, 1995).
[2]K. N. Tu, R. D. Thompson, and B. Y. Tsaur, Appl. Phys. Lett. **38**, 626 (1981).
[3]H. Norde, J. de Sousa Pires, F. d'Heurle, F. Pesavento, S. Petersson, and P. A. Tove, Appl. Phys. Lett. **38**, 865 (1981).
[4]J. E. Baglin, F. M. d'Heurle, and C. S. Petersson, Appl. Phys. Lett. **36**, 594 (1980).
[5]J. A. Knapp and S. T. Picraux, Appl. Phys. Lett. **48**, 466 (1986).
[6]R. D. Thompson, B. Y. Tsaur, and K. N. Tu, Appl. Phys. Lett. **38**, 535 (1981).
[7]R. Baptist, S. Ferrer, G. Grenet, and H. C. Poon, Phys. Rev. Lett. **64**, 311 (1990).
[8]S. Vandré, T. Kalka, C. Preinesberger, and M. Dähne-Prietsch, Phys. Rev. Lett. **82**, 1927 (1999); **82**, 4370 (1999).
[9]I. Manke, H. J. Wen, A. Höhr, A. Bauer, M. Dähne-Prietsch, and G Kaindl, J. Vac. Sci. Technol. B **13**, 1657 (1995).
[10]R. Baptist, S. Ferrer, G. Grenet, and H. C. Poon, Phys. Rev. Lett. **64**, 311 (1990).
[11]D. B. B. Lollmann, T. A. Nguyen Tan, J.-Y. Veuillen, P. Muret, K. Lefki, M. Brunel, and J. C. Dupuy, Appl. Surf. Sci. **65/66**, 704 (1993).
[12]P. Wetzel, C. Pirri, P. Paki, D. Bolmont, and G. Gewinner, Phys. Rev. B **47**, 3677 (1993).
[13]P. Wetzel, S. Saintenoy, C. Pirri, D. Bolmont, and G. Gewinner, Phys. Rev. B **50**, 10 886 (1994).
[14]T. P. Roge, F. Palmino, C. Savall, J. C. Labrune, P. Wetzel, C. Pirri, and G. Gewinner, Phys. Rev. B **51**, 10 998 (1995).
[15]P. Wetzel, S. Saintenoy, C. Pirri, D. Bolmont, G. Gewinner, T. P. Roge, F. Palmino, C. Savall, and J. C. Labrune, Surf. Sci. **355**, 13 (1996).
[16]T. P. Roge, F. Palmino, C. Savall, J. C. Labrune, and C. Pirri, Surf. Sci. **383**, 350 (1997).
[17]S. Vandré, T. Kalka, C. Preinesberger, I. Manke, H. Eisele, M. Dähne-Prietsch, R. Meier, E. Weschke, and G. Kaindl, Appl. Surf. Sci. **123/124**, 100 (1998).
[18]See, e.g., K. A. Gschneidner and L. Eyrine, *Handbook of the Physics and Chemistry of Rare Earths,* Vol. II (North–Holland, Amsterdam, 1979).
[19]C. J. Karlsson, E. Landemark, Y.-C. Chao, and R. I. G. Uhrberg, Phys. Rev. B **50**, 5767 (1994).
[20]D. E. Eastman, F. J. Himpsel, and J. F. van der Veen, Solid State Commun. **35**, 345 (1980).
[21]F. Gerken, J. Phys. F **13**, 703 (1983).
[22]G. Allan, I. Lefebvre, and N. E. Christensen, Phys. Rev. B **48**, 8572 (1993).
[23]See, e.g., G. Ertl, and J. Küppers, *Low-Energy Electrons and Surface Chemistry* (VCH, Weinheim, 1985).
[24]M. T. Cuberes, A. Bauer, H. J. Wen, D. Vandré, M. Prietsch, and G. Kaindl, J. Vac. Sci. Technol. B **12**, 2422 (1994).
[25]M. H. Tuilier, G. Gewinner, C. Pirri, P. Wetzel, D. Bolmont, and O. Heckmann, J. Phys. IV **C9**, 187 (1994).
[26]W. Mönch, Appl. Surf. Sci. **41**, 128 (1989).
[27]M. Prietsch, M. Domke, C. Laubschat, and G. Kaindl, Phys. Rev. Lett. **60**, 436 (1988).
[28]J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[29]F. Flores and C. Tejedor, J. Phys. C **20**, 145 (1987).
[30]F. Bechstedt and M. Scheffler, Surf. Sci. Rep. **18**, 145 (1993).
[31]R. G. Dandrea and C. B. Duke, J. Vac. Sci. Technol. B **11**, 1553 (1993).
[32]J. P. Sullivan, R. T. Tung, D. J. Eaglesham, F. Schrey, and W. R. Graham, J. Vac. Sci. Technol. B **11**, 1564 (1993).
[33]W. Schottky, Naturwissenschaften **26**, 843 (1938).
[34]J. Ortega and F. Flores, Phys. Rev. Lett. **63**, 2500 (1989).
[35]P. Wetzel, C. Pirri, P. Paki, J. C. Peruchetti, D. Bolmont, and G. Gewinner, Solid State Commun. **82**, 235 (1992).
[36]L. Stauffer, A. Mharchi, C. Pirri, P. Wetzel, D. Bolmont, G. Gewinner, and C. Minot, Phys. Rev. B **47**, 10 555 (1993).
[37]J. Y. Duboz, P. A. Badoz, F. Arnaud d'Avitaya, and J. A. Chroboczek, Appl. Phys. Lett. **55**, 84 (1989).
[38]L. Stauffer, C. Pirri, P. Wetzel, A. Mharchi, P. Paki, D. Bolmont, and G. Gewinner, Phys. Rev. B **46**, 13 201 (1992).
[39]T. Rössler and C. Laubschat (private communication).
[40]See, e.g., T. B. Massalski, H. Okamoto, P. R. Subramanian, and L. Kacprzak, *Binary Alloy Phase Diagrams*, 2nd ed. (ASM International, Materials Park, OH, 1990).
[41]C. Preinesberger, S. Vandré, T. Kalka, and M. Dähne-Prietsch, J. Phys. D **31**, L43 (1998).

GaP(001) and InP(001): Reflectance anisotropy and surface geometry

N. Esser
Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstrasse 36, 10623 Berlin, Germany

W. G. Schmidt and J. Bernholc
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

A. M. Frisch,[a] P. Vogt, M. Zorn, M. Pristovsek, and W. Richter
Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstrasse 36, 10623 Berlin, Germany

F. Bechstedt
Institut für Festkörpertheorie und Theoretische Optik, Friedrich-Schiller-Universität, Max-Wien-Platz 1, 07743 Jena, Germany

Th. Hannappel and S. Visbeck
Hahn-Meitner Institut, CD, Glienickerstr. 100, 14109 Berlin, Germany

(Received 20 January 1999; accepted 30 March 1999)

We have investigated the optical anisotropy of GaP(001) and InP(001) surfaces. The samples were prepared by homoepitaxial metalorganic vapor phase epitaxy growth and either directly transferred into ultrahigh vacuum (UHV) or *in situ* capped and, after transfer, decapped in UHV by thermal desorption of a P/As capping layer. Symmetry, composition, and surface optical anisotropy were characterized by low-energy electron diffraction, Auger electron spectroscopy, and reflectance anisotropy spectroscopy. We observe $(2\times1)/(2\times2)$-like reconstructions for the very P-rich and (2×4) reconstructions for the more cation-rich surfaces. No (4×2) reconstruction could be prepared, independent of the preparation method. A comparison of the reflectance anisotropy between GaP(001) and InP(001) surfaces shows similar line shapes for the very cation-rich (2×4) surfaces. For less cation-rich surfaces, however, we observe distinct differences between the spectra of the two systems. In both cases, different line shapes in the reflection anisotropy spectra occur for the (2×4) periodicity, suggesting the existence of different (2×4) geometries. The experimental results are discussed on the background of atomic structures, total energies and reflectance anisotropy spectra obtained *ab initio* from density-functional theory local-density approximation calculations. © *1999 American Vacuum Society.* [S0734-211X(99)03504-0]

I. INTRODUCTION

Much progress has been made in recent years in understanding the microscopic structure of the growth planes of III–V compound semiconductors (for a recent review, see, e.g., Ref. 1). Optical anisotropy at the same time has been proven to be an additional and complementary tool to electron spectroscopic methods for investigating the microscopic surface structure of semiconductors.[2,3] Previous investigations of GaAs(001)[4,5] and InP(001)[6] surfaces have shown that the atomic structure can be successfully clarified in many cases by comparing total energy and band structure calculations for the optical anisotropy on the one hand with experimental data of the optical anisotropy on the other hand. The atomic surface structures of the (2×4)-GaP(001) and -InP(001) surfaces were recently investigated by density functional theory local-density approximation (DFT-LDA) total energy calculations.[7,8] As in the case of InP(001), the atomic structures of GaP(001) were found to be different from the well-investigated GaAs(001) surface,[9–11] due to a large size difference between cations and anions.

While several recent studies have addressed the atomic structure of InP(001),[8,12–14] much less is known about GaP(001). Available experimental studies lead to different and partially contradictive conclusions about the symmetry and structure of the GaP(001) surface. So far it has been suggested mostly that ion bombardment and annealing as well as decapping of GaP(001) results in a $(4\times2)/c(8\times2)$ reconstructed, Ga-rich surface,[15–21] in analogy to the corresponding Ga-rich GaAs(001) surface structure.[9] In a more recent investigation by ion scattering,[22] however, a (2×4) symmetry was reported. Similar ambiguous results for GaP(001) can be found in literature for gas source molecular beam epitaxy (MBE) experiments: In Ref. 23 it is reported that (2×4) and (4×2) reconstruction pattern correspond to P- and Ga-stabilized surfaces, whereas in Ref. 24, a (2×4) reconstruction was observed under Ga as well as under P supply. Additionally, the authors of Ref. 24 found an intermediate (4×4) reconstruction during layer growth and suggested the formation of Ga droplets on a (2×4) reconstructed surface for high amounts of Ga supply. Only few data can be found in literature about the optical anisotropy of GaP(001). To our knowledge, there exist studies of GaP(001) surfaces in MBE, metalorganic vapor phase epitaxy (MOVPE) and chemical beam epitaxy (CBE).[25–28] Apart from Ref. 7 no reflectance anisotropy data for GaP(001) surfaces in ultrahigh vacuum (UHV) are available.

In the present study, clean and well-ordered (001) sur-

[a]Electronic mail: frisch@gift.physik.tu-berlin.de

faces of MOVPE-grown GaP and InP with different stoichiometries are prepared under ultrahigh vacuum conditions. All samples are characterized by low-energy electron diffraction (LEED), Auger electron spectroscopy (AES), and reflectance anisotropy spectroscopy (RAS). Based on total-energy (TE) calculations and by comparison of experimental data to calculated optical anisotropies for various atomic structure models it is shown that several (2×4) reconstructions exist for GaP(001) and InP(001). They realize different stoichiometries. (4×2) structures do not occur in either case.

II. METHOD

The GaP(001) surfaces were prepared by thermal desorption of a protective arsenic/phosphorus double layer (cap) under UHV conditions. For this purpose, homoepitaxial GaP epilayers were grown by MOVPE on highly *n*-doped ($10^{18}\,cm^{-3}$) GaP(001) substrates and capped *in situ* utilizing the photodecomposition of phosphine and arsine by an excimer laser source.[27,29(a)] The thermal desorption of the protective arsenic/phosphorus layer was performed at 690 K in UHV at a base pressure $\leq 10^{-10}$ hPa. The InP(001) samples were grown by MOVPE on *n*-doped ($10^{17}\,cm^{-3}$) InP(001) substrates and transferred under UHV conditions in the analyzing UHV chamber directly after growth, without using the capping technique.[29(b)] Starting with P-rich surface conditions in either case, different reconstructions were prepared by subsequent annealing in UHV. After cooling to room temperature, the surfaces were investigated by LEED, RAS, and AES.

The calculations were based on DFT-LDA. Optical properties are calculated in independent-particle approximation from the DFT-LDA electronic structure. To determine the optical anisotropy, we follow the formalism developed by Del Sole[30] and Manghi *et al.*[31] Computational details can be found in Refs. 6 and 11.

III. RESULTS AND DISCUSSION

A. Surface symmetry and composition

The decapped GaP(001) surfaces (after annealing to 690 K) show a (2×1)/(2×2)-like LEED pattern [see Fig. 1(a)], with clear (2×1) spots plus additional streaks in the [110] half order position. A similar LEED pattern [see Fig. 1(b)] is observed from InP(001) surfaces, directly after transfer into the UHV system. Annealing to 785 and 645 K leads to (2×4) LEED patterns for GaP(001) as well as for InP(001). The LEED pattern of the (2×4) GaP(001) surface does not allow us to distinguish between a (2×4) and a (2×4)/*c*(2×8) reconstruction [Fig. 1(c)] due to not fully resolved fractional order spots in the [110] direction. The LEED pattern of the InP(001) surface, on the contrary, shows a (2×4) periodicity, or a superposition of (2×1)/(2×2) and (2×4) [Fig. 1(d)]. Upon further annealing up to 1000 K the quality of the LEED pattern for GaP(001) is improved and the former streaks develop into sharp spots indicating a pure (2×4) reconstruction, similar to the pattern observed for InP(001) [Figs. 1(e) and 1(f)]. Upon annealing to still higher temperatures [≥1000 K for GaP(001) and ≥725 K for InP(001)] the symmetry of the LEED patterns remains unchanged until the surfaces deteriorate by forming Ga or In droplets, as indicated by a metallic component in the Ga 3*d*/In 4*d* core level photoemission lines (not shown here). AES measurements performed on GaP after decapping and InP after UHV transfer show no detectable carbon or oxygen contaminations. Moreover, the composition of the GaP(001) surfaces was examined. The P/Ga ratio for the (2×1)/(2×2)-like GaP(001) surface shows P-rich surface conditions, but it cannot be exactly determined because of the P desorption induced by the electron beam. On the (2×4) reconstructed GaP(001) surface at low annealing temperatures (only slightly above 785 K), the P/Ga intensity ratio is close to the one of the GaP(110) cleavage face (balanced stoichiometry). Upon annealing to higher temperatures (1000 K) the ratio changes successively by a factor of approximately two towards a Ga-rich surface, suggesting the existence of a more P-rich (2×4) at lower and a Ga-rich (2×4) reconstruction at higher annealing temperatures.

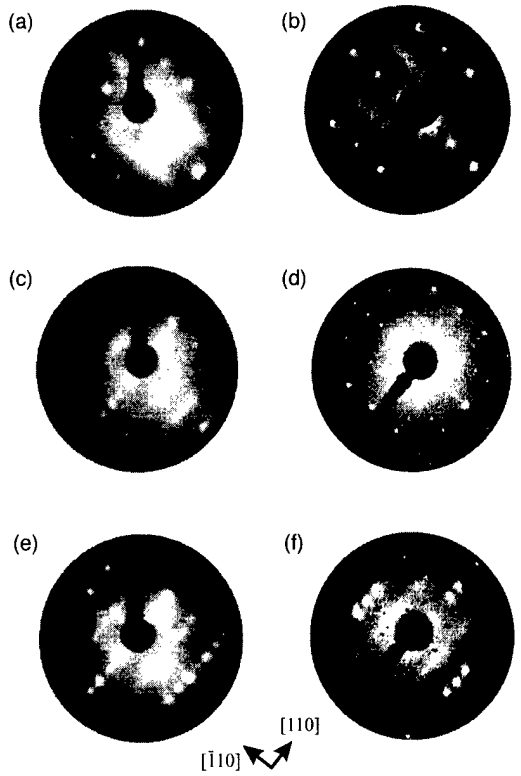


FIG. 1. LEED pattern recorded from GaP(001) and InP(001): (a) (2×1)/(2×2) GaP(001) taken at a electron energy of 53 eV, (b) (2×1)/(2×2) InP(001) (electron energy 50 eV), (c) P-rich (2×4) GaP(001) (electron energy 56 eV), (d) (2×4) InP(001) (electron energy 120 eV), (e) Ga-rich (2×4) GaP(001) (electron energy 52 eV), (f) InP(001) (electron energy 50 eV).

B. Optical anisotropy: Experimental results

Real parts of different RAS spectra recorded from GaP(001) and InP(001) surfaces are shown in Fig. 2. For comparison, a RAS spectrum of the $\beta 2(2\times4)$ GaAs(001) is

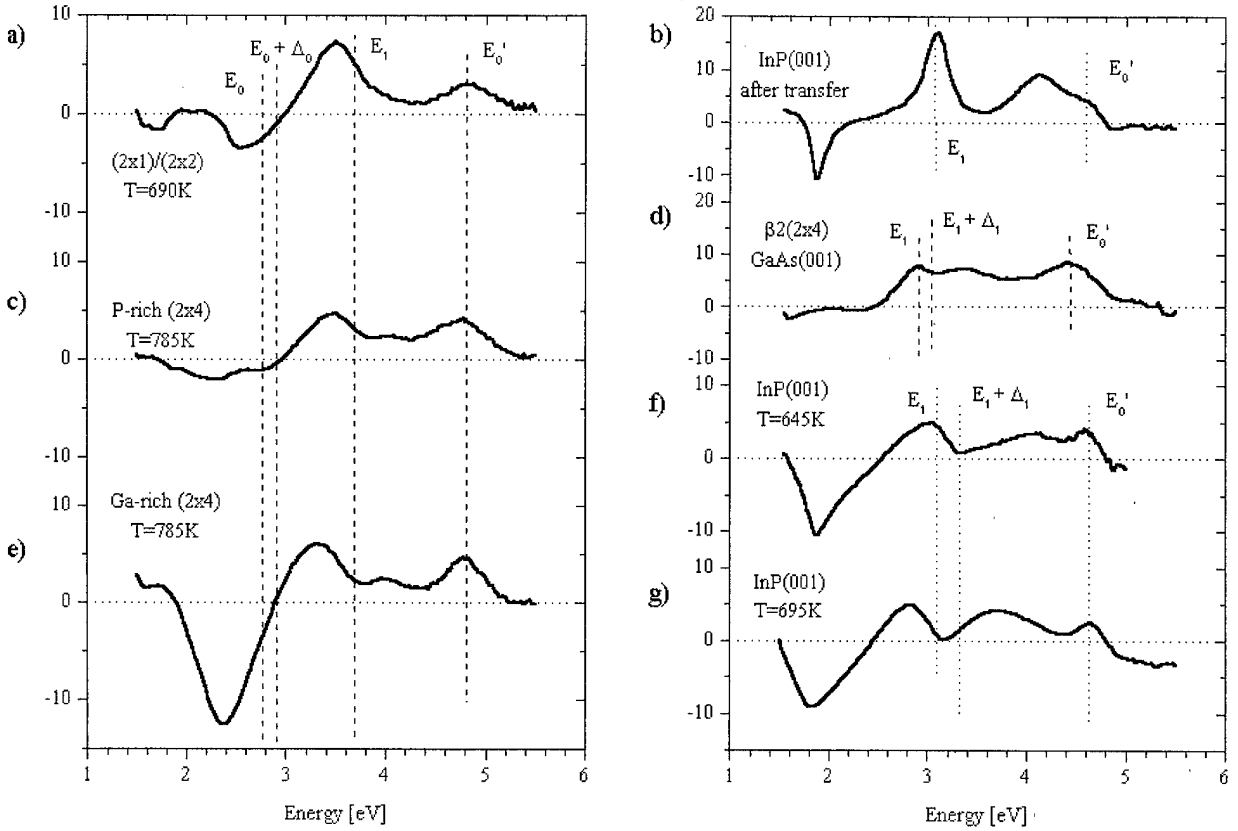


FIG. 2. Real parts of $((r_{[\bar{1}10]}r_{[110]})/\langle r\rangle)\times 10^3$ (LEED patterns and annealing temperatures are indicated in the plot) for: GaP(001) [left panel, (a), (c), (e)] InP(001) [right panel, (b), (f), (g)], and $\beta 2(2\times 4)$ GaAs(001) [right panel, (d)].

shown in Fig. 2(d). Three different spectral line shapes are reproducibly observed on GaP(001) and InP(001), correlated to the surface preparation: a characteristic line shape belongs to the $(2\times 1)/(2\times 2)$-like structures [Figs. 2(a), 2(b)], and two different line shapes are observed in either case for (2×4) surface structures, one correlated to P-rich conditions at lower annealing temperature [Figs. 2(c), 2(f)] and one to cation-rich conditions at higher annealing temperature [Figs. 2(e), 2(g)].

The P-rich GaP(001) (2×4) surface [Fig. 2(b)] shows a spectrum which is qualitatively similar to the one observed for the As dimer terminated $\beta 2(2\times 4)$ GaAs(001) surface.[3–5] It is dominated by pronounced maxima at 3.5 and 4.8 eV, close to the E_1 and E_0' gap energies of GaP (3.69 and 4.77 eV, respectively[32]) and a weak maximum at 4 eV. The RAS spectrum of InP under comparable conditions has a different line shape: In addition to maxima around the bulk critical points E_1 and E_0' (3.16 and 4.68 eV, respectively[33]) a strong minimum at 1.9 eV shows up. We mention that in Ref. 34 a somewhat different spectrum (nearly no negative anisotropy) has been assigned to the P-rich phase of InP (2×4). The reason for that difference is not clear at present. The line shapes shown in Figs. 2(f) and 2(g) are, however, reproducibly observed in our UHV experiments, whereas a line shape comparable to Ref. 34 could not be obtained in spite of repeated attempts. After annealing to higher temperatures the RAS spectra for GaP and InP(001) change and become similar in line shape. New features in the GaP spectrum are a strong minimum at 2.4 eV and a maximum at 3.2 eV, significantly below the E_1 gap of GaP. On InP the first maximum shifts to somewhat lower energies and an additional maximum appears between E_1 and E_0'. On both materials, the maximum at the E_0' point remains unaffected.

Summarizing the experimental results, there are clear similarities between the very anion-rich $(2\times 1)/(2\times 2)$-like and the very cation-rich (2×4) reconstructed GaP(001) and InP(001) surfaces, in particular, if one takes into account the different bulk critical point energies of the two materials. The spectrum of the P-rich (2×4) reconstructed GaP(001) surface, however, is distinct from the corresponding InP spectrum and resembles the one of GaAs(001) $\beta 2(2\times 4)$. The evolution of the RAS spectra of GaP(001) with increasing Ga coverage of the surface implies that at least two distinct (2×4) structures with different stoichiometry exist on the GaP(001) surface. This conclusion cannot be drawn for InP(001) surfaces. Scanning tunneling microscopy images[35] imply that the transition structure corresponds to a two-domain surface composed of P-rich $(2\times 1)/(2\times 2)$ and In-rich (2×4) domains. Also the RAS spectrum [Fig. 2(f)] can be reproduced by a linear superposition of spectra 2(b) and 2(g).

C. Atomic structure of the surfaces

So far, only the (2×4)-surface atomic structures [especially for InP(001)] have been analyzed in detail, while the

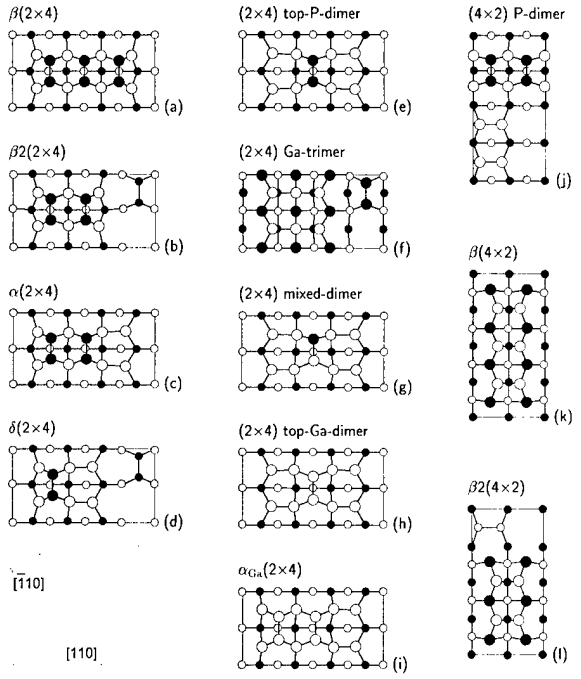


FIG. 3. Top view of relaxed GaP(001) (2×4) and (4×2) surface reconstruction models. Empty (filled) circles represent Ga (P) atoms. Large (small) symbols indicate positions in the first and second (third and fourth) atomic layers. (2×4) and (4×2) reconstructions are ordered by increasing Ga coverage.

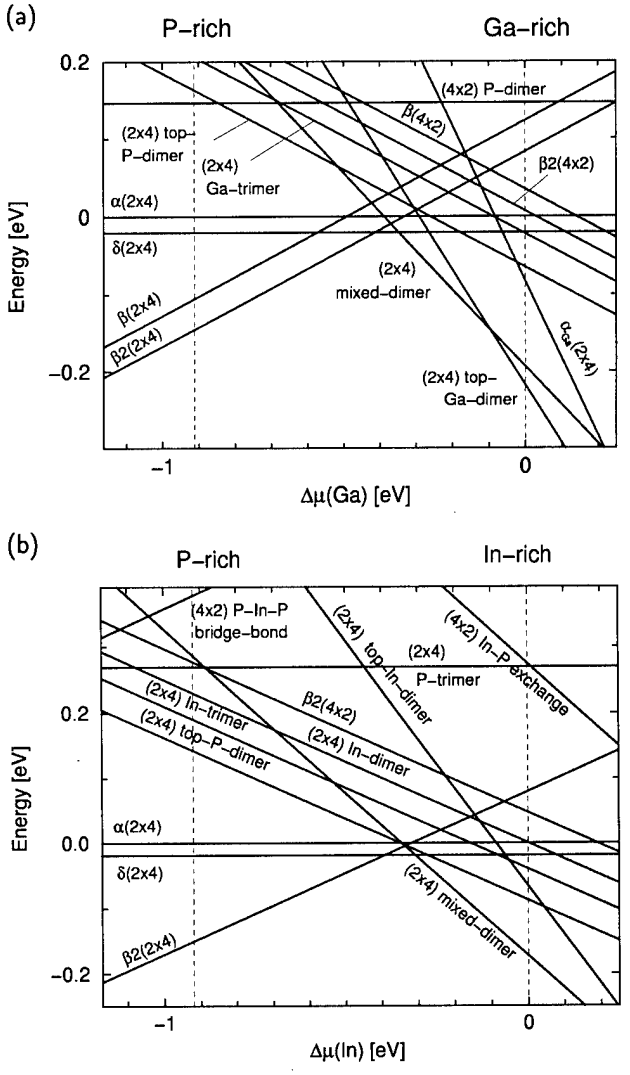


FIG. 4. (a) Relative formation energy per (1×1) unit cell for GaP(001) surface reconstructions vs $\Delta\mu(\mathrm{Ga}) := \mu_{\mathrm{bulk}}(\mathrm{Ga}) - \mu(\mathrm{Ga})$. The approximate thermodynamically allowed range: $-\Delta H_f(\mathrm{GaP}) \leq \Delta\mu(\mathrm{Ga}) \leq 0$ is indicated by dashed lines. (b) Same as in (a), but for InP(001) surface structures.

structure of the (2×1)/(2×2)-reconstructed surfaces is still completely unclear. STM and photoemission data are available for InP(001) (2×4).[8,13,29(a)] TE calculations performed on InP for a large variety of different (2×4)- and (4×2)-reconstruction models revealed a new type of dimer reconstruction involving mixed In–P surface dimers for In-rich surfaces, consistent with the experimental data.[8] Very recently, TE calculations have also been performed for a series of structural models for GaP(001).[7] Figure 3 shows the structural models considered for GaP. One of these models, the $\delta(2\times4)$ surface structure, which is believed to describe the Sb-induced GaAs (2×4) reconstruction,[36] turned out to be energetically favored for GaP(001), but was not considered in earlier studies on InP(001). We included that model in the present work.

The investigated models realize different In or Ga coverages and can therefore be compared energetically only by taking into account the chemical potentials of the surface constituents.[11] In Fig. 4 we show the relative formation energies of the considered surface structures vs the cation chemical potential for both InP(001) and GaP(001) surfaces. The phase diagrams show following fundamental similarities:

(i) (2×4) structures represent the stable surfaces under cation-rich conditions, in contrast to the GaAs(001) case where (4×2) structures are known to exist under cation-rich conditions.

(ii) A new atomic surface structure, consisting of mixed cation–P dimers adsorbed on top of a complete cation layer [Fig. 3(g)] represents the lowest energy structure over a prominent range of chemical potentials.

(iii) For a certain range of preparation conditions, the δ geometry [Fig. 3(d)] may occur. Given the limited accuracy of our calculations,[11] further structures such as α and top P dimer [Figs. 3(c) and 3(e)] cannot be excluded for the transition between cation- and anion-rich surfaces. Finally, under P-rich conditions P dimers in a $\beta2(2\times4)$ geometry [Fig. 3(b)], as known from As-rich GaAs(001) surfaces,[10,37] represent the lowest energy structure. However, this result must be taken with caution, since for P-rich conditions the experimental results indicate the appearance of (2×1)/(2×2) structures on both GaP as well as InP, which were not considered in the TE calculations.

(iv) All equilibrium structures are in agreement with electron counting heuristics.[38]

The similarity between the surface phase diagrams of InP and GaP points towards a common mechanism: The size

difference between anions and cations hinders the accommodation of sp^2-hybridized cation dimer atoms that are typical for cation-rich GaAs(001) surfaces. Instead, the formation of single-dimer structures is favored for cation-rich surfaces.

Besides these similarities, however, there is also an obvious difference: On InP(001) the mixed-dimer structure appears to be the only stable geometry under In-rich conditions, while on GaP(001) the top-Ga dimer model [Fig. 3(h)] is lower in energy in the extreme cation-rich limit.

The phase diagrams of GaP and InP in Fig. 4 demonstrate the existence of different stable (2×4) surface phases depending on the value of the cation chemical potential, i.e., the surface preparation conditions. This finding is in excellent agreement with the qualitative evolution of the surface optical anisotropy on GaP(001).

Since RAS as an optical technique is applicable in UHV as well as gas phase surroundings, the dependence of InP and GaP surface structures on different environments (e.g., growth conditions in MOVPE, MBE, CBE, and UHV) can be addressed by comparing the RAS spectra taken under different conditions. Comparing our spectra with the ones taken in MOVPE and CBE reveal, that for InP as well as for GaP, the same line shapes occur in the different environments.[26–28,39] Moreover, an excellent agreement between symmetry observed by RHEED in CBE[28,39] and LEED in our UHV experiments is found in both InP and GaP cases. There is, however, one exception: A $\beta 2(2\times4)$ seemingly observed in MBE for InP(001)[34] is not showing up in our experiments.

D. Optical anisotropy: Calculation

As shown above, the evolution of the surface optical anisotropy and the calculated equilibrium surface structures is qualitatively consistent. An even more direct comparison between theory and measurement, however, is desirable in order to identify specific surface structures. To this end, we compare the computed surface reflectance anisotropy of the energetically favored structures with the experimental RAS spectra.

Simulated spectra for GaP(001) are shown in the upper panel of Fig. 5.[7] Earlier we have shown that surface related transitions dominate the low-energy part while bulk-related transitions are more pronounced in the high-energy part of the RAS spectra.[7] The top-Ga dimer, mixed dimer, and δ structures show a pronounced negative anisotropy in the low-energy region, with minima between 2.0 and 2.3 eV. The strength of that anisotropy is directly correlated to the number of Ga–Ga bonds along the [110] direction. Its magnitude is the highest for the top-Ga-dimer model with eight bonds, slightly reduced and shifted to lower energies for the mixed-dimer geometry with six cation–cation bonds and flattened for the δ structure with only two such bonds.

The positive anisotropy in the high energy region of the spectra is correlated to the formation of P–P dimers oriented along $[\bar{1}10]$: For the $\beta 2$ geometry with three P–P dimers, we find a relatively strong positive anisotropy between about 2.4 and 4.4 eV. The shape of that anisotropy is roughly preserved for the δ structure, featuring one P–P dimer, however,

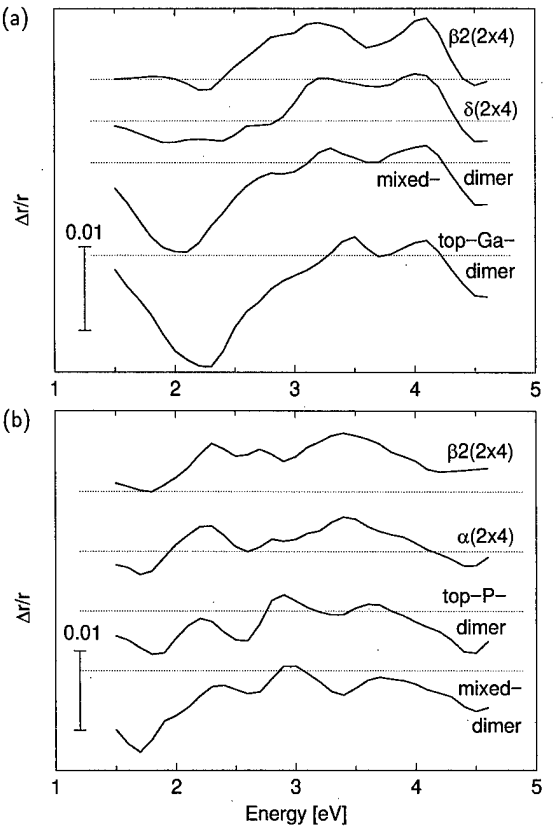


FIG. 5. (a) RAS spectra (same convention as in Fig. 2) calculated for the energetically favored structural models of the GaP(001) (2×4) surface. The zero line in each spectrum is indicated by a horizontal dotted line. (b) Same as in (a), but for InP(001) surface structures.

with reduced magnitude and the spectrum is shifted down. An even further reduction in positive anisotropy occurs for the mixed-dimer and top-Ga-dimer structures (no P–P dimers).

According calculations for InP(001)[6] are shown in the lower panel of Fig. 5. Similar systematic trends as for GaP(001) are observed in the optical anisotropy of InP(001): cation-rich surfaces show a characteristic strong minimum at low energies, P-dimer terminated ones a broad positive anisotropy at higher energies.

In fact, apart from some energy shift due to the neglect of many-body effects in DFT-LDA, the calculated spectra are in rather good agreement with the experimental optical anisotropy. Together with the calculated phase diagrams they allow a conclusive comparison. The RAS spectra measured for the (2×4) structures at higher annealing temperature (In/Ga-rich phases) feature a strong negative peak in the low-energy region. Given the energy position of this peak and its dependence on the preparation conditions, it can be identified with the calculated negative anisotropy arising from cation–cation surface bond related states. Both the top-Ga-dimer model and the mixed-dimer model thus are plausible candidates to explain the Ga-rich surface phase on GaP(001), while on InP the mixed-dimer model represents the only possible structure under In-rich conditions.

The measured spectrum for the GaP(001) (2×4) struc-

ture annealed at lower temperature (more P-rich phase) is dominated by a "camelback" overall spectrum shape with maxima between the energies of the E_0 and E_1 critical points and at the E_0' bulk peak. No negative anisotropy appears. The only computed spectrum with no (or very little) negative anisotropy belongs to the $\beta 2(2\times4)$ structure. Maxima appear at 3.2 and 4.1 eV, close to the calculated energies of the E_1 and E_0' critical points. Our results thus indicate that the less Ga-rich phase of the GaP(001) (2×4) surface corresponds to the $\beta 2(2\times4)$ structure known from As-rich GaAs(001) surfaces. This condition of a vanishing negative anisotropy is clearly not fulfilled in the case of the more P-rich InP(001) (2×4) structure. Consequently, this structure should be different, still containing In–In bonds in addition to P dimers. As mentioned above, a possible explanation would be a surface containing different domains.

IV. CONCLUSIONS

In conclusion, we have presented a comprehensive study of the optical anisotropy of GaP(001) and InP(001) surfaces based on the comparison of experimental results with calculated surface geometries and reflectance anisotropy spectra. Both for a balanced surface stoichiometry and for cation-rich conditions, we find (2×4) reconstructed surfaces that are stabilized by the formation of dimers. Experiment as well as theory suggest the existence of different (2×4) surface phases, depending on the cation content of the surface. Our results indicate that mixed dimers on top of an cation-terminated surface are the ground states of the Ga/In-rich phase of GaP(001) and InP(001) surfaces, consistent with previous findings for the InP(001) (2×4). For the less cation-rich surface phases, we suggest the formation of P–P dimers in a $\beta 2(2\times4)$ geometry for GaP(001). In case of InP, however, cation–cation bonds seem to be still present at the surface.

ACKNOWLEDGMENTS

Financial support by BMBF-Verbundprojekt 22, DFG (Schm 1361/1-1, Es 127/4-1), NSF (DMR 9408437), and ONR (N00014-96-I-0161) is gratefully acknowledged. This work was supported in part by grants of supercomputer time from the DoD Challenge Program and the North Carolina Supercomputer Center.

[1]Q.-K. Xue, T. Hashizume, and T. Sakurai, Prog. Surf. Sci. **56**, 1 (1997), and references therein.
[2]W. Richter and J.-T. Zettler, Appl. Surf. Sci. **100/101**, 465 (1996).
[3]I. Kamiya, D. E. Aspnes, L. T. Florez, and J. P. Harbison, Phys. Rev. B **46**, 15894 (1992).
[4]M. Murayama, K. Shiraishi, and T. Nakayama, Jpn. J. Appl. Phys., Part 1 **37**, 4109 (1998).
[5]A. I. Shkrebtii, N. Esser, W. Richter, W. G. Schmidt, F. Bechstedt, B. O. Fimland, A. Kley, and R. Del Sole, Phys. Rev. Lett. **81**, 721 (1998).
[6]W. G. Schmidt, E. Briggs, J. Bernholc, and F. Bechstedt, Phys. Rev. B **59**, 2234 (1999).
[7]A. M. Frisch, W. G. Schmidt, J. Bernholc, M. Pristovsek, N. Esser, and W. Richter, Phys. Rev. B (to be published).
[8]W. G. Schmidt, F. Bechstedt, N. Esser, M. Pristovsek, Ch. Schultz, and W. Richter, Phys. Rev. B **57**, 14596 (1998); W. G. Schmidt and F. Bechstedt, Surf. Sci. **409**, 474 (1998).
[9]D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L.-E. Swartz, Phys. Rev. **41**, 5701 (1990).
[10]J. E. Northrup and S. Froyen, Mater. Sci. Eng., B **30**, 81 (1995).
[11]W. G. Schmidt, Appl. Phys. A: Mater. Sci. Process. **65**, 581 (1997).
[12]J. Kinsky, Ch. Schultz, D. Pahlke, A. M. Frisch, T. Herrmann, N. Esser, and W. Richter, Appl. Surf. Sci. **123**, 228 (1998).
[13]C. D. MacPherson, R. A. Wolkow, C. E. J. Mitchell, and A. B. McLean, Phys. Rev. Lett. **77**, 691 (1996).
[14]M. M. Sung, C. Kim, H. Bu, D. S. Karpuzov, and J. W. Rabalais, Surf. Sci. **322**, 116 (1995).
[15]Y. Fukuda, M. Shimomura, N. Sanada, and N. Nagoshi, J. Appl. Phys. **76**, 3632 (1994).
[16]N. Sanada, M. Shimomura, and Y. Fukuda, and T. Sato, Appl. Phys. Lett. **67**, 1432 (1995).
[17]A. Watanabe, H. Shimaya, M. Naitoh, and S. Nishigaki, J. Vac. Sci. Technol. B **14**, 3599 (1996).
[18]M. Naitoh, A. Watanabe, A. Konishi, and S. Nishigaki, Jpn. J. Appl. Phys., Part 1 **35**, 4789 (1996).
[19]M. M. Sung and J. W. Rabalais, Surf. Sci. **365**, 136 (1996).
[20]N. Oishi, F. Shoji, A. Konishi, M. Naitoh, and S. Nishigaki, Surf. Rev. Lett. **5**, 223 (1998).
[21]I. M. Vitomirov, A. Raisanen, L. J. Brillson, C. L. Lin, D. T. McInturff, P. D. Kirchner, and J. M. Woodall, J. Vac. Sci. Technol. A **11**, 841 (1993).
[22]M. Naitoh, A. Konishi, H. Inenaga, S. Nishigaki, N. Oishi, and F. Shoji, Surf. Sci. **402-404**, 623 (1998).
[23]J. N. Baillargeon, K. Y. Cheng, and K. C. Hsieh, Appl. Phys. Lett. **56**, 2201 (1990).
[24]M. Yoshikawa, A. Nakamura, T. Nomura, and K. Ishikawa, Jpn. J. Appl. Phys., Part 1 **35**, 1205 (1995).
[25]P. A. Postigo, G. Armelles, and F. Briones, Phys. Rev. B **58**, 9659 (1998).
[26]J. S. Luo, J. F. Geisz, J. M. Olson, and M. C. Wu, J. Cryst. Growth **174**, 558 (1997).
[27]K. Knorr, M. Pristovsek, U. Resch-Esser, N. Esser, M. Zorn, and W. Richter, J. Cryst. Growth **170**, 230 (1997).
[28]M. Zorn, B. Junno, T. Trepk, S. Bose, L. Samuelson, J.-T. Zettler, and W. Richter, Phys. Rev. B (submitted).
[29](a) N. Esser, U. Resch-Esser, M. Pristovsek, and W. Richter, Phys. Rev. B **53**, R13257 (1996); (b) Th. Hannappel, S. Visbeck, K. Knorr, J. Mahrt, M. Zorn, and F. Willig, Appl Phys. A: Mater. Sci. Process. (in press).
[30]R. Del Sole, Solid State Commun. **37**, 537 (1981).
[31]F. Manghi, R. Del Sole, A. Selloni, and E. Molinari, Phys. Rev. B **41**, 9935 (1990).
[32]S. Zollner, M. Garriga, J. Kirchner, J. Humliicek, and M. Cardona, Phys. Rev. B **48**, 11 (1993).
[33]P. Lautenschlager, M. Garriga, and M. Cardona, Phys. Rev. B **36**, 4813 (1987).
[34]K. B. Ozanyan, P. J. Parbrook, M. Hopkinson, C. R. Whitehouse, Z. Sobiesierski, and D. I. Westwood, J. Appl. Phys. **82**, 474 (1997).
[35]P. Vogt, Th. Hannappel, S. Visbeck, K. Knorr, N. Esser, and W. Richter, Phys. Rev. B (to be published).
[36]W. G. Schmidt and F. Bechstedt, Phys. Rev. B **55**, 13051 (1997).
[37]W. G. Schmidt and F. Bechstedt, Phys. Rev. B **54**, 16742 (1996).
[38]M. D. Pashley, Phys. Rev. B **40**, 10481 (1989).
[39]M. Zorn, T. Trepk, J.-T. Zettler, B. Junno, C. Meyne, K. Knorr, T. Wethkamp, M. Klein, M. Miller, W. Richter, and L. Samuelson, Appl. Phys. A: Mater. Sci. Process. **65**, 333 (1997).

In situ reflectance difference spectroscopy of II–VI compounds: A real time study of N plasma doping during molecular beam epitaxy

D. Stifter,[a] M. Schmid, and K. Hingerl
PROFACTOR GmbH, A-4400 Steyr, Austria

A. Bonanni, M. Garcia-Rocha, and H. Sitter
Institute for Semiconductor Physics, Johannes Kepler University Linz, A-4040 Linz, Austria

(Received 19 January 1999; accepted 17 May 1999)

Reflectance difference spectroscopy (RDS) has been performed during the growth and nitrogen-doping process of II–VI layers fabricated by molecular beam epitaxy (MBE). To the MBE chamber, equipped with an electron cyclotron resonance cell for N plasma generation, a RDS system has been attached via a normal incidence viewport which allows the acquisition of spectra during the doping process in the spectral range from 1.5 to 5.5 eV. ZnTe (001) surfaces have been studied under varying exposure conditions, like Zn, Te, and/or N plasma flux onto the sample surface. Furthermore, RDS features in the vicinity of the E_1 and $E_1+\Delta_1$ transitions were used to optimize online the doping performance of the N plasma cell by varying the source parameters, like N pressure and input power. Doping induced surface processes and surface saturation with activated N species have been investigated. *Ex situ* measured spectra are compared with *in situ* acquired data to study the surface Fermi level pinning occurring at ambient pressures. Finally, *in situ* acquired RDS data of δ-doped ZnSe are presented. © *1999 American Vacuum Society.*
[S0734-211X(99)07404-1]

I. INTRODUCTION

During recent years enormous effort has been devoted to the *p*-type doping process of II–VI compounds with nitrogen plasma during molecular beam epitaxy (MBE): the II–VI's show, in general, a preference for one type of doping. Thus, difficulties in achieving high bipolar conductivity are still a major drawback for the fabrication of reliable devices of ZnSe and ZnSe/ZnTe contact layers.[1,2] Carrier compensation and low acceptor activation ratios still have to be overcome[3] and the active N species, which can even depend on the type of plasma cell, is still a subject of research.[4] Therefore, a monitoring and characterization tool is desirable to investigate *in situ* the doping process.

Reflectance difference spectroscopy (RDS) is a surface sensitive method for the optical investigation of surface, bulk, and interface anisotropies.[5,6] RDS measurements can be carried out at ambient pressure as well as in metalorganic chemical vapor deposition (MOCVD) environments and under ultrahigh vacuum (UHV) conditions for MBE. Recently there have been major research efforts to establish RDS in the field of II–VIs,[7] like *in situ* measurements during MBE growth.[8,9] Moreover, the carrier concentration can be determined in a contact free and nondestructive way by evaluating the linear electro-optical (LEO) effect, where a surface built-in electrical field due to surface Fermi level pinning introduces anisotropy into the refractive index, which is proportional to the electric field.[10] In the case of the II–VI's, only *ex situ* RDS measurements have been reported on *n*- and *p*-type doped ZnSe layers[11,12] where the doping concentration has been evaluated by taking advantage of the LEO effect.

[a]Electronic mail: dstift@pr-steyr.ac.at

II. EXPERIMENT

The II–VI MBE chamber which has been used for the experiments presented in this article, is equipped with elementary effusion cells (Zn, Mn, Mg, Cd, Se, and Te) for growth as well as with two self-designed N plasma sources [one discharge current (dc) and one electron cyclotron resonance (ECR) cell] and a $ZnCl_2$ compound source for *p*- and *n*-type doping, respectively. The ECR plasma cell has been constructed after a design presented in Ref. 13. It is built in a compact way using permanent NeFeB magnets instead of magnetic coils, thus facilitating operation and maintenance of the cell. The cell is shuttered by a linear gate valve for modulation or pulse doping of the II–VI layers.

The II–VI layers themselves were grown at a substrate temperature of 350 °C on thermally de-oxidized (100) GaAs substrates in an anion-enriched growth regime (i.e., Se- or Te-rich growth conditions).

For the *in situ* RDS measurements, a commercial RDS system (ISA, Jobin-Yvon) has been attached to the MBE chamber via a normal incidence viewport equipped with a strain free mounted Suprasil window as it is depicted schematically in Fig. 1. RDS spectra can be taken *in situ* from 1.5 to 5.5 eV. In the case of *ex situ* experiments the energy range is expanded to 6 eV. The measured RD spectra are displayed in terms of $\Delta\tilde{r}/\tilde{r}=\Delta r/r+i\Delta\theta$ where $\tilde{r}=re^{i\theta}$ is the complex reflectance. The quantity $\Delta\tilde{r}$ is defined as $\tilde{r}_{[\bar{1}10]}-\tilde{r}_{[110]}$. The directions in the indexes denote the polarization direction of the electric field with respect to the sample's orientation.

For a quantitative analysis, the carrier concentration of the doped layers was also determined by evaporating Au contacts and measuring the Hall effect in van der Pauw geometry.

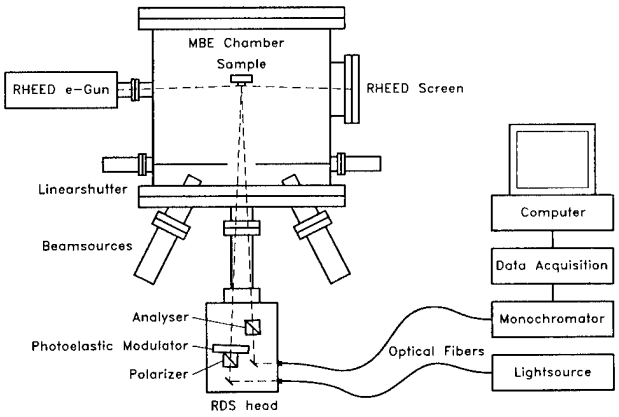


FIG. 1. Schematic of the RDS system attached to the vertical MBE chamber for *in situ* measurements.

In this article we will first focus on RDS results obtained on ZnTe layers and then present data on δ-doped ZnSe.

III. RESULTS AND DISCUSSION

As a starting point for our study, different surface terminations have been investigated for ZnTe layers. *In situ* RD spectra were taken from ZnTe surfaces during growth interruption and under varying exposure conditions, presented in Fig. 2: a surface exposed to Zn, exposed to N plasma, no exposure (vacuum), and exposure to Te. The topmost spectrum was taken after the following procedure: exposure of the surface to N plasma with subsequent growth of an approximately 3 nm thick undoped ZnTe layer, forming a highly δ-doped spike just below the surface.

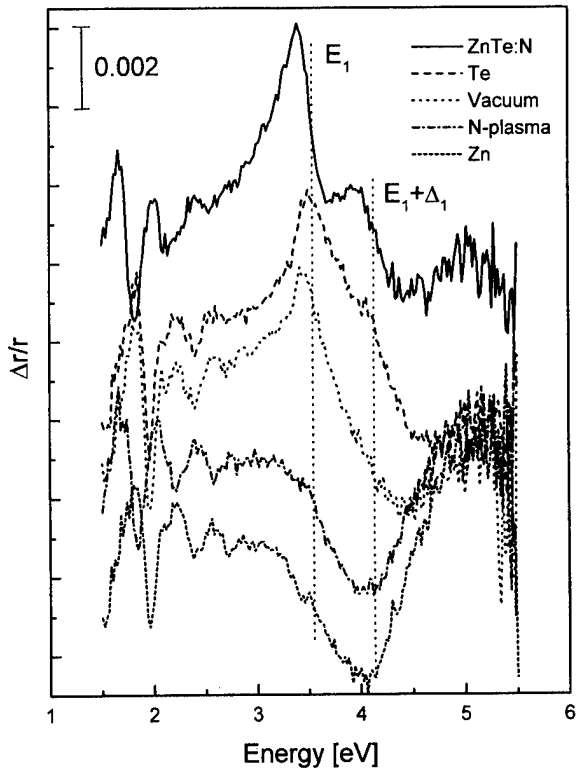


FIG. 2. *In situ* measured RD spectra of a ZnTe (100) surface under different exposure conditions: Zn, Te (undoped and δ doped), N plasma, and vacuum. The spectra were shifted vertically with respect to each other so that they can be distinguished better. The energetic positions of the E_1 and $E_1+\Delta_1$ transitions are indicated.

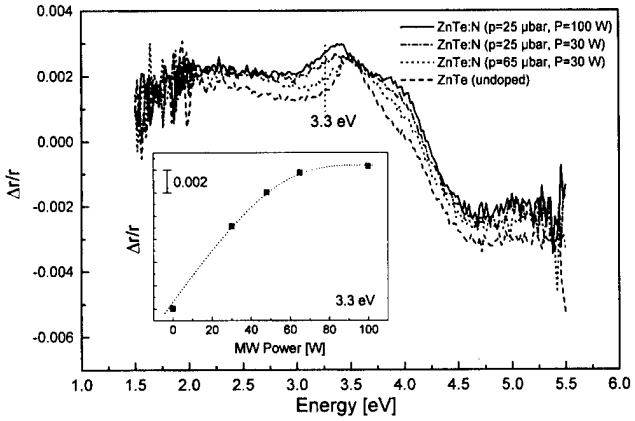


FIG. 3. Dependence of *in situ* measured RD spectra on carrier concentration for varying ECR cell parameters (cell pressure p and MW input power P). The inset gives the $\Delta r/r$ signal at 3.3 eV as a function of the ECR cell's MW power (N pressure: 65 μbar).

In general, all RD spectra are dominated below the energy gap ($E_g=2.32$ eV at room temperature) by interference fringes of the ZnTe layers. The most significant features for Te-terminated surfaces are a peak around 3.5 eV and a shoulder appearing at approximately 4 eV. These features are related to the E_1 and $E_1+\Delta_1$ transitions of ZnTe: $E_1=3.6$ eV and $E_1+\Delta_1=4.17$ eV (at 293 K).[14] In contrast, the spectrum of the Zn-saturated surface is characterized by a broad negative signal in the E_1 and $E_1+\Delta_1$ region. The spectrum of the bare surface (without any particle flux to the surface) is a mixture (linear combination) of the Te- and Zn-saturated surface. When exposing the surface to N plasma, the spectrum changes to that one of a Zn-saturated surface. The same happens if only molecular N_2 is used instead of N plasma. We ascribe this change to the removal of Te dimers from the surface caused by the nitrogen molecules.

When comparing the spectrum of the δ-doped layer with the spectrum taken under Te flux, the enhancement of the peaks around E_1 and $E_1+\Delta_1$ is striking. It has been shown that the LEO effect transfers, in the vicinity of the E_1 and $E_1+\Delta_1$ transitions, oscillator strength from one transition to the other. Thus, the surface built-in field causes an increase (decrease) of the E_1 with respect to the $E_1+\Delta_1$ transition, depending upon the carrier type and the concentration.[10] In the RD spectra characteristic features in the vicinity of these transitions can therefore be used to determine the doping concentration.

In the case of continuous growth and doping the difference between doped and undoped layers is less pronounced, as is seen in Fig. 3. RD spectra were taken from ZnTe surfaces during N plasma doping in order to optimize the performance of the ECR-plasma source. Figure 3 shows the *in situ* RD spectra for four different cases where the parameters

of the ECR cell (p=N_2 cell pressure, P=microwave input power) have been changed. The effective carrier concentration varies in these cases from nominally undoped to 10^{16} cm^{-3} (p=65 μbar, P=30 W), 10^{17} cm^{-3} (p=25 μbar, P=30 W), and 10^{18} cm^{-3} (p=25 μbar, P=100 W), respectively. Thus, the carrier concentration increases with decreasing cell pressure and increasing microwave (MW) power (in agreement with Ref. 15). For these ZnTe samples, the spectrum of the undoped layer differs from the doped ones by a redshift of the peak close or related to the E_1 transition. Furthermore, this peak broadens and increases in intensity when the cell pressure in the ECR source is decreased and the MW input power is increased, resulting in a higher carrier concentration. The shoulder related to the $E_1+\Delta_1$ transition around 4 eV is also raised dramatically as is the baseline of the RDS signal between 2 and 3 eV. In the inset of Fig. 3 the $\delta r/r$ signal is plotted at 3.3 eV as a function of the MW input power of the ECR cell for a fixed N pressure of 65 μbar. The $\Delta r/r$ signal varies proportional to the MW input power in the range from 0 to 60 W and saturates for an input power of ~80 W. These data were taken from kinetic RDS measurements at 3.3 eV and by varying the input power stepwise. For kinetic measurements, one has to take into account the fact that the typical growth rate is in the range of 1 Å/s. Since the analyzing light also "sees" the underlying layers due to the finite penetration depth, it takes some time until the signal stabilizes. During this stabilization time damped oscillations occur (i.e., spirals when plotted on the complex $\Delta\tilde{r}/\tilde{r}$ plane) due to the overgrowth of a layer or interface with different properties from those of the covering layer.[16] Therefore, very pronounced oscillations could also be observed around the E_1 transition during the formation of a δ-doped spike and its subsequent overgrowth (cf. the corresponding spectrum in Fig. 2).

For a quantitative analysis, a series of 0.8 μm thick ZnTe:N layers with varying carrier concentrations was grown on top of 0.2 μm thick undoped ZnTe buffer layers and was analyzed with *ex situ* RDS. Figure 4 demonstrates the evolution of the spectra as a function of the hole concentration. For the undoped sample negative peaks appear at the E_1 and $E_1+\Delta_1$ transitions. In between a positive peak with its maximum at 3.71 eV and a shoulder at 3.87 eV can be observed. All peaks decrease in intensity when the hole concentration is increased to 10^{16} cm^{-3}. When the doping level reaches 2.7×10^{17} cm^{-3}, the $E_1+\Delta_1$ peak changes sign as does the peak at 3.71 eV which is now very pronounced. At the same time the E_1 transition nearly vanishes. The pronounced peak at 3.71 eV increases further in intensity when the concentration is raised to 10^{18} cm^{-3}. The intensity of the $E_1+\Delta_1$ transition also increases; the E_1 peak changes sign but appears only very weakly in the spectrum. A shoulder can also be observed at exactly the same position (3.87 eV) as in the spectrum of the undoped sample. By comparing all the spectra one can conclude that this feature at 3.87 eV does not change in intensity by varying the hole concentration, but is part of the nonlinear RDS baseline due to surface roughness.[12] In order to establish a power law (carrier concentration as a function of characteristic RDS features), like that developed for the case of ZnSe,[11,12] the intensity of the $E_1+\Delta_1$ peak was compared with the pronounced peak at 3.71 eV, resulting in an exponent of approximately 0.5, as predicted by a model presented in Ref. 12. However, one has to state critically that an exact quantitative analysis seems to be difficult; when comparing the *in situ* measured RD spectra with those taken *ex situ*, it is obvious that exposure to the atmosphere has an influence on the Fermi level pinning, a premise for observation of the concentration dependent LEO effect. The pinning may also depend on the surface termination and the doping concentration. So it has been observed that for ZnSe the removal of the surface oxide layer indeed changes the RD spectra.[11] A contribution to this is still open discussion is given in Fig. 5, where *in situ* RDS measurements were performed on a doped ZnTe sample (p = 10^{17} cm^{-3}) during and immediately after growth at a substrate temperature of 350 °C [curve 350 °C (b)]. Afterwards the sample was cooled stepwise in the MBE chamber down to 36 °C. The spectra which were taken during the cooling procedure show that the peaks and features attributed to the E_1 and $E_1+\Delta_1$ transitions (marked by dotted lines) gain in intensity with decreasing temperature. Furthermore, at 350 °C the E_1 and $E_1+\Delta_1$ transitions are redshifted 0.23 eV with respect to data taken at room temperature. In contrast to the observed redshift, the enhancement of the peaks is not a pure temperature effect (i.e., dependence of electrical/optical properties on the temperature): The sample was heated up

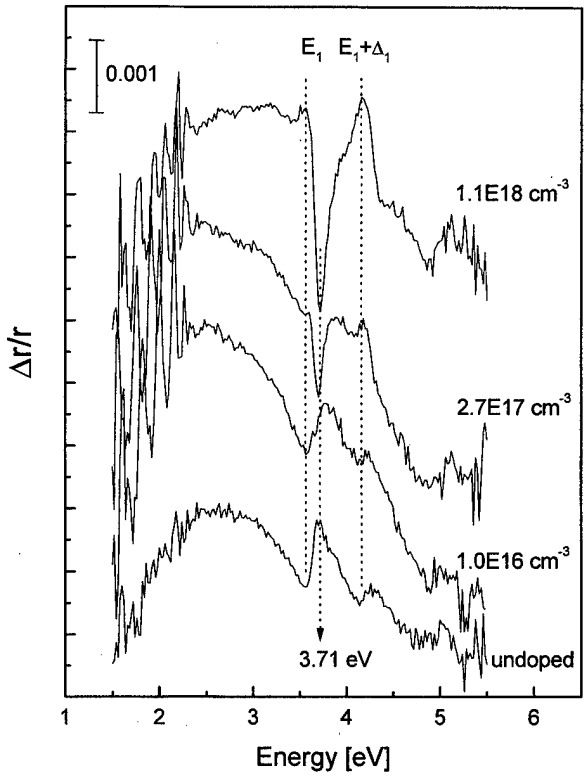


FIG. 4. *Ex situ* RD spectra for an undoped and a p-doped ZnTe layer, demonstrating the evolution of features in the vicinity of the E_1 and $E_1+\Delta_1$ transitions.

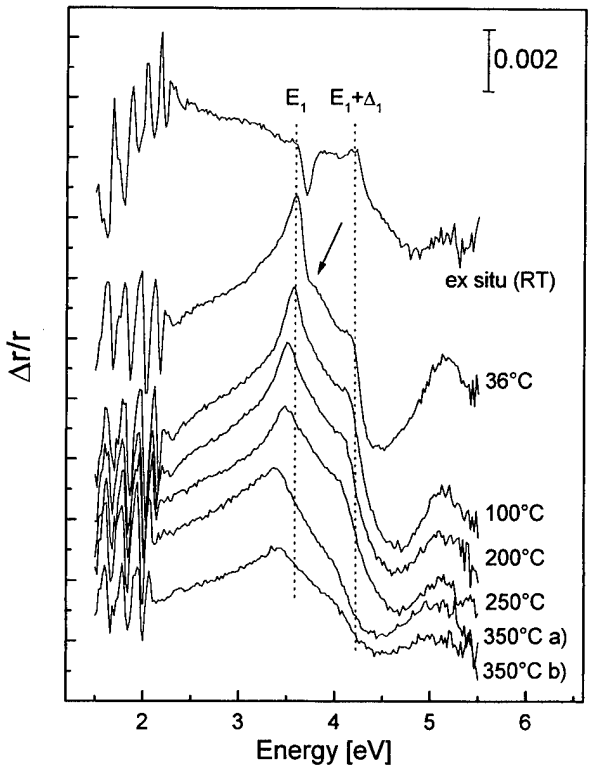


FIG. 5. Development of RD spectra for a doped layer ($p=2\times10^{17}$ cm^{-3}) during a cooling procedure from 350 °C (b) (substrate temperature) to 36 °C. Spectrum 350 °C (a) was taken after cooling when the sample was again heated up to 350 °C. The *ex situ* measured spectrum (room temperature and exposure to air) is given for comparison. The energetic position of the E_1 and $E_1+\Delta_1$ transitions are indicated by dotted lines. The arrow marks a kink at 3.71 eV, which develops into a pronounced peak when the sample is exposed to air.

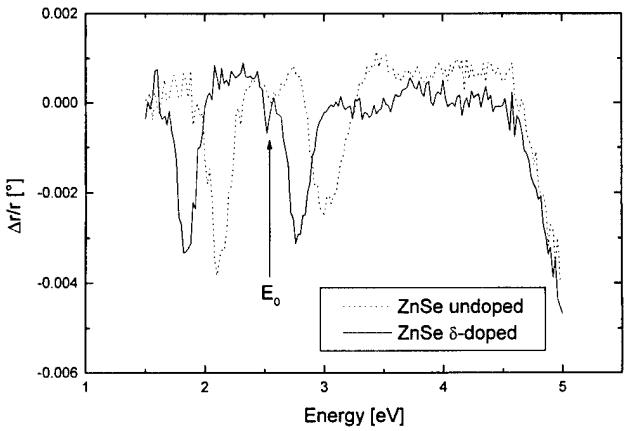


FIG. 6. *In situ* RD spectra of an undoped and a δ-doped ZnSe sample taken at the growth temperature (350 °C). The enhanced feature is related to the E_0 transition (marked by an arrow).

again after some cooling steps. Yet, it was not possible to reproduce the same spectrum for the same temperature when a cooling step has been carried out in between. This is demonstrated by the curves marked 350 °C (a) and 350 °C (b). The spectrum labeled 350 °C (b) was taken immediately after growth. Spectrum 350 °C (a) was taken after the sample was cooled down to 100 °C and then heated up to 350 °C. The energetic position of the E_1 and $E_1+\Delta_1$ transitions is the same for both spectra, indicating that in both cases the same sample temperature has been reached. However, the intensity of the transitions is different. Summarizing, the following could be observed. When the sample was held at one temperature, no changes in the spectrum could be observed. As soon as the temperature was decreased (for example, by 50 °C), the E_1 and $E_1+\Delta_1$ features were enhanced immediately. A quick subsequent heating to the previous temperature resulted in a very slow signal reaction. Kinetic measurements were performed at a fixed energy of 3.4 eV. The sample that was held at 170 °C was heated within 20 s up to 350 °C. In contrast, an exponential decrease of the signal could be observed with a very low exponential time constant of ~140 s. The most probable interpretation for this behavior is that the surface changes during cooling so that Fermi level pinning occurs due to thermal effects which create defects and/or absorption of residual gas particles. A comparison between the spectrum taken *in situ* at 36 °C and the *ex situ* measured data supports this explanation. The spectrum taken *in situ* exhibits, at 3.71 eV, a kink (marked by an arrow) which becomes the pronounced concentration dependent peak in the *ex situ* measured spectra of the air exposed samples. An alternative explanation for the evolution of the spectra, like observation of nitrogen acceptor activation, is less probable but cannot be totally ruled out.

Finally, RDS has been performed *in situ* on doped ZnSe layers for the first time. The data are presented in Fig. 6. The dotted line is the spectrum of an undoped ZnSe layer of ~200 nm thickness. Two interference fringes can clearly be distinguished. In addition a small dip indicates a feature in the vicinity of the E_0 transition (indicated by an arrow), which is modulated by the signal due to interference effects. By choosing the appropriate thickness of the ZnSe layer we could place the feature related to the E_0 transition between two interference fringes. The solid line gives the spectrum of a δ-doped layer as described above for ZnTe: saturation of the surface with N plasma and subsequent overgrowth with a thin undoped layer (30 nm). Enhancement of the E_0 peak can be observed in this case. The influence on the E_1 and $E_1+\Delta_1$ regions (4.8–5.1 eV) is weak (a small decrease of the signal around E_1), in contrast to the *ex situ* results presented in Ref. 11. This is comparable with the behavior of *in situ* and *ex situ* measured ZnTe, where a broadening of the features also takes place for the *in situ* data (cf. Figs. 3 and 4). It can be shown that the appearance of the E_0 transition in ZnSe is due to the presence of anisotropic strain in the layer along the $[\bar{1}10]$ and [110] directions.[17] For the undoped layer, with its thickness around 200 nm, most of the strain, introduced by the mismatch between the GaAs substrate and the ZnSe layer, has been already relaxed. By inserting a δ-doped layer the E_0 transition is enhanced. This transition should be affected neither by the doping itself nor by the LEO effect,[18] but its enhancement is interpreted as a consequence of additionally built-in strain by the δ-doped layer. This can be concluded since the electric field effects show up most strongly at critical points, where one or two joint

masses are negative, not at the fundamental gap transition E_0.[18] Furthermore, as was shown in Ref. 17, anisotropic in-plane lattice mismatch can be detected down to approximately $\Delta a/a = 10^{-5}$. Knowing that the Zn–N bond is approximately 2.1 vs 2.45 Å of the Zn–Se bond would even allow determination of the number of N atoms incorporated. However, this method has to be elaborated on further for definite conclusions.

IV. CONCLUSIONS

In situ RDS is capable of monitoring in real time the changes occurring at the interface of a growing II–VI semiconductor. By studying the RDS response of different surface terminations one can optimize on line the performance of the doping process and determine in a contact-free nondestructive way the carrier concentration: changes in the spectra around the E_1 and $E_1+\Delta_1$ transitions have been analyzed in detail, as for example, the study of the surface Fermi level pinning, which takes place during the cooling procedure of the samples. In the case of δ-doped ZnSe, strain-related effects showing up at the E_0 transition could be observed.

ACKNOWLEDGMENTS

This study was supported financially by the Gesellschaft für Mikroelektronik (GMe), the Fonds zur Förderung der wissenschaftlichen Forschung in Österreich (FWF), and the Jubiläumsfond der Österreichischen Nationalbank. One of the authors (M.G.-R.) would like to thank CONACYT (Mexico) for financial support.

[1]See, for example, R. M. Park, M. B. Troffer, C. M. Rouleau, J. M. DePuydt, and M. A. Haase, Appl. Phys. Lett. **57**, 2127 (1990).
[2]W. Faschinger, S. Ferreira, and H. Sitter, Appl. Phys. Lett. **64**, 2682 (1994).
[3]D. B. Laks, C. G. Van der Walle, G. F. Neumark, and S. T. Pantelides, Appl. Phys. Lett. **63**, 1375 (1993).
[4]M. Grün, N. Sadeghi, J. Cibert, Y. Genuist, and A. Tserepi, J. Cryst. Growth **159**, 284 (1996).
[5]D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).
[6]T. Yasuda, L. H. Kuo, K. Kimura, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, J. Vac. Sci. Technol. B **14**, 3052 (1996).
[7]T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, Phys. Rev. Lett. **77**, 326 (1996).
[8]M. Schmid, K. Hingerl, D. Stifter, A. Bonanni, and H. Sitter, J. Cryst. Growth **184/185**, 218 (1998).
[9]J. T. Zettler, K. Stahrenberg, W. Richter, H. Wehnisch, B. Jobst, and D. Hommel, J. Vac. Sci. Technol. B **14**, 2775 (1996).
[10]H. H. Farrell, M. C. Tamargo, T. J. Gmitter, A. L. Weaver, and D. E. Aspnes, J. Appl. Phys. **70**, 1033 (1991).
[11]N. Kumagai *et al.*, J. Cryst. Growth **184/185**, 505 (1998).
[12]C.-G. Jin, T. Yasuda, K. Kimura, A. Ohtake, L.-H. Kuo, T.-H. Wang, S. Miwa, T. Yao, and K. Tanaka, Jpn. J. Appl. Phys., Part 1 **36**, 6638 (1997).
[13]U. Rössner, D. Brun-Le Cunff, A. Barski, and B. Daudin, J. Vac. Sci. Technol. A **14**, 2655 (1996).
[14]Landolt–Börnstein, *Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung (Springer, Berlin, 1982), Vols. 17b and 22a.
[15]R. P. Vaudo, J. W. Cook, Jr., and J. F. Schetzina, J. Vac. Sci. Technol. B **12**, 1232 (1994).
[16]B. A. Philipps, I. Kamiya, K. Hingerl, L. T. Florez, D. E. Aspnes, S. Mahajan, and J. P. Harbison, Phys. Rev. Lett. **74**, 3640 (1995).
[17]K. Hingerl, T. Yasuda, T. Hanada, S. Miwa, K. Kimura, A. Ohtake, and T. Yao, J. Vac. Sci. Technol. B **16**, 2342 (1998).
[18]D. E. Aspnes and N. Bottka, in *Semiconductors and Semimetals*, edited by R. K. Willardson (Academic, Boston, 1972), Vol. 9.

Specular electron scattering in metallic thin films

W. F. Egelhoff, Jr., P. J. Chen, C. J. Powell, D. Parks, G. Serpa, and R. D. McMichael
National Institute of Standards and Technology, Gaithersburg, Maryland 20899

D. Martien and A. E. Berkowitz
University of California at San Diego, La Jolla, California 92093

(Received 19 May 1999; accepted 21 May 1999)

Specular electron scattering at the surface of metallic thin films is an important phenomenon for a class of magnetic multilayers known as giant magnetoresistance (GMR) spin valves. In the very best GMR spin valves, a significant part of the GMR effect is attributable to specular electron scattering. We have investigated the importance of specular electron scattering by developing surface-modification techniques that produce diffuse electron scattering. We have used these techniques to investigate specular electron scattering in GMR spin valves and in pure metals. Some of the largest effects are found in Au films. It is noted that specular electron scattering will be a highly desirable property in future generations of microelectronic interconnects as the feature size approaches the mean free path of the conduction electrons. In order to meet the Semiconductor Industry Association Roadmap requirements for interconnect resistivity, interconnects with highly specular surfaces must be developed. © *1999 American Vacuum Society.*
[S0734-211X(99)08204-9]

I. INTRODUCTION

The degree to which conduction electrons scatter specularly from the surfaces of ultrathin metal films has been a topic of scientific interest for 60 years.[1] However, the field never really blossomed, and the publication rate over that period averaged only about one article a year. The lack of important technological applications for specular electron scattering in metals appears to be the reason the field developed slowly. It seems likely that this situation is about to change as new technologies of great economic importance will be strongly dependent upon specular electron scattering to achieve high levels of performance.

The giant magnetoresistance (GMR) effect, which was discovered only ten years ago,[2] is already destined to play a key role in a variety of technological products. The next generation of read heads in computer hard disks will use GMR spin valves as the magnetic sensor. Already there is a race on among the hard-disk-drive companies to achieve the largest GMR values and thus the most sensitive heads. Another application is a nonvolatile, radiation-hard memory chip known as magnetoresistive random access memory (MRAM), in which the memory elements are GMR spin valves that store binary data as high or low resistance states. Honeywell is planning to go into production this year with a GMR-MRAM chip containing 10^6 memory elements and scale up rapidly to dynamic random access memory (DRAM) densities over the next few years.[3] Still another application is magnetic position sensors in which an array of GMR sensors and magnets is used with a microprocessor to control a mechanical process. Nonvolatile Electronics is marketing such sensors for a great diversity of applications, including automotive antilock-brake systems and industrial robot control.[4]

At present, GMR values in the range of 5%–7% are readily achievable, and such values are adequate for many initial applications of GMR materials. However, the need for GMR values above 10% will occur over the next few years as hard-disk-drive densities and MRAM densities increase. It now appears that specular electron scattering at the surfaces of GMR spin valves will be crucial to achieving GMR values above 10%.

In the area of microelectronic interconnects, there is a continuing need to lower resistivities as the feature sizes shrink. However, feature sizes are already approaching dimensions at which diffuse electron scattering at the walls of the interconnect are beginning to limit how small resistivities can be. High resistivities cause problems with signal propagation delays and with heat dissipation. The Semiconductor Industry Association (SIA) Roadmap states that there are no known solutions for reaching the target of interconnect resistivities below 1.8 $\mu\Omega$ cm, the value needed for feature sizes below 100 nm.[5] However, in principle, interconnects made of pure Cu with specular walls could achieve values as low as 1.7 $\mu\Omega$ cm. Learning how to achieve this in production will be one of the major challenges of the coming decade for interconnect technology. In our work we are beginning to explore how to fabricate metallic films so that they will scatter electrons more specularly.

II. EXPERIMENT

The NiO substrates used in this work were polycrystalline films ~ 50 nm thick, deposited on 3 in. Si wafers by reactive magnetron sputtering at the University of California at San Diego.[6] At the National Institute of Standards and Technology, the wafers were cleaved into ~1 cm^2 squares, cleaned ultrasonically in a detergent solution, rinsed in distilled water, blown dry, and installed in the deposition chamber.

It is very important to remove the hydrocarbon contamination (several tenths of a nanometer, which accumulates on the NiO from exposure to the laboratory air) prior to the

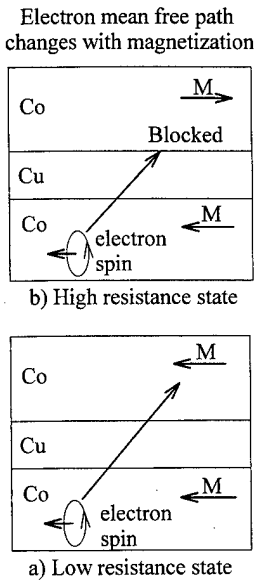


FIG. 1. Illustration of magnetization-dependent electron mean free paths that result in (a) low and (b) high resistance states in GMR materials.

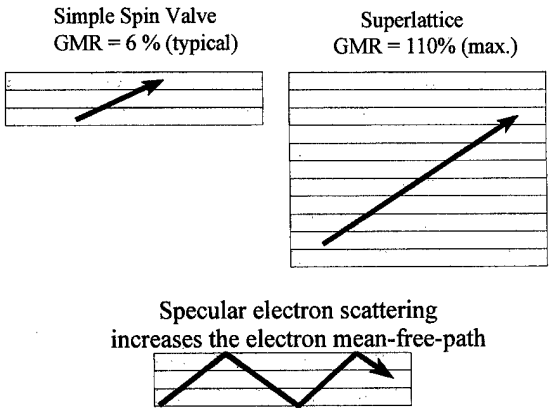


FIG. 2. Illustration of the mean free path of majority spin electrons in a simple spin valve, in a GMR superlattice, and in a hypothetical simple spin valve with specular surfaces (all systems in the parallel magnetization state).

deposition of each metallic film in order to achieve the lowest resistance. NiO substrates were sputtered with a neutralized-beam Ar-ion gun at a beam energy of 100 eV until the carbon was removed, as judged by x-ray photoelectron spectroscopy (XPS) measurements in a chamber connected to the deposition chamber. Ion beam energies of several hundred eV damage the NiO surface leaving it rough.

All films were deposited in a baked, ultrahigh vacuum (base pressure 3×10^{-10} Torr, or 4×10^{-8} Pa) system equipped with nine dc magnetron sputter deposition guns. The metal films were deposited at room temperature (RT) by dc magnetron sputtering in 2 mTorr (~0.3 Pa) Ar. The deposition rates for metal films are typically 0.02–0.05 nm/s. The thicknesses of the deposited films are monitored by two quartz crystal oscillators mounted alongside the substrate. Typical ambient pressures just prior to deposition were 3×10^{-9} Torr, or 4×10^{-7} Pa, of which ~90% was H_2. The low pressure in the chamber was achieved partly by depositing a ~ 1.5 nm Ti film on the inside of the deposition chamber from a centrally mounted Ti filament just prior to introduction of the NiO substrate.

The resistance measurements were made at RT using the four-point probe dc mode in a vacuum chamber connected to the deposition chamber. Resistivities are reported as bulk values, and were determined by multiplying the measured sheet resistance by the thickness.

III. RESULTS AND DISCUSSION

Figure 1 illustrates the basic principle of the high and low resistance states of a GMR-MRAM cell. Two Co films are separated by a Cu film (each ~3 nm thick). When the Co films are magnetized in opposite directions, electrons from one Co film have the wrong spin to enter the other Co film and their path is blocked (i.e., the majority and minority spin band structures of Co match up poorly), and a short mean free path results. However, when the magnetizations are parallel, the bands match and the mean free path is longer.

In device applications other than MRAM, GMR spin valves generally do not switch completely between the high and low resistance states. Instead, the external magnetic field that is to be detected shifts the spin valve only to a certain extent, i.e., from relatively high resistance to relatively low resistance or vice versa.

Figure 2 illustrates the two basic types of GMR films, spin valves, and superlattices. GMR values as large as 110% have been achieved in superlattices,[7] whereas in a simple spin valve (with a single Cu layer) GMR values are typically only 6%. Unfortunately, the magnetization switching process from antiparallel to parallel requires a very large applied field for superlattices (e.g., 3 T or 30k Gauss),[7] which is completely impractical for device applications. In contrast, simple spin valves often switch in fields on the order of a milliTesla (10 G). From our work, it appears that one of the reasons for the much smaller GMR values found in simple spin valves compared to those in superlattices is that the mean free path is limited by diffuse electron scattering at the top and bottom surfaces. If those surfaces could be made to scatter specularly, as illustrated in Fig. 2, and if the level of bulk defects could be made as low as in superlattices, it is very possible that simple spin valves could exhibit GMR values as large as those in superlattices.

We chose Au to begin our investigation of specular electron scattering in metallic thin films. Studies of specular scattering in Au have a long and controversial history.[1,8] Controversies arose largely because it is difficult to separate the bulk and surface contributions to the resistivity.[8] To avoid this problem, we chose to investigate specularity by a different route, that of surface modification. After much experimentation we found that the best way to make the surface of a Au film nonspecular is simply to deposit 0.4 nm Ta on it. It may be expected from many prior studies of the growth of high surface-free-energy metals on low surface-free-energy metals that the highly mobile Au atoms segregate out onto deposited Ta atoms to produce an alloy. Such an alloy should

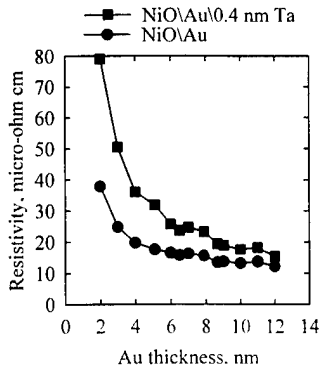


FIG. 3. Resistivity of Au films as deposited and after the deposition of 0.4 nm Ta to suppress the specular scattering at the surface.

be highly resistive, and indeed when we codeposited a $Ta_{50}Au_{50}$ film we found a resistivity of 215 $\mu\Omega$ cm. While we cannot measure the resistivity of the surface alloy generated by the deposition of Ta on Au, if it approached 215 $\mu\Omega$ cm it should very effectively suppress the specular scattering.

Figure 3 presents results on Au films using this method of suppressing specular scattering. Although we cannot prove that suppression of specular scattering is responsible for the increases in resistivity, that does appear to be the most reasonable interpretation. We found by x-ray photoelectron spectroscopy that essentially all the Ta remains in the near-surface region. Deposition of one conductor on another conductor would, in the simplest model, reduce the resistance of the film. Instead, we find a strong effect in the opposite direction. Figure 4 presents the increase in resistivity as a percentage of the initial resistivity. The magnitude of the effect is reasonable, however, from consideration of changes in specularity. For example, at a Au thickness of 3 nm the addition of 0.4 nm Ta approximately doubles the resistivity. The 3 nm Au film has a resistivity about ten times the single crystal value (24.7 $\mu\Omega$ cm vs 2.44 $\mu\Omega$ cm) implying that the mean free path is about 3.5 nm.[8] Since the mean free path and the thickness are similar, an approximate doubling of the resistivity by the addition of Ta would be expected since quenching specular scattering at the top surface is similar to halving the film thickness.

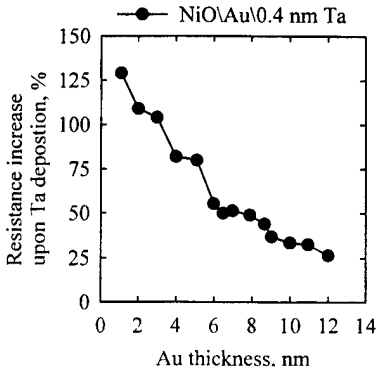


FIG. 4. Percentage increase in the resistance of the Au films upon the deposition of 0.4 nm Ta. The data are from Fig. 3.

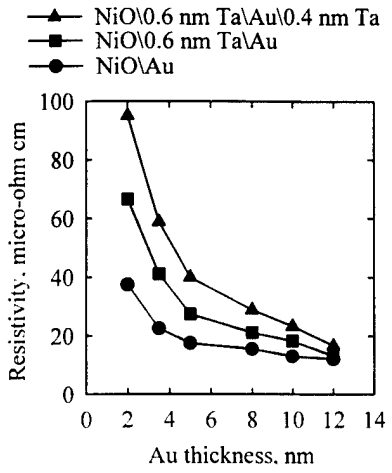


FIG. 5. Comparison of the resistivity of Au films of different thicknesses with no Ta with 0.6 nm Ta as an underlayer to make the bottom surface of the Au film nonspecular, and with 0.6 nm Ta as an underlayer and 0.4 nm Ta as an overlayer to make both the top and bottom surfaces nonspecular.

Figure 5 presents results of our attempts to make both the top and bottom surface of the Au nonspecular. After trying a great variety of methods to roughen the substrate surface, we found that the simple deposition of 0.6 nm Ta gave the highest resistivities for subsequently deposited Au films. As deposited, the 0.6 nm Ta film has a resistivity of ~10 Ω cm, suggesting that it is not continuous. The film probably consists of closely spaced clusters, each containing only a few Ta atoms, since surface diffusion by a refractory metal such as Ta is very limited. Intermixing of Ta and Au is not expected since there is no driving force for Ta surface segregation. Indeed, enthalpy should favor wetting of the Ta by the Au.

The increase in Au resistivity produced by this Ta underlayer is almost as large as that produced in Fig. 3 by suppressing the specular scattering at the top surface. If we interpret this increase to be the suppression of specular scattering at the bottom of the Au film, then it is likely that the bottom of the Au, like the top of the Au, exhibits a high degree of specular scattering in the absence of Ta.

The interpretation that both the top and bottom of the Au films are highly specular is made more plausible by our structural studies. Scanning tunneling microscopy shows that the Au films exhibit a grain size of ~14 nm. The grains are very slightly dome shaped with height differences from dome center to the grain boundary valley averaging only ~0.2 nm. Transmission electron micrographs show the NiO substrate to be remarkably flat. Often the roughness appears to be no greater than ~0.2 nm over distances of 50 nm. Thus, these are extremely flat surfaces.

Also illustrated in Fig. 5 is the effect of depositing 0.4 nm Ta on top of the Au to make the top surface also nonspecular. From these data, the top of the Au film appears to be quite specular scattering even when it is deposited on the Ta underlayer.

Figure 6 presents the percentage increases in resistivity for Fig. 5. An interesting aspect of Fig. 6 is the effect of the

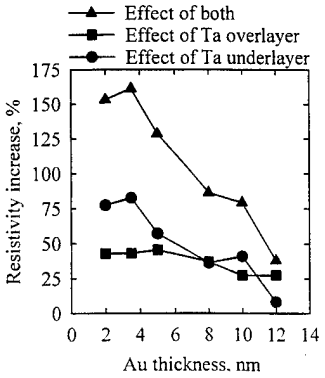


FIG. 6. Percentage increases in resistivity for the data of Fig. 5. The circles are the Ta underlayer case relative to the no Ta case, the squares are the Ta overlayer case relative to the Ta underlayer case, and the triangles are the Ta underlayer and overlayer case relative to the no Ta case.

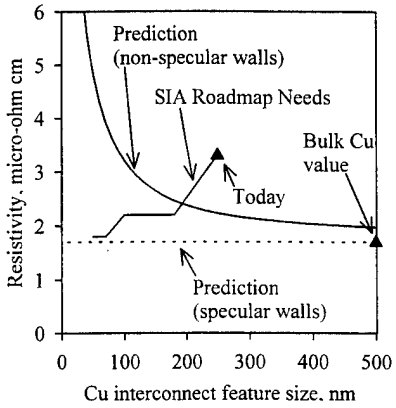


FIG. 7. Illustration of the problem facing future generations of microelectronic interconnects. Today, at a feature size of 250 nm AlCu alloy interconnects with a resistivity of 3.3 $\mu\Omega$ cm are adequate. For the future, the Semiconductor Industry Association (SIA) Roadmap requires declining resistivities as feature sizes shrink, a requirement that can only be met with Cu interconnects. However, even Cu interconnects cannot meet the requirement with nonspecular walls. The illustrated prediction for nonspecular walls is based on the Fuchs–Sondheimer equations and the bulk conductivity of Cu. In contrast, for specular walls the resistivity could, in principle, remain at the bulk Cu value (dotted line) as the feature size shrinks and thereby meet the SIA Roadmap requirements.

Ta overlayer (squares) on Au films that were deposited on the Ta underlayer. For Au thickness of 20–80 nm, the percentage increases are only ~40%, considerably smaller than those of Fig. 4 for which no Ta underlayer was present. These smaller increases are entirely plausible if, in the absence of Ta, some fraction of the electrons travels far enough to scatter specularly from both the top and bottom of the Au. Such multiple specular scattering events would be quenched if either one of the surfaces were nonspecular. Upon making the second surface also nonspecular, there would no longer be any multiple specular scattering events to quench, and a smaller increase in resistivity would be expected.

In the course of this work, we learned many other things about specular scattering. For example, preliminary data indicate that Cu and Ag films exhibit many of the same characteristics as Au films. Presumably, they too tend to have specular surfaces. Furthermore, thin Cu and Ag films deposited on Au films do not quench specular scattering as Ta films do. This result seems quite plausible since all three elements are rather free-electron-like *s*-band metals. Insulators such as NiO or SiO deposited on Au are roughly one third as effective as Ta in quenching the specularity. This result gives some hope that coatings that preserve specularity can be developed. It may be of interest that exposing a 5 nm Au film to air only caused a 5% increase in resistivity, implying that the surface was almost as specular in air as it was in vacuum. Apparently the conduction electrons do not scatter diffusely from the disordered molecular contaminants that condense on metal surfaces exposed to air.

Our experience indicates that diffuse electron scattering will dominate at metal surfaces unless great care is taken to prepare extremely flat surfaces. It appears likely that this result will apply to future generations of microelectronic interconnects. Figure 7 shows the consequences of diffuse electron scattering for this application. In this example, parameters for Cu are used since it appears to be the material of choice for future interconnects. The Fuchs–Sondheimer equations can be used to predict how nonspecular walls make the resistivity increase as the interconnect feature size shrinks.[1] The SIA Roadmap requirements intersect the nonspecular prediction near the 180 nm feature size generation. This generation is targeted for 1999 production in the most recent (1997) SIA Roadmap. For smaller feature sizes, the predicted resistivity rises rapidly as the required resistivity decreases. The adverse consequences of excessive resistivity are increased signal propagation delays and increased heat dissipation.

The best hope of meeting the SIA Roadmap requirements would seem to be to make interconnect walls that scatter electrons specularly. In this case, if the interior of the interconnect is single-crystal Cu (i.e., no additional resistance due to grain boundaries or other defects), the resistivity is predicted to remain at the bulk Cu value as the feature size decreases, and the SIA Roadmap requirements would then be satisfied.

Observations of specular electron scattering in metallic films and GMR spin valves have been reported in recent years.[9,10] However, relatively little research has been conducted into how to prepare suitable metal surfaces or how to modify them to make them highly specular. Presumably there was little motivation to do so in the absence of important technological applications. However, since it appears that this situation will change drastically over the next few years, we have begun to examine the issues involved more seriously.

During the course of our research on GMR spin valves we have made some progress in finding ways to make the spin valve surfaces scatter electrons more specularly and thereby to raise the GMR value (for reasons illustrated in Fig. 2). One method is illustrated in Fig. 8. It is well known in studies of epitaxial growth that smoother surfaces can often be produced if a surfactant is present during film growth.[11] In heteroepitaxy this effect is often the result of modifying the balance of surface and interfacial free energies. However, in

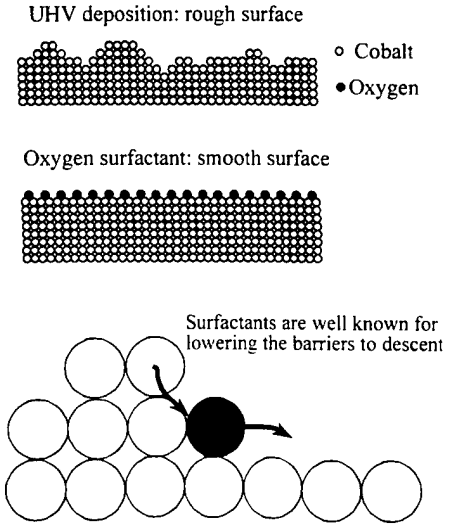


FIG. 8. Illustration of how a surfactant smooths a growing surface by lowering the barrier for descent of an upper terrace atom to a lower terrace.

the case of homoepitaxy, a different mechanism operates, and this is what Fig. 8 illustrates. Energy barriers often exist to inhibit atoms on upper terraces from descending to lower terraces.[11] The random arrival of atoms tends to leave a slightly roughened surface. However, a surfactant layer such as oxygen can lower the energy barrier to descent and produce a smoother surface.[11] In a concerted-motion process, fewer chemical bonds need to be broken in the descent event if the surfactant is present. Oxygen turns out to be an excellent surfactant in spin valves, in part because it is highly mobile and floats out to the growing surface with high efficiency. As a result, it can be used to produce smoother and more specular spin valve surfaces.

Figure 9 presents results for a spin valve of the type: NiO\2.5 nm Co\2 nm Cu\3 nm Co. The best GMR value results from maintaining a constant partial pressure of oxygen of 5×10^{-9} Torr, or 7×10^{-7} Pa, in the deposition chamber for the entire deposition.[10] As seen in Fig. 9, this

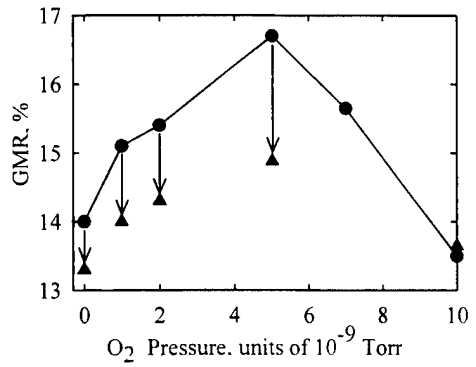


FIG. 9. Plot of the GMR of bottom spin valves as a function of the pressure of O_2 in which they were grown (solid circles). The arrows indicate the drop in GMR which occurs when 2 ML Ta is deposited to suppress the specular scattering at the surface (solid triangles). Note that 1×10^{-9} Torr equals 1.3×10^{-7} Pa.

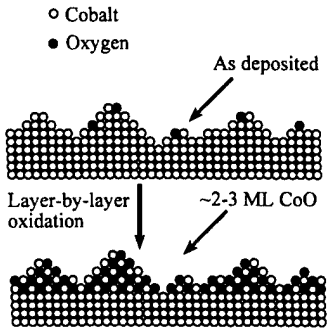


FIG. 10. Proposed model for the surface oxidation process that appears to increase the extent of specular electron scattering at the surface.

situation produces a GMR of 16.7% as opposed to 14.0% without oxygen. We found by x-ray photoelectron spectroscopy that this partial pressure of oxygen results in a surface coverage of oxygen atoms of approximately one half of an atomic layer for most of the spin valve growth (trace amounts of oxygen appear to be left behind at grain boundaries). The beneficial effect of oxygen exists only over a very narrow pressure range. At 1×10^{-8} Torr, or 1.3×10^{-6} Pa, the GMR is already below the no-oxygen value.

The role of specular scattering is illustrated in Fig. 9 by the arrows and triangles. These correspond to the results of depositing 0.4 nm Ta on the spin valve to quench specular scattering. As in the case of Au films, we found after much research that the most effective way to suppress specular scattering was simply to deposit 0.4 nm Ta on the spin valve. The extent of the drop from circle to triangle reflects the importance of specular scattering in each sample. Clearly, there was some specular scattering even in the absence of oxygen, but the right amount of oxygen significantly enhanced it.[10]

Another method we have found for increasing the extent of specular scattering at spin-valve surfaces is illustrated in Fig. 10. After the top Co film was deposited in a NiO\2.5 nm Co\2 nm Cu\3 nm Co spin valve, oxidation of the top 2–3 ML of Co generally produce an increase of about 1%–2% in GMR (absolute, i.e., ΔGMR).[10] Figure 11 presents a typical data set. It appears that if the surface, as deposited, is not very smooth, the oxidation of the surface will begin at high points or protrusions and progress downwards. Whatever

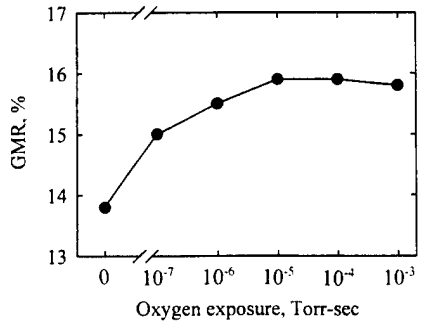


FIG. 11. Plot of the GMR of a bottom spin valve as a function of exposure to O_2 after deposition. Note that 1×10^{-7} Torr s equals 1.3×10^{-5} Pa s.

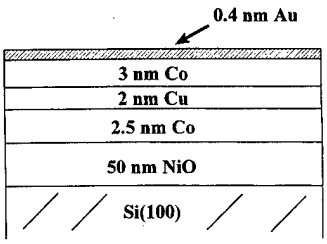


FIG. 12. Illustration of how the deposition for ~2 ML Au can be used to enhance the specular electron scattering at the surface of a spin valve.

bumps exist at the surface are preferentially oxidized first, and as the oxidation front meet a continuous layer of metal atoms, the oxidation rate drops drastically (i.e., as for a self-passivation process). Chemically, it is reasonable that bumps should oxidize more readily than flat layers since broken bonds at the sides of a bump should accelerate the oxidation process. Layer-by-layer oxidation has also been observed in semiconductors.[12] It has long been known that many metal surfaces oxidize very quickly to a depth of a few ML, after which the oxidation slows down dramatically. The idea of Fig. 10 is that oxidizing the high points leaves a metal/insulator interface (CoO/Co) that is flatter than the Co/vacuum interface was, thus increasing specular electron scattering. The process suggested in Fig. 10 may be termed layer-by-layer oxidation. Additional confirmation that the protrusions oxidize preferentially is found in the magnetostatic coupling measurements that we have published previously.[10]

A final method we have found for increasing specular electron scattering at the surface of spin valves is the deposition of thin layers of Au on top of the Co,[10] as is illustrated in Fig. 12. A film of 0.4 nm Au (~2 ML) generally increases the GMR from ~14% to 16%, an effect which is rather similar in magnitude to oxidizing the protrusions. This similarity in magnitude is probably not a coincidence but a consequence of both methods producing similar degrees of specular scattering at the surface. As further evidence that the Au makes the surface more specular, the increase in GMR can be entirely reversed with the deposition of 0.4 nm Ta. As with pure Au films, Ta deposition appears to be an excellent method of revealing the importance of specular scattering by quenching the specular scattering.

IV. CONCLUSIONS

The major points of this article may be summarized as follows:

(1) The importance of specular electron scattering in Au films and in spin valves can be investigated by depositing thin Ta layers to quench specular scattering.
(2) It appears that Au films in the thickness range of 2–12 nm on NiO substrates exhibit a high degree of specular electron scattering at both the top and bottom surfaces.
(3) Microelectronic interconnects are reaching feature sizes that will soon make specular electron scattering at the walls of interconnects a highly desirable property.
(4) Highly specular Co/Cu/Co spin-valve surfaces could well lead to spin valves with GMR valves approaching the 110% value of superlattices.
(5) Relatively little research has been conducted into how to produce metal surfaces that scatter electrons specularly.
(6) An increase in such research would be very timely.

[1]K. Fuchs, Proc. Cambridge Philos. Soc. **34**, 100 (1938); H. Sondheimer, Adv. Phys. **1**, 1 (1952).
[2]E. Velu, C. Dupas, D. Renard, J. P. Renard, and J. Seiden, Phys. Rev. B **37**, 668 (1988); M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas, Phys. Rev. Lett. **61**, 2472 (1988); G. Binasch, P. Grunberg, F. S. Sauerenbach, and W. Zinn, Phys. Rev. B **39**, 4828 (1989).
[3]T. Zhu, in Proceedings of the 43rd Magn. Magn. Mater Conference, 1998, J. Appl. Phys. (to be published).
[4]J. M. Daughton, J. Magn. Magn. Mater. **192**, 334 (1999).
[5]*The National Technological Roadmap for Semiconductors* (Semiconductor Industry Association, San Jose, 1997).
[6]M. J. Carey, F. E. Spada, A. E. Berkowitz, W. Cao, and G. Thomas, J. Mater. Res. **6**, 2680 (1991); M. J. Carey and A. E. Berkowitz, Appl. Phys. Lett. **60**, 3060 (1992); M. J. Carey and A. E. Berkowitz, J. Appl. Phys. **73**, 6892 (1993).
[7]S. S. P. Parkin (unpublished).
[8]J. R. Sambles, K. C. Elsom, and D. J. Jarvis, Philos. Trans. R. Soc. London, Ser. A **304**, 365 (1982); J. R. Sambles, Thin Solid Films **106**, 321 (1983).
[9]M. Jałochowski and E. Bauer, Phys. Rev. B **37**, 8627 (1988); **38**, 5272 (1988); M. Jałochowski, Prog. Surf. Sci. **48**, 287 (1995); E. Z. Luo, S. Heun, M. Kennedy, J. Wollschläger, and M. Henzler, Phys. Rev. B **49**, 4858 (1994); R. Schad, S. Heun, T. Heidenblut, and M. Henzler, *ibid.* **45**, 11430 (1992); D. Wang, J. M. Daughton, C. H. Smith, and E. Y. Chen, IEEE Trans. Magn. **32**, 4728 (1996); H. J. M. Swagten, G. J. Strijkers, P. J. H. Bloemen, M. M. H. Willekens, and W. J. M. de Jonge, Phys. Rev. B **53**, 9108 (1996); Y. Kawawake and H. Sakakima, IEEE Trans. Magn. **33**, 3538 (1997); H. J. M. Swagten, G. J. Strijkers, R. H. J. N. Bitter, W. J. M. de Jonge, and J. C. S. Kools, *ibid.* **34**, 948 (1998); Y. Sugita, M. Satomi, Y. Kawawake, and H. Sakakima, Jpn. J. Appl. Phys., Part 1 **37**, 109 (1998); H. Sakakima, Y. Sugita, M. Satomi, and Y. Kawawake (unpublished).
[10]W. F. Egelhoff, Jr., T. Ha, R. D. K. Misra, Y. Kadmon, J. Nir, C. J. Powell, M. D. Stiles, R. D. McMichael, C.-L. Lin, J. M. Sivertsen, J. H. Judy, K. Takano, A. E. Berkowitz, T. C. Anthony, and J. A. Brug, J. Appl. Phys. **78**, 273 (1995); W. F. Egelhoff, Jr., P. J. Chen, C. J. Powell, M. D. Stiles, R. D. McMichael, J. H. Judy, K. Takano, and A. E. Berkowitz, *ibid.* **82**, 6142 (1997); W. F. Egelhoff, Jr., P. J. Chen, C. J. Powell, M. D. Stiles, R. D. McMichael, J. H. Judy, K. Takano, A. E. Berkowitz, and J. M. Daughton, IEEE Trans. Magn. **33**, 3580 (1997).
[11]R. Kunke, B. Poelsema, L. K. Verheij, and G. Comsa, Phys. Rev. Lett. **65**, 733 (1990); J. Jacobsen, K. W. Jacobsen, P. Stoltze, and J. K. Ndørskov, *ibid.* **74**, 2295 (1995); H. A. van der Vegt, H. M. van Pinxteren, M. Lohmeier, E. Vlieg, and J. M. C. Thornton, *ibid.* **68**, 3335 (1992); B. Poelsema, R. Kunkel, N. Nagel, A. F. Becker, G. Rosenfeld, and G. Comsa, Appl. Phys. A: Solids Surf. **53**, 369 (1991); S. Esch, M. Hohage, T. Michely, and G. Comsa, Phys. Rev. Lett. **72**, 518 (1994); C. W. Oh, E. Kim, and Y. H. Lee, *ibid.* **76**, 776 (1996); H. A. van der Vegt, J. Alvarez, X. Torrelles, S. Ferrer, and E. Vlieg, Phys. Rev. B **52**, 17443 (1995); H. A. van der Vegt, M. Breeman, S. Ferrer, V. H. Etgens, S. Ferrer, V. H. Etgens, X. Torrelles, P. Fajardo, and E. Vlieg, *ibid.* **51**, 14806 (1995); J. Vrijmoeth, H. A. van der Vegt, J. A. Meyer, E. Vlieg, and R. J. Behm, Phys. Rev. Lett. **72**, 3843 (1994); Z. Zhang and M. G. Lagally, *ibid.* **72**, 693 (1994); P. J. Feibelman, *ibid.* **81**, 168 (1998).
[12]N. Miyata, H. Watanabe, and M. Ichikawa, Appl. Phys. Lett. **72**, 1715 (1998).

Spin relaxation of conduction electrons

J. Fabian and S. Das Sarma
Department of Physics, University of Maryland, College Park, Maryland 20742-4111

(Received 20 January 1999; accepted 28 April 1999)

Prospect of building electronic devices in which electron spins store and transport information has revived interest in the spin relaxation of conduction electrons. Since spin-polarized currents cannot flow indefinitely, basic spin-electronic devices must be smaller than the distance electrons diffuse without losing its spin memory. Some recent experimental and theoretical effort has been devoted to the issue of modulating the spin relaxation. It has been shown, for example, that in certain materials doping, alloying, or changing dimensionality can reduce or enhance the spin relaxation by several orders of magnitude. This brief review presents these efforts in the perspective of the current understanding of the spin relaxation of conduction electrons in nonmagnetic semiconductors and metals. © *1999 American Vacuum Society.* [S0734-211X(99)03604-5]

I. INTRODUCTION

Electron spin is becoming increasingly popular in electronics. New devices, now generally referred to as *spintronics*, exploit the ability of conduction electrons in metals and semiconductors to carry spin-polarized current. Three factors make spin of conduction electrons attractive for future technology: (1) electron spin can store information, (2) the spin (information) can be transferred as it is attached to mobile carriers, and (3) the spin (information) can be detected. In addition, the possibility of having long spin relaxation time or spin diffusion length in electronic materials makes spintronics a viable potential technology.

Information can be stored in a system of electron spins because these can be polarized. To represent bits, for example, spin up may stand for one, spin down for zero. But the sheer existence of two spin polarizations is of limited use if we do not have means of manipulating them. Currently used methods of polarizing electron spins include magnetic field, optical orientation, and spin injection. Polarization by magnetic field is the traditional method that works for both metals and semiconductors. Spin dynamics in semiconductors, however, is best studied by optical orientation where spin-polarized electrons and holes are created by a circularly polarized light. Finally, in the spin injection technique a spin-polarized current is driven, typically from a ferromagnet, into the metallic sample. Since spin is both introduced and transferred by current, this method is the most promising for spintronics. Unfortunately, thus far spin injection has not been convincingly demonstrated in semiconductors.

The second factor, the ability of information transfer by electron spins, relies on two facts. First, electrons are mobile and second, electrons have a relatively large spin memory. Indeed, conduction electrons "remember" their spins much longer than they remember momentum states. In a typical metal, momentum coherence is lost after ten femtoseconds, while spin coherence can survive more than a nanosecond. As a result, the length L_1, the spin diffusion length, over which electrons remain spin polarized is much longer than the mean free path distance l over which their momentum is lost. Since L_1 is the upper limit for the size of spintronic elements (in larger elements the spin-encoded information fades away), it is not surprising that significant effort went into finding ways of reducing the spin relaxation. Quite unexpectedly, in quantum wells, but even in bulk semiconductors, donor doping was found to increase the spin memory of conduction electrons by up to three orders of magnitude. In metals one has much less freedom in manipulating electron states. A theoretical study, however, predicts that even there spin memory can be changed by orders of magnitude by band-structure tailoring. Alloying of polyvalent metals with monovalent ones can increase the spin memory by a decade or two. The ability of conduction electrons to transport spin-polarized current over distances exceeding micrometers has now been demonstrated in both metals and semiconductors.

Finally, after the spin is transferred, it has to be detected. In many experiments, the spin polarization is read optically: photoexcited spin-polarized electrons and holes in a semiconductor recombine by emitting circularly polarized light; or the electron spins interact with light and cause a rotation of the light polarization plane. It was discovered, however, that spin can be also measured electronically, through charge-spin coupling. When spin accumulates on the conductor side at the interface of a conductor and a ferromagnet, a voltage or a current appears. By measuring the polarity of the voltage or the current, one can tell the spin orientation in the conductor. Like spin injection, spin-charge coupling has been demonstrated only in metals.

The operational synthesis of spin (information) storage, transfer, and detection can be illustrated on concrete devices. Spin transistor is a trilayer that consists of a nonmagnetic metal (base) sandwiched between two ferromagnets (emitter and collector). Spin-polarized current injected into the base from the emitter causes spin accumulation at the base-collector interface. If the collector magnetic moment is opposite to the spin polarization of the current (and parallel to the emitter magnetic moment, if the injected electrons are from the spin-minority subband), the current flows from the base into the collector. If the collector magnetic moment is parallel to the spin polarization, the current is reversed. In order for the spin accumulation to occur, the current in the

metallic base must remain polarized—the base must be thinner than L_1. Similar principles work in the giant magnetoresistance effect. Multilayer structures with alternating nonmagnetic and ferromagnetic metals have their resistance strongly dependent on the relative orientation of the ferromagnetic moments. The resistance is small if the moments point in the same directions, and large if the directions of neighboring moments are reversed. Again, the information about the moment of a ferromagnetic layer is encoded into electron spins which carry this information through a contiguous nonmagnetic metal into another ferromagnet. Here the information is read and in ideal case the electron is let into the ferromagnet only if its spin is opposite to the direction of the ferromagnetic moment. Otherwise, the electron is scattered at the interface.

Several recent reviews focus on spin-polarized transport. An overview of the subject can be found in Ref. 1. Spin transistors, spin injection, and charge-spin coupling in metallic systems is treated in Ref. 2; a comprehensive account of optical orientation is given in Ref. 3, and recent reviews of giant magnetoresistance are in Refs. 4 and 5. Many suggested spintronic devices have not been demonstrated yet, but their potential seems enormous. Industrial issues related to spintronics can be found in Refs. 6, 7, and Ref. 8 describes some of the recent spintronic schemes and devices.

The present article introduces basic concepts of the spin relaxation of conduction electrons and identifies important unresolved issues in both semiconductors and metals. Particular emphasis is given to the recent experimental and theoretical work that attempts to enhance and/or understand electron spin coherence in electronic materials.

II. MECHANISMS OF SPIN RELAXATION

Spin relaxation refers to the processes that bring an unbalanced population of spin states into equilibrium. If, say, spin up electrons are injected into a metal at time $t=0$ creating a spin imbalance, at a later time, $t=T_1$ (the so-called spin relaxation time), the balance is restored by a coupling between spin and orbital degrees of freedom. Three spin-relaxation mechanisms have been found to be relevant for conduction electrons (Fig. 1): the Elliott–Yafet, D'yakonov–Perel', and Bir–Aronov–Pikus.

The Elliott–Yafet mechanism is based on the fact that in real crystals Bloch states are not spin eigenstates. Indeed, the lattice ions induce the spin-orbit interaction that mixes the spin-up and spin-down amplitudes.[9] Usually the spin-orbit coupling λ is much smaller than a typical band width ΔE and can be treated as a perturbation. Switching the spin-orbit interaction adiabatically, an initially spin-up (down) state acquires a spin-down (up) component with amplitude b of order $\lambda/\Delta E$. Since b is small, the resulting states can be still named "up" and "down" according to their largest spin component. Elliott[9] noticed that an ordinary (spin-independent) interaction with impurities, boundaries, interfaces, and phonons can connect "up" with "down" electrons, leading to spin relaxation whose rate $1/T_1$ is proportional to b^2/τ (τ being the momentum relaxation time

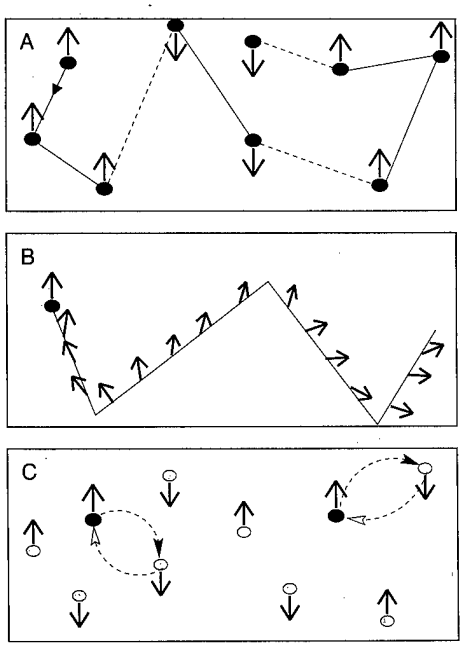


FIG. 1. Relevant spin relaxation mechanisms for conduction electrons. (A) The Elliott–Yafet mechanism. The periodic spin-orbit interaction makes the spin "up" Bloch states contain small spin-down amplitude, and vice versa. Impurities, boundaries, and phonons can induce transitions between spin "up" and "down" leading to spin degradation. (B) The D'yakonov–Perel' mechanism. In noncentrosymmetric crystals spin bands are no longer degenerate: in the same momentum state spin up has different energy than spin down. This is equivalent to having internal magnetic fields, one for each momentum. The spin of an electron precesses along such a field, until the electron momentum changes by impurity, boundary, or phonon scattering. Then the precession starts again, but along a different axis. Since the spin polarization changes during the precession, the scattering acts against the spin relaxation. (C) The Bir–Aronov–Pikus mechanism. The exchange interaction between electrons and holes causes the electron spins to precess along some effective magnetic field determined by hole spins. In the limit of strong hole spin relaxation, this effective field randomly changes before the full precession is completed, reducing the electron spin relaxation.

determined by "up" to "up" scattering). Additional spin-flip scattering is provided by the spin-orbit interaction of impurities, and by the phonon-modulated spin-orbit interaction of the lattice ions (Overhauser[10]). The latter should be taken together with the Elliott phonon scattering to get the correct low-temperature behavior of $1/T_1$.[11] Yafet[11] showed that $1/T_1$ follows the temperature dependence of resistivity: $1/T_1 \sim T$ at temperatures T above the Debye temperature T_D, and $1/T_1 \sim T^5$ at very low T in clean samples (neutral impurities lead to T-independent spin relaxation). Elliott–Yafet processes due to the electron-electron scattering in semiconductors were evaluated by Boguslawski.[12]

In crystals that lack of inversion symmetry (such as zinc-blende semiconductors) the spin-orbit interaction lifts the spin degeneracy: spin-up and spin-down electrons have different energies even when in the same momentum state. This is equivalent to having a momentum-dependent internal magnetic field $\mathbf{B}(\mathbf{k})$ which is capable of flipping spins through the interaction term like $\mathbf{B}(\mathbf{k})\cdot\mathbf{S}$, with $\mathbf{S}$ denoting the electron spin operator. (This term can be further modulated by strain or by interface electric fields.) D'yakonov and

Perel' showed that the lifting of the spin degeneracy leads to spin relaxation.[13] Typically the distance between spin-up and spin-down bands is much smaller than the frequency $1/\tau$ of ordinary scattering by impurities, boundaries, or phonons. Consider an electron with momentum $\mathbf{k}$. Its spin precesses along the axis given by $\mathbf{B}(\mathbf{k})$. Without going the full cycle, the electron scatters into momentum $\mathbf{k}'$ and begins to precess along the direction now given by $\mathbf{B}(\mathbf{k}')$, and so on. The electron spin perceives the scattering through randomly changing precession direction and frequency. The precession angle along the axis of initial polarization (or any other fixed axis) diffuses so its square becomes about $(t/\tau)(\omega\tau)^2$ after time t (ω is the typical precession frequency). By definition T_1 is the time when the precession angle becomes of order one. Then $1/T_1 \approx \omega(\omega\tau)$. The factor $(\omega\tau)$ is a result of motional narrowing as in nuclear magnetic resonance.[14] The spin relaxation rate $1/T_1$ is proportional to the momentum relaxation time τ. We note that in strong magnetic fields the precession along the randomly changing axis is suppressed (spins precess along the external field[15] and electron cyclotron motion averages over different internal magnetic fields[16,17]), leading to a reduction of the D'yakonov–Perel' spin relaxation.

Another source of spin relaxation for conduction electrons was found by Bir, Aronov, and Pikus[18] in the electron-hole exchange interaction. This interaction depends on the spins of interacting electrons and holes and acts on electron spins as some effective magnetic field. The spin relaxation takes place as electron spins precess along this field. In many cases, however, hole spins change with the rate that is much faster than the precession frequency. When that happens the effective field which is generated by the hole spins fluctuates and the precession angle about a fixed axis diffuses as in the case of the D'yakonov–Perel' process. The electron spin relaxation rate $1/T_1$ is then "motionally" reduced and is proportional to the hole spin relaxation time. Similar reduction of $1/T_1$ occurs if holes that move faster than electrons change their momentum before electron spins precess a full cycle.[18] The Bir–Aronov–Pikus spin relaxation, being based on the electron-hole interaction, is relevant only in semiconductors with a significant overlap between electron and hole wave functions.

III. SEMICONDUCTORS

Spin relaxation in semiconductors is rather complex. First, there are different charge carriers to consider. Both electrons and holes can be spin polarized and carry spin-polarized currents. Furthermore, some features of the observed luminescence polarization spectra must take into account excitons, which too, can be polarized. Second, in addition to temperature and impurity content the spin relaxation is extremely sensitive to factors like doping, dimensionality, strain, magnetic and electrical fields. The type of dopant is also important: electrons in p-type samples, for example, can relax much faster than in n-type samples. And, finally, since the relevant electron and hole states are typically very close to special symmetry points of the Brillouin zone, subtleties of the band structure often play a decisive role in determining which spin relaxation mechanism prevails. (Band structure also determines what is polarized—often due to degeneracy lifting, spin and orbital degrees are entirely mixed and the total angular momentum is what is referred to as "spin.") The above factors make sorting out different spin relaxation mechanisms a difficult task.

The first measurement of T_1 of free carriers in a semiconductor was reported in Si by Portis *et al.*[19] and Willenbrock and Bloembergen;[20] these measurements were done by conduction electron spin resonance. Silicon, however, remains still very poorly understood in regards to its spin transport properties. Very little is known, for example, about electronic spin-flip scattering by conventional n and p dopants. Considering that Si may be an important element for spintronics since it is widely used in conventional electronics, its spin relaxation properties should be further investigated.

Much effort was spent on III–V semiconductors where optical orientation[3] enables direct measurement of T_1. In these systems holes relax much faster than electrons because hole Bloch states are almost an equal admixture of spin-up and down eigenstates. The Elliott–Yafet mechanism then gives T_1 of the same order as τ. In quantum wells (QW), however, T_1 of holes was predicted by Uenoyama and Sham[21] and Ferreira and Bastard[22] to be quenched, and even longer than the electron-hole recombination time. This was observed experimentally in n-modulation doped GaAs QWs by Damen *et al.*[23] who measured hole spin relaxation time of 4 ps at 10 K. Hole and exciton spin relaxation was reviewed by Sham.[24]

Compared to holes, electrons in III–V systems remember their spins much longer and are therefore more important for spintronic applications. Typical measured values of electron T_1 range from 10^{-11} to 10^{-7} s. All the three spin relaxation mechanisms have been found contributing to T_1. Although it is difficult to decide which mechanism operates under the specific experimental conditions (this is because in some cases two mechanisms yield similar T_1, but also because experiments often disagree with each other[25]), some general trends are followed. The Elliott–Yafet mechanism dominates in narrow-gap semiconductors, where b^2 is quite large ($\Delta E \approx E_g$ is small). Chazalviel[26] studied n-doped InSb ($E_g \approx 0.2$ eV) and found that Elliott–Yafet scattering by ionized impurities explains the observed $1/T_1$.

If band gap is not too small, the D'yakonov–Perel' mechanism has been found relevant at high temperatures and sufficiently low densities of holes. The D'yakonov–Perel' mechanism can be quite easily distinguished from the Elliott–Yafet one: the former leads to $1/T_1 \sim \tau$ while for the latter $1/T_1 \sim 1/\tau$. The increase in the impurity concentration decreases the efficiency of the D'yakonov–Perel' processes and increases those due to Elliott and Yafet. Another useful test of the D'yakonov–Perel' mechanism is its suppression by magnetic field.[16,17] The first experimental observation of the D'yakonov–Perel' mechanism was reported by Clark *et al.* on moderately doped p samples of GaAs[27] and GaAlAs.[28] Later measurements on less doped samples of

GaAs by Maruschak *et al.*[29] and Zerrouati *et al.*[25] confirmed that the D'yakonov–Perel' mechanism is dominant in GaAs at elevated temperatures.

At low temperatures and in highly p-doped samples (acceptor concentration larger than $10^{17}\,\mathrm{cm}^{-3}$), the Bir–Aronov–Pikus mechanism prevails. As the acceptor concentration increases this mechanism reveals itself at progressively higher temperatures. An increase of $1/T_1$ with increasing p doping signals that the electron-hole spin relaxation is relevant. This was demonstrated in p-type GaAs (for example, Zerrouati *et al.*,[25] Maruschak *et al.*,[29] and Fishman and Lampel[30]) and GaSb (Aronov *et al.*[31]). The physics of spin relaxation in p-doped III–V semiconductors is very rich because several different mechanisms have been shown relevant. More work, however, still needs to be done. It is not clear, for example, what happens at very low temperatures and in very pure samples.[25] There are some indications that at very low temperatures both the D'yakonov–Perel' and the Bir–Aronov–Pikus mechanisms can explain the observed data at whatever doping.[25] Excellent reviews of conduction electron spin relaxation in bulk III–V semiconductors are.[32,33] These references contain both experimental data and many useful formulas of $1/T_1$.

Electron spin relaxation has been studied also in quantum wells. That spin dynamics in quantum wells differs from that in the bulk is obvious from the fact that the relevant spin relaxation mechanisms are very sensitive to factors like mobility (which is higher in QWs), electron-hole separation (smaller in QWs) and electronic band structure (more complicated in QWs because of subband structures and interface effects). Furthermore, the quality of QW samples is very important since $1/T_1$ is strongly influenced by localization and defects. The first measurement of conduction electron T_1 in a QW was reported by Damen *et al.*[23] who studied the dynamics of luminescence polarization in p-modulation doped GaAs/AlGaAs, and obtained $T_1 \approx 0.15$ ns at low temperatures. This relaxation time is three to four times smaller than in a similar bulk sample (the acceptor concentration was $4\times10^{11}\,\mathrm{cm}^{-2}$). It was concluded[23] that the relevant mechanism was Bir–Aronov–Pikus. The recent theoretical study by Maialle and Degani[34] of the Bir–Aronov–Pikus relaxation in QWs indicates that, to the contrary, this mechanism is not efficient enough to explain the experiment. Another possibility is the D'yakonov–Perel' mechanism. Bastard and Ferreira[35] calculated the effectiveness of this mechanism for the case of ionized impurity scattering. Their calculation shows[24] that the D'yakonov–Perel' mechanism is also too weak to explain the experiment. Although some assumptions of the theoretical studies may need to be reexamined (the major difficulty seems to be estimating τ),[24] further experimental work (such as temperature and doping dependence) is required to decide on the relevant mechanism. Recently, Britton *et al.*[36] studied the spin relaxation in undoped GaAs/AlGaAs multiple quantum wells at room temperature. The measured relaxation times vary between 0.07 and 0.01 ns, decreasing strongly with increasing confinement energy. These results seem to be consistent with the D'yakonov–Perel' mechanism.[36]

Spin relaxation studies in quantum wells also promise better understanding of interface effects. In an inversion layer, an electric field arises from the electrostatic confinement. This field induces a spin-orbit interaction which contributes to the spin splitting (the so-called Rashba splitting) of electron bands in addition to the inversion-asymmetry splitting. This should enhance the efficiency of the D'yakonov–Perel' mechanism. Spin precession of conduction electrons in GaAs inversion layers was investigated by Dresselhaus *et al.*[37] using antilocalization. The spin relaxation was found to be due to the D'yakonov–Perel' mechanism, but the spin splitting was identified (by magnetic field dependence) to be primarily due to the inversion asymmetry. This is consistent with an earlier theoretical study of Lommer *et al.*[38] of spin splitting in heterostructures, which predicted that in GaAs/AlGaAs QWs the Rashba term in the Hamiltonian is weak. In narrow-band semiconductors, however, Lommer *et al.* predict that the Rashba term becomes relevant. But this remains a not-yet-verified theoretical prediction. Another interesting study of the interface effects was done recently by Guettler *et al.*[39] following the calculations of Vervoort *et al.*[40] Quantum well systems in which wells and barriers have different host atoms (so-called "no-common-atom" heterostructures) were shown to have conduction electron spin relaxation enhanced by orders of magnitude compared to common-atom heterostructures. In particular, spin relaxation times in (InGa)As/InP QWs were found to be 20 (90) ps for electrons (holes), while the structures with common host atoms (InGa)As/(AsIn)As have spin relaxation times much longer: 600 (600) ps. This huge difference between otherwise similar samples is attributed to the large electric fields arising from the asymmetry at the interface (interface dipolar fields).[39]

Spin relaxation of conduction electrons can be controlled. This was first realized by Wagner *et al.*[41] who δ-doped GaAs/AlGaAs double heterostructures with Be (as acceptor). The measured spin relaxation time was about 20 ns which is two orders of magnitude longer than in similar homogeneously p-doped GaAs. The understanding of this finding is the following. The sample was heavily doped ($8\times10^{12}\,\mathrm{cm}^{-2}$) so the Bir–Aronov–Pikus mechanism was expected to dominate the relaxation. Photogenerated electrons, however, were spatially separated from holes which stayed mostly at the center of the GaAs layer, close to the Be dopants. There was, however, still enough overlap between electrons and the holes for efficient recombination so that the radiation polarization could be studied. The decrease of the overlap between electrons and holes reduced the efficiency of the Bir–Aronov–Pikus mechanism and increased T_1. This experiment can be also taken as a confirmation that the Aronov–Bir–Pikus mechanism is dominant in heavily p-doped heterostructures.

The next important step in controlling spin relaxation was the observation of a large enhancement of the spin memory of electrons in II–VI semiconductor QWs by Kikkawa

et al.[42] Introducing a (two-dimensional) electron gas by n doping, the II–VI QWs was found to increase electronic spin memory by several orders of magnitude. The studied samples were modulation-doped $Zn_{1-x}Cd_xSe$ quantum wells with electron densities 2×10^{11} and $5\times10^{11}\,cm^{-2}$ (an additional insulating sample was used as a benchmark). Spin polarization was induced by a circularly polarized pump pulse directed normal to the sample surface. The spins, initially polarized normal, began to precess along an external magnetic field oriented along the surface plane. After a time δt, a probe pulse of a linearly polarized light detected the orientation of the spins. The major result of the study was that in doped samples electron spin remained polarized for almost three orders of magnitude longer than in the insulating (no Fermi see) sample. The measured T_1 was on the nanosecond scale, strongly dependent on the magnetic field and weakly dependent on temperature and mobility. Although the nanosecond time scales and the increase of the observed polarization in strong magnetic fields (usually a Tesla) could be explained by the D'yakonov–Perel' mechanism,[25] the temperature and mobility (in)dependence remain a puzzle. The overall increase of T_1 by donor doping can be understood in the following way.[42] In insulating samples, photoexcited spin-polarized electrons quickly recombine with holes. This happens in picoseconds. In the presence of a Fermi sea photoexcited electrons do not recombine (there are plenty other electrons available for recombination) so they lose their spins in nanoseconds, which are natural time scales for spin relaxation. There is a caveat, however. The above scenario is true only if holes lose their spins faster than they recombine with electrons. Otherwise only electrons from the Fermi sea with a preferred spin would recombine, leaving behind a net opposite spin that counters that of the photoexcited electrons. The fast hole relaxation certainly happens in the bulk (and similar enhancement of T_1 has been observed in n-doped bulk GaAs by Kikkawa and Awschalom[43]), but not necessarily in quantum heterostructures.[21–23] This issue therefore remains open. Very recent optically pumped nuclear magnetic resonance measurements[44] in n-doped AlGaAs/GaAs multiple quantum well systems indicate unusually long $T_1\gtrsim100\,\mu s$ at temperatures below 500 mK in the two-dimensional electron gas system under the application of a strong ($\gtrsim$12 T) external magnetic field. It is unclear whether this remarkable decoupling (that is, $T_1\gtrsim100\,\mu s$) of the two-dimensional electron gas spins from its environment is an exotic feature of the fractional quantum Hall physics dominating the system, or is a more generic effect which could be controlled under less restrictive conditions.

It was recently demonstrated that spin polarized current can flow in a semiconductor. Hägele *et al.*[45] used a simple but ingenuous setup that consisted of a micrometer i-GaAs block attached to a p-modulation doped GaInAs QW layer. The free surface of the GaAs block was illuminated by a circularly polarized light. The photogenerated electrons then drifted towards the QW under the force of an applied electric field (photoexcited holes moved in the opposite direction towards the surface). The electrons recombined with holes upon hitting the QW, emitting light. By observing the polarization of the emitted light Hägele *et al.* concluded that electrons captured by the QW were polarized. The spin was almost completely conserved after the electrons traveled as long as 4 μm and under the fields up to 6 kV/cm, indicating very long spin diffusion lengths in these experiments.[45]

IV. METALS

Only a dozen elemental metals have been investigated for spin relaxation so far. Early measurements of T_1 were done by the conduction electron spin resonance technique. This technique was demonstrated for metals by Griswold *et al.*,[46] and Feher and Kip[47] used it to make the first T_1 measurement of Na, Be, and Li. This and subsequent measurements established that $1/T_1$ in metals depends strongly on the impurity content (especially in light metals like Li and Be) and grows linearly with temperature at high temperatures. Typical spin relaxation time scales were set to nanoseconds, although in very pure samples T_1 can reach microseconds at low temperatures (for example, in sodium, as observed by Kolbe[48]). Reference 49 is a good source of these early spin relaxation measurements.

The next wave of measurements started with the realization of spin injection in metals. Suggested theoretically by Aronov,[50] spin injection was first demonstrated in Al by Johnson and Silsbee.[51] Later measurements were done on Au[52] and Nb[53] films. The spin injection technique enables measurements of T_1 in virtually no magnetic fields so that T_1 can now be measured in superconductors, spin glasses, or Kondo systems where magnetic field qualitatively alters electronic states. Furthermore, by eliminating the need for magnetic fields to polarize electron spins one avoids complications like inhomogeneous line broadening, arising from g factor anisotropy. Johnson also succeeded in injecting spin-polarized electrons into superconducting Nb films.[53] Spin relaxation of electrons (or, rather, quasiparticles) in superconductors is, however, poorly understood and the experiments, now done mostly on high-T_c materials[54,55] only manifest the lack of theoretical comprehension of the subject. Still waiting for its demonstration is spin injection into semiconductors. Although it was predicted long ago by Aronov and Pikus,[56] it still remains a great experimental challenge.

The observations that $1/T_1\sim T$ at high temperatures, the strong dependence on impurities, and characteristic nanosecond time scales has led to the general belief that conduction electrons in metals lose their spins by the Elliott–Yafet mechanism. Although simple estimates and even some analytical calculations were done for simplest metals like Na (Yafet[11]), careful numerical calculations are lacking. Experimental data are usually analyzed to see if the simple relation suggested by Yafet,[11]

$$1/T_1\sim b^2\rho, \tag{1}$$

where ρ is resistivity, is obeyed. The spin-mixing b^2 is the fitting parameter so the temperature behavior of $1/T_1$ is determined solely by ρ. At high temperatures $1/T_1\sim\rho\sim T$ as observed. At low temperatures the spin relaxation should

obey the Yafet law $1/T_1 \sim T^5$ (in parallel to the Bloch law $\rho \sim T^5$), but so far this has not been observed, mainly due to the large contribution from impurity and boundary scattering. (Even after subtracting this temperature independent background the uncertainties of the measurements prevent a definite experimental conclusion about the low T behavior.)

Equation (1) suggests that dividing $1/T_1$ by b^2 one obtains resistivity, up to a multiplicative (material independent) constant. Resistivity, divided by its value ρ_D at T_D and expressed as function of reduced temperature T/T_D follows a simple Grüneisen curve, the same for all simple metals. Monod and Beuneu[49] applied this reasoning to then available experimental data of T_1. For the spin mixing b^2, they substituted values obtained from atomic parameters of the corresponding elements. The resulting (revised) scaling is reproduced in Fig. 2. (The original scaling[49] has Γ_s divided by b^2, not by $b^2\rho_D$.) The picture is surprising. While some metals (the "main group") nicely follow a single Grüneisen curve, others do not. There seems to be no obvious reason for the observed behavior. Metals Na and Al, for example, are quite similar in that their atomic b^2 differ by less than 10%.[49] Yet the spin relaxation times at T_D are 0.1 ns for Al and 20 ns for Na.[57]

The solution to this puzzle can be found by recognizing[57] that the main group is formed by monovalent alkali and noble metals, while the metals with underestimated b^2, Al, Pd, Mg, and Be are polyvalent (no other metals have been measured for T_1 in a wide enough temperature region). Monovalent metals have their Fermi surfaces inside Brillouin zone boundaries so that distance between neighboring bands, ΔE is quite uniform and of the order of the Fermi energy E_F. The spin mixing is then $b^2 \approx (\lambda/E_F)^2$ for all states on the Fermi surface. Polyvalent metals, on the other hand, have Fermi surfaces which cross Brillouin zone boundaries, and often also special symmetry points and accidental degeneracy lines. When this happens the electron spin relaxation is significantly enhanced. This was first noted by Silsbee and Beuneu[58] who estimated the contribution to Al $1/T_1$ from accidental degeneracy lines. Later the present authors gave a rigorous and detailed treatment of how not only accidental degeneracy, but all the band anomalies contribute to $1/T_1$.[57] This treatment led to the spin-hot-spot model[57] which explains why all the measured polyvalent metals have spin relaxation faster than expected from a naive theory. In addition to explaining experiment, the spin-hot-spot model predicts the behavior of other polyvalent metals. The model is illustrated in Fig. 3.

As an example, consider a metal whose Fermi surface crosses a single Brillouin zone boundary.[57,59] The distance between energy bands ΔE is about E_F for all Fermi surface states except those close to the boundary. There $\Delta E \approx 2V$,[14] where V is the Fourier component of the lattice potential associated with the boundary. Since in most cases $V \ll E_F$ the spin mixing $b^2 \approx (\lambda/V)^2$ is much larger than on average. If an electron jumps into such states, the chance that its spin will be flipped is much enhanced. Similarly if the electron jumps from these "spin hot spots." But how much the states with $\Delta E \approx 2V$ contribute to spin relaxation depends on how many they are relative to the number of states on the Fermi

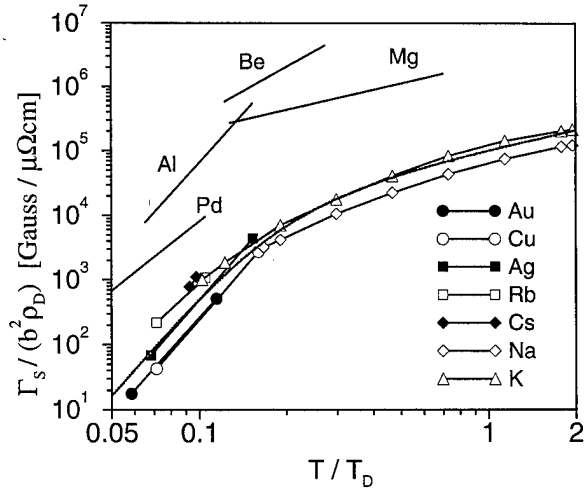


FIG. 2. Revised Monod–Beuneu scaling. The measured width $\Gamma_s = \text{const} \times (1/T_1)$ of the conduction electron spin resonance signal is divided by the effective spin-mixing probability b^2 obtained from atomic parameters, and by resistivity ρ_D at Debye temperature T_D. This should follow a Grüneisen curve when plotted as function of reduced temperature T/T_D. The alkali metals fall onto a single curve while Al, Pd, Be, and Mg do not, indicating that their b^2 is much larger than estimated from atomic parameters.

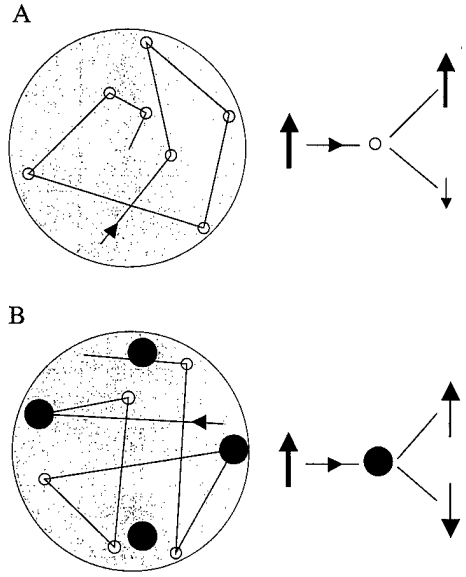


FIG. 3. Spin-hot-spot model. (A) Monovalent metals. As electrons scatter and change momentum, they perform a random walk on the Fermi surface. At each jump the electrons have a small chance of flipping their spin (the Elliott–Yafet mechanism), indicated on the right. In monovalent metals, this chance is uniform over the Fermi surface and is roughly equal to $(\lambda/E_F)^2$. (B) Polyvalent metals. Fermi surfaces of polyvalent metals contain spin hot spots (back stains), which are states at Brillouin zone boundaries, special symmetry points, or accidental degeneracy lines. If an electron jumps into such a state, its chance of flipping spin is much enhanced. Although spin hot spots form a small part of the Fermi surface and the probability that an electron jumps there is quite small, they nevertheless dominate the electron spin relaxation.

surface. A single electron experiences thousands of jumps due to momentum scattering before its spin flips. Therefore the spin relaxation rate $1/T_1$ is determined by the average $\langle b^2 \rangle$ of b^2 over the Fermi surface. The majority of states with $\Delta E \approx E_F$ contribute $(\lambda/E_F)^2 \times 1$ (the value of b^2 times the probability of occurrence, which in this case is close to one) to $\langle b^2 \rangle$. The probability of finding a state with $\Delta E \approx 2V$ on the Fermi surface turns out to be about V/E_F,[59] so the spin hot spots contribute about $(\lambda/V)^2 \times (V/E_F)$, which is $(\lambda/E_F)^2 \times (E_F/V)$. This is larger by E_F/V than the contribution from ordinary states. Typically $E_F/V \approx 10$, and considering that in reality the Fermi surface crosses more than one Brillouin zone boundary, the spin relaxation can be enhanced up to two orders of magnitude. Electron jumps that include at least one spin-hot-spot state dominate spin relaxation to the extent that the majority of scattering events (those outside the spin hot spots) can be neglected.

The spin-hot-spot picture not only solves a long-standing experimental puzzle, but also shows a way to tailor the spin relaxation of electrons in a conduction band. Spin relaxation of a monovalent metal, for example, can be enhanced by alloying with a polyvalent metal. This brings more electrons into the conduction band. As the Fermi surface increases, it begins to cross Brillouin zone boundaries and other spin-hot-spot regions. The enhancement of $1/T_1$ can be significant. Similarly, $1/T_1$ can be reduced by orders of magnitude by alloying polyvalent metals with monovalent. Applying pressure, reducing the dimensionality, or doping into a semiconductor conduction bands as well as any other method of modifying the band structure should work. The rule of thumb for reducing $1/T_1$ is washing the spin hot spots off the Fermi surface. (Another possibility would be to inhibit scattering in or out the spin hot spots, but this is hardly realizable.)

The most important work ahead is to catalog $1/T_1$ for more metallic elements and alloys. So far only the simplest metals have been carefully studied over large enough temperature ranges, but even in these cases it is not clear, for example, as to how phonon-induced $1/T_1$ behaves at low temperatures. It is plausible that understanding $1/T_1$ in the transition metals will require new insights (such as establishing the role of the s-d exchange). Another exciting possibility is that the measurements at high enough temperatures will settle the question of the so-called "resistivity saturation"[60] which occurs in many transition metals. Indeed, the two competing models of this phenomenon imply different scenarios for $1/T_1$: the "phonon ineffectiveness" model[61] implies saturation of $1/T_1$, while the model emphasizing the role of quantum corrections to Boltzmann theory[62] apparently does not.[63] Finally, theory should yield probabilities of various spin-flip processes in different metals. Empirical pseudopotential and density functional techniques seem quite adequate to perform such calculations. Some work in this direction is already under way.[64]

V. CONCLUSION

We have provided a brief informal review of the current understanding of spin relaxation phenomenon in metals and semiconductors. Although studying spin relaxation through electron spin resonance measurements and developing its microscopic understanding through quantitative band structure analyses were among the more active early research areas in solid state physics (dating back to the early 1950s), it is surprising that our current understanding of the phenomenon is quite incomplete and is restricted mostly to bulk elemental metals and some of the III–V semiconductor materials (both bulk and quantum well systems). There is a great deal of renewed current interest in the subject because of the potential spintronics applications offering the highly desirable possibility of monolithic integration of electronic, magnetic, and optical devices in single chips as well as the exciting prospect of using spin as a quantum bit in proposed quantum computer architectures. It should, however, be emphasized that all of these proposed applications necessarily require comprehensive quantitative understanding of physical processes controlling spin coherence in electronic materials. In particular, there is an acute need to develop techniques which can manipulate spin dynamics in a controlled coherent way which necessitates having long spin relaxation times and/or spin diffusion lengths. Our understanding of spin coherence in small mesoscopic systems and more importantly, at or across interfaces (metal/semiconductor, semiconductor/semiconductor) is currently rudimentary to nonexistent. Much work (both theoretical and experimental as well as materials and fabrication related) is needed to develop a comprehensive understanding of spin coherence in electronic materials before the spintronics dream can become a viable reality.

ACKNOWLEDGMENTS

This work is supported by the U.S. ONR and the DOD. The authors thank P. B. Allen and M. Johnson for useful discussions.

[1]G. Prinz, Phys. Today **48**, 58 (1995); Science **282**, 1660 (1998).
[2]M. Johnson, Science **260**, 320 (1993); J. Magn. Magn. Mater. **140-144**, 21 (1995); **156**, 321 (1996); Mater. Sci. Eng., B **31**, 199 (1995).
[3]*Optical Orientation, Modern Problems in Condensed Matter Science*, edited by F. Meier and B. P. Zakharchenya (North-Holland, Amsterdam, 1984), Vol. 8.
[4]J-Ph. Ansermet, J. Phys.: Condens. Matter **10**, 6027 (1998).
[5]P. B. Allen, Solid State Commun. **102**, 127 (1997).
[6]J. Bond, Solid State Technol. **36**, 39 (1993).
[7]J. L. Simonds, Phys. Today **48**, 26 (1995).
[8]J. Gregg, W. Allen, N. Viart, R. Kirschman, Ch. Sirisathitkul, J.-P. Shille, M. Gester, S. Thompson, P. Sparks, V. Da Costa, K. Ounadjela, and M. Skvarla, J. Magn. Magn. Mater. **175**, 1 (1997).
[9]R. J. Elliott, Phys. Rev. **96**, 266 (1954).
[10]A. W. Overhauser, Phys. Rev. **89**, 689 (1953).
[11]Y. Yafet, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1963), Vol. 14.
[12]P. Boguslawski, Solid State Commun. **33**, 389 (1980).
[13]M. I. D'yakonov and V. I. Perel', Sov. Phys. JETP **33**, 1053 (1971); Sov. Phys. Solid State **13**, 3023 (1972).
[14]C. Kittel, *Introduction to Solid State Physics*, 7th ed. (Wiley, New York, 1996).
[15]M. I. D'yakonov and V. I. Perel', Sov. Phys. JETP **38**, 177 (1973).
[16]E. L. Ivchenko, Sov. Phys. Solid State **15**, 1048 (1973).
[17]B. P. Zakharchenya, E. L. Ivchenko, A. Ya. Ryskin, and A. V. Varfolomeev, Fiz. Tverd. Tela (St. Petersburg) **18**, 230 (1976) [Sov. Phys. Solid State].

[18]G. L. Bir, A. G. Aronov, and G. E. Pikus, Sov. Phys. JETP **42**, 705 (1976).
[19]A. M. Portis, A. F. Kip, C. Kittel, and W. H. Brattain, Phys. Rev. **90**, 988 (1953).
[20]F. K. Willenbrock and N. Bloembergen, Phys. Rev. **91**, 1281 (1953).
[21]T. Uenoyama and L. J. Sham, Phys. Rev. Lett. **64**, 3070 (1990); Phys. Rev. B **42**, 7114 (1990).
[22]R. Ferreira and G. Bastard, Phys. Rev. B **43**, 9687 (1991).
[23]T. C. Damen, L. Viña, J. E. Cunningham, J. Shah, and L. J. Sham, Phys. Rev. Lett. **67**, 3432 (1991).
[24]L. J. Sham, J. Phys.: Condens. Matter **5**, A51 (1993).
[25]K. Zerrouati, F. Fabre, G. Bacquet, J. Bandet, J. Frandon, G. Lampel, and D. Paget, Phys. Rev. B **37**, 1334 (1988).
[26]J.-N. Chazalviel, Phys. Rev. B **11**, 1555 (1975).
[27]A. H. Clark, R. D. Burnham, D. J. Chadi, and R. M. White, Solid State Commun. **20**, 385 (1976).
[28]A. H. Clark, R. D. Burnham, D. J. Chadi, and R. M. White, Phys. Rev. B **12**, 5758 (1975).
[29]V. I. Maruschak, T. V. Lagunova, M. N. Stepanova, and A. N. Titkov, Fiz. Tverd. Tela (St. Petersburg) **25**, 2140 (1983) [Sov. Phys. Solid State].
[30]G. Fishman and G. Lampel, Phys. Rev. B **16**, 820 (1977).
[31]A. G. Aronov, G. E. Pikus, and A. N. Titkov, Sov. Phys. JETP **57**, 680 (1983).
[32]G. E. Pikus and A. N. Titkov, in *Optical Orientation, Modern Problems in Condensed Matter Science*, edited by F. Meier and B. P. Zakharchenya (North-Holland, Amsterdam, 1984), Vol. 8, p. 73.
[33]G. E. Pikus, V. A. Marushchak, and A. N. Titkov, Sov. Phys. Semicond. **22**, 115 (1988).
[34]M. Z. Maialle and M. H. Degani, Appl. Phys. Lett. **70**, 1864 (1997); Phys. Rev. B **55**, 13 771 (1997).
[35]G. Bastard and R. Ferreira, Surf. Sci. **267**, 335 (1992).
[36]R. S. Britton, T. Grevatt, A. Malinowski, R. T. Harley, P. Perozzo, A. R. Cameron, and A. Miller, Appl. Phys. Lett. **73**, 2140 (1998).
[37]P. D. Dresselhaus, C. M. A. Papavassiliou, R. G. Wheeler, and R. N. Sacks, Phys. Rev. Lett. **68**, 106 (1992).
[38]G. Lommer, F. Malcher, and U. Rössler, Phys. Rev. Lett. **60**, 728 (1988).
[39]T. Guettler, A. L. Triques, L. Vervoort, R. Ferreira, Ph. Roussignol, P. Voisin, D. Rondi, and J. C. Harmand, Phys. Rev. B **58**, R10179 (1998).
[40]L. Vervoort, R. Ferreira, and P. Voisin, Phys. Rev. B **56**, R12744 (1997).
[41]J. Wagner, H. Schneider, D. Richards, A. Fischer, and K. Ploog, Phys. Rev. B **47**, 4786 (1993).
[42]J. M. Kikkawa, I. P. Smorchkova, N. Samarth, and D. D. Awschalom, Science **277**, 1284 (1997).
[43]J. M. Kikkawa and D. D. Awschalom, Phys. Rev. Lett. **80**, 4313 (1998).
[44]N. N. Kuzma, P. Khandelwal, S. E. Barrett, L. N. Pfeiffer, and K. W. West, Science **281**, 686 (1998).
[45]D. Hägele, M. Oestreich, W. W. Rühle, N. Nestle, and K. Eberl, Appl. Phys. Lett. **73**, 1580 (1998); J. M. Kikkawa and D. D. Awschalom, Nature (London) **397**, 139 (1999).
[46]T. W. Griswold, A. F. Kip, and C. Kittel, Phys. Rev. **88**, 951 (1952).
[47]G. Feher and A. F. Kip, Phys. Rev. **98**, 337 (1955).
[48]W. Kolbe, Phys. Rev. B **3**, 320 (1971).
[49]P. Monod and F. Beuneu, Phys. Rev. B **19**, 911 (1979); F. Beuneu and P. Monod, *ibid.* **18**, 2422 (1978).
[50]A. G. Aronov, Sov. Phys. JETP **24**, 32 (1976).
[51]M. Johnson and R. H. Silsbee, Phys. Rev. Lett. **55**, 1790 (1985); Phys. Rev. B **37**, 5326 (1988).
[52]M. Johnson, Phys. Rev. Lett. **70**, 2142 (1993); J. Appl. Phys. **75**, 6714 (1994); A. Y. Elezzabi, M. R. Freeman, and M. Johnson, Phys. Rev. Lett. **77**, 3220 (1996).
[53]M. Johnson, Appl. Phys. Lett. **65**, 1460 (1994); J. Magn. Magn. Mater. **148**, 349 (1995).
[54]V. A. Vas'ko, V. A. Larkin, P. A. Kraus, K. R. Nikolaev, D. E. Grupp, C. A. Nordam, and A. M. Goldman, Phys. Rev. Lett. **78**, 1134 (1997).
[55]Z. W. Dong, R. Ramesh, T. Venkatesan, M. Johnson, Z. Y. Chen, S. P. Pai, V. Talyansky, R. P. Sharma, R. Shreekala, C. J. Lobb, and R. L. Greene, Appl. Phys. Lett. **71**, 1718 (1997).
[56]A. G. Aronov and G. E. Pikus, Sov. Phys. Semicond. **10**, 698 (1976).
[57]J. Fabian and S. Das Sarma, Phys. Rev. Lett. **81**, 5624 (1998).
[58]R. H. Silsbee and F. Beuneu, Phys. Rev. B **27**, 2682 (1983).
[59]J. Fabian and S. Das Sarma, J. Appl. Phys. **85**, 5075 (1999).
[60]Z. Fisk and G. W. Webb, Phys. Rev. Lett. **36**, 1084 (1976).
[61]P. J. Cote and L. V. Maisel, Phys. Rev. Lett. **40**, 1586 (1978).
[62]B. Chakraborty and P. B. Allen, Phys. Rev. Lett. **42**, 736 (1979); P. B. Allen and B. Chakraborty, Phys. Rev. B **23**, 4815 (1981).
[63]P. B. Allen (private communication).
[64]J. Fabian and S. Das Sarma (unpublished).

Kinetics of MnAs growth on GaAs(001) and interface structure

F. Schippan, A. Trampert, L. Däweritz,[a] and K. H. Ploog
Paul-Drude-Institut für Festkörperelektronik, Hausvogteiplatz 5-7, D-10117 Berlin, Germany

(Received 17 January 1999; accepted 30 April 1999)

On different As-rich GaAs(001) templates, well characterized by reflectance difference spectroscopy, nucleation and growth of NiAs-type MnAs is investigated in real time by reflection high-energy electron diffraction. Using very high As_4/Mn flux ratios and low growth rates, one of the two occurring azimuthal alignments of the $(\bar{1}100)$ orientation can be nearly suppressed even in the nucleation stage, and it vanishes completely with further growth. Annealing is found to be very effective in surface smoothing. In dependence on the As/Mn ratio the MnAs$(\bar{1}100)$ surface develops different reconstructions. This finding is important for further investigations in the growth of double heterostructures. High-resolution transmission electron microscopy of as-grown MnAs/GaAs samples reveals an abrupt interface. The lattice mismatch accommodation is anisotropic with regularly arranged misfit dislocations along the $[\bar{1}10]$ direction and less localized coherency strain in the [110] direction, consistent with a near-coincidence-site lattice model. © *1999 American Vacuum Society.* [S0734-211X(99)02904-2]

I. INTRODUCTION

The heteroepitaxial growth of magnetic layers on semiconducting substrates has drawn considerable attention because of its potential to develop novel device structures.[1] MnAs as a promising candidate satisfies material requirements for growth of such structures by molecular beam epitaxy (MBE) on GaAs(001) and it is interesting as a model system for investigating the epitaxy of dissimilar materials.[2,3] Although several MnAs phases exist[4]—paramagnetic NiAs-type γMnAs above 125 °C, orthorhombic MnP-type βMnAs between 125 and 45 °C, and ferromagnetic NiAs-type αMnAs below 45 °C—and the lattice mismatch between MnAs and GaAs is extremely large, epitaxial ferromagnetic αMnAs films have been successfully grown on GaAs(001).[2,3,5] The different epitaxial orientations found are related to template effects.[3] A problem inherent in MBE growth of lattice-mismatched systems is three-dimensional (3D) islanding. For αMnAs films it has been shown that surface roughening due to 3D island growth can be reduced by increasing the As_4:Mn beam equivalent pressure (BEP) ratio and decreasing the growth rate.[6] MBE growth in the presence of atomic hydrogen has also been applied to improve the surface morphology.[6]

It is obvious, that for successful growth of MnAs/GaAs heterostructures an exact control of the epitaxial orientation during nucleation and growth is essential. A single epitaxial orientation and high surface smoothness are a prerequisite for the development of structurally perfect growth faces which eventually reconstruct. The different structure and the different atomic bonding of the materials involved in the envisaged heterostructures lead to interfaces which are of particular importance because they do not only determine the epitaxial alignment, but also the way of lattice mismatch accommodation and, additionally, the character and density of extended defects in the heterosystem.

In this work, we study the nucleation and growth of MnAs on different GaAs(001) templates that are well characterized regarding structure and stoichiometry by reflectance difference spectroscopy (RDS). Reflection high-energy electron diffraction (RHEED) is applied to monitor the evolution of differently oriented domains with film thickness, in the dependence on the template structure and growth conditions. On surfaces of high smoothness the MnAs$(\bar{1}100)$ reconstruction is investigated as function of the As_4/Mn ratio at static and dynamic conditions. High-resolution transmission electron microscopy (HRTEM) is used to analyze the structure of the as-grown interface.

II. EXPERIMENT

The MnAs layers were grown by solid-source MBE on semi-insulating GaAs(001). After growing a GaAs buffer at a substrate temperature T_s of 550 °C the substrate was cooled down to 250 °C. The As_4 shutter was closed at $T_s \sim 500$ °C. To prepare different GaAs(001) templates with $d(4\times4)$, $c(4\times4)$, and (2×4) reconstruction, the As_4 BEP during cooling was either kept constant at 2×10^{-6} Torr, as used during growth, or reduced to 3×10^{-7} Torr. MnAs was grown at growth rates between 5 and 20 nm h^{-1}. The Mn flux was calibrated by secondary ion mass spectrometry and by cross-sectional thickness measurements by HRTEM.

The template structure was characterized by RHEED and RDS. The recorded signal in the RDS measurements is the real part of the RD:

$$\mathrm{Re}(\Delta r/r) = 2(r_{[\bar{1}10]} - r_{[110]})/(r_{[110]} + r_{[\bar{1}10]}), \quad (1)$$

where $r_{[\bar{1}10]}$ and $r_{[110]}$ are the near-normal-incidence reflectances for light polarized parallel to the principal $\langle 110\rangle$ axes of GaAs.

To monitor the MnAs nucleation and growth by RHEED, a system consisting of a CCD camera, video recorder, and

[a]Author to whom correspondence should be addressed; electronic mail: daweritz@pdi-berlin.de

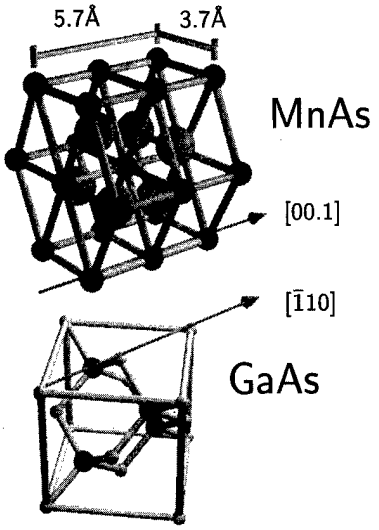


FIG. 1. Epitaxial relationship between NiAs-type MnAs($\bar{1}$100) and GaAs(001) in the orientation A. From bulk lattice data the (1×1) surface unit mesh of MnAs($\bar{1}$100) has a size of 5.7×3.7 Å.

image processing was used. The electron beam was in the GaAs[110] direction which is the most instructive azimuth in these measurements.

III. RESULTS AND DISCUSSION

A. Nucleation and growth

By analyzing RHEED patterns taken in the [110]GaAs and [$\bar{1}$10] GaAs azimuths during MnAs growth, we find as principal epitaxial orientation for NiAs-type MnAs on very As-rich GaAs(001) surfaces:

($\bar{1}$100)MnAs‖(001)GaAs and [0001]MnAs‖[$\bar{1}$10]GaAs

(orientation A).

This epitaxial relationship is schematically represented in Fig. 1. Depending on the GaAs substrate surface stoichiometry, the MnAs growth conditions and layer thickness, a simultaneous formation of domains with 90° azimuthal rotation can occur as a second orientation of minor importance:

($\bar{1}$100)MnAs‖(001)GaAs and [0001]MnAs‖[110]GaAs

(orientation B).

The MnAs layer formation was analyzed in detail by recording RHEED linescans (indicated at the top of Fig. 2), with the incident electron beam along [110]GaAs, and plotting them as a function of layer thickness. Figure 2 presents such plots for MnAs growth on different GaAs(001) templates resulting from different As_4 pressures at a constant substrate temperature of 250 °C. These templates are characterized by their RDS spectra (right-hand side column). Since the As_4 pressure was held constant before and during the supply of Mn atoms, the experiments considered in Fig. 2 differ, however, not only concerning the GaAs template but also in the

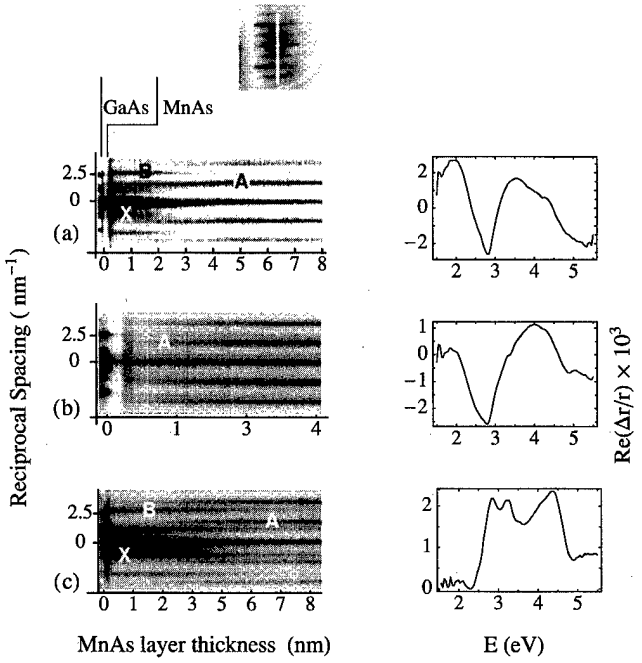


FIG. 2. RHEED linescan (shown at the top) and its plot as function of layer thickness showing the evolution of the A and B orientation during MnAs growth on different GaAs(001) templates, electron beam along [110]GaAs (left-hand side column), and RDS spectra of the template (right-hand side column). Growth conditions: (a) $c(4\times4)$ template, $R=9$ nm h^{-1}, BEP ratio =14; (b) $d(4\times4)$ template, $R=5$ nm h^{-1}, BEP ratio=50; (c) (2×4) template, $R=20$ nm h^{-1}, BEP ratio=5.

MnAs growth conditions (growth rate, As_4:Mn BEP ratio), i.e., we cannot clearly distinguish between template and growth condition related effects.

The RDS spectrum shown in Fig. 2(a) is typical for a substrate temperature of 250 °C and an As_4 pressure of 2×10^{-6} Torr and corresponds to a $c(4\times4)$ reconstruction (cf. Ref. 7). Exposing the surface to the Mn flux, the GaAs diffraction spots in the RHEED plot disappear immediately due to the random incorporation of Mn atoms. After deposition of less than 0.3 nm MnAs, corresponding roughly to 1 ML MnAs, new diffraction spots appear. The measured interplanar spacings d of 0.58 nm (labeled A) and 0.37 nm (B) relate to the epitaxial orientations A and B, respectively. The additionally observed d value of 0.84 nm (X) is slightly higher than the GaAs(001)-$c(4\times4)$ unit mesh dimension. Therefore, it is tentatively ascribed to the substrate surface with bond stretching caused by Mn incorporation. The RHEED plot reveals that A- and B-oriented domains nucleate simultaneously. The orientation A dominates, however, after ~2 nm deposition and exists exclusively after ~7 nm deposition. The RHEED intensity of the X spot also vanishes at ~2 nm MnAs.

As revealed by the line shape of the RDS spectrum of Fig. 2(b), in this particular experiment with nominally the same substrate temperature and As_4 pressure as in the case considered previously, the template was more As-rich. The spectrum is typical for the $d(4\times4)$ reconstruction (cf. Ref. 7). In addition, the Mn flux was reduced, leading to a BEP ratio of 50 and a MnAs growth rate of 5 nm h^{-1} as compared to a

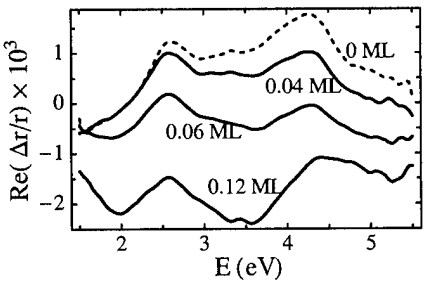


FIG. 3. RDS spectra taken at different Mn coverages for deposition on GaAs(001)-(2×4) at 550 °C. The coverage data are based on the assumption that the Mn atoms occupy one (001)Ga plane with 6.26×10^{14} atoms cm^{-2}.

BEP ratio of 14 and a growth rate of 9 nm h^{-1} in the first experiment. The RHEED plot suggests that under these conditions the MnAs layer grows (nearly) exclusively in the A orientation from the very beginning.

The effect of a less As-rich template on the epitaxial growth was investigated in a third experiment [Fig. 2(c)] with a reduced As_4 pressure of 3×10^{-7} Torr. Additionally, the Mn flux was increased, leading to a BEP ratio of 5 and a MnAs growth rate of 20 nm h^{-1}. Under these conditions, the RHEED pattern of the template indicates a reconstruction slightly beyond the $c(4\times4)\rightarrow(2\times4)$ transition. In contrast to the foregoing experiments, the feature at ~2.8 eV in the RDS spectrum now has a positive sign, indicating that the surface is terminated by a single layer of As dimers oriented along $[\bar{1}10]$. The concentration of As dimers with [110] orientation atop this layer, typical for the $c(4\times4)$ and $d(4\times4)$ structure,[7] is at least drastically reduced if not completely removed. On such a template, the B orientation of MnAs clearly dominates in the nucleation stage. The contribution of B-type domains to the surface and simultaneously the RHEED intensity of the diffraction spot X only slowly reduce with layer thickness, although finally, the orientation A becomes dominating. The persistence of the X-spot intensity suggests that the formation of a closed film is delayed by the coexistence of A and B domains.

As demonstrated, a very high As coverage of the substrate and a high As_4/Mn BEP ratio are necessary to avoid (nearly) completely the formation of B-type domains in the nucleation stage. A possible explanation for this is that during Mn incorporation a local desorption of As atoms occurs, as is observed during deposition of Si atoms on GaAs(001)-(2×4).[8] To prove this assumption, Mn was deposited with a flux of 2×10^{11} atoms cm^{-2} s^{-1} on GaAs(001)-(2×4) at a higher substrate temperature of 550 °C and an As_4 pressure of 1×10^{-6} Torr. Indeed, the RDS spectrum (Fig. 3) shows the typical behavior expected for such a process that is similar to the transition from the As-terminated (2×4) to the Ga-terminated (4×2) reconstruction.[7] The spectrum shifts as a whole to a negative level and develops a negative feature at ~2 eV which is related to the existence of Ga dimers. Summarizing, these experiments suggest that for optimizing the MnAs growth on GaAs a two-stage approach with different As_4 pressure during nucleation and growth should be applied.

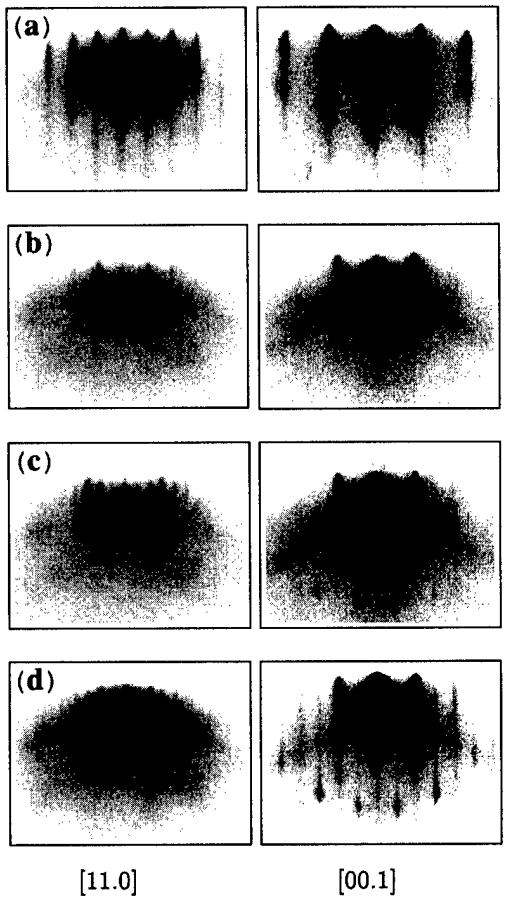


FIG. 4. RHEED patterns taken in the $[11\bar{2}0]$ and [0001] azimuths of differently reconstructed static MnAs$(\bar{1}100)$ surfaces, (a) as-grown (1×1) structure after growth at 250 °C, Mn shutter closed, As shutter open; (b) (1×1) structure after 10 min sample annealing at 400 °C under constant As_4 flux; (c) (2×1) reconstruction after As_4 flux interruption, substrate temperature 400 °C, (d) (2×2) reconstruction after Mn deposition at 350 °C. The As_4 and Mn fluxes applied in these experiments correspond to a BEP ratio of 90.

B. Surface reconstruction

The starting surface to study the reconstruction under nongrowing conditions was a 50 nm thick, type-A oriented MnAs film, grown on GaAs(001)-$c(4\times4)$ at a substrate temperature of 250 °C and an As_4:Mn BEP ratio of about 90. This high BEP ratio was chosen to avoid the formation of B-type domains. Figure 4(a) shows RHEED patterns of the as-grown surface after closing the Mn shutter with the As shutter open. The distances of the RHEED streaks in the orthogonal $[11\bar{2}0]$MnAs and [0001]MnAs azimuths correspond to a surface unit mesh with the lattice parameters of 5.8±0.1 and 3.8±0.1 Å, respectively. In the following, this surface unit mesh indicated schematically in Fig. 1 is referred to a (1×1) structure. Figure 5 shows typical AFM images observed for an as-grown surface and for a surface after annealing of the sample at 400 °C under constant As_4 flux, respectively, corresponding to the RHEED patterns shown in Figs. 4(a) and 4(b). The image of the as-grown film reveals 3D islands with a root-mean-square (rms) roughness of 14 Å, measured over a lateral scale of 2.5 μm. The surface morphology improves remarkably during annealing with a

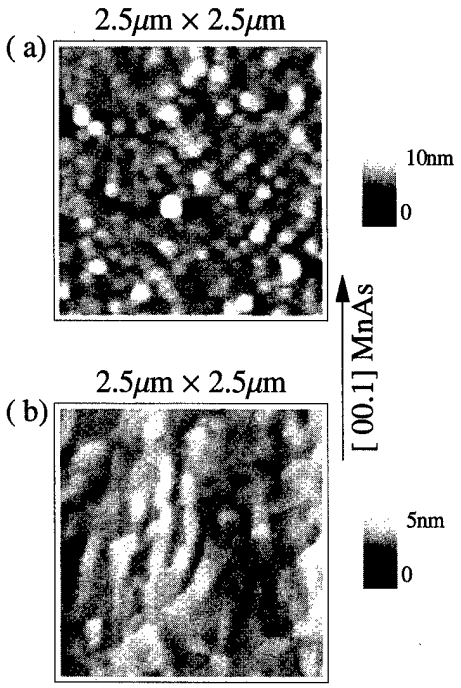


FIG. 5. AFM images of (a) the as-grown MnAs($\bar{1}100$) surface (250 °C, As_4 BEP=5×10^{-6} Torr, 20 nm h^{-1}, BEP ratio=90) and (b) the surface after an additional annealing at 400 °C, annealing time 10 min.

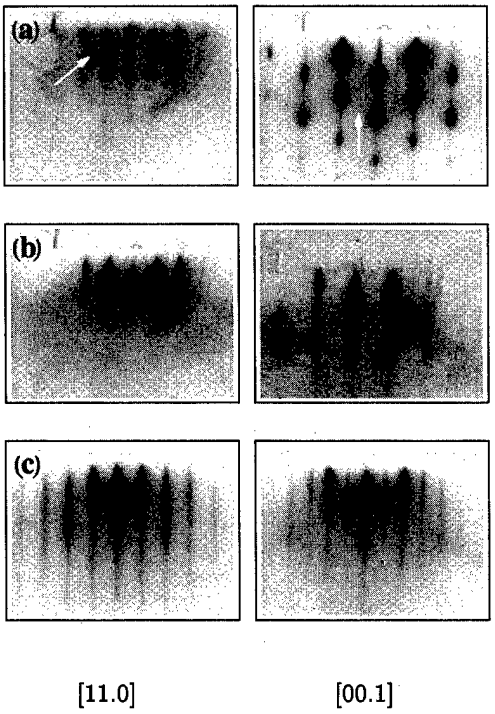


FIG. 6. RHEED patterns taken in the [11$\bar{2}$0] and [0001] azimuths of the reconstructed MnAs($\bar{1}$10) surface during growth. (a) (1×2) reconstruction during buffer layer growth with a relatively rough surface; (b) (1×1) structure of the smoothed surface of the annealed buffer layer; (c) (1×2) reconstruction during growth on a well-annealed, smooth surface. Growth parameters: 250 °C, As_4=5×10^{-6} Torr, 20 nm h^{-1}, BEP ratio=90. The arrows in (a) indicate additional spots occasionally observed on the rough buffer layer (left-hand side) and faint half-order streaks (right-hand side).

reduction of the rms roughness to 5.5 Å after 10 min. This smoothing is clearly reflected in the RHEED patterns [Fig. 4(b)] by the loss of the transmission-like features. The (1×1) RHEED symmetry is maintained during this procedure. After interrupting the As_4 flux, however, half-order spots are observed in the [11$\bar{2}$0] azimuth whereas the periodicity observed in the [0001] azimuth is unchanged or eventually doubled [Fig. 4(c)]. This phase transition from the (1×1) to the (2×1) [or (2×2)] structure is reversible when the As_4 flux is reinstated. It also occurs when the substrate temperature is reduced down to 350 °C. Whereas the (2×1)→(1×1) transition after As_4 exposure is abrupt over the whole investigated substrate temperature range between 400 and 350 °C, the reverse (1×1)→(2×1) transition after As_4 flux interruption becomes increasingly delayed and incomplete with temperature reduction. At 350 °C only weak half-order spots are observed with a delay time of about 2 min. This clearly shows that the (1×1)→(2×1) transition is related to As desorption, which is slowed down at reduced substrate temperature due to the high As_4 background pressure in the growth chamber, even after closing the As shutter. However, the As adsorption occurs rapidly after opening of the As shutter, leading to the abrupt (2×1)→(1×1) transition.

If the (2×1) reconstructed surface is exposed to the Mn flux, with the As shutter closed, a (2×2) reconstruction develops [Fig. 4(d)]. The delay time of the appearance of the new half-order spots in the [0001] azimuth is about 10 s for a substrate temperature of 350 °C and increases to about 45 s when the substrate temperature is reduced to 325 °C. Again, this can be explained by the slow As desorption due to the high As_4 background pressure in the growth chamber.

After having discussed the static MnAs($\bar{1}100$) surface we now focus on the growing surface. MnAs growth under typical conditions (250 °C, As_4=5×10^{-6} Torr, As_4:Mn BEP =90, R=20 nm/h) leads to a (1×2) reconstruction. This structure is well developed when, a buffer layer is first grown which is smoothed by 10 min annealing at 400 °C. The RHEED patterns presented in Fig. 6(c) for MnAs growth on a 50 nm thick annealed buffer layer are streaky in both azimuths and show clear half-order spots in the [0001] azimuth. Although less ordered, this (1×2) structure already exists during buffer layer growth. However, for thicker layers the surface can become relatively rough, as demonstrated in Fig. 6(a), showing RHEED patterns of a 60 nm thick buffer layer. The patterns are spotty in both azimuths and only faint half-order streaks are observed in the [0001] azimuth. The RHEED patterns shown in Fig. 6(b) for the same surface after annealing evidence that the annealing procedure is very effective in improving the surface morphology and structure. Even if the RHEED pattern of the initial surface shows some additional spots in the [11$\bar{2}$0] azimuth [Fig. 6(a), left-hand side], as occasionally observed, the streaky pattern after annealing evidences a smooth template with (1×1) structure that is maintained during cooling to the growth temperature of 250 °C.

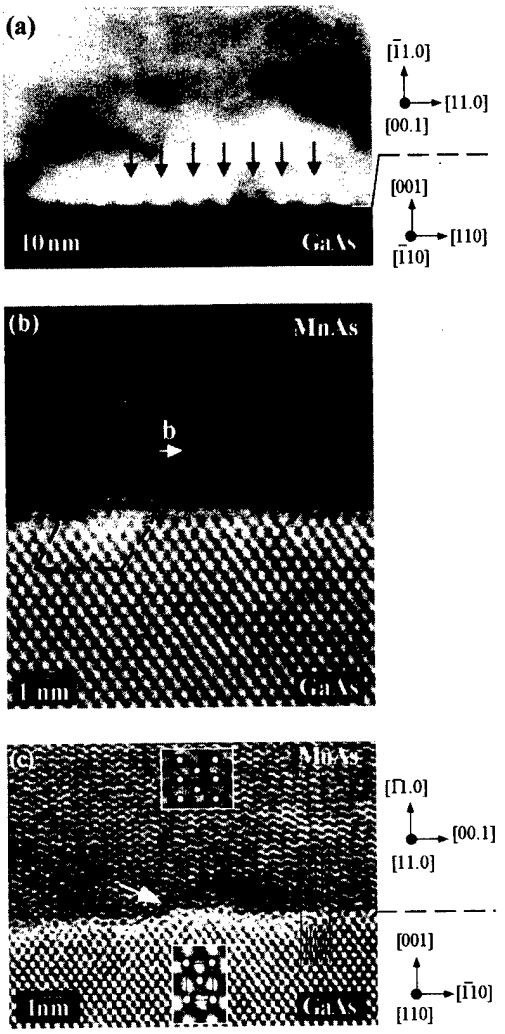


FIG. 7. Cross-sectional dark field micrograph (a) and the HRTEM image (b) of the MnAs/GaAs interface with the incident beam parallel to [0001]MnAs ‖ [1̄10]GaAs. Note the array of periodic strain contrast along the interface (arrows) in (a) and the Burgers circuit around a dislocation core in (b); (c) shows a cross-sectional HRTEM image of the heterostructure in [11$\bar{2}$0]MnAs ‖ [110]GaAs projection. Three {220}GaAs planes fit to two {0002}MnAs planes. The interface appears structurally abrupt with monoatomic steps (arrow). The inset shows magnified parts of the image including models of the underlying crystal lattices.

C. Interface structure

In spite of the marked difference in both the crystal symmetry and the lattice constants of MnAs and GaAs, MnAs grows epitaxially. Because of the symmetry breaking, however, the structure of the interface is more complicated. In the following we analyze the interface of the sample for which the real-time RHEED data are presented in Fig. 2(a). Figure 7(a) shows a cross-sectional dark field micrograph of the MnAs/GaAs heterostructure imaged along the [0001]MnAs ‖ [1$\bar{1}$0]GaAs direction for which the lattice misfit is ~8%. The abrupt change in contrast reflects a smooth and chemically sharp boundary with no indication of an extended interfacial phase. A periodic array of strain contrast features along the interface (marked by arrows) is clearly observed as expected for an array of misfit dislocations. In fact, the high-resolution image [Fig. 7(b)] confirms the semicoherent description of the interface where regions of lattice matching are separated by localized misfit dislocations, which are characterized by a strong lattice plane bending perpendicular and parallel to the interface. A Burgers circuit around such a dislocation core determines the Burgers vector to $\mathbf{b}=1/3[11\bar{2}0]$, a typical lattice dislocation in hexagonal materials. This Burgers vector is located parallel to the boundary plane and, therefore, most efficient in strain relief. However, by measuring the mean distance D between the dislocations, a residual strain ϵ in the epilayers is calculated if applying the equation:

$$\epsilon = f_0 - \frac{b}{D}, \tag{2}$$

where f_0 defines the lattice misfit. Assuming that the mismatch between the MnAs layer and the GaAs substrate is completely relaxed at the growth temperature, a residual compressive strain arises during sample cooling to room temperature, mainly due to a discontinuous increase of the lattice constant a of about 1% at the ferromagnetic transition temperature $T_c \sim 45\,^{\circ}\mathrm{C}$. Magnetization measurements have proved the films to be ferromagnetic.[5]

Figure 7(c) shows a HRTEM micrograph along the GaAs[110] projection, where the lattice mismatch is about 33% between the {220}GaAs and the {0002}MnAs planes. The HRTEM contrast of the MnAs lattice imaged in the [11$\bar{2}$0] direction is wavy-like with a period corresponding to the hexagonal lattice constant c in agreement with image simulations. The interface appears atomically abrupt and steps of monolayer height are observed. No localized misfit dislocation or strong coherence strain features are visible in this image. At first glance, the interface appears completely incoherent as expected for heterosystems with a large lattice mismatch and weak interfacial bond strength. However, a more careful inspection reveals an interface structure, where the mismatch accommodation becomes plausible by employing the near-coincidence-site lattice model: every fourth {0002}MnAs plane passes into every sixth {220}GaAs plane forming a "geometric" misfit dislocation. This 4/6-ratio reduces the actual lattice mismatch to about 5%, a reasonable value to guarantee epitaxial growth.

It should to be noted that the HRTEM investigation did not reveal an interfacial layer with different epitaxial orientations, as deduced from our *in situ* experiments. In view of the above demonstrated annealing effects, we conclude that such an interlayer must already be rearranged during extended growth or by phase transitions during cooling.

IV. SUMMARY AND CONCLUSIONS

The GaAs(001)/MnAs interface formation and MnAs layer growth were studied in detail by characterizing the GaAs template by RDS and monitoring MnAs nucleation and growth by *in situ* RHEED. At 250 °C and As-rich conditions MnAs grows with NiAs structure in ($\bar{1}$100) orientation on GaAs(001)-$c(4\times4)$ with the hexagonal c-axis parallel to [$\bar{1}$10]GaAs (orientation A) or parallel to [110]GaAs (orienta-

tion B. Real-time measurements show that the relative probability of nucleation in type A and in B orientation can be varied in a controlled manner. It drastically increases in favor of the A orientation when the As coverage of the GaAs template and the As_4:Mn BEP ratio are increased. The formation of type-B domains in the MnAs nucleation stage at moderate BEP ratios, which is probably due to As desorption from the substrate during incorporation of Mn atoms, vanishes after the growth of several ML MnAs. The surface, which roughens at a layer thickness of several 10 nm, can be very effectively smoothed by annealing at temperatures up to 400 °C. Interestingly, such surfaces develop different reconstructions at static as well as dynamic conditions. In analogy to the compound semiconductors they depend on the As_4/Mn ratio. This is very promising for future work since the surface reconstruction can be used as system-independent phenomena to reproduce growth conditions and to control the quality of surfaces, interfaces, and bulk layers.[9,10]

As shown by HRTEM, the differences between MnAs and GaAs in crystal symmetry and lattice constant lead to an anisotropic lattice mismatch accommodation. Along the $[\bar{1}10]$ direction the misfit strain is relieved by regularly arranged misfit dislocations, whereas along the perpendicular interface direction, the four times larger value of the lattice misfit produces no localized misfit strain. The interfacial atomic arrangement is explained by a near-coincidence-site lattice model.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the technical support in the growth experiments by P. Schützendübe. We thank the Forschungszentrum Jülich for the use of their microscopic facilities.

[1]G. A. Prinz, Science **250**, 1092 (1990).
[2]M. Tanaka, J. P. Harbison, M. C. Park, Y. S. Park, T. Shin, and G. M. Rothberg, J. Appl. Phys. **76**, 6278 (1994).
[3]M. Tanaka, J. P. Harbison, and G. M. Rothberg, J. Cryst. Growth **150**, 1132 (1995).
[4]*Binary Alloy Phase Diagrams*, 2nd ed., edited by T. B. Massalski, H. Okamoto, P. R. Subramanian, and L. Kacprzak (American Society of Metals, Metals Park, OH, 1990), Vol. 1, p. 293.
[5]F. Schippan, A. Trampert, L. Däweritz, K. H. Ploog, B. Dennis, K. U. Neumann, and K. R. A. Ziebeck, J. Cryst. Growth **201/202**, 674 (1999).
[6]Y. Morishita, K. Iida, A. Tsuboi, H. Taniguchi, and K. Sato, J. Cryst. Growth **187**, 228 (1998).
[7]I. Kamiya, D. E. Aspnes, L. T. Florez, and J. P. Harbison, Phys. Rev. B **46**, 15894 (1992), and references given therein.
[8]L. Däweritz, K. Stahrenberg, P. Schützendübe, J.-T. Zettler, W. Richter, and K. H. Ploog, J. Cryst. Growth **175/176**, 310 (1997).
[9]L. Däweritz and R. Hey, Surf. Sci. **236**, 15 (1990), and references given therein.
[10]O. H. Hughes, D. Korakakis, T. S. Cheng, A. V. Blant, N. J. Jeffs, and C. T. Foxon, J. Vac. Sci. Technol. B **16**, 2237 (1998).

Reflectance difference spectroscopy and magneto-optical analysis of digital magnetic heterostructures

A. Bonanni,[a] G. Prechtl, W. Heiss, F. Schinagl, S. Holl, H. Krenn, and H. Sitter
Institute of Semiconductor Physics, Johannes Kepler University, A-4040 Linz, Austria

D. Stifter and K. Hingerl
Profactor GmbH, Wehrgrabengasse 5, A-4040 Steyr, Austria

(Received 19 January 1999; accepted 17 May 1999)

We report on the investigation of $CdTe/Cd_{0.75}Mg_{0.25}Te$ quantum wells grown by molecular-beam epitaxy and containing a digital distribution of magnetic MnTe fractional monolayers in the well. The influence of electronic confinement and of exciton primary localization on the formation of ferromagnetic ordering on the scale of small spin clusters (polarons) has been investigated by performing selective excitation of the exciton luminescence. The calculated temperature dependence of the free magnetic polarons binding energy F is found to qualitatively agree with the photoluminescence excitation data and the ferromagnetic clusters result to be more stable in these magnetic structures of reduced dimensionality, with respect to the three-dimensional case. The magneto-optical Kerr effect is discussed in the case of reflectance difference spectroscopy (RDS) and from the angular dependence of the RDS spectra in magnetic field, signals due to interband Faraday rotation and signals originating from crystal anisotropy could be distinguished. The Faraday rotation in these structures is found to follow the Curie–Weiss law with a transition temperature of -6 K corresponding to the nominal ion spin concentration in the magnetic barriers. © *1999 American Vacuum Society.* [S0734-211X(99)07304-7]

I. INTRODUCTION

Diluted magnetic semiconductors (DMS), an ideal system for the analysis of magneto-optical effects, have lately attracted considerable attention.[1–3] The strong s,p-d exchange interaction between the spins of the extended band states and the localized magnetic ions gives rise to peculiar phenomena. The spin splitting of the electronic states is enhanced in an applied magnetic field or, equivalently, the difference between the refractive indices corresponding to the two circular polarizations of positive and negative heliticity is increased. This effect is particularly striking for DMS systems and the linearly polarized light passing through a film of these materials can show a considerable rotation of the polarization plane (Faraday effect).

An intense effort has been recently devoted to the fabrication of digital magnetic heterostructures (DMH)[4] in order to study the wide spectrum of tunable carrier-ion spin interactions, in comparison with the behavior of conventional diluted magnetic semiconductors. The insertion of two-dimensional (2D) magnetic monolayers (ML) in semiconductor quantum wells (QWs), enables one to investigate the rearrangement of the magnetic moments within the 2D planes and the effect on their magnetic interaction. In this way the Zeeman splitting and the spin-flip scattering rate can be varied over a wide range.

The strong exchange interaction between carrier spins and the spins of magnetic ions in semimagnetic semiconductors leads to a ferromagnetic alignment of the ion spins within the Bohr radius of a localized exciton and therefore to the formation of exciton magnetic polarons (EMP). The spin ordering produces a decrease in the energy of the carriers and related phenomena like the Stokes shift between the maximum of the photoluminescence (PL) spectrum and the maximum of the PL excitation line.

Stability criteria,[5] binding energies,[6] and formation dynamics[7,8] have been studied in a large variety of diluted magnetic semiconductors. EMP binding energy depends on the concentration of paramagnetic ions and on the conditions of the primary exciton localization in the QW.[9] A direct way to obtain the EMP energy is the selective excitation of excitons in the band of localized states[7] where the Stokes shift between the PL and the luminescence excitation energy peaks is considered as produced by the EMP binding energy.

In this work we report on magneto-optical and reflectance difference spectroscopy (RDS) analysis of DMHs.

II. EXPERIMENT

All the DMHs in our study were synthesized in a vertical molecular-beam epitaxy reactor equipped with *in situ* 20 keV reflection high-energy electron diffraction system for monitoring of the surface reconstructions during growth. Visible interferometry with He–Ne laser light has been applied in order to control the nucleation process. Single CdTe QWs embedded in $Cd_{0.75}Mg_{0.25}Te$ barriers, were grown lattice matched with $Cd_{0.96}Zn_{0.04}Te$ (001)-oriented substrates and in the well region equispaced MnTe layers have been digitally inserted (see inset of Fig. 7), resulting in 2D spin distributions.

We investigated the exact composition and thickness of the individual layers performing high-resolution x-ray dif-

[a]Corresponding author; electronic mail: alberta.bonanni@jk.uni-linz.ac.at

TABLE I. Description of the heterostructures and EMP binding energies experimentally obtained, in comparison with the values for alloy materials.

Sample	Well width (Å)	Digital layers in the QW (ML)	Total Mn concentration (%)	EMP binding energy (meV)	EMP binding energy[a] (meV)
DMH1	50	2×1/2	6	17	10[b]
DMH2	50	4×1/4	6	15	10[b]
DMH3	70	4×1	18	23	23
DMH4	70	4×1/4	4.5	0	0

[a]Diluted case from Ref. 7.
[b]Values for a 50 Å wide QW and $x=10\%$.

fraction measurements and comparing the results with simulated spectra. The obtained sample properties like well width, thickness and number of inserted MnTe layers, average Mn concentration, are summarized in Table I. In order to verify the digital distribution of the MnTe layers in the CdTe matrix, we performed conventional transmission electron microscopy and high-resolution transmission electron microscopy on a reference sample. For magneto-optical measurements, the samples were mounted in Faraday configuration in the bore of a superconducting split-coil magnet with fields up to 8 T and cooled to temperatures between 1.7 and 30 K. A tunable Coherent CR599 dye laser operating in the wavelength range 620–800 nm was used for PL and luminescence excitation experiments. A commercially available (ISA Jobin Yvon) system for RDS[10] has been utilized and spectra in the range between 1.5 and 5.5 eV are displayed in terms of the variation of polarization direction ($\Delta\theta$).

III. RESULTS AND DISCUSSION

A. Magnetic polarons in DMH

Binding energies of the EMP were determined from the Stokes shift between the laser energy and the luminescence maximum under selective excitation of localized excitons. Figure 1 shows PL and luminescence excitation spectra for sample DMH3 (see Table I). The maximum of the luminescence excitation spectrum corresponds to the energy of the free heavy-hole exciton. Under selective excitation, the PL is shifted toward lower energy by $\Delta E=23$ meV. Since this energy shift appears to be independent of the excitation energy, we ascribe it to the magnetic polaron formation.

The alignment of the Mn^{2+} spins in the case of an external magnetic field [Fig. 2(a)] and a decrease of the magnetic susceptibility with the increment of the temperature [Fig. 2(b)], lead to a suppression of ΔE and validate our assumption.[7] In Table I we summarize, together with the structural parameters, the EMP binding energies obtained from our experiments and, for comparison, experimental values for the diluted magnetic material $Cd_{1-x}Mn_xTe$[9] with the same nominal average Mn concentration as the DMH under study.

The formation of a stable EMP is essentially a question of balance between kinetic and carrier-ion exchange energy, therefore the dimension of the system plays a crucial role. A stability range larger than in the three-dimensional (3D) case, has been recently predicted for quasi-2D EMPs in confining structures.[11] Concerning the bound magnetic polarons formed from excitons bound to shallow impurities, the Coulomb potential of the impurity center provides the initial localization and MPs are detectable at low Mn content.[12] In the case of excitons not bound to impurities, however, the situation is considerably different: free magnetic polarons (FMP) are not stable in diluted materials and primary localization on the potential of the alloy and/or magnetic fluctuation is required. In the case of alloys, the polaron energy decreases and the formation time increases by reducing the concentration x of magnetic ions.

It has been shown[7] that the conditions for the formation of localized magnetic polarons in the 3D case are fulfilled for Mn contents higher than 10%: for concentrations below this threshold, the formation process is interrupted by exciton recombination and the lifetime of the exciton, shorter than the polaron formation time, hinders the direct observation of the shift due to magnetic polarons in PL experiments. A

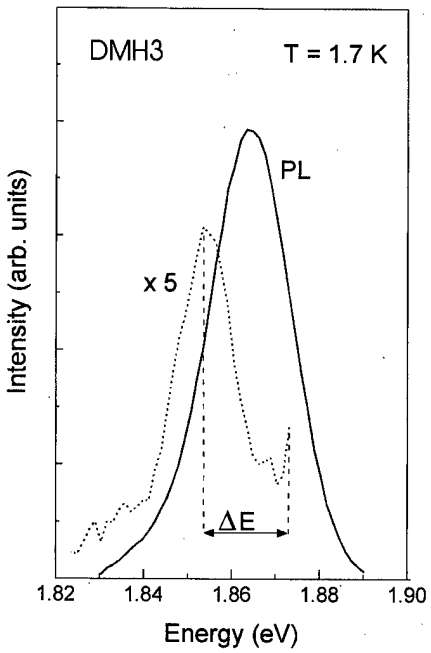


FIG. 1. PL and luminescence excitation spectra for sample DMH3. ΔE is the Stokes shift between laser energy and luminescence maximum under selective excitation.

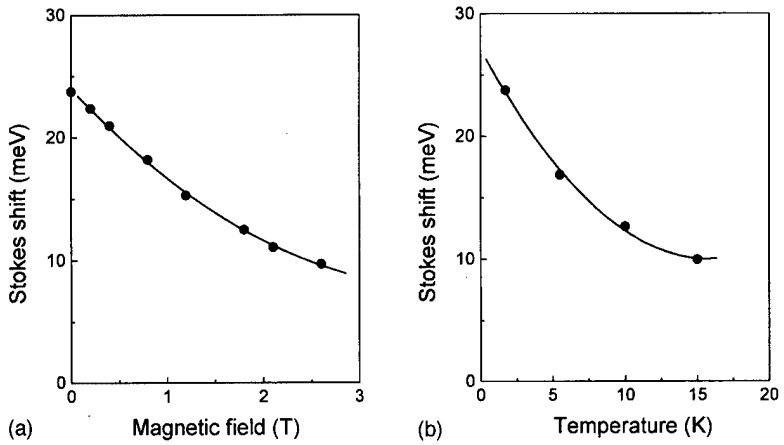


FIG. 2. External magnetic field (a) and temperature (b) dependence of the Stokes shift ΔE.

reduction of the dimensionality of the carrier system, however, enhances the stability of the EMPs, as theoretically predicted[11] and experimentally[13] shown for diluted magnetic semiconductors. Polaron binding energy, as a function of confinement, is expected first to increase as the electronic function becomes compressed, and subsequently to decrease as the effective magnetic concentration seen by the carriers diminishes with the thinning of the confining planes.[3]

Our results (see Table I) are in agreement with these assumptions: DMH samples with 70 Å QW behave, from the point of view of the EMP binding energy, like the equivalent alloy. Polaron formation at Mn concentrations below 10% is not observed (see sample DMH4) and at 18% we find a value of 23 meV for sample DMH3, in complete conformity with the data reported in literature[6] for a $Cd_{0.82}Mn_{0.28}Te$ well of the same width. Regarding the 50 Å QW, on the other hand, we observe the formation of EMP at Mn concentrations below 10% and higher values of the EMP binding energies (15 and 17 meV) in comparison with the diluted case (10 meV) obtained at the threshold value of 10%.

These effects are ascribable to the enhancement of MP stability related to the increment in the quantum confinement and to the effects of primary localization. By considering the same Mn concentration (samples DMH1 and DMH2), but a different arrangement in the magnetic layers (2×1/2 ML for DMH1 and 4×1/4 ML for DMH2), we can notice a decrease in the binding energy (from 17 meV for DMH1 to 15 meV for DMH2) with the thinning of the MnTe layers, due to the decreasing of effective magnetic concentration seen by the carriers.

B. Free energy of FMP

The presence of electrons in quantum confined states modifies, thus, in an effective sense the otherwise antiferromagnetic ion–ion interaction to turn over to a net small size ferromagnetism (polaron state). As mentioned earlier, the exchange interaction between extended electrons and waves and the magnetic ions in digital structures is direct and produces a redshift of the luminescence line which can be explained in terms of an enhanced binding energy of exciton coupled polarons or by means of the energetically preferred state of free magnetic polarons in contrast to plane wave states, respectively.

In order to describe the temperature dependence of the free energy of FMP, we visualize the system as a cloud of Mn spins, within the spatially separated MnTe layers in the DMHs, polarized by free carriers. We assume, then, a Hamiltonian with contact exchange interaction and, because of the fractional occupation of sites within the magnetic layers, we neglect ion–ion interactions.[14] These assumptions lead to

$$H = T + V(z) + \gamma \sum_i s S_i \delta(\bar{r} - \bar{R}_i), \tag{1}$$

with $T = -\hbar^2/2m^* \nabla^2$ the kinetic energy operator, the $V(z)$ the quantum well potential energy, according to Fig. 3 and γ the exchange constant between a free spin s at site $\bar{r}$ and the spatially localized Mn spins S_i at sites $\bar{R}_i$. Since the density of states of light holes is small, excitons can be considered as consisting of heavy holes and electrons solely and is possible

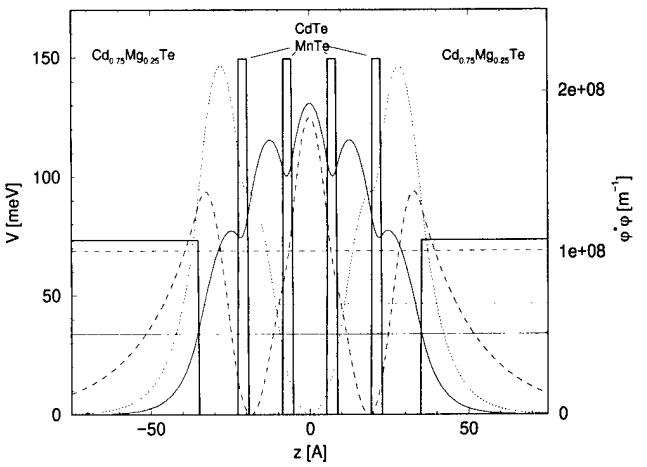


FIG. 3. CdTe valence band potential well with 4 MLs of MnTe embedded in $Cd_{0.75}Mg_{0.25}Te$ barriers. $\varphi^*(z)\varphi(z)$ and the corresponding energies of the ground (solid), first excited (dotted), and second excited (dashed) state are shown.

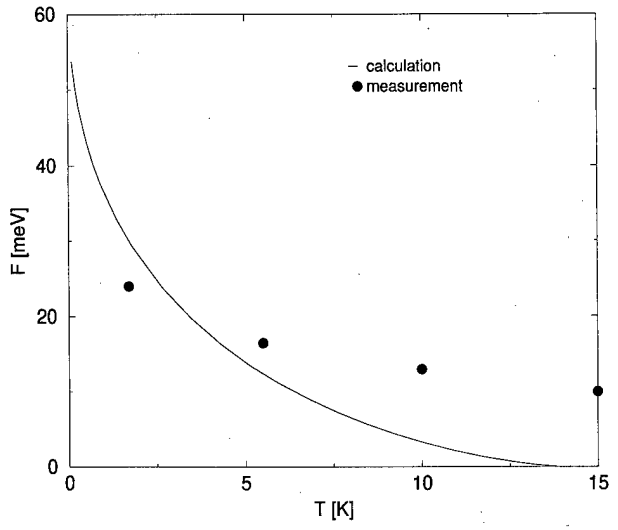


FIG. 4. Calculated temperature dependence of the FMP binding energy F compared with experimental PL excitation data for sample DMH2.

to choose an adiabatic approximation to solve for the eigenstates and eigenfunctions ψ of the Hamiltonian. We assume a cylindrical symmetric and separable wave function

$$\psi(\bar{r})=\varphi(z)\phi(\rho) \tag{2}$$

which represents a first order approximation in terms of a perturbational approach. φ is calculated according to the well potential $V(z)$ with $m^*=0.72m_e$ and assumes a gaussian-like wave function $\phi(\rho)=1/\sqrt{\pi}\rho_0 e^{-\rho^2/2\rho_0^2}$. The free energy is minimized according to the polaron radius $\rho_0(z)$ which is approximated to be finite only for z coordinates within the magnetic layers. Then the free energy is given by

$$F(\rho_0)=-\beta^{-1}\ln\prod_i\sum_{S_i}\frac{1}{2S+1}\times\exp(-\beta\gamma s S_i\psi^*(\bar{R}_i)\psi(\bar{R}_i)), \tag{3}$$

with $\beta=1/k_BT$ and $S=5/2$ for Mn ions. By applying a continuum approximation and sum up over all magnetic layers

$$F(\rho_0)=-\beta^{-1}\sigma\sum_z\int 2\pi\rho d\rho\ln\sum_{S_i}\frac{1}{2S+1}\times\exp(-\beta\gamma s S_i\psi^*(R_i)\psi(R_i)), \tag{4}$$

where $\sigma=x2/a^2$ denotes the density of magnetic sites per MnTe layer with $0\leqslant x\leqslant 1$.

By solving the integral by series expansion, an analytic expression for $F(\rho_0)$ is obtained. We use $N_0\gamma=0.366$ eV as an effective value of the coupling constant between ions and excitons[15] with $N_0=4/a^3$ and $a=6.466$ Å the lattice constant of the barrier material $Cd_{0.75}Mg_{0.25}Te$.[16] In Fig. 4, the temperature dependence of the minimum of the free energy F is shown to follow qualitatively the behavior of the experimental data.

C. Magneto-optical analysis of DMHs

1. Magneto-optical Kerr effect in RDS

Empirically the effect of a static or low frequency $\bar{B}$ field with its components B_k produces an additional polarization P_i of an isotropic medium, described in SI units by

$$P_i(\omega)=\varepsilon_0\gamma_{ijk}E_j(\omega)B_k, \tag{5}$$

where γ_{ijk} is an antisymmetric tensor of imaginary numbers.[17] If the polarization direction of light E_j is chosen to be perpendicular to the magnetic field $\bar{B}$, as it is the case of Faraday geometry, $\bar{B}=B\bar{z}$, the dielectric tensor is modified by the magnetic field

$$\varepsilon_{ij}(B)=\varepsilon_{ij}(0)+\begin{pmatrix}0 & -i\gamma B & 0\\ i\gamma B & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \tag{6}$$

with a real γ. The quantity γ can be frequency dependent and thereby show the magnetic moments of the electrons via their band structure. In order to describe this frequency dependence, microscopic models for γ are used to explain the measured data.

In order to determine the optical response in RDS we use the Berreman approach[18,19] to calculate the components of the Jones Matrix describing the reflection (r) of p- and s-polarized light by the incident $E(i)$ field

$$\begin{pmatrix}E_{rp}\\ E_{rs}\end{pmatrix}=\begin{pmatrix}r_{pp} & r_{sp}\\ r_{ps} & r_{ss}\end{pmatrix}\begin{pmatrix}E_{ip}\\ E_{is}\end{pmatrix}. \tag{7}$$

Without magnetic field it is always possible, in materials with twofold rotation symmetry, to find a polarization direction where r_{sp} and r_{ps} are zero. Considering thin epitaxial layers as expansions to the first order in the thickness d, we obtain that r_{pp} and r_{ss} are not modified by the $\bar{B}$ field, whereas

$$r_{sp}=r_{ps}=-\frac{4\pi d n_a\gamma B}{\lambda(n_s+n_a)^2} \tag{8}$$

with n_s, n_a the refractive index of the isotropic substrate and ambient ($n_a=1$) at the wavelength λ. For the structures under study, in the energy range around the gap of CdTe the imaginary part of the refractive index for the substrate is small with respect to the real part (up to 2.5 eV), thus r_{sp} is a real number.

The real and imaginary part of the RDS signal are $\Delta\tilde{r}/\tilde{r}=\Delta r/r+i\Delta\theta$ with $r_{ss}=\tilde{r}=r\exp(i\theta)$ and $r_{pp}=\tilde{r}+\Delta\tilde{r}=(r+\Delta r)\exp(i(\theta+\Delta\theta))$.[10] We define $\Delta\tilde{r}$ as $\tilde{r}_{110}-\tilde{r}_{110}$, where the indices denote the incident the incident polarization vectors. In case of an magneto-optical effect in thin epitaxial layers, $\Delta\theta$ and $\Delta r/r$ arising from an isotropic crystal (surface) are replaced by

$$\Delta\theta(B)=\Delta\theta(0)+2\frac{r_{sp}\sin(\theta)}{r}, \tag{9}$$

$$\frac{\Delta r}{r}(B)=\frac{\Delta r}{r}(0)-2\frac{r_{sp}\cos(\theta)}{r} \tag{10}$$

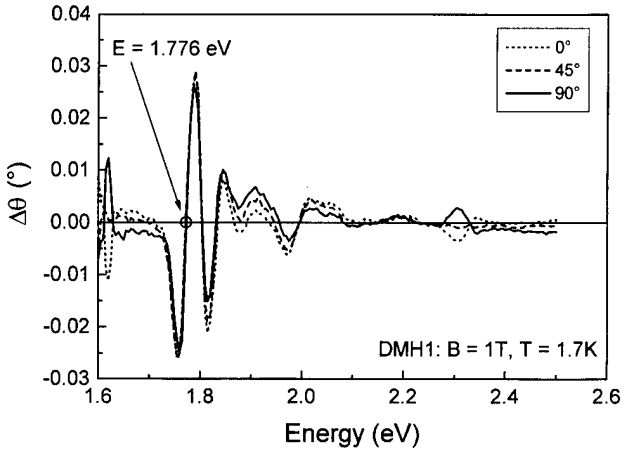


FIG. 5. RDS spectra ($T=1.7$ K and $B=1$ T) for sample DMH1 at different angles of polarization α.

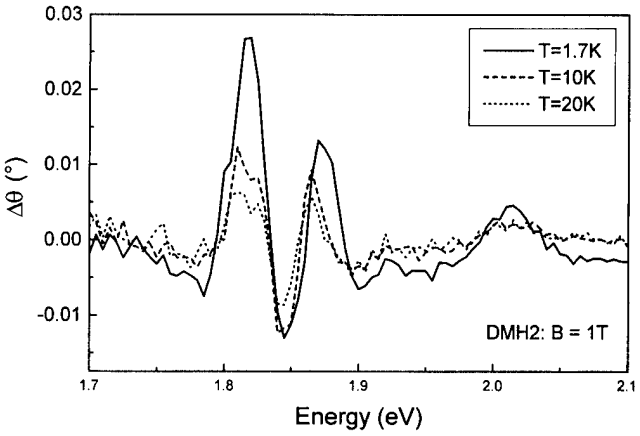


FIG. 6. Temperature dependence of the RDS spectrum for sample DMH2.

when r_{sp} is real. Although it appears in this formula that the absolute phase θ can be determined, it has to be pointed out, that at $B=0$ the additional term vanishes. In order to determine γ the absolute value as well as the phase are calculated from the known epilayer structure.

Since in presence of magnetic field the dielectric tensor is not symmetric, it cannot be diagonalized and the eigenstates are elliptical (circular if the material is isotropic). The rotation of the polarization under magnetic field is due to the difference in the reflection coefficients of the left and right handed eigenstates. The RDS setup applied to our structures is therefore a longitudinal Kerr effect setup where the substrate is used as a mirror and the basic description is the same as for the Faraday effect, where the light passes twice through the epilayer.

2. RDS of DMH

It has been theoretically predicted and experimentally verified by a comparison of PL excitation data with Faraday rotations measured in transmission,[20] that vanishing Faraday angles (θ_F) angles coincide with discrete electron states.

In our RDS experiments, the incident radiation is transmitted through the nonmagnetic CdMgTe barriers and reflected by the CdZnTe substrate. Hence, the light is passing the cw region twice giving a value for $\Delta\theta$ corresponding to $2\theta_F$.

For sample DMH1, we show in Fig. 5 the RDS spectrum ($T=1.7$ K in an external magnetic field $B=1$ T) for different angles of polarization (α) with respect to the [110] direction of the crystal. Maxima and minima between 1.75 and 2.1 eV appear to be independent of α, whereas the two distinct minima at 1.613 and 2.33 eV develop towards a maximum when α is tilted from 0° to 90°. The observed angular dependence suggests that these features are due to the crystal anisotropies solely while the other features are magnetic field induced. In particular the angular dependent structure at 1.613 eV coincides with the energy band gap of the CdZnTe substrate and the peak at 2.33 eV is due to the CdMgTe barrier.

The Stokes shift for the zero transition of the dominant feature (1.776 eV) in Fig. 5, with respect to the PL maximum (1.759 eV) from the same sample,[21] is produced by the exciton localization due to the magnetic polaron effect.

Considering the interband Faraday rotation proportional to the magnetization M[1] and assuming a paramagnetic behavior ($\chi=M/H$, where χ is the magnetic susceptibility), the Curie–Weiss law $\chi=\text{const}/(T+T^*)$ can be used to determine a temperature T which is proportional to the sum of all the exchange interaction integrals. In order to evaluate this critical parameter, we studied the temperature dependence of the RDS spectrum for sample DMH2, observing a decrement of the signal with the increase of T (Fig. 6). For a quantitative determination, $1/\Delta\theta$ at 1.797 eV is shown as a function of T in Fig. 7. A linear extrapolation of these data gives a value of $T^*=T_{AF}=-6$ indicating antiferromagnetic interactions between the Mn spins.

By comparing this antiferromagnetic temperature with values experimentally determined for bulk $Cd_{1-x}Mn_xTe$,[1] we can deduce an effective concentration of 20% for the magnetic ions. This value is in good agreement with the Mn

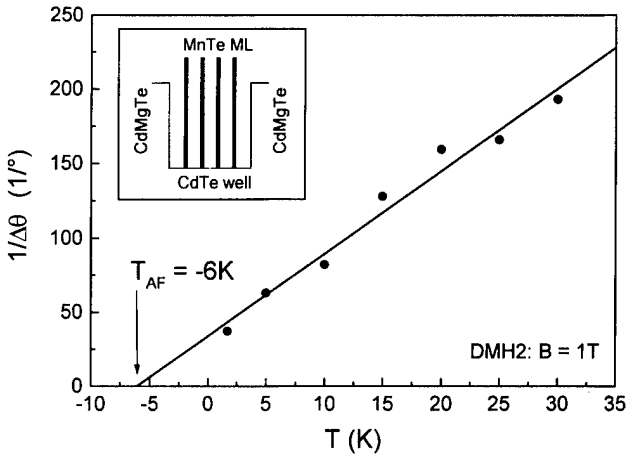


FIG. 7. $1/\Delta\theta$ at 1.797 eV vs T. A linear extrapolation of these data gives a value for $T_{AF}=-6$, indicating ferromagnetic interactions between the Mn spins. The inset shows a sketch of the potential profile for DMHs.

content expected for magnetic barriers with a thickness of 1/4 ML.

IV. CONCLUSION

RDS and magneto-optical analysis have been carried out on CdTe/MnTe-based DMHs. The binding energy of FMPs, probed via selective excitation of the PL, results to be higher in these structures, compared to the 3D case of alloys and its calculated temperature dependence results is in agreement with the experimental data. A model for the magneto-optical Kerr effect in the case of RDS is presented and from RDS measurement performed in magnetic field it was possible to discriminate between signals due to Faraday rotation and signals due to crystal anisotropy.

ACKNOWLEDGMENTS

Support by, Fonds zur Förderung der wissenschaftlichen Forschung in Österreich, FWF is acknowledged. The basic operation of the laboratory infrastructure used, was in part founded by the "Gesellschaft für Mikroelektronik" GME.

[1]D. U. Bartholomev, J. K. Furdyna, and A. K. Ramdas, Phys. Rev. B **34**, 6943 (1986).
[2]K. Ando, K. Takahashi, and T. Okuda, Phys. Rev. B **46**, 12 289 (1992).
[3]D. D. Awschalom, M. R. Freeman, N. Samarth, H. Luo, and J. K. Furdyna, Phys. Rev. Lett. **66**, 1212 (1991).
[4]S. A. Crooker, D. A. Tulchinsky, J. Levy, D. D. Awschalom, R. Garcia, and N. Samarth, Phys. Rev. Lett. **75**, 505 (1995).
[5]I. A. Merkulov, G. R. Pozina, D. Coquillat, N. Paganotto, J. Siviniat, J. P. Lascaray, and J. Cibert, Phys. Rev. B **54**, 5727 (1996).
[6]E. O. Giel, R. Hellmann, G. Mackh, D. R. Yakovlev, W. Ossau, A. Waag, and G. Landwehr, Mater. Sci. Forum **182/184**, 519 (1995).
[7]G. Mackh, W. Ossau, D. R. Yakovlev, A. Waag, G. Landwehr, R. Hellmann, and E. O. Göbel, Phys. Rev. B **49**, 10 248 (1994).
[8]V. V. Rossin, F. Henneberger, and J. Puls, Phys. Rev. B **53**, 16 444 (1996).
[9]G. Mackh *et al.*, Phys. Rev. B **50**, 14 069 (1994).
[10]D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).
[11]A. V. Kavokin and K. V. Kavokin, Semicond. Sci. Technol. **8**, 191 (1993).
[12]M. Bugajski, P. Becla, P. A. Wolff, D. Heiman, and L. R. Ram-Mohan, Phys. Rev. B **38**, 10 512 (1988).
[13]G. Mackh, W. Ossau, D. R. Yakovlev, A. Waag, T. Litz, and G. Landwehr, Solid State Commun. **88**, 221 (1993).
[14]A. Kavokin, B. Gil, and P. Bigenwald, Phys. Rev. B **57**, R4261 (1998).
[15]J. A. Gaj, R. Planel, and G. Fishman, Solid State Commun. **29**, 435 (1979).
[16]A. Waag, J. Cryst. Growth **131**, 607 (1993).
[17]S. Huard, *Polarization of Light* (Wiley, New York, 1997).
[18]W. Berreman, J. Opt. Soc. Am. **62**, 502 (1972).
[19]R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1977).
[20]B. Buda, M. Dahl, N. Von Truchness, and A. Waag, J. Cryst. Growth **138**, 652 (1994).
[21]A. Bonanni, W. Heiss, G. Prechtl, D. Stifter, M. Schmid, K. Hingerl, W. Jantsch, and H. Sitter, J. Cryst. Growth **184/185**, 921 (1998).

Coherent soft x-ray scattering from InP islands on a semiconductor substrate

M. Adamcyk, C. Nicoll, T. Pinnington, and T. Tiedje[a)]
Advanced Materials and Process Engineering Lab, Department of Physics and Astronomy, University of British Columbia, Vancouver V6T 1Z4, Canada

S. Eisebitt, A. Karl, R. Scherer, and W. Eberhardt
Forschungszentrum Jülich, Jülich D-52428, Germany

(Received 20 January 1999; accepted 10 April 1999)

Coherent soft x-ray scattering experiments from a semiconductor sample consisting of InP islands on a smooth semiconductor substrate are described. The soft x-ray scattering was performed with 266 eV photons produced by an undulator source. Using a position sensitive detector, we are able to detect diffusely scattered x rays in the vicinity of the specular reflection, with an in-plane momentum transfer of up to 6 μm^{-1}. Using Huygens–Fresnel theory and atomic force microscope images of the surface structure, we simulated the scattering assuming a finite lateral coherence length for the incident radiation. The lateral coherence length of the incident beam was found to be 20 μm from a fit to the observed diffraction pattern from a pinhole. The effect of changes in the surface morphology on the speckle pattern was simulated to explore the potential of coherent soft x-ray scattering for the study of surface structure dynamics. © *1999 American Vacuum Society.* [S0734-211X(99)03704-X]

The random intensity variation observed in the specular and near specular reflection of coherent radiation from random surfaces is referred to as speckle. The speckle is caused by random path differences between radiation scattered from different locations on the sample, which interfere at the detector in the far field to produce a random interference pattern. Speckle patterns are sensitive to the particular configuration of the random surface whereas incoherent light scattering is sensitive to statistical averages of the structure. Therefore, one can learn about the dynamics of surface structure in thermal equilibrium even when the average structure is constant by measuring the time dependence of the speckle pattern. The use of coherent visible radiation to study dynamical phenomena is an established technique.[1] However, the small wavelengths available with synchrotron radiation in principle allow one to probe much smaller length scales than is possible with visible light. Undulators in third generation synchrotron radiation sources produce radiation that is coherent enough to observe speckle patterns. Recently, the first coherent hard x-ray scattering from random bulk and surface structures have been carried out.[2–5] However, reflection coefficients and diffuse scattering from surfaces and interfaces are very small in the hard x-ray wavelength range except near grazing incidence. This reduces the sensitivity of hard x rays for large angle scattering from surfaces. Large angle scattering is needed to probe structure at high spatial frequencies (short length scales). Except at grazing incidence, reflection coefficients are orders of magnitude larger for soft x rays than for hard x rays. A much larger range of scattering angles is therefore experimentally accessible for diffuse reflection in the soft x-ray range. In addition, material specific scattering experiments on low-Z elements are possible with soft x rays. For these reasons, soft x-ray scattering is a potentially useful technique for studying the structure of surfaces.

In this article, we report experiments with undulator radiation, measuring soft x-ray speckle patterns from InP islands grown on smooth InAs-coated InP substrates. Numerical simulations of the partially coherent scattering have been carried out according to Huygens–Fresnel scattering theory in the Fraunhofer approximation. These calculations are used to simulate the change in the speckle pattern resulting from small variations in the surface morphology of the sample.

Our experiments were carried out at the X1B beamline at the National Synchrotron Light Source at the Brookhaven National Lab. The experimental setup is shown in Fig. 1. The dimensions that are indicated in Fig. 1 correspond to a typical setup, but were varied between experiments. In order to block stray light from the beamline, a 3 mm diameter pinhole was placed in the beam downstream from the monochromator exit slit. To increase the lateral coherence length of the incident x rays, the synchrotron radiation beam was apertured using a commercial (Melles Griot) 20 μm diameter pinhole. To minimize diffraction effects on the size of the illuminated area on the sample, the downstream-pinhole-to-sample distance was kept as short as possible or about 30 mm. The sample was mounted on a single axis rotation stage so that the angle of incidence could be varied. Generally, the detector was placed 540 mm away from the sample in the horizontal direction approximately parallel to the direction of the incident beam (see Fig. 1). Stepper motors were employed to reposition the detector into the specular beam when the angle of incidence was varied. The position sensitive detector consisted of a multichannel plate stack with a two-dimensional (2D) spatially resolving resistive anode readout. The active area consists of 1024×1024 pixels,

[a)]Also at Department of Electrical and Computer Engineering.

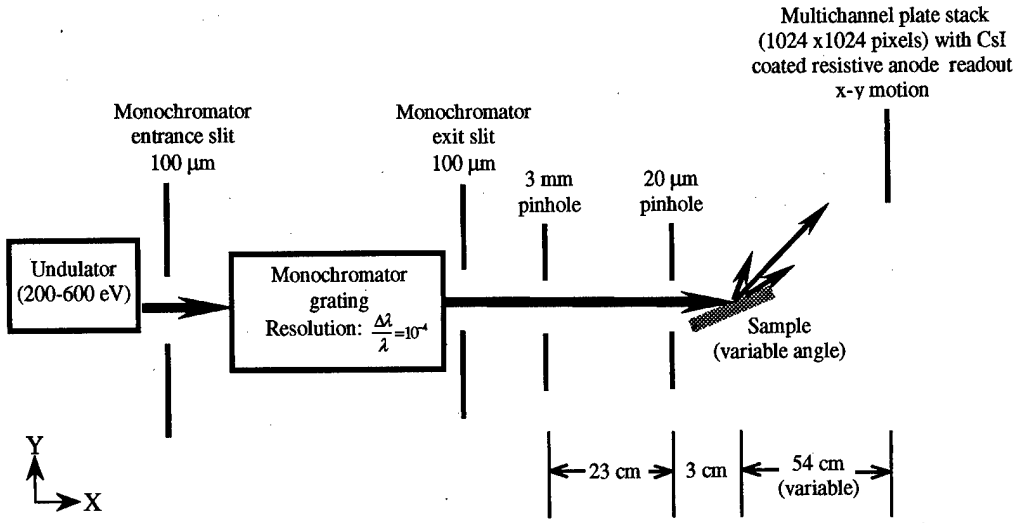


FIG. 1. Configuration of soft x-ray scattering experiment. Slit openings and pinhole sizes are variable, with typical dimensions indicated.

which are $40\,\mu m \times 20\,\mu m$ in size. The image acquisition times were typically a few minutes.

Numerical simulations of the scattering were made using atomic force microscope (AFM) images of the samples and a Huygens–Fresnel scattering theory.[6] Figure 2 shows the geometry considered in the simulations along with the definition of the momentum transfer vectors parallel and perpendicular to the plane of the surface. Following the theory of statistical optics we can write the mathematical expression for the light intensity in terms of the autocorrelation function of the field. We compute the scattering from the sample by following the propagation of the autocorrelation function of the incoming light through an optical system described by a complex pupil function. The later takes into account the size of the illuminated area on the sample and the random phases associated with the path difference resulting from the surface height profile.

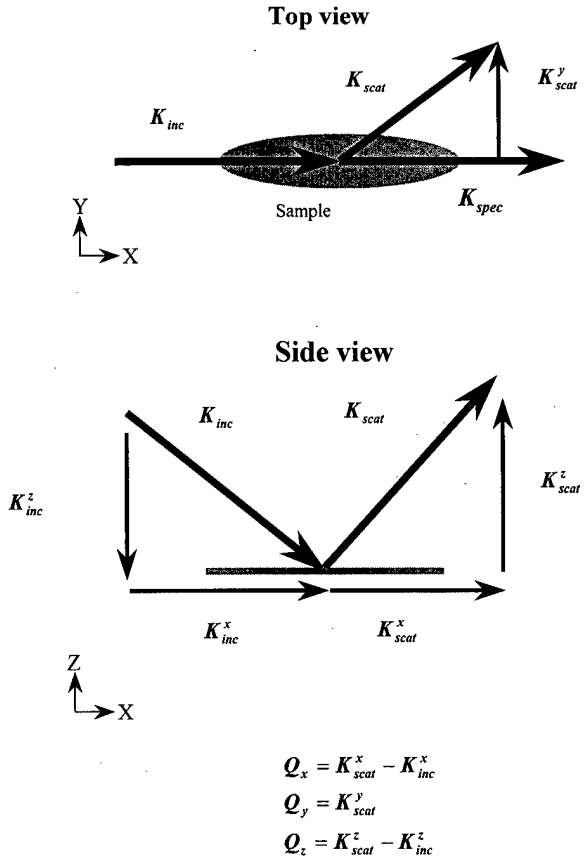


FIG. 2. Diagram showing the experimental geometry and the momentum transfer vectors.

In the following discussion, all prefactors in the scattering intensity are omitted in order to simplify the mathematical expressions. We define the complex pupil function as follows:

$$P(x,y)=p(x,y)e^{iQ_z\cdot h(x,y)}, \tag{1}$$

where $p(x,y)$ is an elliptical aperture function which simulates the footprint of the beam on the sample. The aperture function is unity inside the elliptical illuminated region on the sample surface and is zero elsewhere. $h(x,y)$ is the surface height function relative to a reference plane, which was obtained from AFM measurements. The real-space autocorrelation of the complex pupil function is given by

$$\bar{\mathbf{P}}(\Delta x,\Delta y)=\int\int P\left(x-\frac{\Delta x}{2},y-\frac{\Delta y}{2}\right) \times P^{*}\left(x+\frac{\Delta x}{2},y+\frac{\Delta y}{2}\right)dx\,dy. \tag{2}$$

The real-space autocorrelation function of the incoming light is assumed to be Gaussian and is given by

$$\mu(\Delta x,\Delta y)=\exp\left(-\frac{(\sin\theta\Delta x)^2+\Delta y^2}{L_c^{\,2}}\right), \tag{3}$$

where L_c is the lateral coherence length along the cross section of the beam and θ is the angle of incidence. The $\sin\theta$

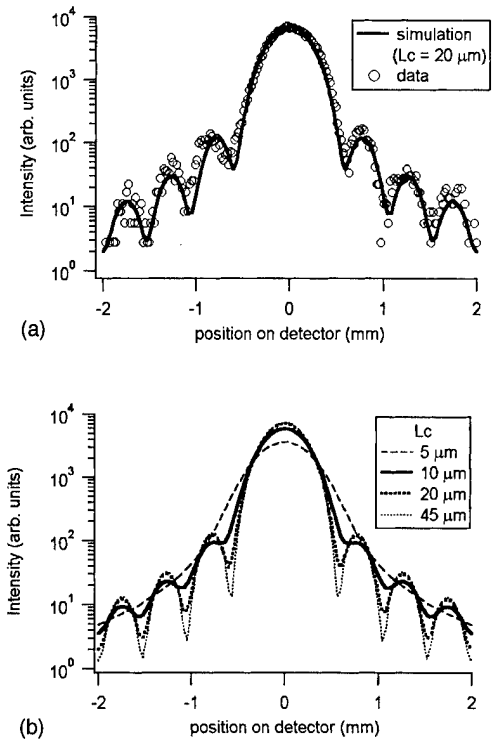


FIG. 3. Cross section of an Airy pattern in the x-ray intensity measured 800 mm from a 10 μm diameter pinhole at a photon energy of 215 eV. (a) Fit to the experimental data using a lateral coherence length of 20 μm and (b) simulations for various values of the lateral coherence length.

takes into account the projection of the cross section of the beam onto the sample. The two autocorrelation functions are Fourier transformed and convolved together to yield the pattern of the intensity of the radiation on the detector:

$$\hat{I}(Q_x,Q_y)=\hat{P}(Q_x,Q_y)*\hat{\mu}(Q_x,Q_y). \tag{4}$$

Intuitively, one can think of the calculation as summing the complex phase factors arising from the path difference between each point on the sample and the considered pixel on the detector. The longitudinal coherence length was not taken into account in these calculations and is therefore assumed to be infinite.

The Fraunhofer diffraction pattern (Airy pattern) produced by a 10 μm diameter pinhole was measured by placing the detector 800 mm behind the aperture. In order to maximize the wavelength and diffraction ring contrast in this measurement, the energy of the incident photons was set close to the long wavelength limit of the beamline at 215 eV (5.8 nm wavelength). Comparing this measurement with simulations done using the previously discussed model allowed us to evaluate the lateral coherence length of the incident x-ray beam. A cross section of the measured diffraction pattern is shown in Fig. 3(a). The experimental data are in good agreement with the calculated diffraction pattern for a lateral coherence length, $L_c=20\,\mu$m. Decreasing the coherence length diminishes the contrast of the interference rings, however, the average intensity remains relatively unchanged. Destructive interference is reduced by the finite coherence

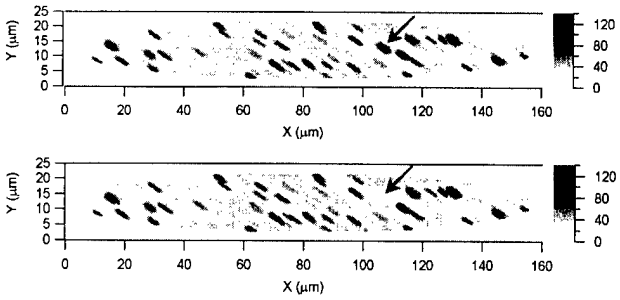


FIG. 4. Atomic force microscope images of InP islands on an InP substrate with an InAs wetting layer. The height scale is in nanometers. The x-y axes are aligned along the [100] crystal directions. In the lower image, the island indicated by the arrow in top image has been erased in order to simulate the effect of a change in the surface morphology.

length of the x rays as illustrated at the bottom of Fig. 3 where we show simulations for various coherence lengths.

The sample used in the experiments described here consists of InP islands grown on an InAs-terminated InP wafer by solid source molecular beam epitaxy. The substrate's native oxide was thermally desorbed under an As_4 overpressure at a temperature of 540 °C, measured by optical band-gap thermometry. The reflection high-energy electron diffraction (RHEED) pattern was strong but spotty after the oxide desorption. The growth was carried out at a substrate temperature of 500 °C under a phosphorus overpressure (BEP$\approx 6\times10^{-5}$ mbar). Rotation of the RHEED pattern at the beginning of the growth suggested that P was substituted with As during the oxide desorption. Presumably InP does not wet the InAs layer, which terminated the surface prior to growth. This is our explanation for why the InP grows as islands on this sample. This sample is an appropriate test sample for coherent soft x-ray scattering since the surface is rough enough to give strong diffuse scattering yet not so rough that the specular reflection is lost. Also the InP islands are large enough that they scatter in the $<6\ \mu m^{-1}$ in-plane spatial frequency range that is readily accessible with our experimental geometry and detector size. InAs quantum dots, for example, scatter primarily outside the spatial frequency range that we are sensitive to here. The AFM measurements were made on an area of the sample close to the incident spot for the x-ray beam but not at the identical location. The simulation results can therefore only agree with the measured scattering data in a statistical average sense.

X-ray scattering measurements were done on this sample at a photon energy of 266 eV (4.7 nm wavelength) since this energy has the maximum photon flux for the beamline. By keeping the photon energy below the carbon edge, we minimize the effects of loss of throughput due to absorption by carbon contamination on the beamline optics. The configuration of the experimental setup is shown in Fig. 1, with the angle of incidence ranging between 3.4° and 7.4°. The AFM image of the sample used for the scattering simulations is shown in Fig. 4. The elliptical border of the AFM image indicates the area of the x-ray spot on the sample at an incidence angle of 7.4°. The corresponding measured speckle pattern is given in Fig. 5 along with coherent ($L_c=\infty$) and

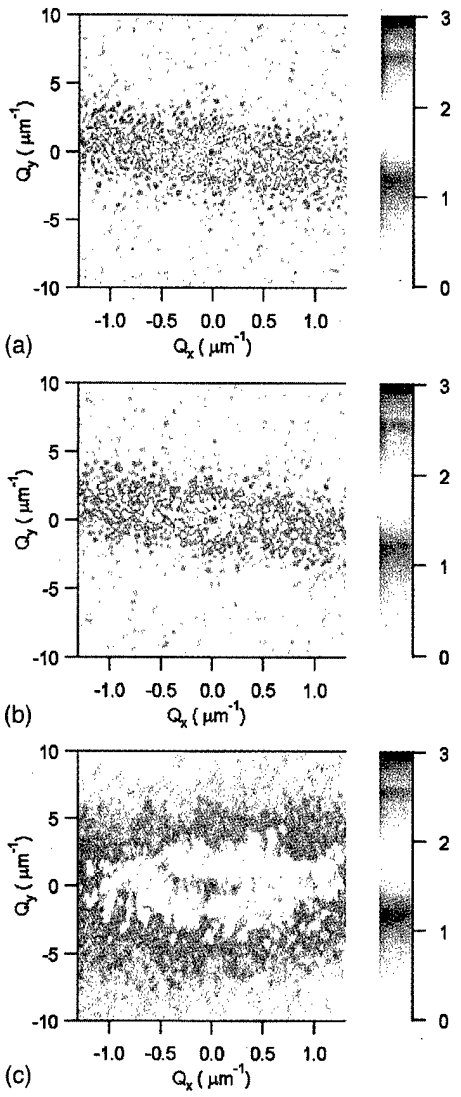


FIG. 5. Two-dimensional scattering from the InP island sample. The angle of incidence is 7.4° with respect to the surface and the photon energy is 266 eV. The intensity colormap is on a logarithmic scale with the same dynamic range for all three images. The three images show: (a) the simulated speckle pattern calculated from the atomic force microscope image in Fig. 4 with an infinite lateral coherence length, (b) the same simulation except with a lateral coherence length of 20 μm, and (c) experimental scattering data for the sample whose AFM image was used in the simulations. The longitudinal coherence length is assumed to be infinite in all of the simulations.

partially coherent ($L_c = 20\,\mu$m) simulation results.

We refer to the number of decades that is spanned by the intensities in an image as the "dynamic range" of the image. To facilitate comparison, all three results were plotted on a color scale with the same dynamic range as the experimental data. The relative intensities of the simulations were scaled to match the experimental data. No attempt was made to simulate the absolute intensity. The overall shape of the simulated intensity and the size of the speckles agree with the experimental data. The intensity of the central specular reflection relative to the surrounding diffuse reflection is approximately correct. However, the simulations have a larger dynamic range than the data and therefore many of the pixels in the calculated images are saturated white in the figure. The simulation also has a higher speckle contrast. The finite longitudinal coherence of the x-ray beam, which is neglected in our simulations may also contribute to the difference between the measured and simulated scattering. The AFM does not necessarily produce a true image of the sample since an AFM image is a complex convolution of the profile of the end of the tip with the surface of the sample. This could change the scattering distribution as a function of spatial frequency.

The longitudinal coherence length (spectral purity) for our experimental conditions was

$$L_{\text{long}} = \frac{\lambda^2}{\Delta\lambda} \approx 50\ \mu\text{m}. \tag{5}$$

This is on the order of the maximum path difference between light beams scattered from either end of the illuminated region on the sample. The finite longitudinal coherence has not been taken into account in our simulations. Second, the simulations were carried out using the lateral coherence length obtained by fitting the Airy rings produced by a 10 μm diameter pinhole. In the measurements on the InP sample, the pinhole was larger, and the photon energy was higher. The lateral coherence length depends on the solid angle of the pinhole seen by the exit slit of the monochromator and will decrease with increasing pinhole diameter. Also, the shorter wavelength output of the undulator used for the reflectivity experiments will also reduce the coherence length relative to the coherence length obtained from the pinhole diffraction. Both of these effects (finite longitudinal coherence and overestimate of the lateral coherence length) will cause the measured speckle contrast to be smaller than the simulated speckle contrast as observed in Fig. 5.

The incident angle and the size of the detector limit the maximum range in Q_x that can be measured without moving the detector. In order to expand the range of Q_x, the detector was scanned in overlapping steps. By scanning the detector in steps of one half the detector length with the angle of incidence kept constant, we were able to explore scattering with values for the in-plane momentum transfer Q_x ranging

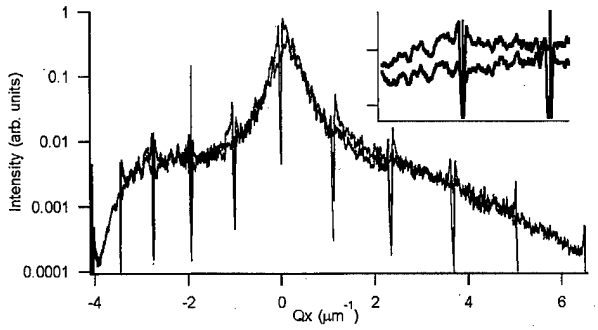


FIG. 6. X-ray scattering for large values of the in-plane momentum transfer obtained by moving the position sensitive detector in steps of one-half of its' length. The photon energy is 266 eV, and the angle of incidence is 5.4°. The inset shows an enlargement of part of the graph in order to show the reproducibility of the data between experiments (scan lines are offset for clarity).

up to 6 μm^{-1}. In Fig. 6, we present the result of an experiment to measure a large Q_x range for 266 eV photons incident on the sample at an angle of 5.4°. The intensity was integrated along an 80 pixel wide line of slope $\Delta Y/\Delta X = 0.15$ with respect to the edge of the detector. The individual image intensities were normalized to the intensity of the incoming synchrotron radiation and the data acquisition time before plotting the spectra on a common intensity scale. The intensity peaks at the edge of the detector are artifacts typical of this type of imaging detector. A comparison of the overlap regions between two consecutive detector positions shows the reproducibility of the speckle as illustrated in the inset in Fig. 6. This confirms that the observed fluctuations in intensity are speckle due to the surface scattering and are not detector noise.

In the simulation, the pattern of the axis of the x-ray intensity pattern is tilted with respect to the frame of the image [see Fig. 5(a)]. This is caused by the orientation of the InP islands on the sample. In the AFM images in Fig. 4, the islands are elongated and aligned with the [110] crystal direction on the substrate. The orientation of the islands creates directional scattering, which is responsible for the tilt in the diffuse scattering. The slope along which the detector intensity was integrated (see Fig. 6) is slightly smaller than the tilt observed in the simulated data, which indicates that the sample [100] crystal axes were slightly misaligned with respect to the plane of incidence during the measurements.

The effect of a change in the surface morphology on the speckle pattern was simulated by erasing one InP island from the AFM image used to simulate the intensity distribution. The location of the missing island is indicated with an arrow before and after being removed in Fig. 4. The scattering simulation was computed with the modified AFM image, and the resulting speckle pattern was subtracted from the original simulation. The absolute value of the difference image is shown in Fig. 7 using the same logarithmic grayscale as in Fig. 5. There is a measureable difference between the two images with a spatial frequency distribution that mirrors the distribution of scattering in the original simulation. There are about 20 large InP islands in the original image. Since the scattering is dominated by the islands, one would expect that removing one island would reduce the average nonspecular scattering intensity by about 5% spread over the same spatial frequency range as the original speckle pattern. This is consistent with the difference image in Fig. 7: near the specular position, the intensity of the difference is about 1.3 orders of magnitude (i.e., a factor of 20) smaller than the average intensity of the scattering.

Changing the position of one of the islands will have a similar effect on the speckle as removing one island, except that the overall scattering intensity will not change. The individual speckles in the experimental data cover about 25 pixels with approximately 150 counts/pixel [see Fig. 5(b)].

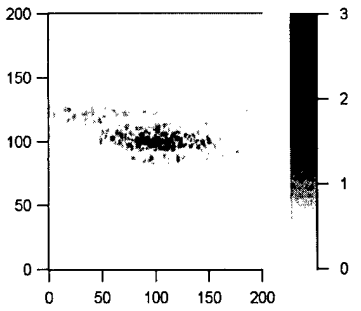


FIG. 7. Absolute value of the difference between the speckle patterns calculated for the two images in Fig. 4 for 20 μm lateral coherence length. Other parameters are indicated in Fig. 5.

Assuming that the uncertainty in the measurement of the intensity in one speckle is given by the square root of the total number of counts, we should be able to detect changes in the intensity on the order of $\sim 1/\sqrt{4000} = 1.6\%$. Thus, the current experiment has enough sensitivity to detect the motion or disappearance of one of the InP islands. This indicates that coherent soft x-ray scattering can be used to study the dynamics of the InP islands in the test sample used in this experiment. Initial experiments to explore the island motion/evaporation at elevated temperature were only partially successful due to the movement of the x-ray beam on the sample, caused by thermal expansion of the sample and sample holder. A brighter undulator source and higher count rate detector would improve the sensitivity and make it possible to detect smaller changes with better time resolution. Many orders of magnitude increase in source brightness are possible before sample heating becomes a problem.

In conclusion, we have described measurements of speckle in the scattering of coherent soft x rays from InP islands on a flat semiconductor substrate. These experiments and simulations of the scattering using atomic force microscope images of the surface structure, demonstrate the feasibility of coherent soft x-ray scattering as a means for studying the dynamics of the structure of surfaces in equilibrium or near equilibrium.

The authors thank Carolina Pistonesi for help with the speckle simulations at the beginning of this project and NSERC for financial support.

[1]Ed. J. C. Dainty, *Laser Speckle and Related Phenomena*, 2nd ed. (Springer, New York, 1984).
[2]S. Bauer, G. B. Stephenson, M. Sutton, R. Brüning, E. Dufresne, S. G. J. Mochrie, G. Grübel, J. Als-Nielsen, and D. Abernathy, Phys. Rev. Lett. **74**, 2010 (1995).
[3]Z. H. Cai, B. Lai, W. B. Yun, I. McNulty, K. G. Huang, and T. P. Russel, Phys. Rev. Lett. **73**, 82 (1994).
[4]M. Sutton, S. G. J. Mochrie, T. Greytak, S. E. Nagler, L. E. Berman, G. A. Heald, and G. B. Stephenson, Nature (London) **352**, 608 (1991).
[5]J. L. Libbert, R. Pindak, S. B. Dierker, and I. K. Robinson, Phys. Rev. B **56**, 6454 (1997).
[6]J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985).

Infrared study of Si surfaces and buried interfaces

A. Milekhin[a] and M. Friedrich
Institute für Physik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany

K. Hiller, M. Wiemer, and T. Gessner
Zentrum für Mikrotechnologien, Technische Universität, D-09107 Chemnitz, Germany

D. R. T. Zahn
Institute für Physik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany

(Received 30 March 1999; accepted 20 April 1999)

Multiple internal reflection and transmission IR spectra of hydrophobic and hydrophilic Si wafers, Si wafers with thermally grown SiO_2 layers, and Si wafers bonded at high and room temperature were investigated. It was found that the surface of the as-prepared hydrophobic wafer is terminated by hydrogen and water molecules, while the IR spectra of hydrophilic wafer demonstrate only the presence of water molecules at the surface. IR spectra of Si wafers covered by a thermally grown SiO_2 layer exhibit a number of the strong absorption bands assigned to combinational phonon bands in SiO_2. The wafer bonding leads to the appearance of siloxane and hydroxyl groups at the buried interface whose absorption bands were observed in IR spectra. A rearrangement of atoms at the buried interface takes place after annealing of Si bonded wafers. IR spectra of room temperature bonds show a large number of water molecules and presence of the hydrogen in the oxide layer at the interface. © *1999 American Vacuum Society.* [S0734-211X(99)03804-4]

I. INTRODUCTION

Silicon wafer bonding is a promising technique in microelectronics for the monolithic integration of micromechanical and electronic devices. Recently, significant progress has been achieved in optimizing the bonding process and the understanding of the Si–Si bonding mechanism.[1–4] However, the information about the nature and evolution of the interface in bonded wafers, especially, with buried thermally grown SiO_2 is still lacking. Characterization and the quality control of the Si–SiO_2 interface is of great importance for silicon-on-insulator related microelectronic devices, sensors and actuators. One of the reasons is that conventional surface science techniques (electron diffraction, photoemission, Raman scattering, photoluminescence, etc.) cannot access the buried interface.

IR spectroscopy traditionally used for the investigation of vibrational properties of various coatings, surfaces and interfaces[5–11] is also particularly suited for the characterization of buried interfaces and can access these interfaces over a wide spectral range (2000–8000 cm^{-1}). Recently, IR spectroscopy was applied successfully for the study of chemical nature of bonded Si wafers.[1,2,12–14] It is established that the interface of joined wafers depends strongly on the preliminary surface treatment, i.e., chemical preparation of wafers, ion implantation, plasma treatment, oxidation and annealing process, etc.

In this article, we present the results of a systematic IR spectroscopical investigation of bonded silicon wafers with buried thermally grown SiO_2 layer. IR response of Si wafers is considered after each technological step and as a result the contributions of the outer silicon surfaces and the inner interface of Si bonds to absorption are determined. The absorption lines in IR spectra of hydrophobic and hydrophilic Si wafers, Si wafers with thermally grown SiO_2 and the bonded Si wafers are identified.

II. EXPERIMENTAL DETAILS

Several types of experimental geometries can be applied for measuring IR spectra. Multiple internal reflection is usually used for the characterization of surface layers as thin as a monolayer.[5] In this geometry, the film is deposited on an infrared transparent substrate with trapezoidal cross section.[15] Multiple internal transmission is the technique for the study of the buried interface region between two joined substrates. One of the variants of this geometry is shown in Fig. 1(a). Infrared radiation trapped inside the sample is transmitted through the interface many times before exiting the opposite bevel. An alternative approach is the use of single-pass multiple internal reflection elements. These elements prepared from infrared transparent material are pressed gently to the whole single or bonded wafer to get optical contact and to allow the IR light to enter the sample and exit from another element as it is shown in Fig. 1(b). The distance between elements can be varied thus allowing to find the optimum between the whole signal magnitude and the intensity of the IR response from the interface region. This technique does not require any treatment of the sample and allows to study the IR response from the surface or the buried interface just after any technological process step or the bonding process. IR spectra of the Si bonded wafers carry information about the Si surface and buried interface. In order to distinguish these two contributions the following method can be used. The Si wafer is bonded with two Si wafers covered by SiO_2 layers. The double bonded Si wafer is used to produce the structure shown schematically in Fig.

[a]Also at Institute of Semiconductor Physics, 630090, Novosibirsk, Russia; electronic mail: a.milekhin@physik.tu-chemnitz.de

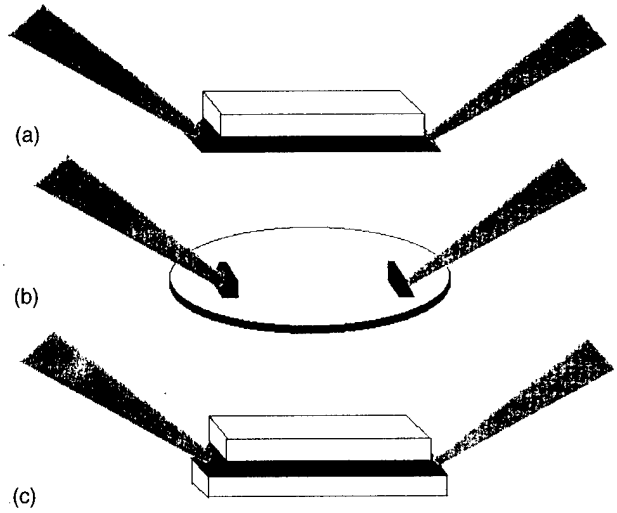


FIG. 1. Sketch of the geometry of experiments. (a) Geometry of multiple internal transmission of Si bond. (b) Si wafer with single-pass multiple internal reflection elements. (c) Geometry of multiple internal transmission of double bonded Si wafers.

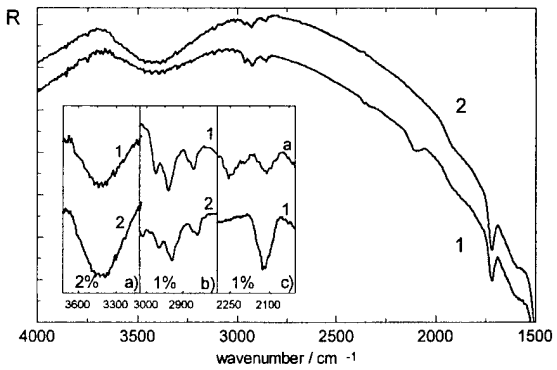


FIG. 2. IR *p*-polarized multiple internal reflection spectra of the as-prepared hydrophobic and hydrophilic Si wafers (curves 1, 2, respectively). The insert shows OH and C–H absorption bands of as-prepared hydrophilic and hydrophobic Si wafers [inserts (a) and (b)] and the spectra of as-prepared and stored in air hydrophobic Si wafers [insert (c), curves 1 and a, respectively].

1(c). The refraction coefficient for SiO_2 is lower than for Si, therefore, the IR light undergoes total internal reflection in the intermediate Si wafer. As a result only the IR response from the buried interfaces is measured.

In our study, we used all experimental geometries shown in Fig. 1. The samples with a size of $0.4\times20\times50$ mm^3 and with 54.7° bevels were prepared from Si(100) wafers to allow the IR radiation to pass through the samples 100 times. The bevels were prepared using selective etching in KOH (30%) solution in water at a temperature of 80 °C. The etching velocity is minimal in (111) direction which forms an angle of 54.7° with the (100) plane. Two types of samples were investigated. Part of the samples were covered with a thermal oxide of 1 μm thickness grown at 1100 °C in a wet atmosphere (O_2+H_2 with a ratio of 1:1.65) at normal pressure. The samples were either hydrophilic by treatment in H_2O_2:HCl:H_2O=1:1:5 and NH_4OH:H_2O_2:H_2O=1:1:5 solutions or hydrophobic by additional treatment in HF (49%):NH_4(40%)=1:4 solution. IR spectra of the wafers were recorded just after the treatment and after five days storage in air. After that the chemical treatment was repeated and samples were annealed at T=450 °C and T=1100 °C during 5 h in the same regime which is used for the wafer bonding process. After annealing the IR spectra were again measured.

The silicon wafer treated by RCA cleaning and two Si wafers covered by SiO_2 layers with a thickness of 1 μm were bonded as shown in Fig. 1(c) and annealed at a temperature of 450 °C in nitrogen. Both external surfaces of the double bond were left unpolished in order to enhance diffuse scattering and thus exclude any signals from external surfaces. In addition, we probed the Si bonded wafers without SiO_2 layers joined at high (800 °C) and low (200 °C) temperature using the single-pass multiple internal reflection elements technique.

IR spectra were measured with *p*-polarized light using a Fourier transform infrared spectrometer IFS-66 in the spectral range 400–4000 cm^{-1} with a spectral resolution of 4 cm^{-1}. The choice of *p*-polarization for multiply internal reflection geometry is explained by the greatly enhanced perpendicular component of absorption relative to the parallel component.[4] The transmission spectrum of the empty channel of the spectrometer without reflection unit was used as a reference spectrum. The number of scans was 1000.

III. RESULTS AND DISCUSSION

Typical IR spectra of Si substrates with hydrophobic and hydrophilic surfaces using trapezoidal prisms are presented in Fig. 2 (curves 1 and 2, respectively) in the spectral range 4000–1500 cm^{-1}. These spectra are characterized by the broad absorption band centered at 3400 cm^{-1} assigned to O–H stretch vibrations of adsorbed water molecules[16] and by the absorption lines at 2850–2950 cm^{-1} attributed to the CH_2 and CH_3 stretch vibrational modes of organic species on the Si surface caused by chemical cleaning.[17] A comparison of intensity of these modes is presented in the inserts (a) and (b) (curves 1 and 2, respectively). No noticeable influence of the C–H_x species on the surface arrangement of the Si wafers was found. The feature at 1720 cm^{-1} is caused by the vibrational mode due to interstitial oxygen in silicon.[18] The IR spectra of the hydrophobic Si wafer reveal in addition a feature at 2100 cm^{-1} [insert (c), curve 1] assigned to the Si–H stretch modes in the dihydride at the (100) silicon surface.[10] This feature is modified by weak vibrational modes at 2090 and 2145 cm^{-1} which are assigned to the Si–H stretch in monohydride and trihydride.[19] The Si–H bonds appear due to passivation of Si surface by the H atoms after HF treatment.

The IR spectra of hydrophilic samples did not change after storing of the samples in air during five days while the spectra of hydrophobic samples demonstrate an alternation of the surface arrangement. One can see from the insert in Fig. 2 [insert (c), curves 1 and a] that the intensity of Si–H modes decreases and a new mode at 2250 cm^{-1} related to Si–H stretching vibration in O_3–Si–H species appears. The weak mode at 2200 cm^{-1} is assigned to the intermediate

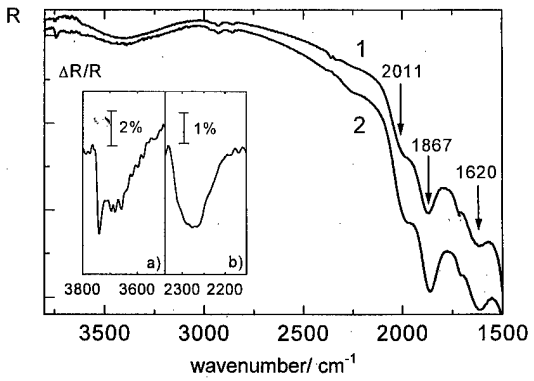


FIG. 3. IR *p*-polarized IR spectra of the Si wafer covered by SiO_2 layers (curve 1), and double bonded Si structure (curve 2). Insert shows the difference between IR spectra of double bonded Si wafer and Si wafer with SiO_2 layers.

oxidized species $SiH_2(O_2)$. This confirms previous results[11] that the oxygen molecules preferentially attack the Si–Si bonds of the surface Si atoms and that the surface Si–H bonds are quite inert to oxygen molecules.

In the next step, the samples were chemically treated by the same way as it was done before and annealed as it was described above. The IR spectra of the hydrophobic and hydrophilic samples annealed at a temperature of 450 and 1000 °C for 5 h are virtually identical. A typical IR spectrum of an annealed sample is presented in Fig. 3 (curve 1). The native oxide layer grown at the surface of the sample is responsible for the appearance of new absorption lines at frequencies of 1620, 1867, and 2011 cm^{-1}. These lines are related to multiphonon processes in silicon dioxide. In order to determine the frequency position of optical phonons in thermally grown silicon dioxide *p*- and *s*-polarized single transmission IR spectra of the same sample were measured and a curve fitting procedure was employed. The IR spectra of silicon dioxide layers reveal phonon bands in the region between 1050 and 1223 cm^{-1} (TO_1 and LO_1) and 816 and 826 cm^{-1} (TO_2 and LO_2). The silicon dioxide film exhibits additionally a TO-LO phonon pair in IR spectra at 457 and 507 cm^{-1} (TO_3 and LO_3).[19] The results of the mode identification are given in Table I. Our assignment of the TO_1+TO_2 and the LO_1+TO_2 modes is in agreement with conclusions of Han and Aydil[20] based on the isotopic shift of combinational modes in SiO_2 films produced by plasma enhanced chemical deposition. On the base of the phonon frequencies obtained from the analysis of *p*- and *s*-polarized single transmission IR spectra of the Si bonds with buried SiO_2 the absorption line at 1620 cm^{-1} is related to $2TO_2$ mode.

TABLE I. Identification of multiphonon absorption lines in SiO_2.

Observed peak frequency (cm^{-1})	Identification	Calculated peak frequency (cm^{-1})
1620	$2TO_2$	1632
1867	TO_1+TO_2	1866
2011	LO_1+TO_2	2039

The decreasing absorption in the region 2850–2950 cm^{-1} manifests the lower carbon concentration in annealed samples (Fig. 3, curve 1) in comparison with the as-prepared samples (Fig. 2, curves 1 and 2). It means that the annealing causes partial desorption of the CH_x species from Si surface.

The use of the double bonded Si sample geometry allows the influence of signals from external surfaces to be excluded because the light undergoes a multiple internal reflection in the middle silicon wafer cladded by silicon oxide layers. Figure 3 shows *p*-polarized IR spectrum of this structure (curve 2) in comparison with the spectra of the Si wafers covered by thermal SiO_2 layers. The observed absorption bands are shown in details in the inserts (a) and (b) where normalized spectra taken as

$$\frac{\Delta R}{R}=\frac{R_b-R_0}{R_0}, \tag{1}$$

where R_b and R_0 are the IR spectra of the double bonded Si wafer and a Si wafer with SiO_2 layer, respectively. A new absorption line at 2250 cm^{-1} appears in the spectra of Si wafer bond which was attributed to the Si–H stretching vibrations in the O_3–Si–H complexes. This mode is not observed in the IR spectra of the annealed structure (Fig. 3, curve 1) and thus it originates from Si–SiO_2 interface region located between bonded wafers. It does not contradict previous results[21] where this mode is assigned to Si–H stretch modes of hydrogen inside the oxide layer because the thickness of this layer was of order of 1 nm. We propose that the wafer molecules adsorbed at the surfaces of wafers before bonding hydrolyses to siloxane groups at the interface of Si wafer bonds.[4,19,21] It also leads to a decrease in absorption up to 2% in the range of 3000–3600 cm^{-1} caused by the decreasing amount of hydroxyl species at the interface after bonding. The absorption line at 3739 cm^{-1} of the wafers with SiO_2 layers is due to hydrogen-bonded hydroxyl (–OH) groups at the two oxide surfaces.[4] The reason of increasing absorption by combinative SiO_2 modes in Si wafer bonds is not clear up to now. It can probably be caused by a change of stoichiometric structure of interface silicon oxide in comparison with the thermally grown SiO_2 layer. To clarify this effect, further investigations are needed.

The frequency position of the absorption feature at 2250 cm^{-1} observed in the spectra of Si wafer bonds is shifted towards higher frequency with annealing of the sample. Figure 4 shows this shift at the spectral range of Si–H vibrations. One can see that with increasing annealing temperature the mode frequency is shifted from 2250 cm^{-1} (450 °C) to 2280 cm^{-1} (1100 °C). It is known that the hydrogen in the oxide layer causes the absorbance in the spectral range of 2260–2300 cm^{-1}.[4] We propose that the hydrogen atoms located near the buried interface diffuse into the oxide layer producing Si–H bonds in SiO_2 in the vicinity of the interface.

The arrangement of chemical species at the buried interface of the Si wafers bonded at high (800 °C) and low

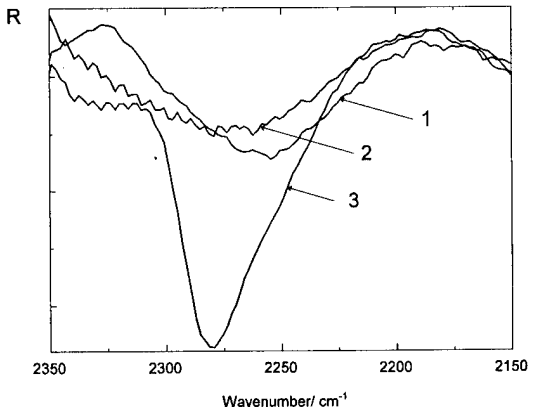


FIG. 4. IR *p*-polarized multiple internal transmission spectra of the double bonded Si structure annealed at 400 °C (curve 1), 800 °C (curve 2), 1100 °C (curve 3) measured at the spectral range of the Si–H vibrations.

(200 °C) temperature was probed using the single-pass multiple internal reflection elements. The number of passes through the interface was about 100. The IR spectrum of Si wafers bonded at high temperature (dotted curve in Fig. 5) is very similar to the spectrum of the hydrophilic Si wafer (curve 2 in Fig. 2). No evidence of Si–H modes at the interfaces was found. A weak feature at 3750 cm^{-1}, characteristic of isolated silanol groups at the partially oxidized interface,[12] probably formed during rinsing prior bonding or appeared due to formation of Si–O–Si bridges between joined wafers after the bonding procedure. The IR spectrum of Si wafers bonded at low temperature (solid curve in Fig. 5) changed drastically in comparison with the spectrum of the hydrophilic Si wafer and the high temperature bonds. One can see from Fig. 5 that absorbance at the frequency range from 3200 to 3600 cm^{-1} due to OH vibrations at the buried interface is a few times larger than in the case of the single Si wafer and the high temperature Si bonds. This is in accordance with previous experiments[12,14] and the model proposed by Tong and Gösele[3] which demonstrates the decreasing interface layer of the physically absorbed water molecules

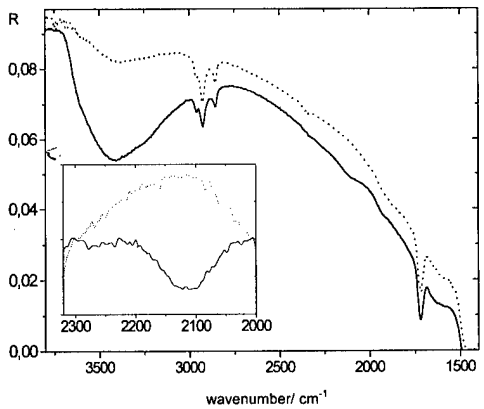


FIG. 5. IR *p*-polarized transmission spectra of the Si bonded wafers joined at high (800 °C) and low (200 °C) temperature measured with the single-pass multiple internal reflection elements.

from 1 up to 0.15 nm with an increasing bonding temperature from 20 up to 800 °C. In addition, new modes occur in the IR spectrum of the high temperature Si bond at 2110 and 2250 cm^{-1} (see the insert of Fig. 5). The mode at 2110 cm^{-1} was assigned to Si–H stretch mode in dihydride ($Si–H_2$) species[22] at the interface. The partial oxidation of the buried interface during bonding leads also to formation of Si–H stretch modes in silicon dioxide at the buried interface.

IV. CONCLUSION

In conclusion, we present the results of a systematic IR investigation of hydrophilic and hydrophobic Si wafers, Si wafers covered by thermally grown SiO_2 layer and bonded Si wafers. It was shown that the nature of interface vibrational states on the Si surface strongly depends on the surface treatment. IR spectra of the samples show the presence of the physically absorbed water molecules and $C–H_x$ species at the interfaces. The surface of hydrophobic Si wafers is terminated by hydrogen. After storing in air oxidation of the hydrophobic samples takes place and as a result the absorption lines due to Si–H vibrations in the $Si–H–(O_3)$ and $SiH_2(O_2)$ complexes appear. Annealing of the hydrophilic and hydrophobic Si wafers leads to the same surface arrangement for both samples. The hydrogen atoms are desorbed and the oxidation process takes place. A number of combinational phonon modes in SiO_2 observed in the IR spectra of Si wafer with thermally grown SiO_2 layer was identified.

IR spectra of Si wafers covered by SiO_2 layers bonded at high temperature demonstrate the presence of physically adsorbed water molecules and C–H species, hydroxyl and siloxane groups at the buried interface. It was shown that the bonding process leads to the appearance of hydrogen-bonded hydroxyl (–OH) and siloxane groups in silicon dioxide at the interfaces. IR spectra of the low temperature bonds reveal in addition the Si–H stretch modes at the interface. This is probably due to partial formation of the silicon oxide at the buried interface which takes place at the low temperature bonding.

ACKNOWLEDGMENT

This research was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 379.

[1]U. Gösele and H. Stenzel, *Proceedings of the 3rd International Symposium on Semiconductor Wafer Bonding: Science, Technology, and Application* (The Electrochemical Society, Pennington, NJ, 1995), Vol. 95-7, p. 33.
[2]M. K. Weldon, V. E. Marsico, Y. J. Chabal, A. Agarwal, D. J. Eaglesham, J. Sapjeta, W. L. Brawn, D. C. Jacobson, Y. Caudano, S. B. Christman, and E. E. Chaban, J. Vac. Sci. Technol. B **15**, 1065 (1997).
[3]Q.-Y. Tong and U. Gösele, Ref. 1, p. 78.
[4]M. K. Weldon, V. E. Marsico, Y. J. Chabal, D. R. Hamann, S. B. Christman, and E. E. Chaban, Surf. Sci. **368**, 163 (1996).
[5]Y. J. Chabal, Surf. Sci. Rep. **8**, 211 (1988).
[6]L. Ling, S. Kuwabara, T. Abe, and F. Shimura, J. Appl. Phys. **73**, 3018 (1993).
[7]M. Niwano, J. Kageyama, K. Kinashi, and N. Miyamoto, J. Vac. Sci. Technol. A **12**, 465 (1994).
[8]M. K. Weldon, B. B. Stefanov, K. Raghavachari, and Y. J. Chabal, Phys. Rev. Lett. **79**, 2851 (1997).

[9]Y. J. Chabal, G. S. Higashi, K. Raghavachari, and V. A. Burrow, J. Vac. Sci. Technol. A **7**, 2104 (1989).
[10]Y. J. Chabal, Surf. Sci. **168**, 594 (1986).
[11]M. Niwano, J. Kageyama, K. Kurita, K. Kinashi, I. Takahashi, and N. Miyamoto, J. Appl. Phys. **76**, 2157 (1994).
[12]D. Feijoo, Y. J. Chabal, and S. B. Christman, Appl. Phys. Lett. **65**, 2548 (1994).
[13]Y. J. Chabal, D. Feijoo, S. B. Christman, and C. A. Goodwin, Ref. 1, p. 305.
[14]M. K. Weldon, Y. J. Chabal, S. B. Christman, E. E. Chaban, L. C. Feldman, and D. R. Hamann, J. Vac. Sci. Technol. B **14**, 3095 (1996).
[15]N. J. Harrik, *Internal Reflection Spectroscopy* (Wiley, New York, 1967).
[16]R. K. Iler, *The Chemistry of Silica* (Wiley, New York, 1979), pp. 622–729.
[17]*Sadtler Handbook of Infrared Spectra*, edited by W. W. Simons (Sadtler, Philadelphia, PA, 1978).
[18]K. Krishnan, P. J. Stout, and M. Watanabe, in *Practical Fourier Transform Infrared Spectroscopy*, edited by J. R. Ferraro and K. Krishnan (Academic, San Diego, 1990), pp. 285–349.
[19]P. Dumas, Y. J. Chabal, and P. Jakob, Surf. Sci. **269/270**, 867 (1992).
[20]S. M. Han and E. S. Aydil, Appl. Phys. Lett. **70**, 3269 (1997).
[21]H. Ogawa and T. Hattori, Appl. Phys. Lett. **61**, 577 (1992).
[22]Y. J. Chabal, M. A. Hines, and D. Feijoo, J. Vac. Sci. Technol. A **13**, 1719 (1995).

Infrared and Raman studies of confined and interface optical phonons in short-period GaAs/AlAs superlattices with a grating coupler

A. Milekhin[a)]
Institut für Physik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany and Institute of Semiconductor Physics, 630090 Novosibirsk, Russia

M. Rösch and E. Batke
Physikalisches Institut der Universität Würzburg, D-97074 Würzburg, Germany

D. R. T. Zahn
Institut für Physik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany

K. Köhler and P. Ganser
Fraunhofer-Institut für Angewandte Festkörperphysik, D-79108 Freiburg, Germany

V. Preobrazhenskii and B. Semyagin
Institute of Semiconductor Physics, 630090 Novosibirsk, Russia

(Received 19 January 1999; accepted 17 May 1999)

The optical phonon modes in short-period GaAs/AlAs superlattices were investigated by means of Raman spectroscopy and far-infrared reflection spectroscopy using a grating coupler technique. Grating couplers provide a wave vector transfer parallel to the layers, enabling study of in-plane dispersion of the phonon modes in the superlattice structure. The mixing of confined and interface optical modes was observed as the in-plane phonon wave vector was varied. The experimental frequencies of the superlattice optical modes obtained from infrared and Raman spectra are in good agreement with calculations performed using a dielectric continuum model. © *1999 American Vacuum Society.* [S0734-211X(99)07204-2]

I. INTRODUCTION

The properties of phonons in semiconductor quantum wells and superlattices continue to generate significant theoretical and experimental interest.[1,2] Short-period superlattices composed of the polar semiconductors GaAs and AlAs essentially display two characteristic differences in the behavior of optical vibrations compared with their bulk constituents.

The optical phonons become localized within the corresponding superlattice layer due to the large difference between cationic masses in GaAs and AlAs. This leads to quantization of the phonon spectrum and a set of confined modes within each compound material.

Interface effects on the confined modes arise due to different dielectric functions of the participating layer material. The dielectric function of the superlattice is presented as an anisotropic tensor, and confined mode frequencies show dispersions when the orientation of the phonon wave vector q is tilted from the layer structure growth axis z toward the in-plane direction x. These modes propagating along the layers are referred to as interface modes.

The behavior of confined optical phonons in superlattices is well documented and understood (see Ref. 1) but experiments that directly probe in-plane dispersion are rare.[3–6] It was shown that the odd confined modes disperse strongly with an increasing in-plane wave vector and anticross with one another, while even confined modes remain dispersionless.[4,6] This agrees with continuum models predicting the intermixing of confined longitudinal optical (LO), transverse optical (TO), and interface modes.[7,8] Up to now only Raman data were available for the study of in-plane dispersion of optical phonons in superlattices.

In this article we present a new approach for the characterization of optical phonon in-plane dispersion in GaAs/AlAs superlattices by grating coupler assisted infrared (IR) reflection spectroscopy, a technique traditionally used to investigate electronic properties of low-dimensional systems.[9]

The optical vibrational modes propagating along the layers (interface modes) do not usually appear in the infrared because of the small component of the electric field perpendicular to the superlattice layer plane. However, they can be observed in IR experiments due to additional in-plane wave vector transfer on the phonon system which is provided by means of a grating coupler deposited onto the superlattice (Fig. 1). A grating coupler modulates the far-infrared field in the vicinity of the grating producing a significant component of the electric field perpendicular to the layer plane. It allows one to probe the excitations propagating in plane of the layers with the wave vector

$$q_x = \frac{\omega}{c}\sin\varphi + \frac{2\pi n}{d}, \tag{1}$$

where φ is the angle of incidence of the light, c is the light velocity, d is the period of the grating coupler, and n is an integer.[10] Thus, IR reflection spectra of GaAs/AlAs superlattices using grating couplers with different periods and/or

a)Electronic mail: a.milekhin@physik.tu-chemnitz.de

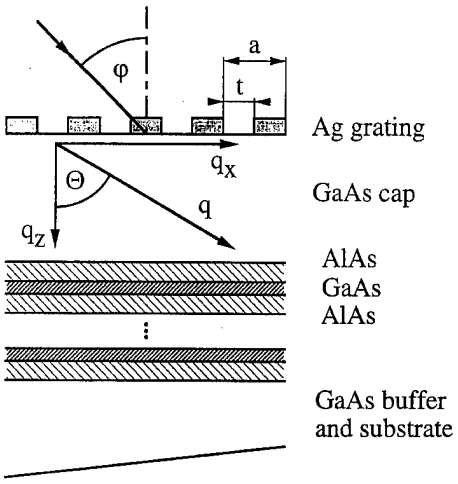


FIG. 1. Schematic cross section of the sample with a grating coupler. The GaAs substrate is wedged in order to avoid the influence of Fabry–Pérot resonances in the experiments.

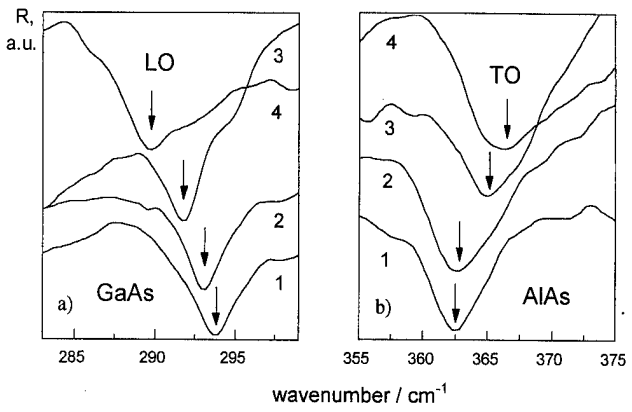


FIG. 2. Reflection spectra of sample A measured in the range of (a) LO vibrations of GaAs and (b) TO phonons of AlAs at $T=9$ K. Curve (1) was measured without a grating coupler whereas curves (2)–(4) were measured using a grating coupler with periods of 20, 6, and 2 μm. The vertical arrows indicate the position of the confined mode frequencies.

higher Fourier components make it possible to investigate the in-plane wave vector dispersion of optical phonons in a wide range.

II. EXPERIMENTAL DETAILS

Figure 1 displays a schematic cross section of our molecular beam epitaxy (MBE) grown samples. Two samples, A and B, were studied with a $(GaAs)_5(AlAs)_7$ superlattice period on doped GaAs substrates (sample A: $N_{Si}\approx 1\times10^{18}\,cm^{-3}$; sample B: $N_{Si}\approx 2\times10^{17}\,cm^{-3}$) covered by an undoped GaAs buffer (sample A: 1.5 μm, sample B: 1 μm). Sample A was grown without rotating the substrate during growth whereas sample B was rotated. The numbers of periods were 8 (sample A) and 55 (sample B). On the top of structure A, Ag grating couplers with periods $d=2$, 6, and 20 μm were fabricated. On sample B only a 6 μm grating was prepared.

The infrared reflection spectra were recorded using a Fourier transform IR (FTIR) spectrometer Bruker IFS-113 V equipped with an Oxford Instruments flow cryostat at a temperature 9 K. The resolution was 0.5 cm^{-1} over the entire spectral range. The reflection spectra of the superlattice with and without the grating couplers were measured at oblique incidence ($\varphi\approx70°$) with p-polarized spectra to increase the in-plane component of the wave vector.

The Raman scattering experiments were performed at $T=300$ K using the 514.5 nm line of an Ar^+ laser with a power of 100 mW. The scattered light was analyzed in backscattering geometry using a Dilor XY triple monochromator equipped with a charge coupled device (CCD) camera for multichanel detection. The scattering geometries employed were $z(xx)$-z and $z(xy)$-z with x, y, and z parallel to the [100], [010], and [001] directions, respectively. The monochromator slits were set to a spectral resolution of 2.9 cm^{-1}.

III. RESULTS AND DISCUSSION

Figure 2 shows the experimental reflection spectra of sample A without and with grating couplers of different periods, measured at off-normal incidence conditions in the range of the LO phonons of GaAs [Fig. 2(a)] and the TO phonons of AlAs [Fig. 2(b)]. Without the grating coupler, i.e., small q_x, the fundamental confined phonon (LO_1) of GaAs is observed and identified by its frequency, as is the TO_1, phonon of AlAs. These frequencies can be found from the bulk dispersion for the effective wave vectors:[1,11]

$$k_m=\frac{\pi}{(n_0+\delta)a_0}m,\quad m=1,2,...,n_0, \tag{2}$$

where m is the index of the confined mode, a_0 is the monolayer thickness, and n_0 is the number of monolayers in the corresponding sample constituent. Parameter δ [$\delta\approx0.8$ for the GaAs/AlAs system (see Ref. 11)] reflects the penetration depth of the confined modes in the neighboring layers. Only the odd numbered confined modes are accessible by IR spectroscopy due to their nonzero macroscopic dipole moment. In our experiments, the existence of high-index ($m\geq3$) odd modes was not detected.

For a larger wave vector transfer q_x via the grating coupler, only the higher frequency confined mode LO_1 is still observable in sample A. In Fig. 2 it is clearly seen that the frequency of this LO phonon mode of GaAs decreases with a decrease in grating period a, i.e., with increasing wave vector transfer q_x. No phonon modes at higher n values [see Eq. (1)] of q_x were observed in these experiments.

In the frequency range of the optical vibrational modes of AlAs, the confined TO_1 phonon and LO_1 (not shown here) are identified for the limit $q_x\rightarrow0$. Again, only the fundamental confined phonons were detected. If wave vector transfer parallel to the layers is imposed on the sample via the grating coupler, in sample A the TO_1 phonon still remains observable and the frequency of this phonon increases with an increase of q_x.

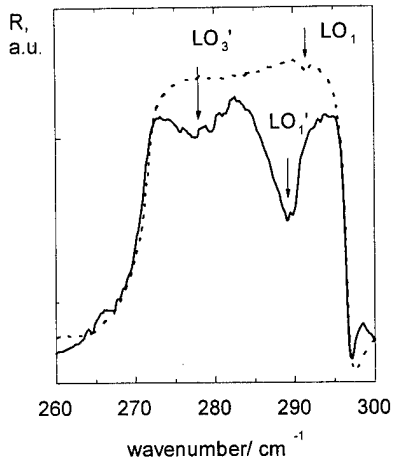


FIG. 3. Reflection spectra of sample B measured in the range of optical vibrations of GaAs at $T=9$ K. The dotted line presents the IR spectrum of the sample without a grating coupler. The spectrum of the same sample with a grating coupler period of 6 μm is shown by a solid line.

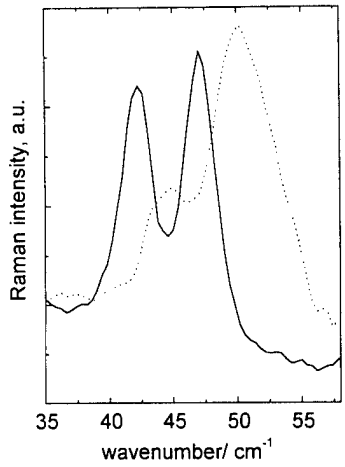


FIG. 5. Raman spectra of sample A (solid line) and sample B measured parallel in $z(x,x)$-z geometry (dotted line) in the spectral region of acoustic phonons.

The large number of superlattice periods makes it possible to observe the high-order confined optical modes (with $m=3$) in the IR reflection spectra of sample B. Figure 3 shows IR spectra of sample B with and without the grating coupler measured in the spectral range of the "reststrahlen" band of GaAs. The IR spectrum without the grating (dashed line) reveals only a weak feature at a frequency near 292 cm^{-1} which was attributed to the LO_1 confined phonon in GaAs layers. Dramatic changes occur in the spectra with the grating coupler. The frequency position of the LO_1 mode was shifted towards lower frequency (the feature denoted as LO_1' in Fig. 3) and this mode becomes much stronger. Moreover, a new mode assigned to the LO_3 mode with $q_x \neq 0$ (LO_3' in Fig. 3) was observed at a frequency near 277 cm^{-1}. The even confined modes cannot be observed in infrared spectra due to their zero dipole momentum.

The selection rules for Raman spectroscopy allow the even modes localized in these structures to be observed in $z(x,x)$-z geometry while odd confined modes are active in $z(x,y)$-z geometry. Figures 4(a) and 4(b) show Raman spectra of samples A and B measured in both geometries in the spectral range of optical phonons of GaAs, respectively. The dominant features in the Raman spectra of sample A, shown by arrows, are due to scattering by the LO_1 and LO_2 confined phonons. The shoulder in the spectra of sample B denoted as LO_b corresponds to LO phonon scattering from bulk GaAs. The broad feature at a frequency of about 276 cm^{-1} was attributed to scattering by GaAs interface phonons. The strong scattering by LO phonons from the substrate prevents the observation of the LO_1 phonons from the GaAs superlattice layers of sample B. In order to obtain the frequency position of optical phonons a fitting of Raman spectra by Lorentzian peaks, shown by dotted lines in Fig. 4, was performed.

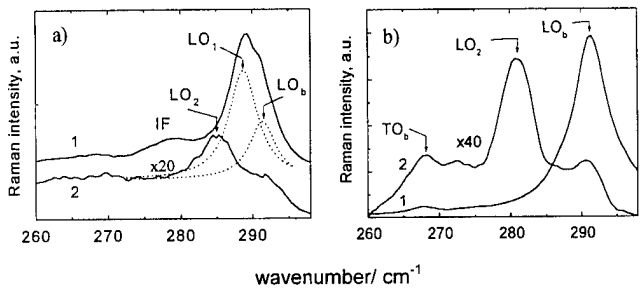


FIG. 4. Raman spectra of (a) sample A and (b) sample B in the GaAs phonon region using backscattering geometry. Curves (1) and (2) correspond to crossed $z(x,y)$-z and parallel $z(x,x)$-z geometries, respectively. The positions of arrows correspond to the frequencies of the optical phonons which were determined by fitting of the Raman spectra by the Lorentzian curves. The Lorentzian lines used for fitting are shown by dotted lines.

The comparison of Raman spectra of samples A and B shows that the frequency position of the GaAs LO_2 phonon in sample A is about 4 cm^{-1} higher than that in sample B. This fact can be explain by a deviation in thickness of GaAs layers in the samples with respect to the nominal one. To verify this the Raman scattering was measured in the acoustic spectral range in $z(x,x)$-z geometry. The Raman spectra of samples A and B presented in Fig. 5 show doublets centered at frequencies of 44 and 46 cm^{-1}, respectively. These doublets correspond to an excitation of folded longitudinal-acoustic (LA) phonons propagating perpendicular to the superlattice layers. Using the dispersion relation for acoustic phonons[12] we have determined the periods of the superlattices in samples A and B as 14 and 13 monolayers (ML), respectively. An analysis of Raman spectra of folded acoustic and confined optical phonons (LO_1, LO_2, and LO_3) obtained from Raman and IR measurements was performed in the same way as in Ref. 13. It allowed us to determine the thicknesses of the GaAs and AlAs layers as $d_1=6$ ML and $d_2=8$ ML and $d_1=5$ ML and $d_2=8$ ML for samples A and B respectively. The difference between the nominal and the actual parameters of sample A may be caused by monolayer fluctuations of the layer thickness because the sample was grown without being rotated. In the calculations we have used the optical phonon frequencies measured by Raman

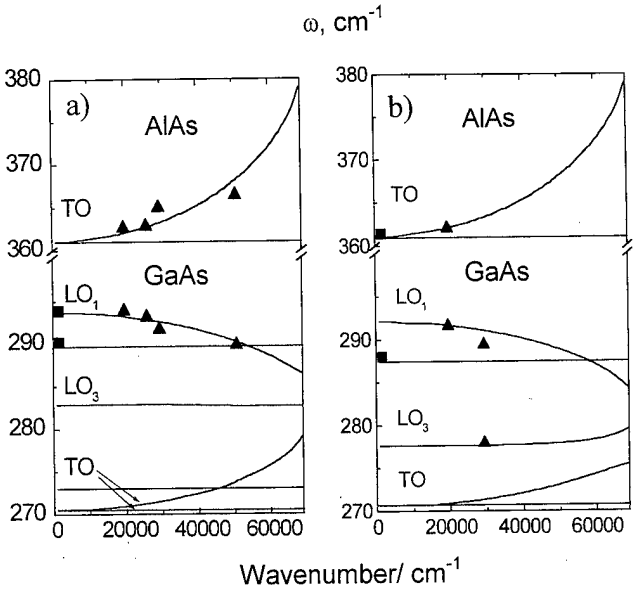


FIG. 6. Dependence of the optical phonon frequencies with in-plane wave vector q_x for (a) sample A and (b) sample B. The experimental Raman and IR data are presented by squares and triangles, respectively.

scattering corrected for a change in temperature (300–10 K) simply by adding 5.2 cm^{-1}, a procedure justified by the known temperature dependence of the bulk modes.[14]

In order to explain the behavior of optical phonons in the structures we have calculated the frequency dependence of optical modes on the in-plane wave vector using a continuum model proposed by Chamberlain *et al.*[15] In this model the optical modes of the superlattice are described as a linear combination of the longitudinal-optical, transverse-optical, and interface phonon modes which satisfy both electrostatic and mechanical conditions at the interfaces.

Figure 6 shows the calculated in-plane wave vector dispersion of the optical modes for, samples A and B with the actual thickness of the layers. The experimental data are shown here by symbols. The value of the in-plane wave vector corresponding to the experimental optical modes was calculated using Eq. (1). For Raman experiments in backscattering geometry q_x is zero. The parameters used for the calculations were taken from Ref. 15. As one can see from Fig. 6, the odd-order confined modes, in particular LO_1, undergo significant dispersion while the even-order modes remain dispersionless. Strictly speaking, classification of the mixed modes in terms of polarization and mode number is valid only for $q_x=0$. Nevertheless, for identification purposes we label them by their number at $q_x=0$. With an increasing in-plane wave vector the frequency of the LO_1 modes decreases whereas the TO_1 mode splits into two branches, one dispersionless (polarized along the y direction) and the other with an upward dispersion (polarized along the z direction). Interaction does not occur for modes with different symmetry.

One can see from Fig. 6 that good agreement between the calculated dispersion and experimental data was obtained within experimental error.

IV. CONCLUSION

We have presented the results of IR and Raman experiments in short-period GaAs/AlAs superlattices. Raman experiments allow one to observe zone-center confined optical phonons and folded acoustic phonons in the superlattices. The actual thickness of the layers was obtained from an analysis of frequencies of confined optical and folded acoustic phonons. The in-plane wave vector dispersion of optical phonons in short-period superlattices was studied by grating coupler assisted infrared reflection spectroscopy. The experimental data are in good agreement with the calculations performed with in the frame work of a continuum model.

ACKNOWLEDGMENTS

One of the authors (A.M.) gratefully acknowledges the hospitality of the members of the Lehrstuhl für Technische Physik at the Physikalisches Institut der Universität Würzburg and also thanks Dr. D. Tenne for fruitful discussions. The experimental work was partly supported by the Deutsche Forschungsgemeinschaft.

[1]M. Cardona and G. Güntherodt, *Light Scattering in Solids V*, Topics in Applied Physics, Vol. 66 (Springer, Berlin, 1989).
[2]T. Ruf, *Phonon Raman Scattering in Semiconductors, Quantum Wells and Superlattices: Basic Research and Application*, Springer Tracts in Modern Physics, Vol. 142 (Springer, Berlin, 1998).
[3]A. Huber, T. Egeler, W. Ettmüller, H. Rothfritz, G. Tränkle, and G. Abstreiter, Superlattices Microstruct. **9**, 309 (1991).
[4]R. Hessmer, A. Huber, T. Egeler, H. Haines, G. Tränkle, G. Weimann, and G. Abstreiter, Phys. Rev. B **46**, 4071 (1992).
[5]G. Scamarcio, M. Heines, G. Abstreiter, E. Molinari, S. Baroni, A. Fisher, and K. Ploog, Phys. Rev. B **47**, 1483 (1991).
[6]M. Zunke, R. Schorer, G. Abstreiter, W. Klein, G. Weimann, and M. P. Chamberlain, Solid State Commun. **93**, 847 (1995).
[7]N. C. Constantinou, O. Al-Dossary, and B. K. Ridley, Solid State Commun. **86**, 191 (1993).
[8]C. Trallero-Giner, F. Garcia-Moliner, V. R. Velasko, and M. Cardona, Phys. Rev. B **45**, 11944 (1992).
[9]E. Batke and D. Heitmann, Infrared Phys. **24**, 189 (1984).
[10]L. Zheng, W. L. Schaich, and A. H. Mac Donald, Phys. Rev. B **41**, 8493 (1990).
[11]B. Samson, T. Dumelow, A. A. Hamilton, T. J. Parker, S. R. P. Smith, D. R. Tilley, C. T. Foxon, D. Hilton, and K. J. Moore, Phys. Rev. B **46**, 2375 (1992).
[12]M. Rytov, Akust. Zh. **2**, 71 (1956) [Sov. Phys. Acoust. **2**, 68 (1956)].
[13]V. A. Haisler, D. A. Tenne, N. T. Moshegov, A. I. Toropov, A. P. Shebanin, and A. A. Yaskin, Phys. Solid State **38**, 1235 (1996).
[14]B. Jusserand and J. Sapriel, Phys. Rev. B **24**, 7194 (1981).
[15]M. P. Chamberlain, M. Cardona, and B. K. Ridley, Phys. Rev. B **48**, 14356 (1993).

Spontaneous and piezoelectric polarization effects in III–V nitride heterostructures

E. T. Yu,[a] X. Z. Dang, P. M. Asbeck, and S. S. Lau
Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, California 92093-0407

G. J. Sullivan
Rockwell Science Center, Thousand Oaks, California 91358

(Received 19 January 1999; accepted 3 May 1999)

The role of spontaneous and piezoelectric polarization in III–V nitride heterostructures is investigated. Polarization effects and crystal polarity are reviewed in the context of nitride heterostructure materials and device design, and a detailed analysis of their influence in nitride heterostructure field-effect transistors is presented. The combined effects of spontaneous and piezoelectric polarization are found to account well for carrier concentrations observed in AlGaN/GaN transistor structures with low to moderate Al concentrations, while the data for higher Al concentrations are consistent with defect formation in the AlGaN barrier. Theoretical analysis suggests that incorporation of In into the barrier and/or channel layers can substantially increase polarization charge at the heterojunction interface. The use of polarization effects to engineer Schottky barrier structures with large enhancements in barrier height is also discussed, and electrical characteristics of transistors with conventional and polarization-enhanced Schottky barrier gates are presented. The polarization-enhanced barrier is found to yield a marked reduction in gate leakage current, but to have little effect on transistor breakdown voltage. © *1999 American Vacuum Society.* [S0734-211X(99)05804-7]

I. INTRODUCTION

III–V nitride heterostructures are of outstanding current interest for a wide range of device applications, including blue and ultraviolet light-emitting diodes and lasers,[1] high-temperature/high-power electronics,[2–6] visible-blind ultraviolet photodetectors,[7,8] and field-emitter structures.[9,10] In addition, these materials, by virtue of their wurtzite crystal structure and high degree of ionicity, exhibit a variety of material properties that either are not found or are of considerably reduced importance in conventional zincblende III–V semiconductors. Of particular interest are piezoelectric and spontaneous polarization effects, which recent experimental and theoretical investigations have revealed to be of great importance in the design and analysis of nitride heterostructure devices,[11–14] and which can be exploited to advantage in nitride materials and device engineering.

In this article, we review the basic phenomena of spontaneous and piezoelectric polarization in nitride semiconductors and discuss their role in nitride heterostructure device physics, with particular emphasis on the design and analysis of nitride-based heterostructure field-effect transistors (HFETs). Section II provides an overview of piezoelectric and spontaneous polarization effects in nitride heterostructures, the influence of crystal polarity, and the role of polarization effects in the analysis and design of various nitride heterostructure devices. Section III focuses specifically on the role of spontaneous and piezoelectric polarization in the analysis and design of nitride HFETs; the use of polarization effects to engineer Schottky barrier structures in nitride HFETs is addressed in Sec. IV. Section V concludes the article.

II. OVERVIEW OF POLARIZATION EFFECTS

The wurtzite crystal structure of nitride semiconductors combined with epitaxial growth that is performed typically in the (0001) orientation leads to the existence of a piezoelectric polarization field and associated electrostatic charge densities in strained material that have been shown to influence carrier distributions, electric fields, and consequently a wide range of optical and electronic properties of nitride materials and devices. In addition, recent theoretical results have indicated that nitride semiconductors also possess a large spontaneous polarization,[15] associated with which will be electrostatic charge densities analogous to those produced by piezoelectric polarization fields. In bulk material it is assumed that rearrangement of surface charges nullifies spatially uniform piezoelectric and spontaneous polarization fields. In heterostructures or inhomogeneous alloy layers, however, variations in composition are expected to create nonvanishing spontaneous and piezoelectric polarization fields and associated charge densities that can dramatically influence material properties and device behavior.

The piezoelectric polarization field $\mathbf{P}_{pz}$ is determined by the piezoelectric coefficients e_{ij} and the strain tensor ϵ_j, and is given by

$$\mathbf{P}_{pz}=e_{ij}\epsilon_j=d_{ij}c_{jk}\epsilon_k, \quad j,k=xx,yy,zz,yz,zx,xy, \tag{1}$$

where d_{ij} are the piezoelectric coefficients relating the polarization to the stress tensor $\sigma_j=c_{jk}\epsilon_k$, and c_{jk} is the elastic tensor. Contracted matrix notation is used, and summation

[a]Electronic mail: ety@ece.ucsd.edu

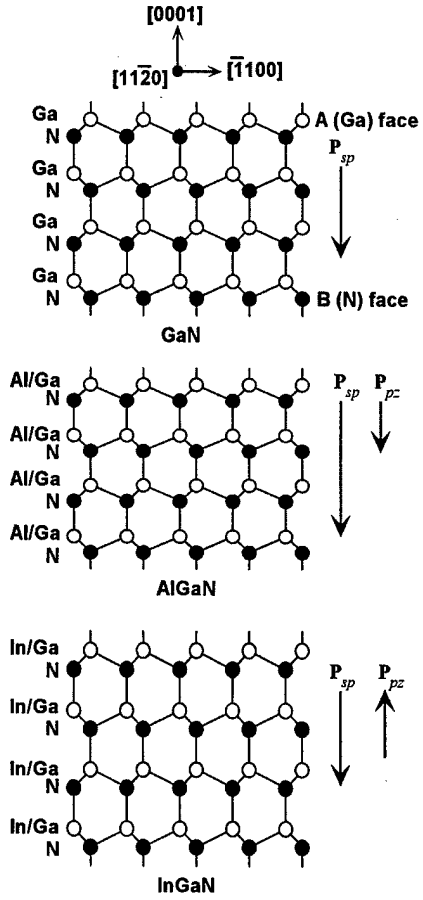


FIG. 1. Crystal structure, spontaneous polarization fields ($\mathbf{P}_{sp}$) and piezoelectric polarization fields ($\mathbf{P}_{pz}$) for GaN (top), and for $Al_xGa_{1-x}N$ and $In_yGa_{1-y}N$ coherently strained to GaN (0001).

over repeated indices is assumed. The total polarization field $\mathbf{P}$ is then the sum of $\mathbf{P}_{pz}$ and the spontaneous polarization $\mathbf{P}_{sp}$, and is related to the polarization charge density ρ_{pol} according to

$$\nabla \mathbf{P} = \nabla \cdot (\mathbf{P}_{sp} + \mathbf{P}_{pz}) = -\rho_{pol}. \tag{2}$$

From Eq. (2) it is evident that at an abrupt heterojunction interface, a polarization sheet charge density will be present arising from the difference in spontaneous polarization between the two heterojunction constituents and from the abrupt change in strain ϵ_j that gives rise to a discontinuity in $\mathbf{P}_{pz}$. In compositionally graded alloy layers, the spatially varying polarization field will produce either positive or negative volume charge densities that will in many respects act as donors or acceptors, respectively. Finally, it should be noted that very gradual changes in $\mathbf{P}$, as would arise from uniform or nearly uniform stress in an epitaxial layer structure, will produce only very small polarization charge densities.

Closely related to the investigation of polarization and its influence on nitride heterostructure device physics is the issue of crystal polarity. Figure 1 shows a schematic diagram of the GaN wurtzite crystal structure. The lack of inversion symmetry in the (0001) plane gives rise to two possible polarities for (0001) epitaxial growth. The upper (0001) surface in the figure is referred to as the A (or Ga) face, and the lower (000$\bar{1}$) surface as the B (or N) face. It has been found empirically that the orientation of high-quality nitride films grown by metalorganic chemical vapor deposition (MOCVD) is typically (0001), whereas molecular-beam epitaxy (MBE) can yield high-quality films with either the (0001) or the (000$\bar{1}$) orientation. (0001) and (000$\bar{1}$) films can be distinguished from each other by a variety of techniques including wet chemical etching, characterization of physical morphology, and convergent beam electron diffraction.[16]

Polarity in epitaxially grown nitride films is significant in the context of polarization because the spontaneous and piezoelectric polarization fields have a well-defined orientation with respect to the A and B crystal faces. Theoretical calculations[15] indicate that for GaN, AlN, and InN the spontaneous polarization field is in the [000$\bar{1}$] direction, i.e., pointing from the A face to the B face as indicated in Fig. 1. Experimental[12,13] and theoretical[15] results have indicated that the signs of the relevant piezoelectric coefficients in the nitrides are such that for (0001) films grown under tensile or compressive strain, $\mathbf{P}_{pz}$ is in the [000$\bar{1}$] or [0001] direction, respectively. The orientation of the piezoelectric polarization field with respect to the crystallographic axes in the nitride semiconductors is opposite that found in other III–V semiconductors and is the same as that found in II–VI semiconductors, a consequence of the greater ionicities of III–V nitrides compared to those of other III–V semiconductors.[13,15,17]

For strained III–V nitride epitaxial layers grown in the (0001) orientation, a piezoelectric polarization will be present aligned along the [0001] direction and given by

$$P_{pz,z} = 2\left(e_{31} - \frac{c_{13}}{c_{33}} e_{33}\right)\epsilon_1. \tag{3}$$

The values of the spontaneous polarization and of the relevant piezoelectric and elastic constants in III–V nitrides have in many cases not been definitively established. Table I shows values for several relevant physical quantities for GaN, AlN, and InN.[18–32] In this work, we have used theoretical values[15] for spontaneous polarization and for the piezoelectric coefficients e_{31} and e_{33}. For spontaneous polarization, only a single set of theoretical values, and no experimental values, are available. For the piezoelectric coefficients, only theoretical values are available for InN; for AlN, experimental and theoretical values appear to be in good agreement. In the case of GaN, there is a significant discrepancy between the experimental and theoretical values, particularly for e_{31}; for consistency, and given the large uncertainty likely to exist for both the experimental and theoretical values, we have used the theoretical values for e_{31} and e_{33}. For the elastic constants c_{13} and c_{33}, we have used the averages of the values listed in Table I. Values for ternary and quaternary alloys are then obtained by simple linear interpolation.

Recent work has demonstrated that polarization effects can exert a pronounced influence on heterostructure material

TABLE I. Selected physical constants for GaN, AlN, and InN.

	GaN	AlN	InN	Reference
a (Å)	3.189	3.112	3.548	18
c (Å)	5.185	4.982	5.760	18
e_{31} (C/m^2)	−0.32	⋯	⋯	20
e_{31} (C/m^2)	−0.22	⋯	⋯	21
e_{31} (C/m^2)	−0.36	⋯	⋯	19, 22
e_{31} (C/m^2)	⋯	−0.58	⋯	23
e_{31} (C/m^2)	−0.49	−0.60	−0.57	15
e_{33} (C/m^2)	0.65	⋯	⋯	20
e_{33} (C/m^2)	0.44	⋯	⋯	21
e_{33} (C/m^2)	1	⋯	⋯	19, 22
e_{33} (C/^{2})	⋯	1.55	⋯	23
e_{33} (C/m^2)	0.73	1.46	0.97	15
c_{13} (GPa)	158	⋯	⋯	24
c_{13} (GPa)	70	⋯	⋯	25
c_{13} (GPa)	106	⋯	⋯	26
c_{13} (GPa)	114	⋯	⋯	27
c_{13} (GPa)	110	100	⋯	28
c_{13} (GPa)		120	⋯	29
c_{13} (GPa)		99	⋯	30
c_{13} (GPa)	100	127	94	31
c_{13} (GPa)	103	108	92	32
c_{33} (GPa)	267	⋯	⋯	24
c_{33} (GPa)	379	⋯	⋯	25
c_{33} (GPa)	398	⋯	⋯	26
c_{33} (GPa)	209	⋯	⋯	27
c_{33} (GPa)	390	390	⋯	28
c_{33} (GPa)	⋯	395	⋯	29
c_{33} (GPa)	⋯	389	⋯	30
c_{33} (GPa)	392	382	200	31
c_{33} (GPa)	405	373	224	32
$P_{sp,z}$ (C/m^2)	−0.029	−0.081	−0.032	15

properties and device physics, and furthermore that these effects can be used to achieve substantial improvements in various aspects of device performance. In $In_yGa_{1-y}N$/GaN quantum well structures, polarization fields in the $In_yGa_{1-y}N$ quantum wells lead to substantial redshifts in luminescence energy that are screened at high carrier concentrations.[33–37] Recently, it has also been shown that local strain fields associated with threading dislocations can create substantial electrostatic sheet charge densities, and consequently electric fields, at a free surface or a heterojunction interface in the vicinity of the dislocation.[38] In $Al_xGa_{1-x}N$/GaN HFETs, positive polarization charge at the $Al_xGa_{1-x}N$/GaN interface contributes to the formation of a two-dimensional electron gas (2DEG) with extremely high carrier concentrations.[12,13] Furthermore, polarization charges can be used in a nitride HFET to achieve a dramatic increase in barrier height, and consequently a reduction in gate leakage current, without increasing the total barrier thickness or the Al concentration in the $Al_xGa_{1-x}N$ layer.[39] Finally, polarization charges in compressively strained $In_yGa_{1-y}N$ deposited on GaN can be used to lower the surface barrier, and consequently decrease turn-on voltage, for GaN field emitters.[40]

III. POLARIZATION EFFECTS IN HFETs

It has emerged from studies of $Al_xGa_{1-x}N$/GaN HFETs that polarization charges at the $Al_xGa_{1-x}N$/GaN interface are

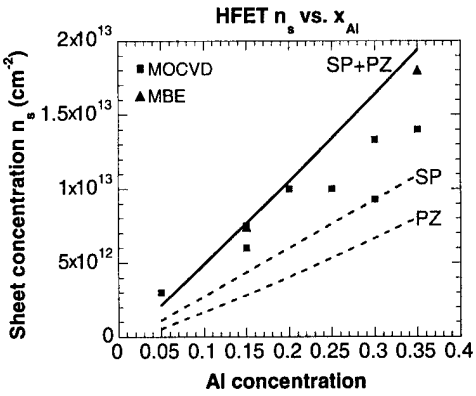


FIG. 2. Measured (symbols) and calculated (lines) sheet carrier concentrations in $Al_xGa_{1-x}N$/GaN HFET structures grown by both MBE and MOCVD as a function of Al concentration. Calculated values show the separate and combined contributions of piezoelectric and spontaneous polarization.

a dominant factor in the formation of a 2DEG with extremely high carrier concentration at the heterojunction interface.[12,13] For an $Al_xGa_{1-x}N$/GaN HFET epitaxial layer structure, the $Al_xGa_{1-x}N$ will be under tensile strain and, for growth on the (0001) surface (i.e., A or Ga face), a positive polarization charge density will be present at the $Al_xGa_{1-x}N$/GaN heterojunction interface. For growth on the (0001) surface, the polarization sheet charge density σ_{pol} at the heterojunction interface and the electron sheet concentration n_s in the 2DEG will be given approximately by

$$n_s = \sigma_{pol}/e - (\epsilon_{AlGaN}/de^2)(e\phi_b + E_F - \Delta E_c) + \tfrac{1}{2}N_d d, \quad (4)$$

$$\sigma_{pol}/e = -2[e_{31} - (c_{13}/c_{33})e_{33}](a_{GaN}/a_{AlN} - 1)x + P_{sp,z}^{GaN} - P_{sp,z}^{AlGaN}, \quad (5)$$

where e_{31}, e_{33}, c_{13}, and c_{33} are the relevant piezoelectric and elastic constants for $Al_xGa_{1-x}N$, a_{GaN} and a_{AlN} are the lattice constants of GaN and AlN, respectively, $P_{sp,z}^{GaN}$ and $P_{sp,z}^{AlGaN}$ are the spontaneous polarizations of GaN and $Al_xGa_{1-x}N$, ϵ_{AlGaN} is the dielectric constant of $Al_xGa_{1-x}N$, ϕ_b is the $Al_xGa_{1-x}N$ Schottky barrier height, E_F and ΔE_c are the Fermi energy and conduction-band offset, respectively, at the heterojunction interface, and d and N_d are the thickness and donor concentration, respectively, in the $Al_xGa_{1-x}N$ barrier layer.

Figure 2 shows n_s (obtained from Hall measurements) as a function of Al concentration for a series of 300 Å $Al_xGa_{1-x}N$/GaN HFET structures grown by both MBE and MOCVD. Also shown are values of n_s calculated from Eqs. (4) and (5) using values for e_{31}, e_{33}, c_{13}, and c_{33} determined in the manner described above, and assuming $N_d = 1 \times 10^{18}$ cm^{-3}. For Al concentrations below ~20%, the agreement between the calculated and measured values is very good, provided that the contributions of both spontaneous and piezoelectric polarization charge are included. For higher Al concentrations, the measured values for n_s are significantly below the calculated values. Estimates of strain relaxation in $Al_{0.25}Ga_{0.75}N$/GaN HFETs[41] have indicated that

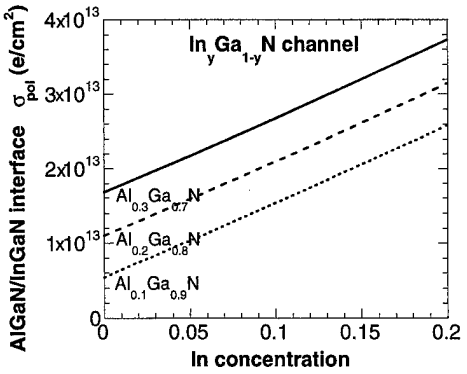


FIG. 3. Polarization sheet charge concentration at the 2DEG interface of an HFET structure containing an $Al_xGa_{1-x}N$ barrier and $In_yGa_{1-y}N$ channel, shown for Al concentrations of 10%, 20%, and 30% as a function of In concentration.

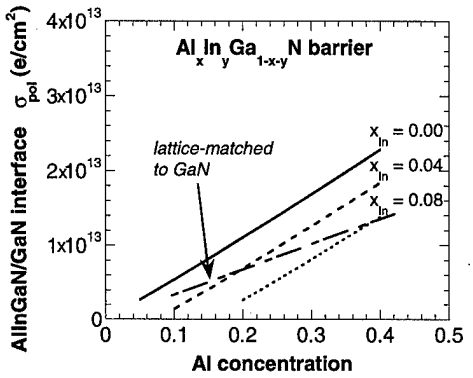


FIG. 4. Polarization sheet charge concentration at the 2DEG interface of an HFET structure containing an $Al_xIn_yGa_{1-x-y}N$ barrier layer, shown for In concentrations of 0%, 4%, and 8% as a function of Al concentration. For barriers lattice matched to GaN, the polarization charge density increases with In concentration.

strain relaxation initiates for $Al_{0.25}Ga_{0.75}N$ layer thicknesses between 20 and 40 nm, suggesting that the discrepancy between our calculated and measured values at Al concentrations above ~20% may be due partly to strain relaxation and, consequently, a reduction in piezoelectric polarization charge, or possibly to formation of compensating defects in the $Al_xGa_{1-x}N$ layer.

The analysis of sheet carrier concentration as a function of composition can be extended to include In-containing alloys. Because of the relatively large lattice mismatch between InN and GaN compared to that between AlN and GaN (11.3% and 2.4%, respectively), low In concentrations in an $In_yGa_{1-y}N$ alloy will lead to relatively high piezoelectric charge densities at an $In_yGa_{1-y}N$/GaN heterojunction interface. Because $In_yGa_{1-y}N$ grown on GaN will be under compressive strain, a negative piezoelectric sheet charge density will be present at the bottom of the $In_yGa_{1-y}N$ layer and a positive sheet charge density at the top, for growth on the (0001) surface. Incorporation into an HFET structure of an $In_yGa_{1-y}N$ channel layer with an $Al_xGa_{1-x}N$ barrier should therefore lead to a substantial increase in polarization charge density at the 2DEG interface, as shown in Fig. 3. The major additional contribution to polarization charge is expected to be piezoelectric, since the calculated values of spontaneous polarization for GaN and InN differ by only a small amount.

An additional possibility is incorporation of an $Al_xIn_yGa_{1-x-y}N$ barrier into an HFET structure. Incorporation of In into the barrier layer will allow higher Al concentrations to be attained without strain relaxation. Because the discontinuity in spontaneous polarization is expected to provide a larger contribution to polarization charge at the heterojunction interface than piezoelectric polarization for a given change in Al concentration, an $Al_xIn_yGa_{1-x-y}N$ barrier structure may be expected to provide a substantial polarization sheet charge (due to spontaneous polarization) even at compositions lattice matched to GaN as shown in Fig. 4. Furthermore, for a given degree of tensile strain the polarization charge at the $Al_xIn_yGa_{1-x-y}N$/GaN interface is expected to increase with In concentration, and consequently the maximum polarization charge density that can be achieved without strain relaxation in the barrier layer will also increase with In concentration.

IV. SCHOTTKY BARRIER ENHANCEMENT IN NITRIDE HFETs

As an example of the use of polarization effects to improve various aspects of device performance, we describe recent studies in which polarization effects were used to achieve large increases in Schottky barrier height in an $Al_xGa_{1-x}N$/GaN HFET structure,[39] and in which the influence of this increase on transistor characteristics was investigated. Figure 5(a) shows a schematic energy band diagram for an HFET epitaxial layer structure in which the barrier incorporates a thin GaN layer grown on top of the conventional $Al_xGa_{1-x}N$ barrier. Figure 5(b) shows the correspond-

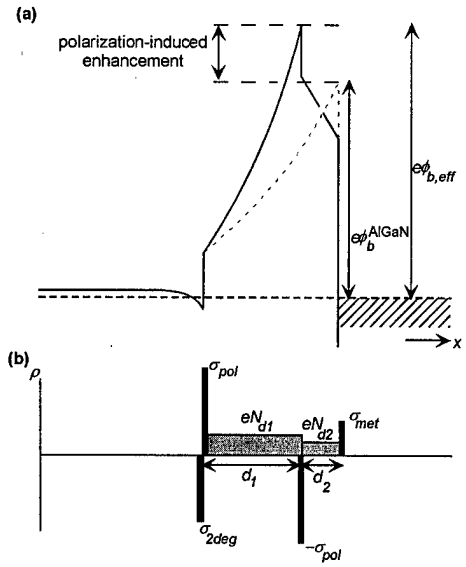


FIG. 5. (a) Schematic energy band diagram and (b) charge distribution for an HFET structure in which the barrier consists of GaN grown atop $Al_xGa_{1-x}N$. The negative polarization sheet charge at the upper $Al_xGa_{1-x}N$/GaN interface leads to a large enhancement in effective barrier height.

ing charge distribution in this structure, assuming growth on the (0001) surface. As shown in Fig. 5, incorporation of a GaN layer above the $Al_xGa_{1-x}N$ barrier layer allows the negative polarization charge at the top of the $Al_xGa_{1-x}N$ layer to be positioned within the Schottky barrier structure, thereby increasing the effective barrier height. This approach is analogous to the use of a thin p^+ layer near the metal–semiconductor interface of an n-type Schottky diode to increase the effective barrier height electrostatically. In the structure shown in Fig. 5(a), however, control over layer thickness and composition in epitaxial growth allows the magnitude and position of the polarization charge to be controlled very precisely.

A straightforward electrostatic analysis shows that the 2DEG sheet concentration n_s and the effective barrier height $\phi_{b,\mathrm{eff}}$ are given by

$$n_s=\frac{\sigma_{\mathrm{pol}}-(\epsilon_{\mathrm{AlGaN}}/d_1)(\phi_b^{\mathrm{GaN}}+E_F/e-V)+eN_{d1}d_1/2+(\epsilon_{\mathrm{AlGaN}}/\epsilon_{\mathrm{GaN}})(eN_{d1}d_2+eN_{d2}d_2^2/2d_1)}{e[1+(\epsilon_{\mathrm{AlGaN}}/\epsilon_{\mathrm{GaN}})(d_2/d_1)]}, \tag{6}$$

$$\phi_{b,\mathrm{eff}}=\Delta E_c/e+\phi_b^{\mathrm{GaN}}+\frac{ed_2}{\epsilon_{\mathrm{GaN}}}(n_s-N_{d1}d_1-N_{d2}d_2), \tag{7}$$

where σ_{pol} is the GaN/$Al_xGa_{1-x}N$ interface polarization sheet charge density (including both spontaneous and piezoelectric polarization), d_1 and d_2 are the $Al_xGa_{1-x}N$ and GaN layer thicknesses, N_{d1} and N_{d2} are the $Al_xGa_{1-x}N$ and GaN layer donor concentrations, $\epsilon_{\mathrm{AlGaN}}$ and ϵ_{GaN} are the dielectric constants, ϕ_b^{GaN} is the GaN Schottky barrier height, ΔE_c is the conduction-band offset, and E_F is the Fermi level at the 2DEG interface. In deriving Eqs. (6) and (7) we have assumed that $\phi_b^{\mathrm{AlGaN}}=\phi_b^{\mathrm{GaN}}+\Delta E_c$, as suggested by direct measurements of n-$Al_xGa_{1-x}N$ Schottky barrier heights.[42]

Figures 6(a) and 6(b) show, respectively, the effective barrier heights $\phi_{b,\mathrm{eff}}$ and sheet carrier concentrations n_s calculated as functions of GaN layer thickness d_2 for a total barrier thickness d_1+d_2 of 300 Å, assuming $N_{d1}=1\times10^{18}\,\mathrm{cm}^{-3}$, $N_{d2}=5\times10^{17}\,\mathrm{cm}^{-3}$, and $\phi_b^{\mathrm{GaN}}=1.0\,\mathrm{V}$, and showing both the separate and combined effects of the spontaneous and piezoelectric contributions to polarization charge. The total barrier thickness was kept at 300 Å because in an HFET device it is desirable to maintain a small barrier thickness to ensure a high gate capacitance and, consequently, high channel conductance and transconductance. Figure 6 shows that extremely large increases in barrier height may be expected, particularly for higher Al concentrations, when a thin GaN layer is incorporated at the top of the barrier. However, the sheet concentration decreases with increasing GaN layer thickness, necessitating a tradeoff between n_s and $\phi_{b,\mathrm{eff}}$ in selecting an optimum GaN layer thickness.

Measurements of n_s by capacitance–voltage profiling and of $\phi_{b,\mathrm{eff}}$ by photoresponse have shown that for an HFET structure consisting of 225 Å $Al_{0.25}Ga_{0.75}N$/75 Å GaN grown by MOCVD on nominally undoped GaN on a sapphire substrate, an effective barrier height of 1.89 ± 0.05 V and a sheet carrier concentration of $4.5\times10^{12}\,\mathrm{cm}^{-2}$ are obtained. In comparison, a barrier height of 1.52 ± 0.05 V and a sheet carrier concentration of $5.0\times10^{12}\,\mathrm{cm}^{-2}$ are obtained for a control HFET structure with a 300 Å $Al_{0.25}Ga_{0.75}N$ barrier.[39] The extremely large increase in effective barrier height of 0.37 V and the sheet carrier concentrations observed in these structures are consistent with those expected when either spontaneous or piezoelectric polarization effects are included, but are smaller than those expected from the combined effects of spontaneous and piezoelectric polarization. Studies of similarly grown HFET structures with barriers

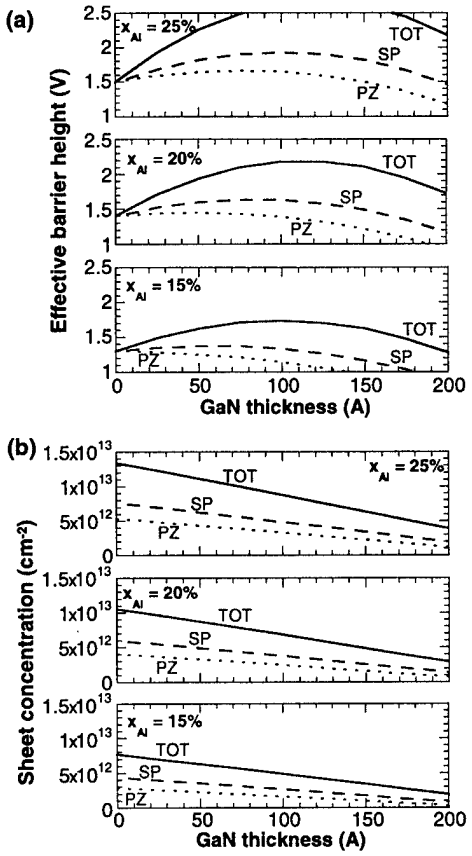


FIG. 6. Calculated values for (a) effective barrier height and (b) electron sheet concentration in a polarization-enhanced HFET Schottky barrier structure, shown as functions of GaN layer thickness for a total barrier thickness of 300 Å, for Al concentrations of 15%, 20%, and 25%. The separate and combined influences of spontaneous (SP) and piezoelectric (PZ) polarization are shown.

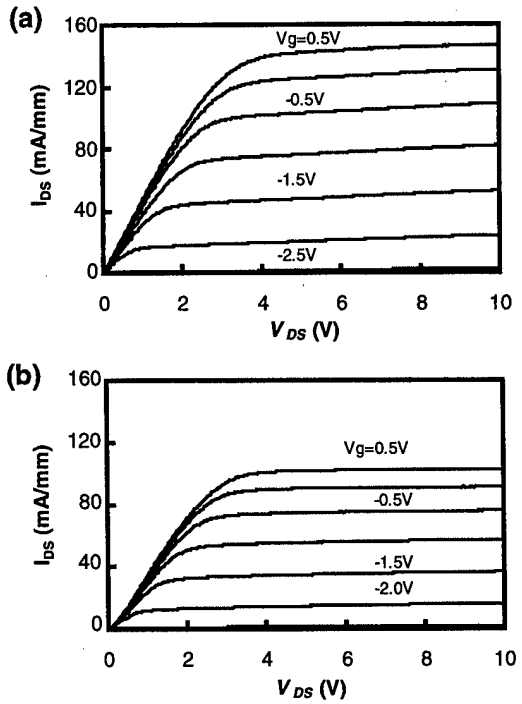


FIG. 7. Current–voltage characteristics for (a) an HFET with a conventional 300 Å $Al_{0.25}Ga_{0.75}N$ barrier and (b) an HFET with a polarization-enhanced 225 Å $Al_{0.25}Ga_{0.75}N$/75 Å GaN barrier.

consisting of 225 Å $Al_{0.30}Ga_{0.70}N$/75 Å GaN and 300 Å $Al_{0.30}Ga_{0.70}N$ have shown that the barrier height changes from 1.56 ± 0.05 V to 1.83 ± 0.05 V with incorporation of the top GaN layer, an increase of 0.27 V. These observations suggest that at these Al concentrations, partial strain relaxation in the $Al_xGa_{1-x}N$ layers may occur, reducing the contribution of piezoelectric polarization to the charge distribution in the epitaxial layer structure. An additional possibility is formation of compensating defects in the $Al_xGa_{1-x}N$ layer at high Al concentrations.

To determine the influence of the observed barrier-height enhancement on HFET device characteristics, transistor structures were fabricated from the epitaxial layer structures described above with barriers consisting of 300 Å $Al_{0.25}Ga_{0.75}N$ (conventional barrier) and 225 Å $Al_{0.25}Ga_{0.75}N$/75 Å GaN (enhanced barrier). Transistors were fabricated with gate lengths of 1 μm and gate widths of 25 and 50 μm, and source-gate and gate-drain spacings of 1 μm each. Ti/Al metallization annealed at 950 °C for 30 s was used for ohmic contacts, and Ni/Au for the Schottky gate contacts.

Figure 7 shows transistor characteristics for the conventional- and enhanced-barrier HFET structures. The enhanced-barrier HFET is characterized by slightly lower drain current and transconductance than the conventional-barrier HFET, features which we attribute primarily to the slightly lower sheet concentration and Hall mobility observed in the enhanced-barrier epitaxial layer structure (4.5×10^{12} cm^{-2} and 620 cm^2/V s at 300 K, in comparison to 5.0×10^{12} cm^{-2} and 800 cm^2/V s for the conventional-barrier

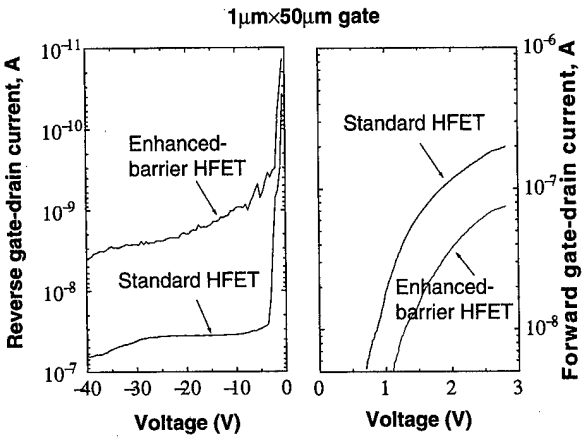


FIG. 8. Gate-drain diode characteristics for the conventional- and enhanced-barrier HFET structures.

structure). Figure 8 shows the gate-drain diode characteristics for these structures. As expected for the enhanced-barrier structure, there is a strong suppression in reverse-bias leakage current—a factor of approximately ten to several hundred—and an increase in forward-bias turn-on voltage of approximately 0.3 V.

A further benefit that might be expected from an increased barrier height is that, if transistor breakdown were to occur primarily via tunneling current through the gate, the enhanced-barrier structure should exhibit a substantially increased breakdown voltage. Detailed investigations of this possibility have been performed using the conventional- and enhanced-barrier HFET device structures. Figure 9 shows the drain and gate currents as functions of the drain-source voltage V_{ds} for the conventional- and enhanced-barrier HFET structures. Defining breakdown to occur at a current of 1 mA/mm, essentially identical breakdown voltages of 98 and 100 V are measured for the enhanced- and conventional-barrier devices. Subsequent analysis of breakdown voltage as a function of temperature has shown that the breakdown voltage increases with increasing temperature, indicating that the primary breakdown mechanism is impact ionization rather than gate tunneling, and explaining the similarity in breakdown voltages measured for the different device structures.[43]

V. CONCLUSIONS

We have investigated the role of spontaneous and piezoelectric polarization in nitride heterostructures, with particular emphasis on the design, characterization, and analysis of nitride HFETs. A detailed understanding of polarization effects and of the closely related issue of crystal polarity is essential in nitride heterostructure materials and device engineering. For $Al_xGa_{1-x}N$/GaN HFETs we have found that, at low to moderate Al concentrations, the 2DEG carrier concentrations observed are in very good agreement with those expected to arise from the combined effects of spontaneous and piezoelectric polarization. At higher Al concentrations, the data are consistent with partial strain relaxation in the

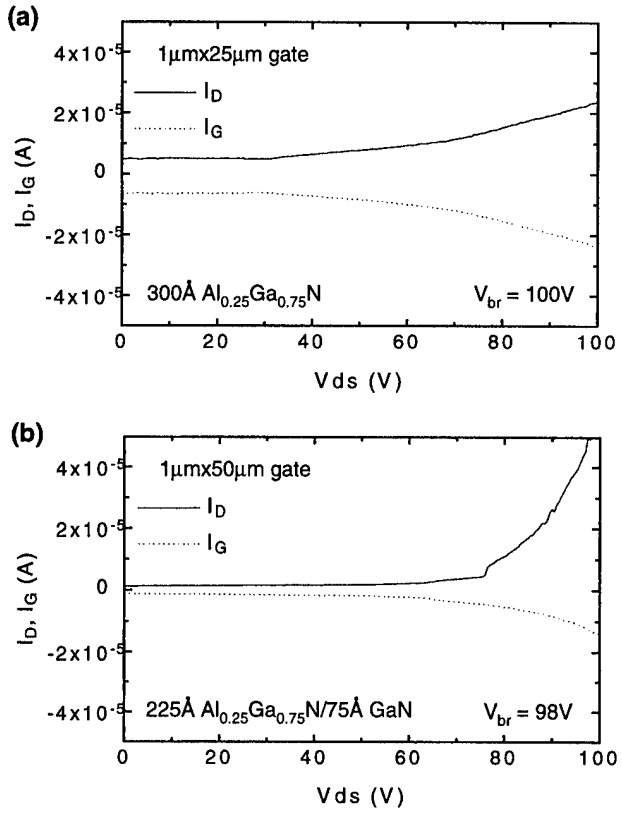


FIG. 9. Off-state gate and drain current as functions of source-drain voltage for (a) conventional-barrier and (b) enhanced-barrier HFET structures, yielding breakdown voltages of 100 and 98 V, respectively.

$Al_xGa_{1-x}N$ barrier layer, which would be accompanied by a reduction in the piezoelectric contribution to polarization charge at the heterojunction interface; an additional possibility is formation of compensating defects in the $Al_xGa_{1-x}N$ layers. Theoretical analysis has indicated that incorporation of In into the barrier and/or channel layers can substantially increase the polarization charge present at the 2DEG heterojunction interface. The use of polarization effects to engineer Schottky barrier structures with substantially increased barrier heights has also been discussed, and current–voltage characteristics of HFETs incorporating either conventional $Al_xGa_{1-x}N$ barriers or polarization-enhanced barriers have been presented. The increase in barrier height in the polarization-enhanced structures leads to a marked suppression in gate leakage current but has a negligible effect on transistor breakdown voltage, as breakdown in these transistor structures was found to be dominated by impact ionization rather that gate tunneling.

ACKNOWLEDGMENTS

The authors are grateful to Dr. J. M. Van Hove of SVT Associates, Inc. and Dr. J. M. Redwing and Dr. K. S. Boutros of Epitronics for providing data concerning sheet carrier concentrations in $Al_xGa_{1-x}N$/GaN HFETs grown, respectively, by MBE and MOCVD. The authors would like to acknowledge financial support from BMDO (Dr. Kepi Wu). One of the authors (E.T.Y.) would like to acknowledge financial support from the Alfred P. Sloan Foundation.

[1]S. Nakamura and G. Fasol, *The Blue Laser Diode: GaN Based Light Emitters and Lasers* (Springer, Berlin, 1997).
[2]M. A. Khan, Q. Chen, M. S. Shur, B. T. McDermott, J. A. Higgins, J. Burm, W. J. Schaff, and L. F. Eastman, IEEE Electron Device Lett. **17**, 584 (1996).
[3]O. Aktas, Z. F. Fan, A. Botchkarev, S. N. Mohammad, M. Roth, T. Jenkins, L. Kehias, and H. Morkoç, IEEE Electron Device Lett. **18**, 293 (1997).
[4]Y. F. Wu, B. P. Keller, S. Keller, D. Kapolnek, P. Kozodoy, S. P. DenBaars, and U. K. Mishra, Appl. Phys. Lett. **69**, 1438 (1996).
[5]G. J. Sullivan, M. Y. Chen, J. A. Higgins, J. W. Yang, Q. Chen, R. L. Pierson, and B. T. McDermott, IEEE Electron Device Lett. **19**, 198 (1998).
[6]S. C. Binari, J. M. Redwing, G. Kelner, and W. Kruppa, Electron. Lett. **33**, 242 (1997).
[7]J. M. Van Hove, R. Hickman, J. J. Klaassen, P. P. Chow, and P. P. Ruden, Appl. Phys. Lett. **70**, 2282 (1997).
[8]Q. Chen, J. W. Yang, A. Osinsky, S. Gangyopadhyay, B. Lim, M. Z. Anwar, M. A. Khan, D. Kuksenkov, and H. Temkin, Appl. Phys. Lett. **70**, 2277 (1997).
[9]R. D. Underwood, S. Keller, U. K. Mishra, D. Kapolnek, B. P. Keller, and S. P. DenBaars, J. Vac. Sci. Technol. B **16**, 822 (1998).
[10]T. Kozawa, M. Suzuki, Y. Taga, Y. Gotoh, and J. Ishikawa, J. Vac. Sci. Technol. B **16**, 833 (1998).
[11]A. Bykhovski, B. Gelmont, and M. S. Shur, J. Appl. Phys. **74**, 6734 (1993).
[12]P. M. Asbeck, E. T. Yu, S. S. Lau, G. J. Sullivan, J. Van Hove, and J. M. Redwing, Electron. Lett. **33**, 1230 (1997).
[13]E. T. Yu, G. J. Sullivan, P. M. Asbeck, C. D. Wang, D. Qiao, and S. S. Lau, Appl. Phys. Lett. **71**, 2794 (1997).
[14]R. Gaska, J. W. Yang, A. D. Bykhovski, M. S. Shur, V. V. Kaminski, and S. M. Soloviov, Appl. Phys. Lett. **72**, 64 (1998).
[15]F. Bernardini, V. Fiorentini, and D. Vanderbilt, Phys. Rev. B **56**, R10024 (1997).
[16]E. S. Hellman, MRS Int. J. Nitride Semicond. Res. **3**, 11 (1998).
[17]R. M. Martin, Phys. Rev. B **5**, 1607 (1972).
[18]S. Strite, M. E. Lin, and H. Morkoç, Thin Solid Films **231**, 197 (1993).
[19]G. D. O'Clock, Jr. and M. T. Duffy, Appl. Phys. Lett. **23**, 55 (1973).
[20]M. A. Littlejohn, J. R. Hauser, and T. H. Glisson, Appl. Phys. Lett. **26**, 625 (1975).
[21]A. D. Bykhovski, V. V. Kaminski, M. S. Shur, Q. C. Chen, and M. A. Khan, Appl. Phys. Lett. **68**, 818 (1996).
[22]A. D. Bykhovski, B. L. Gelmont, and M. S. Shur, J. Appl. Phys. **81**, 6332 (1997).
[23]J. G. Gualtieri, J. A. Kosinski, and A. Ballato, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 53 (1994).
[24]V. A. Savastenko and A. U. Sheleg, Phys. Status Solidi A **48**, K135 (1978).
[25]Y. Takagi, M. Ahart, T. Azuhata, T. Sota, K. Suzuki, and S. Nakamura, Physica B **219&220**, 547 (1996).
[26]A. Polian, M. Grimsditch, and I. Grzegory, J. Appl. Phys. **79**, 3343 (1996).
[27]R. B. Schwarz, K. Khachaturyan, and E. R. Weber, Appl. Phys. Lett. **70**, 1122 (1997).
[28]C. Deger, E. Born, H. Angerer, O. Ambacher, M. Stutzmann, J. Hornsteiner, E. Riha, and G. Fischerauer, Appl. Phys. Lett. **72**, 2400 (1998).
[29]K. Tsubouchi and N. Mikoshiba, IEEE Trans. Sonics Ultrason. **SU-32**, 634 (1985).
[30]L. E. McNeil, M. Grimsditch, and R. H. French, J. Am. Ceram. Soc. **76**, 1132 (1993).
[31]K. Kim, W. R. L. Lambrecht, and B. Segall, Phys. Rev. B **53**, 16310 (1996).
[32]A. F. Wright, J. Appl. Phys. **82**, 2833 (1997).
[33]T. Takeuchi, S. Sota, M. Katsuragawa, M. Komori, H. Takeuchi, H. Amano, and I. Akasaki, Jpn. J. Appl. Phys., Part 2 **36**, L382 (1997).
[34]M. B. Nardelli, K. Rapcewicz, and J. Bernholc, Appl. Phys. Lett. **71**, 3135 (1997).
[35]T. Takeuchi, C. Wetzel, S. Yamaguchi, H. Sakai, H. Amano, I. Akasaki,

Y. Kaneko, S. Nakagawa, Y. Yamaoka, and N. Yamada, Appl. Phys. Lett. **73**, 1691 (1998).
[36]S. F. Chichibu, A. C. Abare, M. S. Minsky, S. Keller, S. B. Fleischer, J. E. Bowers, E. Hu, U. K. Mishra, L. A. Coldren, S. P. DenBaars, and T. Sota, Appl. Phys. Lett. **73**, 2006 (1998).
[37]H. S. Kim, J. Y. Lin, H. X. Jiang, W. W. Chow, A. Botchkarev, and H. Morkoç, Appl. Phys. Lett. **73**, 3426 (1998).
[38]C. C. Shi, P. M. Asbeck, and E. T. Yu, Appl. Phys. Lett. **74**, 573 (1999).
[39]E. T. Yu, X. Z. Dang, L. S. Yu, D. Qiao, P. M. Asbeck, S. S. Lau, G. J. Sullivan, K. S. Boutros, and J. M. Redwing, Appl. Phys. Lett. **73**, 1880 (1998).
[40]R. D. Underwood, P. Kozodoy, S. Keller, S. P. DenBaars, and U. K. Mishra, Appl. Phys. Lett. **73**, 405 (1998).
[41]A. D. Bykhovski, R. Gaska, and M. S. Shur, Appl. Phys. Lett. **73**, 3577 (1998).
[42]L. S. Yu, D. J. Qiao, Q. J. Xing, S. S. Lau, K. S. Boutros, and J. M. Redwing, Appl. Phys. Lett. **73**, 238 (1998).
[43]X. Z. Dang, R. J. Welty, D. Qiao, P. M. Asbeck, S. S. Lau, E. T. Yu, K. S. Boutros, and J. M. Redwing, Electron. Lett. **35**, 602 (1999).

Electric force microscopy of induced charges and surface potentials in GaN modified by light and strain

P. M. Bridger, Z. Z. Bandić, E. C. Piquette, and T. C. McGill[a)]
Thomas J. Watson, Sr. Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125

(Received 19 January 1999; accepted 3 May 1999)

We have studied molecular beam epitaxy grown GaN films using electric force microscopy to detect sub-1 μm regions of electric field gradient and surface potential variations associated with GaN extended defects. The large piezoelectric coefficients of GaN together with strain introduced by crystalline imperfections produce variation in piezoelectrically induced electric fields around these defects. The consequent spatial rearrangement of charges can be detected by electrostatic force microscopy, and can be additionally modified by externally applied strain and illumination. The electron force microscopy signal was found to be a function of the applied tip bias, showed reversal under externally applied strain, and was sensitive to above band gap illumination. © *1999 American Vacuum Society.* [S0734-211X(99)05704-2]

I. INTRODUCTION

Nitride based devices have been of great interest in the last few years, notably due to their success in optoelectronics, where blue light emitting lasers and diodes have been demonstrated, and later successfully commercialized.[1] Further applications of nitrides are expected in the arena of high power and high temperature devices,[2–4] as well as solar blind ultraviolet detectors.[5] It has been recently demonstrated that the large intrinsic piezoelectric coefficients of GaN and AlN are responsible for an anomalously large concentration of two-dimensional electron gas at the AlGaN/GaN interface in GaN/AlGaN heterojunction field effect transistors (HFET).[6,7] Other possibilities exist for the enhancement of electric properties of contacts to nitrides by piezoelectric engineering as recently demonstrated in the case of Schottky contacts.[8] While most of the recent research has emphasized electronic device aspects of the piezoelectric effect,[6–8] comparatively little work has concentrated on the investigation of fundamental properties and nanoscale characterization of piezoelectrically induced phenomena. One consequence of the piezoelectric effect is that it allows electrostatic force imaging of charge redistribution around defects due to local variations in strain caused by crystalline imperfections. Albeit nonquantitative, electric force microscopy (EFM) of electric field gradients can provide interesting insight into the nature of both defects and piezoelectric effect in nitrides, while surface potentiometry can quantitatively map the change in surface potentials due to charge redistribution.[9,10]

II. EXPERIMENT

The gallium nitride layers studied here were grown on *c*-plane sapphire substrates by radio frequency plasma assisted molecular beam epitaxy (MBE). Details of the growth conditions are presented elsewhere.[11,12] The GaN films were nucleated using AlN buffer layers, and are predominantly Ga polar as determined by reflection high energy electron diffused diffraction (RHEED) reconstruction at low temperature,[13] and by KOH etching.[14,15] Growth conditions were slightly Ga rich, leading to locally flat (0001) Ga-face films which contain pits on the surface induced by dislocations or $(000\bar{1})$ inversion domains threading along the growth direction. These pits are readily observed in atomic force microscopy (AFM) images.[16]

The EFM data were collected using a Digital Instruments Nanoscope IIIa controller and a Bioscope scanning probe microscope operating in tapping mode. EFM was performed in two ways: by detecting the electric field gradient and by detecting the surface potential. To detect the electric field gradient, a voltage is applied to a metallic coated AFM tip which is scanned across the surface at a constant tip–sample separation. Phase differences induced by electrostatic forces on the oscillating tip are detected and give a qualitative measurement of local electric field gradients. To detect the local surface potential, an oscillating voltage is applied directly to the AFM tip; $V_{\text{applied}}=V_0\cos(\omega t)$. The tip feels a force of $F=(dC/dz)(V_{\text{tip}}-V_{\text{sample}})V_{\text{applied}}$ where (dC/dz) is the vertical derivative of the tip–sample capacitance. In order to determine the surface potential of the sample, the tip voltage is adjusted to equal the sample potential so that the tip feels no force. In all cases the tip–sample separation was set at 50 nm.

III. RESULTS AND DISCUSSION

The electric field gradient was measured as a function of tip voltage to rule out topographical artifacts. Topography can affect the EFM image for films with a permanent polarization, $\mathbf{P}$, since the surface charge $\sigma_s=\mathbf{P}\cdot\hat{\mathbf{n}}$ where $\hat{\mathbf{n}}$ is the surface normal, will not be constant over a rough surface. Variation in the induced surface charges result in a force differential between the tip and the surface that increases with tip voltage which can be observed in the series of EFM images in Fig. 1. It was also found that the force gradient was a function of the magnitude of the tip voltage and not the sign, consistant with the theoretical V^2 dependence.[10,17]

[a)]Electronic mail: tcm@ssdp.caltech.edu

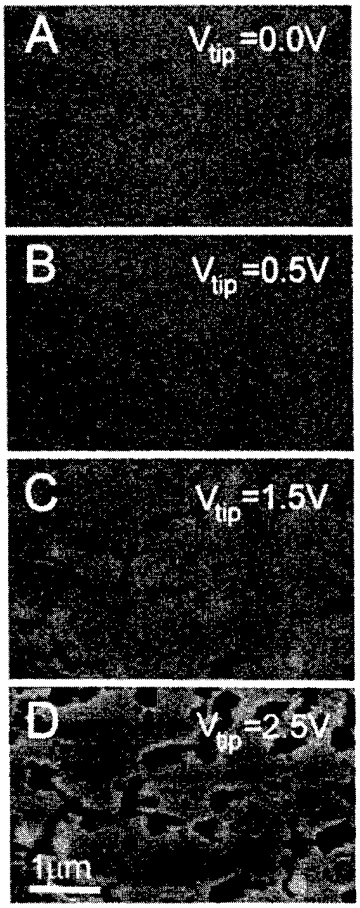


FIG. 1. Electric field gradient image as a function of tip applied voltage to an unstrained, unilluminated sample. The AFM tip bias increases from A–D.

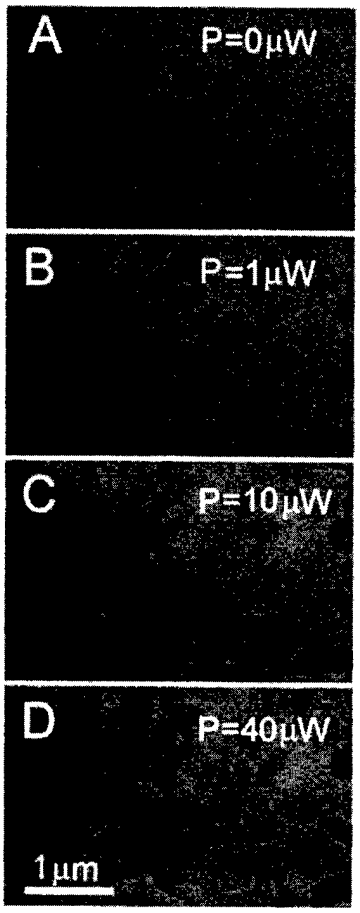


FIG. 3. Electric field gradient image as a function of optical power at 325 nm. Optical power increases from A to D. Tip bias=5 mV, and the tip-sample separation was 50 nm.

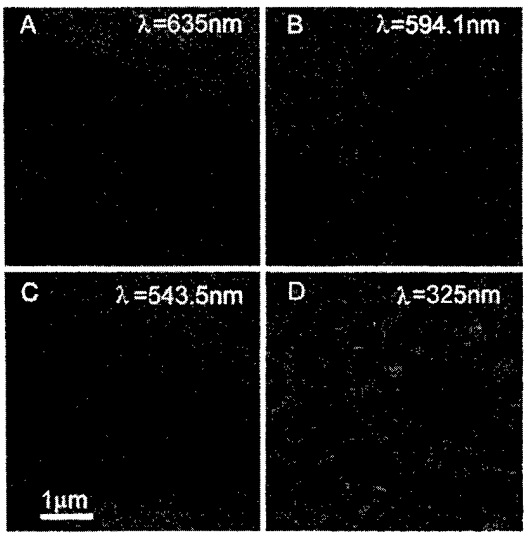


FIG. 2. Electric field gradient images at different illuminating wavelengths indicated in the upper right of each image. The tip bias was held at 5 mV with 50 nm tip–sample separation in all cases. A: 20 mW red diode laser; B: 1 mW yellow HeNe laser; C: 1 mW green HeNe laser; D: 10 μW from an UV HeCd laser.

The electric force was also found to be light sensitive and was measured at several discrete wavelengths above and below the GaN band gap energy (3.4 eV = 365 nm) using lasers at 635, 594.1, 543.5, and 325 nm as shown in Fig. 2. The sample was illuminated at a small glancing angle to eliminate interference with the AFM. The tip voltage was held constant at 5 mV, creating very weak contrast in the EFM image without illumination as observed in Fig. 1A or 3A. When the sample was illuminated with a photon energy below band gap (Figs. 2A–2C) no apparent difference in EFM contrast could be observed, even with optical powers above 1 mW. However, there is a significant increase in EFM contrast when the sample is exposed to light with a photon energy above the band gap (Fig. 2D). This increase is associated with the generation of electron-hole pairs, which cannot be obtained with photons of energy smaller than the GaN band gap. We speculate that the separation of generated charges by internal polarization fields, as well as the increase in the sample conductivity are responsible for the observed change.

Figure 3 shows the EFM images obtained with tip voltage held constant at 5 mV, illuminated with the 325 nm laser, as

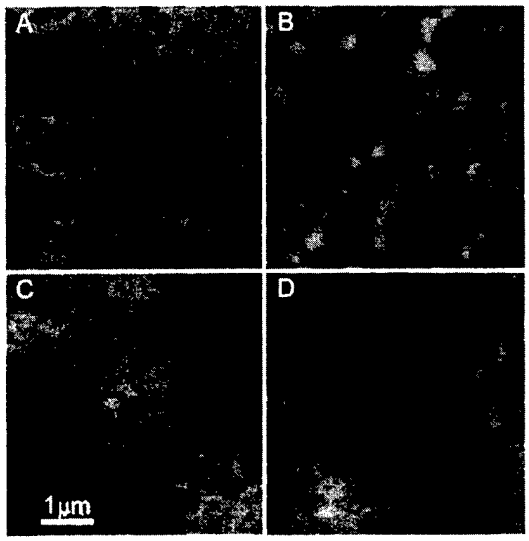


FIG. 4. A: Electric field gradient of an unstrained sample; B: electric field gradient of a strained sample. In both cases, the tip voltage was held at 2.0 V; C: surface potential of an unstrained sample; D: surface potential of a strained sample. In all cases the sample was unilluminated and the tip-sample separation was 50 nm.

a function of the optical power. The optical power required to produce a visible change in the EFM image was found to be as small as 1 μW. No change could be observed for the longer wavelengths even for optical powers up to three orders of magnitude greater.

Since the piezoelectric effect will change the magnitude of the internal fields and consequently the surface charge distribution/potential, a homebuilt stage was used to externally apply strain to GaN films by bending the sapphire substrate. The induced strain is tensile and approximately 1%. Subsequently, both electric field gradient and surface potential measurements were made on unstrained and strained samples and are shown in Fig. 4. Contrast reversal is observed in the electric field gradient signal of the strained sample as shown in Figs. 4A and 4B. Strained samples also showed regions which had changes in the surface potential of approximately 0.1 V as indicated in Figs. 4C and 4D. The strain induced polarization fields are believed to be responsible for the observed changes in electric field gradients and surface potential.

These experimental results clearly show that defects in GaN can induce electric field and potential inhomogeneities that are detectable by EFM. The exact nature of these inhomogeneities, however, remains unclear. We can speculate on several possibilities, including the effects of inversion domains and threading dislocations. First, in the case of inversion domain defects, the spontaneous and strain-induced polarization within these domains is antiparallel to that of the bulk crystal, and obviously would induce opposite surface charging effects. The presence of inversion domains can explain the contrast reversal observed for samples under applied strain. In the case of threading dislocations, localized strain will result in piezoelectric charging where the dislocation intersects the surface. Other possibilities include Fermi level pinning effects and local band bending.

IV. CONCLUSIONS

In summary, we have successfully demonstrated that EFM technique can be used to detect local variations in piezoelectrically induced charge and potential on the sub-1 μm scale. We have found that electron-hole pair generation by light illumination can significantly enhance observed contrast in the EFM images, which we associate with the charge separation due to builtin polarization fields. External application of strain caused contrast reversal in EFM signal which is attributed to the increase in induced polarization fields over the builtin polarization fields.

ACKNOWLEDGMENTS

The authors would like to thank R. A. Beach for his useful comments and suggestions. This work was supported by the Advanced Research Project Agency, and monitored by the Office of Naval Research under Grant No. N00014-92-J-1845.

[1]S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, Y. Sugimoto, and H. Kiyoku, Appl. Phys. Lett. **70**, 1417 (1997).
[2]M. A. Khan and M. S. Shur, Mater. Sci. Eng., B **46**, 69 (1997).
[3]S. M. Mohammad and H. Morkoç, Prog. Quantum Electron. **20**, 361 (1996).
[4]Z. Z. Bandić, E. C. Piquette, P. M. Bridger, R. A. Beach, T. F. Kuech, and T. C. McGill, Solid-State Electron. **42**, 2289 (1998).
[5]J. M. Van Hove, R. Hickman, J. J. Klaassen, P. P. Chow, and P. P. Ruden, Appl. Phys. Lett. **70**, 2282 (1997).
[6]G. Zandler, J. A. Majewski, M. Stadele, P. Vogl, and F. Compagnone, Phys. Status Solidi B **204**, 133 (1997).
[7]A. Bykhovski, B. Gelmont, and M. Shur, J. Appl. Phys. **74**, 6734 (1993).
[8]E. T. Yu, G. J. Sullivan, P. M. Asbeck, C. D. Wang, D. Qiao, and S. S. Lau, Appl. Phys. Lett. **71**, 2794 (1998).
[9]Y. Martin, D. W. Abraham, and H. K. Wickramasinghe, Appl. Phys. Lett. **52**, 1103 (1988).
[10]D. Sarid, *Scanning Force Microscopy* (Oxford University Press, New York, 1991).
[11]E. C. Piquette, P. M. Bridger, Z. Z. Bandić, and T. C. McGill, J. Vac. Sci. Technol. B **17**, 1241 (1999).
[12]E. C. Piquette, P. M. Bridger, Z. Z. Bandić, and T. C. McGill, Mater. Res. Soc. Symp. Proc. **512**, 387 (1998).
[13]A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrup, Appl. Phys. Lett. **72**, 2114 (1998).
[14]M. Seelmann-Eggebert, J. L. Weyher, H. Obloh, H. Zimmermann, A. Rar, and S. Porowski, Appl. Phys. Lett. **71**, 2635 (1997).
[15]J. M. Van Hove, M. F. Rosamond, R. Hickman II, J. J. Klaassen, C. Polley, A. Wowchak, and P. P. Chow, J. Vac. Sci. Technol. B (submitted).
[16]P. M. Bridger, Z. Z. Bandić, E. C. Piquette, and T. C. McGill, Appl. Phys. Lett. **72**, 3166 (1998).
[17]F. Saurenbach and B. D. Terris, Appl. Phys. Lett. **56**, 1703 (1990).

Piezoelectric fields in nitride devices

R. A. Beach and T. C. McGill
T. J. Watson, Sr. Laboratories of Applied Physics 128-95, California Institute of Technology, Pasadena, California 91125

(Received 19 January 1999; accepted 3 May 1999)

We have calculated the piezoelectric field and charge distribution for various III-nitride heterostructures. Our calculations include strain energy minimization and doping effects, and are presented to show the magnitude of piezoelectric effects in strained layers. We compare our calculated results to device results where available. These include the two-dimensional electron gas in heterojunction field effect transistors, Schottky diodes with strained layers for Schottky height engineering, and III-nitride single quantum wells. Calculations that included energy considerations resulted in good agreement between predicted and observed field and charge distributions for the heterojunction fields-effect transistors structure. © *1999 American Vacuum Society.* [S0734-211X(99)07104-8]

I. INTRODUCTION

Column III-nitride semiconductors have unique material properties that make them extremely attractive for optoelectronic and high-power, high-frequency applications. Their band gaps range from 1.9 eV for InN to 6.2 eV for AlN. They have small effective masses, typically 0.2 for electrons and 0.5 for holes, and high optical phonon energies in the range of 600–900 cm^{-1}. In addition they possess a spontaneous electric polarization, and have the largest piezoelectric constants of the semiconductors. This allows for interesting possibilities in strain field engineering. Numerous papers have been published cataloging piezoelectric effects in nitride structures. These include the piezoresistive effect in GaN/AlN/GaN semiconductor–insulator–semiconductor (SIS) structures,[1] in which small deformations of the SIS structure lead to a reduction in the in-plane resistance. A similar effect is seen in AlGaN/GaN heterojunction field-effect transistors (HFETs). In this case the change in spontaneous polarization and strain induced piezoelectric polarization induce a compensating two-dimensional electron gas (2DEG) at the AlGaN–GaN interface.[2–4] Piezoelectric fields have also been shown to effect the optical properties of InGaN/GaN and GaN/AlGaN quantum wells. The strain induced electric field tilts the conduction and valence bands within the well, resulting in charge separation and increased recombination times as well as a redshift in the emission verse absorption spectrum.[5–7] In the last few years, substantial improvement in the structural and electronic properties of GaN has been achieved. This has resulted in major advances in blue-green light-emitting diodes and lasers as well as HFET power devices. To further improve the performance of these devices, the effects of piezoelectricity must be accounted for. In addition, the use of piezoelectric fields can result in a number of interesting devices such as piezoelectrically enhanced Shottky diodes and photodetectors. We present calculations of the electric fields produced in a number of III-nitride structures. We have used a strain tensor formalism and we compare the results to experiment where available.

II. BASICS OF PIEZOELECTRICITY

Piezoelectricity is the result of two processes. One is the rearrangement of electronic states in the crystal. The other is displacement of charged constituents within the lattice. Displacement of charged constituents is the dominating process in GaN. A quantitative description of the strain induced polarization is given by

$$P_i = \sum_j e_{ij}\epsilon_j . \tag{1}$$

Here P_i is the polarization in the i direction, ϵ_j is the j component of strain, and e_{ij} is the ij component of the piezoelectric tensor. The strain generated in pseudomorphic epilayers is determined by the lattice mismatch for the in-plane strain and energy minimization for the perpendicular component of strain. With the three-axis set as the growth direction, ϵ_1 and ϵ_2 are given by

$$\epsilon_{1,2} = \frac{a_s - a_l}{a_l} . \tag{2}$$

Here, a_s (a_1) is the substrate (epilayer) lattice constant. The lattice constants were assumed to follow Vegard's law. A recent article by Angerer[8] suggests that Vegard's law underestimates the Al content in hexagonal AlGaN. This may introduce some slight error in the following analysis, but is unlikely to outweigh the other uncertainties such as Shottky heights of contacts and strain relaxation in epilayers. The strain energy, U, is given by

$$U = \frac{1}{2}\sum_{ij} C_{ij}\epsilon_i\epsilon_j . \tag{3}$$

The C_{ij}'s are the elastic constants of the strained layer. The perpendicular component of strain is determined by minimizing this strain energy. The piezoelectric polarization can then be calculated using Eq. (1). For all alloys a linear interpolation of piezoelectric constants and elastic moduli has been used. Due to the lack of experimental data for bulk InN, we have used the elastic coefficients of GaN in place of InN. We have used the spontaneous polarization and piezoelectric constants of Ref. 9, along with the elastic coefficients of

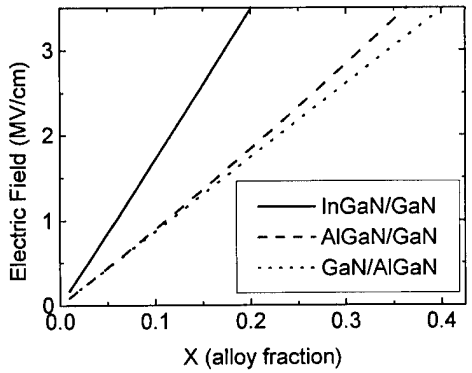


FIG. 1. Comparison of fields in III–N structures

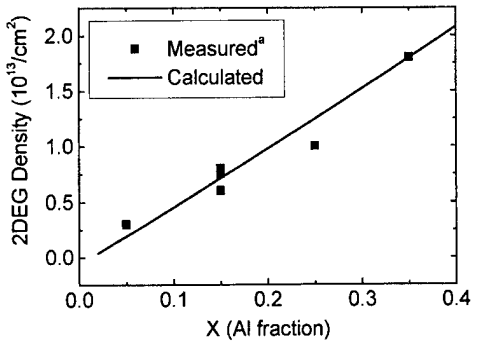


FIG. 2. Comparison of calculated and measured 2DEG densities for AlGaN(300 A) GaN HFETs. Data from Refs. 4 and 13.

Refs. 10 and 11. The spontaneous polarization of the bulk is assumed to be canceled by charging of surface states. If this were not the case, an enormous field, on the order of a few MV/cm, would be present across the entire sample. The effect of changes in spontaneous polarization from the bulk to epilayer are accounted for in these calculations by assuming a linear relationship between alloy fractions and spontaneous polarizations. Figure 1 shows the calculated fields generated in InGaN/GaN, GaN/AlGaN, and AlGaN/GaN epilayers for undoped structures. The direction of these fields depends on the polarity of the growth. For InGaN grown on GaN and GaN grown on AlGaN, there is a compressive strain in the epilayer. This results in a piezoelectric field in the [000–1] direction. For AlGaN grown on GaN the strain is tensile, and results in a piezoelectric field in the [0001] direction. The effect of bulk spontaneous field cancellation is to increase the electric field for GaN/AlGaN and AlGaN/GaN and to decrease the field for InGaN/GaN. It is interesting to note the similarities in electric field for the AlGaN/GaN and GaN/AlGaN structures. From a quick look at the piezoelectric coefficients, one would expect the AlGaN/GaN structure to have a 20% larger field than the GaN/AlGaN structure due to AlN's larger piezoelectric coefficient. This is not the case, however. The change in spontaneous electric field has the same magnitude, but opposite sign, for both structures. Although not shown, the spontaneous contribution represents 2/3 of the resulting field. This results in the closely matched curves for AlGaNGaN and GaN/AlGaN in Fig. 1. Very good agreement between calculated and observed electric fields was obtained for the GaN/AlGaN structure. For $GaN/Al_{0.15}Ga_{0.85}N$ a field of 0.46 MV/cm was calculated while measurements employing photoluminescence resulted in a value of 0.42 MV/cm.[6] The electric field is given approximately by $E=-18.6\,\mathrm{MV/cm}\,x(-3.0\,\mathrm{MV/cm}\,x)$ for $In_xGa_{(1-x)}N/GaN$ [$GaN/Al_xGa_{(1-x)}N$]. The negative sign indicates an electric field in the [000–1] direction. This is toward the substrate in metalorganic chemical vapor deposition (MOCVD) and halide vapor phase epitaxy (HVPE) grown samples.[12]

III. HFETs

Nitride-based HFETs have emerged as attractive candidates for high power devices operating at microwave frequencies. These devices benefit from a high 2DEG that arises at the AlGaN–GaN interface. This 2DEG has been shown to arise from a large electric field present in the strained AlGaN barrier layer.[3] The expected 2DEG density for a Schottky/AlGaN/GaN structure is given by

$$n_{2d}=\frac{\sigma_p}{e}-\frac{\epsilon_{\mathrm{AlGaN}}}{de^2}(e\varphi_b+E_F-\Delta E_c)+0.5N_l d. \tag{4}$$

Here, σ_p is the polarization charge at the AlGaN–GaN interface caused by the change in spontaneous polarization and the piezoelectric polarization. $\epsilon_{\mathrm{AlGaN}}$ is the dielectric constant of the AlGaN layer, N_l is the doping density in the AlGaN layer, and d is the AlGaN layer thickness. We have used a value of 1.0 eV for the Schottky barrier height of the contact on GaN. We have assumed an increase in Schottky height for AlGaN identical to the conduction band offset, thus reducing $e\varphi_b$-ΔE_c to $e\varphi_{\mathrm{GaN}}$. E_F is the Fermi level with respect to the conduction band edge at the AlGaN/GaN interface. We have used a value of 0.1 eV for E_F and 10^{18}/cm^3 for N_l in these calculations. Figure 2 shows the results of this calculation for variable Al content in the AlGaN layer, and compares the results to experimental measurements found in the literature.[4,13] The thickness of the AlGaN layer is also an important parameter in these HFETs. Figure 3 shows the dependence of the 2DEG density on AlGaN barrier thickness for three Al contents. As can be seen in the figure, a distinct

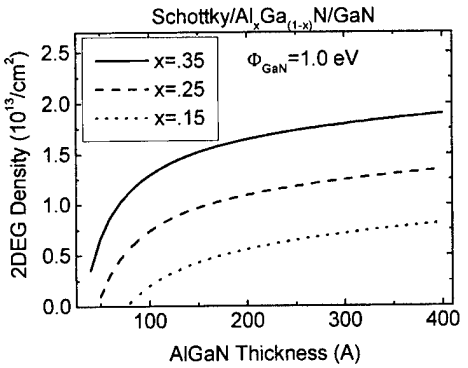


FIG. 3. Thickness dependence of 2DEG in HFETs.

role off occurs in the 2DEG density at ~150 A. After this point, the electron density begins to flatten out with little increase for added thickness.

IV. QUANTUM WELLS

Light emitting diodes (LEOs) based on GaN, AlGaN, and InGaN have already been produced and marketed with great success.[14] It has been well documented that these quantum well structures possess a large redshift in the emission versus absorption spectrum. This has been shown to result from the piezoelectric fields generated by strain in the well layer.[7] These piezoelectric fields result in charge separation and decreased oscillator strength for the quantum well structure, resulting in decreased efficiency for light emitting devices. Here we incorporate the effects of doping on the fields present in the quantum well, and present a method for determining the most advantageous doping profiles for light emitting devices. Due to the detrimental effects the electric field has on oscillator strength, and hence, quantum efficiency, minimization of these fields is of particular interest to the LED and laser community. Of first order importance is the doping. The polarity of the growth determines the sign of the resulting piezoelectric polarization charge at the well interface. As mentioned earlier, MOCVD and HVPE have been shown to result in Ga faced GaN. This results in a strain generated electric field that points towards the substrate. For the depletion region field to reduce the piezoelectric field a growth sequence resulting in p-GaN/InGaN/n-GaN/substrate is required for Ga faced material. For complete cancellation of the electric field in the well region, the charge in the donor depletion region plus the piezoelectric polarization charge at the well-donor region interface must equal zero. A mathematical statement of this is

$$P = qL_dN_d = \sqrt{\frac{2\epsilon_s qN_dN_aV_{bi}}{(N_d+N_a)}}. \tag{5}$$

Here, P is the polarization in the well, N_d is the donor density in the n region, N_a is the acceptor density in the p region, L_d is the depletion length in the donor region, ϵ_s is the dielectric constant in the bulk, and V_{bi} is the built in voltage. Due the high doping concentrations required for p-type GaN, a large disparity in doping density occurs in GaN p-n structures. Setting the ratio of acceptors in the p region to donors in the n region equal to α, i.e., $N_a = \alpha N_d$, and solving for N_d, this reduces to

$$N_d = \frac{P^2(\alpha+1)}{2\epsilon_s qV_{bi}\alpha}. \tag{6}$$

Figure 4 shows the results of this calculation for a range of alloy compositions with α set to 1. Due to the approximately linear nature of the polarization charge on alloy composition, the doping density can be well approximated by a second order polynomial in x. For InGaN/GaN $N_d \sim 6.7 \times 10^{20}/\text{cm}^3x^2$ for GaN/AlGaN $N_d \sim 2.0 \times 10^{19}/\text{cm}^3x^2$. These numbers are well within the region of doping densities achieved in GaN and should not cause difficulties in growth of appropriately doped structures. One caveat that must be discussed is the dependence of the built-in voltage, V_{bi}, on doping density. The built-in voltage is the energy difference between the fermi levels in the n and p layers. We have used a constant V_{bi} for these calculations of 3.2 eV. This will require modification based on Fermi levels determined for actual devices, but is not expected to vary significantly. In addition, the on state voltage will decrease the depletion region and result in a lower cancellation field. Reduction of the value used here by the voltage drop across the the device in the on state will be required when determining the doping density of the structure.

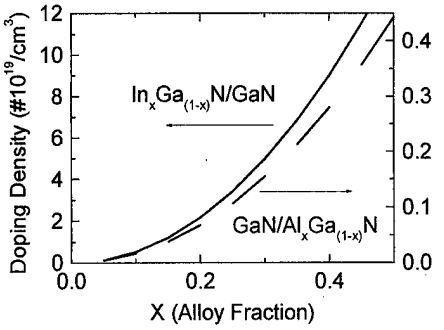


FIG. 4. Graph of required doping density for field cancellation in InGaN/GaN and GaN/AlGaN quantum wells for α equal to 1.

V. SCHOTTKY DETECTORS

Currently, GaN is predominantly grown with Ga faced polarity. N faced material, however, can result in a wide range of new and interesting devices. One is a ultraviolet detector employing piezoelectric fields to reject leakage current. Figure 5 shows the structure and band diagram for a field enhanced detector. We present the barrier enhancement and rejection field for this device. A simple electrostatic analysis results in the following equation connecting the Schottky height at the metal-AlGaN interface to the depletion length in the GaN bulk.

$$\varphi_{\text{AlGaN}} = \frac{0.5N_deL^2}{\epsilon_{\text{GaN}}} + \frac{(N_deL - \sigma_p)d}{\epsilon_{\text{AlGaN}}} + \frac{0.5N_1ed^2}{\epsilon_{\text{AlGaN}}} + E_F + \Delta E_c. \tag{7}$$

In Eq. (7), N_d (N_1) is the doping density in the GaN (AlGaN), L is the depletion width, d is the strained layer thickness, E_F is the Fermi level to conduction band edge energy

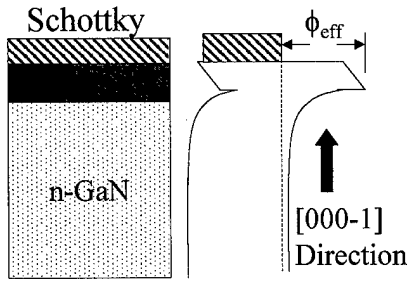


FIG. 5. Schematic of epitaxial layer structure and band-edge energy diagram for Schottky detector.

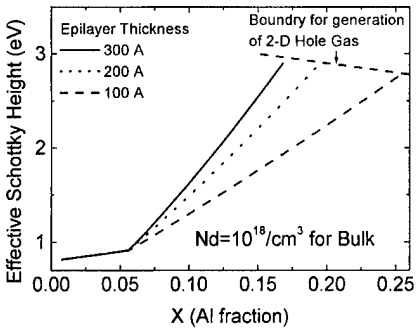


FIG. 6. Graph showing effective Schottky heights as a function of Al mole fraction for three values of epilayer thickness.

separation, and ΔE_c is the conduction band offset. After solving this equation for the depletion width (L), the effective Schottky height can be calculated from

$$\varphi_{\text{eff}}=\frac{0.5N_d eL^2}{\epsilon_{\text{GaN}}}+E_F+\Delta E_c. \tag{8}$$

Figure 6 shows the calculated effective Schottky height for three values of barrier thickness versus aluminum content in the barrier region. We have used $10^{18}/\text{cm}^3$($10^{16}/\text{cm}^3$) for the bulk (strained layer) doping density.

For low Al mole fraction the barrier height is equal to the Shottky height of the metal-AlGaN interface. In this region the polarization field is not large enough to overcome the depletion region field and produce added Schottky height. The Al molar fraction at which this kink occurs is dependent on the doping in the GaN. The maximum effective Schottky height is reached when the GaN valence band at the AlGaN/GaN interface moves above the Fermi level. At this point, a 2D hole gas is formed at the AlGaN/GaN interface. This is analogous to the electron gas formed in the HFET structure. Increasing the polarization field further results in increased hole accumulation, but little additional height is achieved. It has been found that coherently strained $Al_{0.25}Ga_{.075}N$ can be grown up to ~300 A on GaN.[15] This indicates that strain relaxation should not cause a significant shift from these calculated Schottky heights. Doping density is found to have a significant effect on the effective Schottky height. Figure 7 illustrates the changes in effective height with doping in both the barrier layer and bulk. It was found that doping resulted in a later turn on and lower over all enhancement of the Schottky height. Doping in the bulk layer had the greatest effect on Schottky heights, however, doping in the barrier layer caused significant reduction in effective barrier height as well.

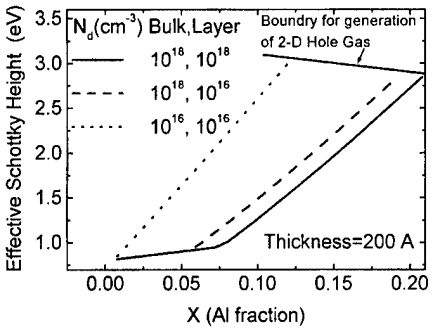


FIG. 7. Effective Schottky height vs Al mole fraction for three sets of doping parameters.

VI. CONCLUSIONS

In summary, we have calculated the piezoelectric fields generated in a few III-nitride structures. These fields were found to be approximately linear with alloy composition and in agreement with experimental photoluminescence measurements. The 2DEG in HFET devices was shown to accumulate quickly after a critical barrier thickness and then level off after approximately 150 A. It was shown that by the appropriate choice of doping, piezoelectric fields in quantum wells may be reduced or eliminated. The doping densities required for field minimization were shown to be acceptably moderate, and a method for determining the doping density requirements for a given structure was presented. The calculated effective Schottky height for N-polar AlGaN/GaN photodetectors was presented. The AlGaN thickness and Al mole fraction for significant Schottky height increases were found to be moderate and within acceptable growth parameters. Doping was found to decrease the effective Schottky heights in this structure.

ACKNOWLEDGMENTS

This work has been supported by the Defense Advanced Research Projects Administration and Electric Power Research Institute monitored under ONR Grant No. MDA972-98-1-0005.

[1]A. D. Bykhovski, V. V. Kaminski, M. S. Shur, Q. C. Chen, and M. A. Khan, Appl. Phys. Lett. **68**, 818 (1996).
[2]A. Bykhovski, B. Gelmont, and M. Shur, J. Appl. Phys. **74**, 6734 (1993).
[3]E. T. Yu, G. J. Sullivan, P. M. Asbeck, C. D. Wang, D. Qiao, and S. S. Lau, Appl. Phys. Lett. **71**, 2794 (1997).
[4]P. M. Asbeck, E. T. Yu, S. S. Lau, G. J. Sullivan, J. Van Hove, and J. Redwing, Electron. Lett. **33**, 1230 (1997).
[5]J. S. Im, H. Kollmer, J. Off, A. Sohmer, F. Scholz, and A. Hangleiter, Phys. Rev. B **57**, R9435 (1998).
[6]J. S. Im, H. Kollmer, J. Off, A. Sohmer, F. Scholz, and A. Hangleiter, Mater. Res. Soc. Symp. Proc. **482**, 513 (1998).
[7]T. Taakeuchi, S. Sota, M. Katsuragawa, M. Komori, H. Takeuchi, H. Amano, and I. Akasaki, Jpn. J. Appl. Phys., Part 2 **36**, L382 (1997).
[8]H. Angerer, Appl. Phys. Lett. **71**, 1504 (1997).
[9]F. Bernardini, V. Fiorentini, and D. Vanderbilt, Phys. Rev. B **56**, R10024 (1997).
[10]A. Polian, M. Grimsditch, and I. Grzegory, J. Appl. Phys. **79**, 3343 (1996).
[11]L. E. McNeil, M. Grimsditch, and R. H. French, J. Am. Ceram. Soc. **76**, 1132 (1993).
[12]E. S. Hellman, MRS Internet J. Nitride Semicond. Res. **3**, 11 (1998); electronic mail: http//nrs.mij.mrs.org
[13]J. M. Redwing, J. S. Flynn, M. A. Tischler, W. Mitchel, and A. Saxler, Mater. Res. Soc. Symp. Proc. **395**, 201 (1996).
[14]S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, Y. Sugimoto, and H. Kiyoku, Appl. Phys. Lett. **70**, 1417 (1997).
[15]H. S. Kim, J. Y. Lin, H. X. Jiang, W. W. Chow, A. Botchkarev, and H. Morkoc, Appl. Phys. Lett. **73**, 3426 (1998).

Minigaps in strained silicon quantum wells on tilted substrates

T. J. Thornton,[a] F. Ge, A. Andresen, D. Pivin, J. Bird, and D. K. Ferry
Department of Electrical Engineering and Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-5706

(Received 19 January 1999; accepted 3 May 1999)

The two-dimensional electron gas formed at the inverted surface of a tilted silicon substrate shows unusual magnetotransport properties due to the presence of a minigap in the density of states. For metal–oxide–semiconductor inversion layers the strong scattering at the interface limits the mobility to values $\mu < 10-20\,000$ cm^2/V s. To achieve mobilities approaching 10^5 cm^2/V s we have used strained Si:SiGe quantum wells grown on substrates tilted away from the (001) normal by 0°, 2°, 4°, 6°, and 10°. Their transport properties have been measured in the temperature range of 20–500 mK. All the samples show strong Shubnikov–de Haas oscillations. For the 2° and 4° samples the envelope of the fast oscillations is modulated by a longer period oscillation at low magnetic fields. We attribute the slow oscillation in the 2° and 4° samples to the presence of a minigap. For the 6° and 10° samples the minigap is higher than the Fermi energy and is not expected to influence the transport properties. © *1999 American Vacuum Society.*
[S0734-211X(99)05604-8]

I. INTRODUCTION

Tilted substrates have been used to explore a number of issues related to semiconductor interfaces. Vicinal GaAs substrates, i.e., those tilted a few degrees from high symmetry planes such as the (001) and (111) have been used to study the growth of quantum wires[1] and lateral surface superlattices[2,3] and for investigating dopant incorporation.[4] Vicinal silicon substrates have an even longer history, having been used in the early years of metal–oxide–semiconductor (MOS) technology to find Si:SiO_2 interfaces with the lowest interface state density.[5] Silicon wafers with high index surfaces such as $(11n)$ are readily available and correspond to offcut angles of a few degrees between the surface normal and the [001] direction. Surprisingly, silicon MOS inversion layers on (115) and (118) substrates have shown curious transport anomalies,[6,7] which were later explained by the formation of a minigap in the density of states.[8] Recently, Si:SiGe quantum wells on tilted substrates have been used to reduce the density of threading dislocations[9] present in these strained layers as well as to explore the nucleation of dislocations[10] via the modified Frank–Read mechanism. Strained silicon quantum wells on tilted substrates also show interesting surface morphology as well as anisotropic transport properties.[11,12] In this article we describe recent transport measurements of modulation doped Si:SiGe quantum wells on various offcut substrates at temperatures down to 20 mK. The results are consistent with the formation of a minigap in the density of states. We begin with a discussion of the layer structure and surface morphology of these layers.

II. LAYER STRUCTURE AND SURFACE MORPHOLOGY

The strained silicon quantum wells used in this work all had similar layer structures (see Fig. 1) and were grown by gas source molecular beam epitaxy (MBE), the details of which have been presented elsewhere.[13] Approximately 1% tensile strain is required in the silicon channel to produce the necessary conduction band offset. This can be achieved by growing a thin silicon layer on a relaxed SiGe substrate with a germanium concentration in the range of 20%–30%. In our case the germanium concentration in the buffer layer is increased linearly from 0% to 29% over a thickness of 2 μm and is followed by 1 μm of $Si_{0.72}Ge_{0.28}$. The 100 Å thick quantum well is modulation doped, i.e., it is separated from 500 Å of heavily doped $Si_{0.72}Ge_{0.28}$ by a 150 Å spacer layer with the same alloy concentration. The layers are capped with 50 Å of silicon.

During growth of the graded buffer layer, threading dislocations nucleate and are free to glide on the four equivalent {111} planes producing strain relieving misfit segments (60° dislocations). Surface steps due to the 60° misfit dislocations[14] and reduced growth rates near to them[15] lead to the well known "cross-hatch" pattern on the sample surface. The {111} glide planes intersect the surface of standard (001) substrates along ⟨110⟩ directions and the cross hatching is orthogonal. However, on tilted substrates such as $(11n)$ the situation is different.[11,12] Two of the {111} planes still intersect the surface along $[\bar{1}10]$. The other two planes align at a small angle $\alpha/2$ on either side of [110], given by

$$\cos\alpha = \frac{2\cos^2\phi + \sin^2\phi}{2\cos^2\phi + 3\sin^2\phi}, \tag{1}$$

where $\phi = \arccos(n/\sqrt{n^2+1})$ is the offcut angle between the surface normal and the [001]. The different surface morphologies for an on-axis and a 6° off-axis sample are shown in Fig. 2 and measurements of α for different values of ϕ confirm the relationship in Eq. (1).[12]

The surface roughness of the cross hatching shown in Fig. 2 has a typical correlation length of 5–10 μm with a root mean square (rms) amplitude of ~10 nm for the on-axis

[a]Electronic mail: t.thornton@asu.edu

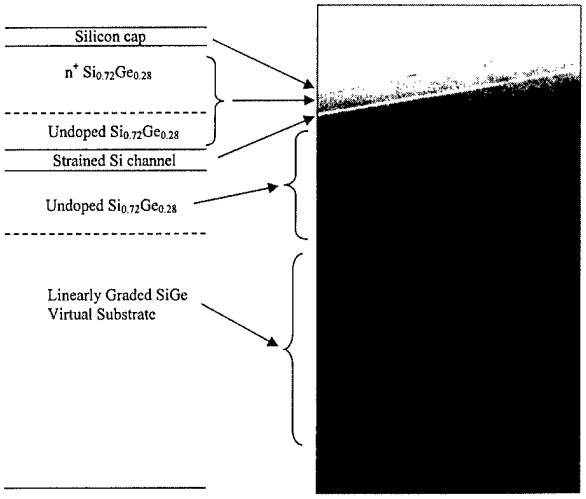


FIG. 1. Layer structure and cross-sectional transmission electron microscopy (TEM) micrograph of one of the strained silicon quantum wells. All of the samples used had similar layer structures. The TEM cross section confirms that the dislocations are buried below the quantum well.

sample, decreasing to ~5 nm for the higher off-cut angle samples. Despite the fact that the surface roughness is comparable to the thickness of the quantum well it does not seem to unduly influence the electron mobility, with values as high as 95 000 $cm^2/V\,s$ being obtained for the quantum wells grown on (001) substrates. However, as the tilt angle is increased, additional surface features with a much smaller length scale develop due to the formation of terraces on the $(11n)$ surface. In Fig. 3 an atomic force microscopy (AFM) image of the 6° sample shows the terraces running along the $[\bar{1}10]$ direction. The surface of Si$(11n)$ substrates consists of single and double steps[16] and assuming an even mixture the average terrace height and width can be calculated. For a 6° off-cut substrate the average terrace width would be ~2 nm. In fact, the average terrace width and height in Fig. 3 is 17 and 2 nm, respectively, suggesting that considerable step bunching has occurred.

III. ELECTRON TRANSPORT ON TILTED SUBSTRATES

The presence of the terraces leads to anisotropic transport properties for the quantum wells grown on tilted substrates.[11,12] For transport perpendicular to the step edges

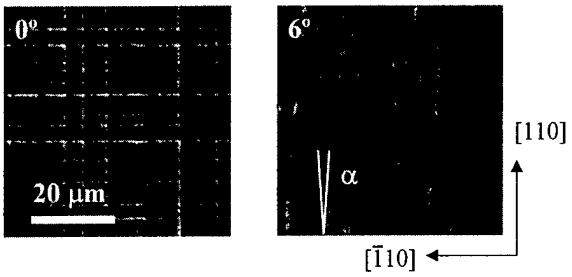


FIG. 2. Nomarski optical micrographs of the surface of the on-axis and 6° tilted samples.

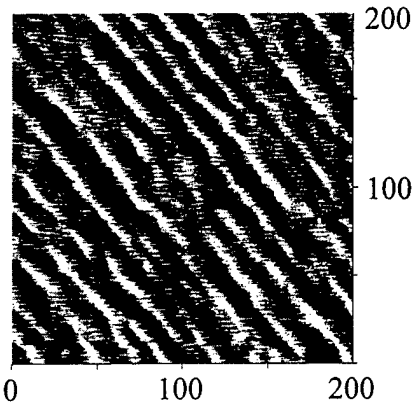


FIG. 3. 200 nm×200 nm AFM image of the surface of the 6° sample. The terraces are parallel to the [110] direction.

extra electron scattering reduces the mobility compared to values measured parallel to the terraces. Figure 4 shows the sheet resistance as a function of magnetic field for a 10° off-cut sample at a temperature of 0.4 K. The samples have standard Hall bar geometries aligned parallel and perpendicular to the step edges. the Shubnikov–de Haas (SdH) oscillations arise from Landau quantization of the density of states as a result of the large magnetic field. They have the same periodicity in each case confirming that the measured sheet density is the same. The higher resistance measured with a current flowing perpendicular to the terraces confirms the lower mobility of this orientation. The mobility measurements are summarized in Table I. The results show a general trend towards increasing anisotropy in the mobility for increasing off-cut angle.

When cooled to lower temperatures the amplitude of the SdH oscillations increases as expected from the reduction in thermal broadening. The 2° and 4° samples, however, show an additional slow oscillation at low fields. The amplitude of the new oscillation is small and easily obscured by the background magnetoresistance and the much larger amplitude of the fast oscillations. In Fig. 5 we show the low field SdH

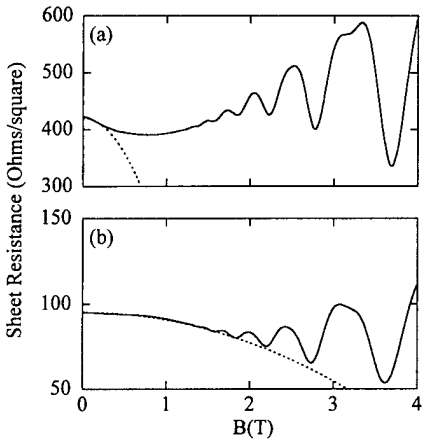


FIG. 4. Magnetoresistance of the 10° sample measured (a) perpendicular and (b) parallel to the terrace edges. The measurement temperatures were (a) 400 and (b) 600 mK.

TABLE I. Electron mobilities and sheet densities of the samples used in this work.

Off-cut angle (deg)	Electron mobility ($cm^2/V\,s$) $\mu_\parallel$	$\mu_\perp$	Sheet density (10^{11} cm^{-2})
0	95 000	95 000	6.9
2	69 400	69 800	9.7
4	45 200	35 300	7.4
6	43 300	40 000	7.1
10	62 600	14 200	10.5

oscillations of the 0°, 2°, 4°, and 6° samples after subtraction of a quadratic term describing the background magnetoresistance.[17] For a uniform two-dimensional electron gas (2DEG) with a sheet density N_s the oscillations have a periodicity that is given by $\Delta(1/B)=4e/N_s\hbar$ and another of twice the frequency if the field is large enough that the spin splitting can be resolved. The data in Fig. 5 are plotted against inverse magnetic field to highlight the $1/B$ periodicity. The slow oscillation is weaker in the 2° sample than in the 4° sample and is not present at all in the 6°, 10°, and on-axis samples. The Fourier transforms of the data from the 0°, 2°, and 4° samples plotted against sheet density are shown in Fig. 6. All of the fast Fourier transform (FFT) spectra have a distinct peak corresponding to sheet densities of 6.9, 9.7, and 7.4×10^{11} cm^{-2} for the 0°, 2°, and 4° samples, respectively. However, only the 2° and 4° samples have a distinct peak at frequencies corresponding to lower electron concentrations. This additional peak is due to the slow oscillation present in the data of Figs. 5(b) and 5(c).

Results similar to those in Figs. 5 and 6 have been reported for silicon MOS inversion layers on 1.2° and 10° off-cut substrates.[6,18] In both these cases a slow oscillation evolved into a faster oscillation with increasing magnetic

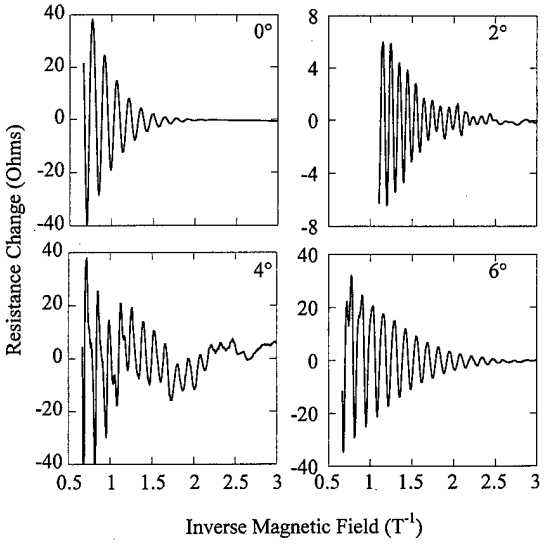


FIG. 5. Magnetoresistance of the 0°, 2°, 4°, and 6° samples plotted against inverse magnetic field. The measurement temperature for the on-axis sample was 80 mK and was less than 20 mK for the others. The background magnetoresistance was subtracted from the data.

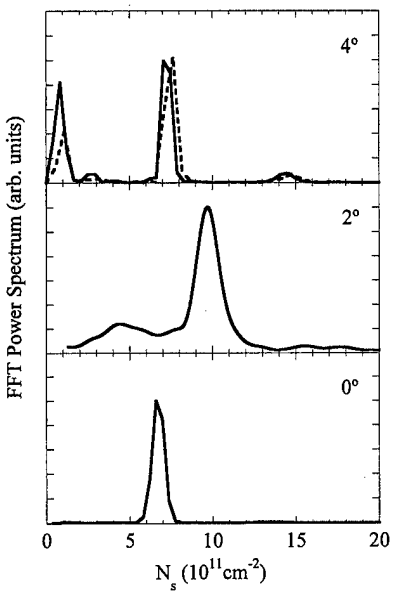


FIG. 6. FFT power spectra for the 0°, 2°, and 4° data of Fig. 5 (solid lines). The dashed line shows the FFT from the 4° sample at a temperature of 400 mK.

field and the results were interpreted in terms of the so-called valley projection model (VPM) which predicts the existence of a minigap in the density of states.[8] For a 2DEG on a (001) substrate there are two low energy valleys in the E-k dispersion with minima at $(0,0,\pm k_0)$ where $k_0=0.85(2\pi/a)$, a being the lattice constant, which is 5.43 Å for silicon. When projected onto the (001) surface the two valleys are coincident and the valley degeneracy is 2. However, when projected onto a low angle vicinal substrate such as $(11n)$ the valleys no longer coincide but cross at points given by $k=0.15k_0\sin\varphi$ where φ is the off-cut angle. The degeneracy is lifted at the crossing points due to valley–valley interactions and a minigap is formed. The width of the minigap increases with the sheet density and off-cut angle and is typically[8] a few meV for densities of a few 10^{12} cm^{-2}.

Figure 7 illustrates the occupation of the conduction band in the 2°, 4°, and 6° samples and can be used to explain the results of Fig. 5. For both the 2° and 4° samples the minigap is located below the Fermi energy, E_F. The lower band is completely full and for small magnetic fields only the upper band contributes to the SdH oscillations with a periodicity corresponding to a sheet density N_{slow}, i.e., the slow oscillations observed in the 2° and 4° samples. As the magnetic field increases, tunneling can occur through the minigap and for fields such that $E_g=\hbar\omega_c$ magnetic breakdown[19] occurs and the minigap cannot be resolved. At this point the SdH oscillations are faster with a periodicity corresponding to a sheet density N_{fast}. The situation for the 4° sample is similar except now the minigap is higher in energy, because of the larger tilt angle, and the Fermi energy in this sample is smaller. For the 6° sample the minigap is higher than the Fermi energy and only the faster oscillations are observed. Although this argument explains the qualitative behavior we observe, the position of the minigap is not in quantitative

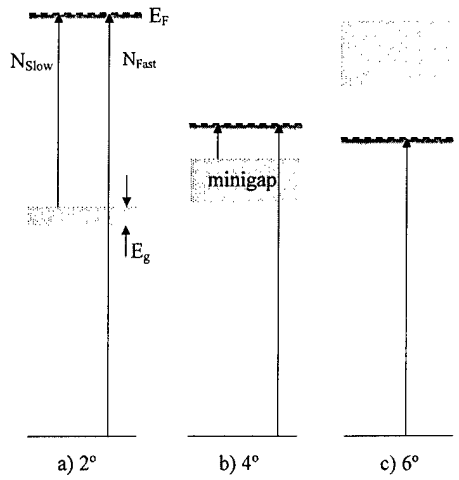


FIG. 7. Schematic illustration of the electron occupancy in the conduction band for the 2°, 4°, and 6° samples. The vertical arrows represent the carriers contributing to the fast and slow oscillations observed in the magnetoresistance.

agreement. For vicinal MOS field effect transistors (MOSFETs) the lowest minigap would occur at an energy corresponding to a sheet density of 1.2 and 4.7×10^{11} cm^{-2} for the 2° and 4° samples, respectively. The values of N_{fast} and N_{slow} that we obtain would place the energy gap at an occupancy of approximately 4 and 6×10^{11} cm^{-2} although, without knowing the exact width of the minigaps in each case, assigning accurate values is difficult.

We have measured the low field SdH oscillations in the 4° sample as a function of temperature in an attempt to determine the width of the minigaps. The amplitude of the slow oscillation peak in the FFT power spectrum decreases with temperature as is shown in Fig. 6. By extrapolation to zero amplitude the slow oscillation component vanishes at a temperature of ~0.75 K corresponding to a thermal energy of 65 μeV. If thermal broadening were reducing the width of the minigap and consequently reducing the amplitude of the slow SdH component, then 65 μeV would be a lower estimate for the width of the minigap. However, this value is smaller than we expect from comparison with vicinal MOSFETs, and thermal broadening may be reducing the amplitude via other mechanisms.

The minigap can also be estimated from the magnetic field at which breakdown occurs, i.e., when the fast oscillations first appear. For the 4° sample this value is approximately 0.4 T and assuming $E_g=\hbar\omega_c$ we get an energy gap of 0.24 meV. Again, this value is somewhat smaller than we would expect from a similar MOSFET sample but the additional strain in the silicon quantum well may be influencing the width of the minigap.

IV. CONCLUSION

We have measured the transport properties of strained silicon quantum wells on substrates tilted away from the (001) normal by 0°, 2°, 4°, 6°, and 10°. All the layers shows cross hatching with a characteristic length scale of 10 μm while the tilted substrates show terracing with much smaller characteristic length scales. The extra scattering induced by the terraces leads to anisotropic electron mobility when measured parallel and perpendicular to the step edges. At the lowest temperatures an additional slow oscillation is seen in the magnetoresistance of the 2° and 4° samples that we attribute to the presence of a minigap.

ACKNOWLEDGMENT

The authors thank Dr. Juan Fernandez of Imperial College, London, for growing the strained silicon samples used in this work.

[1]B. A. Joyce, J. H. Neave, J. Zhang, D. D. Vvedensky, S. Clakre, K. J. Hugill, T. Shitara, and A. K. Myersbeaghton, Semicond. Sci. Technol. **5**, 1147 (1990).
[2]M. S. Miller, H. Weman, C. E. Pryor, M. Krishnamurthy, P. M. Petroff, H. Kroemer, and J. L. Merz, Phys. Rev. Lett. **68**, 3464 (1992).
[3]T. Fukui, K. Tsubaki, H. Saito, M. Kasu, and S. Honda, Surf. Sci. **267**, 588 (1992).
[4]M. R. Fahy, J. H. Neave, M. J. Ashwin, R. Murray, R. C. Newman, B. A. Joyce, Y. Kadoya, and H. Sakaki, J. Cryst. Growth **127**, 871 (1993).
[5]H. Sakaki (private communication).
[6]P. J. Stiles, T. Cole, and A. A. Lakhani, J. Vac. Sci. Technol. **14**, 969 (1997).
[7]T. Cole, A. A. Lakhani, and P. J. Stiles, Phys. Rev. Lett. **38**, 722 (1997).
[8]L. J. Sham, S. J. Allen, A. Kamgar, and D. C. Tsui, Phys. Rev. Lett. **40**, 472 (1978).
[9]G. L. Zhou, Z. Ma, M. E. Lin, L. H. Allen, and H. Markoç, Appl. Phys. Lett. **63**, 2094 (1993).
[10]P. M. Mooney, F. K. LeGoues, J. Tersoff, and J. O. Chu, J. Appl. Phys. **75**, 3968 (1994).
[11]T. J. Thornton, J. M. Fernandez, S. Kaya, P. W. Green, and K. Fobelets, Appl. Phys. Lett. **70**, 1278 (1997).
[12]P. Waltereit, J. M. Fernandez, S. Kaya, and T. J. Thornton, Appl. Phys. Lett. **72**, 2262 (1998).
[13]J. M. Fernandez, A. Matsumura, X. M. Zhang, M. H. Xie, L. Hart, T. J. Thornton, and B. A. Joyce, J. Mater. Sci. **6**, 330 (1995).
[14]S. Y. Shirayaev, F. Jensen, and J. W. Peterson, Appl. Phys. Lett. **64**, 3305 (1994).
[15]E. A. Fitzgerald, Y. H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, A. R. Kortan, F. A. Thiel, and B. E. Weir, J. Vac. Sci. Technol. B **10**, 1807 (1992).
[16]J. Wasserful and W. Ranke, Surf. Sci. **315**, 227 (1994).
[17]T. J. Thornton, A. Matsumura, and J. Fernandez, Surf. Sci. **361/362**, 547 (1996).
[18]S. M. Goodnick, J. R. Sites, K. S. Yi, D. K. Ferry, and C. W. Wilmsen, Phys. Lett. **97A**, 111 (1983).
[19]A. B. Pippard, Proc. R. Soc. London, Ser. A **207**, 1 (1962).

DX centers in $Al_{0.37}Ga_{0.63}As/GaAs$ and $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ heterostructures

H. H. Wieder[a] and Huseyin Sari
Department of Electrical and Computer Engineering, University of California, La Jolla, California 92092-0407

(Received 19 January 1999; accepted 17 May 1999)

Electrical and galvanomagnetic measurements, made on modulation δ-doped-(Si) $Al_{0.37}Ga_{0.63}As/GaAs$ and $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ heterostructures, fabricated into modulation doped field effect transistor-like gated Hall bars, were used to determine their *DX* center energies and densities in the normal and persistent photoconductive mode. Self-consistent Poisson/Schrödinger simulations of the gate voltage dependence and of the temperature dependence of the charge transport parameters of these heterostructures provide the position of the Fermi levels in the barrier layers, E_{FB}, relative to the conduction band minima and the electron densities in their quantum wells. The energy, $E_{DX}=0.14$ eV of $Al_{0.37}Ga_{0.63}As$ determined from the temperature independent equilibrium position of E_{FB} is consistent with the average of the three lowest *DX* center energies of this alloy. It is also consistent with that determined by others, on epitaxial $Al_{0.37}Ga_{0.63}As$ layers, using conventional Hall measurement, deep level transient spectroscopic and hydrostatic pressure measurements. Strain relaxed $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ heterostructures, grown on GaAs substrates, were used for similar measurements and simulations. These yield the energy of the *DX* centers in $In_{0.34}Al_{0.66}As$, relative to its Γ-valley minimum, $E_{DX}=0.18$ eV. This value in conjunction with the previously determined *DX* center energies of AlAs and that of $In_{0.52}Al_{0.48}As$, above its conduction band edge, $E_{DX}(x)$ is correlated with, but not identical to, the composition dependence of the *L* band of $In_xAl_{1-x}As$. © *1999 American Vacuum Society.*
[S0734-211X(99)07004-3]

I. INTRODUCTION

The Si donor in compound semiconductors such as $Al_xGa_{1-x}As$ or $In_xAl_{1-x}As$ can exist in either one of two configurations: a shallow, hydrogenic-like donor (d^+,d^0) at a substitutional site, or a deep, interstitial state, displaced along ⟨111⟩, leaving a vacancy at the substitutional site.[1] The large lattice relaxation which takes place at (or below) a critical temperature, T_c, is accompanied by the localization of two electrons on this deep donor defined as a *DX* center.[2,3] Metastability of such *DX* centers appears in the persistence of their charge below T_c and it is interpreted in terms of configurational (electron+elastic) barriers which affect the emission and capture of electrons. As long as the thermal energy is small compared to the height of these configurational barriers emission and capture of electrons is negligibly small. However, optical excitation of the *DX* centers can induce the transfer of their electrons to the conduction band where they remain, producing a persistent photoconductivity (PPC) because they cannot overcome this configurational barrier.

For any $Al_xGa_{1-x}As$ composition, the energy of the *DX* centers depends on the molar Al concentration and their density depends on the Si doping density, N_d. In $Al_{0.2}Ga_{0.8}As$ where the *DX* center energy level is slightly above the conduction band minimum (CBM) (at $x=0.22$ the thermal binding energy of the *DX* centers becomes equal to that of the shallow donors) the ratio of the shallow donors to *DX* centers depends[4] on N_d and a PPC is observed below $T_c \sim 150$ K. In higher molar Al concentrations the Fermi level, E_F, rises with increasing N_d but becomes pinned[5] on the energy, E_{DX}, of the *DX* center, while in GaAs, where E_{DX} lies above the conduction band edge, it has been estimated[6] that between 1.5 and 2×10^{19} cm^3 Si donors might be required for E_F to become resonant with the *DX* center energy. An alternative approach to raising E_F is to lower E_{DX} relative to the CBM by the application of hydrostatic pressure. In GaAs a hydrostatic pressure of at least 20 kbar is required to bring E_{DX} below E_F.Qualitatively, hydrostatic pressure has the same effect as an increase in the Al molar density. Furthermore, the *DX* center is sensitive not only to the number of nearest neighbor Al atoms but also to their atomic configuration.[7,8]

In $Al_xGa_{1-x}As/GaAs$ modulation doped heterostructures electrons can transfer from the *DX* centers into the lower lying states of the GaAs quantum well and the electric field between the positively charged doping layer and the two-dimensional (2D) negative electrons causes a band bending which limits the transfer of electrons when the Fermi level reaches the *DX* energy. In that case only a fraction of the *DX* centers are occupied and the position of the Fermi level depends on the *DX* center density of states. If the *DX* centers are located[9] in the barrier layer of the modulation doped field effect transistor (MODFET), then the Fermi level in the barrier layer, E_{FB}, can be raised by raising the gate voltage, V_g. The electron density in the quantum well, n_s, rises with V_g and then saturates; however, gate current leakage may limit the maximum applicable V_g. Galvanomagnetic measure-

[a]Electronic mail: wieder@ece.ucsd.edu

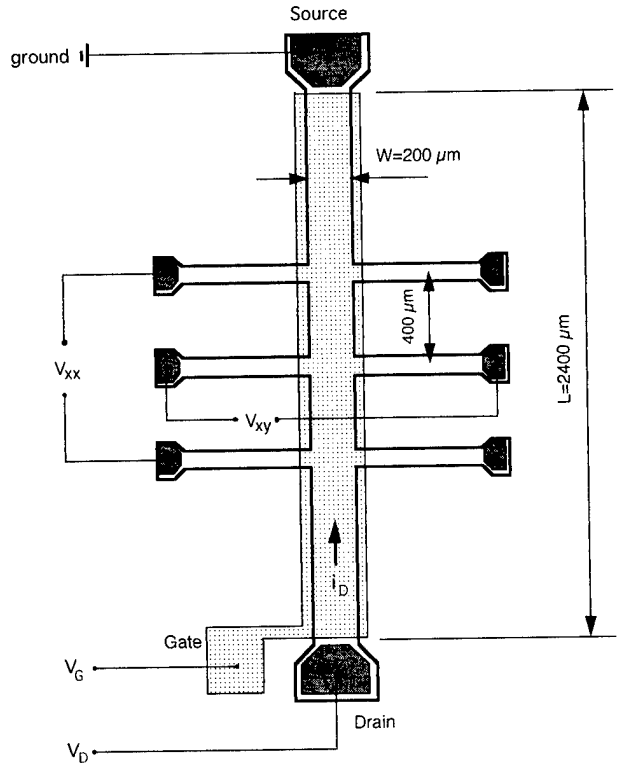


FIG. 1. Configuration of MODFET-like, δ-doped eight-arm gated Hall bar structure used for investigating the properties of the *DX* centers in the barrier layers of $Al_xGa_{1-x}As$ and $In_xAl_{1-x}As$ heterostructures.

ments were used previously by ourselves and others[10–13] for investigating the composition dependence and other properties of $Al_xGa_{1-x}As$ and $In_xAl_{1-x}As$-based modulation-doped heterostructures containing *DX* centers. Such measurements were used, in the vicinity of the direct to indirect transition of these alloys, to determine the energy and the Fermi level dependence of the *DX* centers in two Si δ-doped heterostructures: $Al_{0.37}Ga_{0.63}As/GaAs$ and $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$, described in this article.

II. $Al_{0.37}Ga_{0.63}As/GaAs$ HETEROSTRUCTURES

$Al_{0.37}Ga_{0.63}As/GaAs$ heterostructures were grown by means of solid source molecular beam epitaxy on semi-insulating (100)-oriented GaAs substrates. An $Al_{0.37}Ga_{0.63}As$ barrier layer, with a nominal thickness of 70 nm, δ doped with Si to $\sim 7.5\times 10^{12}$ cm^2 was deposited on top of a 300-nm-thick, undoped, GaAs layer with an intermediate, 10-nm-thick, undoped, $Al_{0.37}Ga_{0.63}As$ spacer layer between them. A two-dimensional electron gas (2DEG) is induced in the quantum well of the GaAs layer by modulation doping. Such heterostructures were used for making, by means of photolithographic procedures, eight-arm gated Hall bar MODFET-like test structures, such as shown in Fig. 1. The electron density, n_s, in the quantum well as a function of V_g, measured with temperature as a constant parameter, between 25 and 160 K, shown in Fig. 2, is similar to that obtained by Buks *et al.*;[10] $n_s(V_g)$ consists of an extended linear region which becomes sublinear with increasing V_g as n_s tends toward saturation between 25 and 120 K. The $n_s(V_g)$ curve at

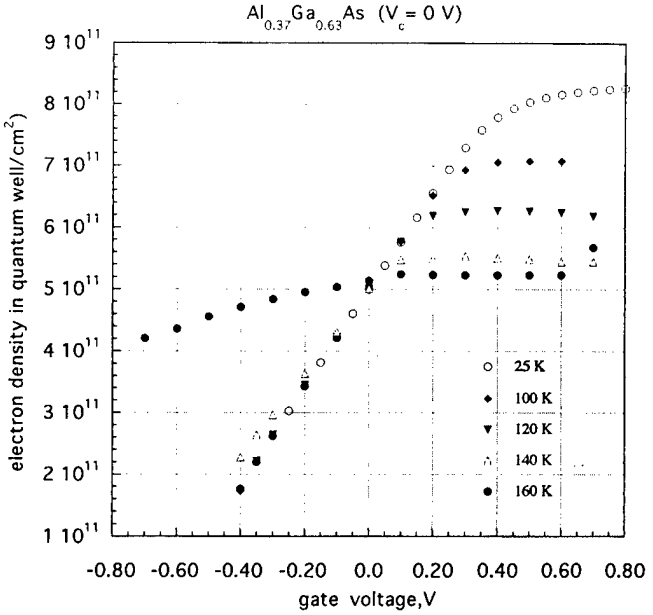


FIG. 2. Electron density measured as a function of gate voltage and temperature in the quantum well of an $Al_{0.37}Ga_{0.66}As/GaAs$ heterostructure.

160 K has a negligible linear range and suggests that the Fermi level is pinned at or near this temperature.

Figure 3 shows that $n_s(V_g)$ also depends on the constant "freeze-in" potential, V_c, applied to the gate while it is cooled below T_c. A negative V_c applied during cooldown is considered to increase the shallow donor density; consequently, more electrons can be transferred into the 2DEG with increasing V_g than for $V_c=0$ or for positive V_c values. The $n_s(V_g)$ in Fig. 3 is considered here in terms of a two-piece approximation: a linear range in which the MODFET

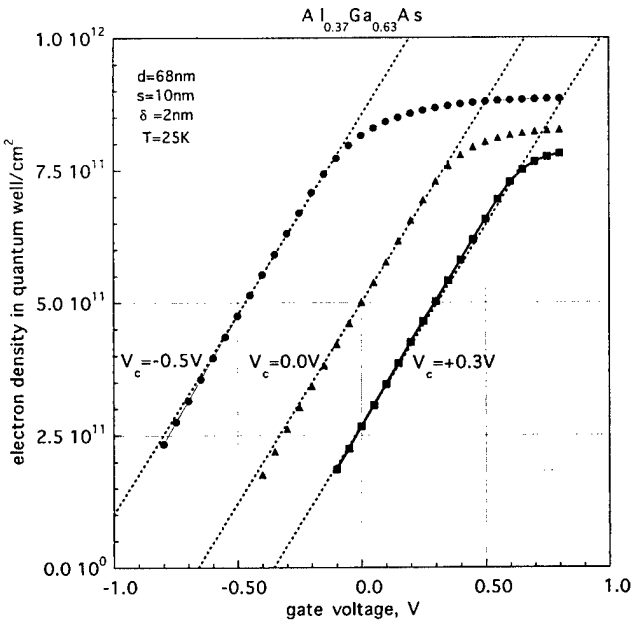


FIG. 3. Electron density, n_s, measured as a function of gate voltage, in the same quantum well shown in Fig. 2, as a function of the freeze-in potential, V_c at constant temperature.

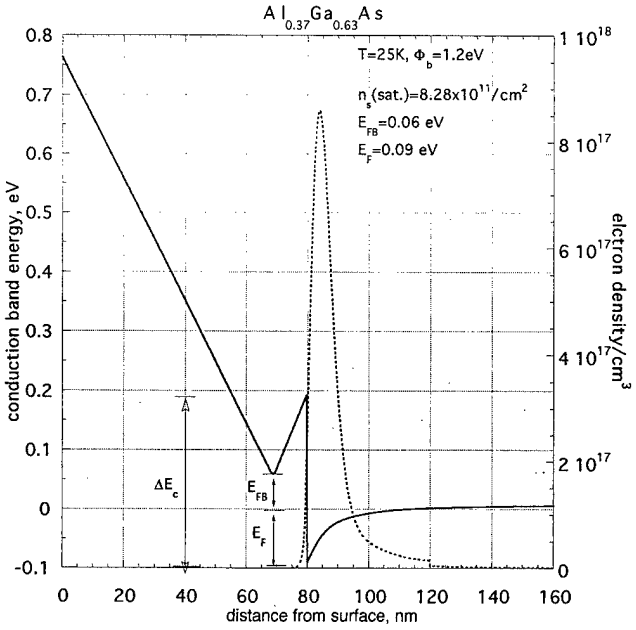


FIG. 4. Poisson/Schrödinger self-consistent simulation of the conduction band edge dependence on position within the heterostructure, shown in Fig. 2, at 25 K, with n_s at saturation; ΔE_c=conduction band offset between $Al_{0.37}Ga_{0.66}As$ and GaAs, E_{FB}=Fermi level relative to conduction band minimum in barrier layer, E_F=Fermi level relative to conduction band minimum of quantum well. Dashed line shows position dependence of electron density in quantum well.

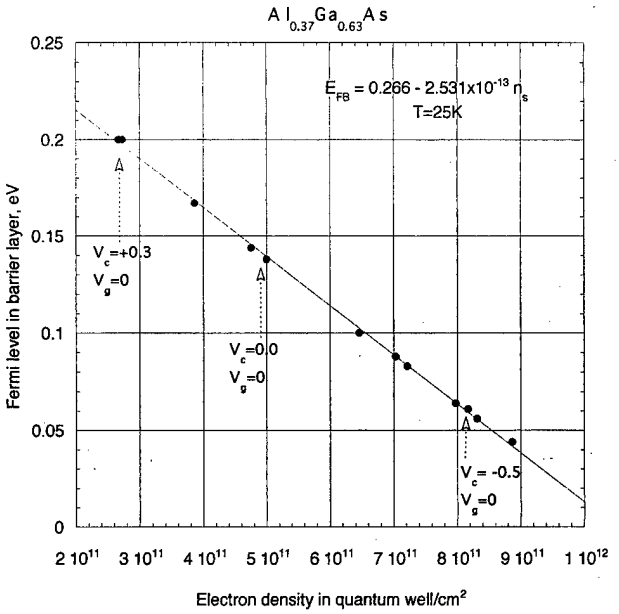


FIG. 5. Fermi level in barrier layer calculated as a function of gate voltage and freeze-in potential at 25 K for the same specimen shown in Figs. 2 and 3.

charge control equation, $qn_s = C_i(V_g - V_T)$, is applicable. The capacitance between the δ-doping plane and the center of charge of the 2DEG electrons, $C_i = 1.21\times10^{-7}$ F/cm^2, is essentially the same for each V_c. The *DX* centers are considered to remain so for $V_c = 0$; a threshold, $V_T(n_s = 0) = -0.66$ V, is obtained by a linear extrapolation, and full saturation is assumed at $V_g(th) \geq 0.44$ V with $n_s(\text{sat}) = 8.28\times10^{11}$ cm^2.

In order to fit the linear $n_s(V_g)$ portion of the $V_c = 0$ curve in Fig. 3, δ doping is modeled as $N_d = 7.6\times10^{18}$ cm^3 contained within the last 2 nm, at the edge of the barrier layer, providing an effective sheet donor density, $N_T = (N^+ - N^X) = 1.52\times10^{12}$ cm^2. Here N^+=the shallow ionized donor density and N^X is the occupied *DX* center density. N_T was derived from the relation

$$V_T = \phi_b - \Delta E_c - qN_T/C_a \quad (1)$$

with the surface barrier, $\phi_b = 1.2$ eV, the conduction band offset between $Al_{0.37}Ga_{0.63}As$ and GaAs, $\Delta E_c = 0.287$ eV, and with $C_a = 1.55\times10^{-7}$ F/cm^2 representing the ratio of the permittivity of $Al_{0.37}Ga_{0.63}As$, $\epsilon_s = 11.94$, to the barrier layer thickness, $d = 68$ nm. For the other two curves shown in Fig. 3, $N_T(V_c = -0.5\ \text{V}) = 19.98\times10^{12}$ cm^2 and $V_T = -1.134$ V while for $N_T(V_c = +0.3\ \text{V}) = 1.23\times10^{12}$ cm^2 $V_T = -0.354$ V. These parameters were used for solving self-consistently[14] the Poisson/Schrödinger equations which provide the position dependence of the conduction band energy within the heterostructure relative to its surface, the specific energies and the electron densities in the occupied subbands of the quantum well (whose sum is in good agreement[12] with n_s derived from the Hall measurements). Figure 4 shows the result of such a calculation performed for $n_s(\text{sat}) = 8.28\times10^{11}$ cm^2 which yields a Fermi level, $E_{FB} = 0.06$ eV, relative to the conduction band minimum of the barrier layer. By means of similar calculations we have obtained $n_s(V_g)$ data consistent with the measurements shown in Fig. 3 and a dependence of E_{FB} on $n_s$such as shown in Fig. 5.

Figure 2 shows that for $V_c = 0$, $n_s(V_g = 0) = 5\times10^{11}$ cm^2, it is essentially independent of temperature for $T \leq T_c$, and from Fig. 5 that it correspond to $E_{FB} = 0.14$ eV. We consider this E_{FB} to be in equilibrium with the average of the energies of the three lowest discrete *DX* centers[15] of this alloy: $DX1 = 117$ meV, $DX2 = 138$ meV, and $DX3 = 189$ meV, effectively, $E_{DX} = 146$ eV. The *DX* center nearest to the conduction band edge, $DX0 = 61$ meV, is considered to be empty. Raising E_{FB} by raising V_g, initially, ionizes additional shallow donors increasing n_s while keeping the density of the "frozen-in" electrons in the *DX* centers constant. For higher V_g a fraction of the electron emitted from the shallow donors are trapped, possibly on *DX* centers, making $n_s(V_g)$ sublinear; saturation sets as E_{FB} reaches $DX0$.

Persistent photoconductivity (PPC) obtained on an identical but ungated Hall bar test structure, subjected to light, generated by an adjacent GaAs light emitting diode (LED), is shown in Fig. 6; one branch represents $n_s(T)$ while cooling from room temperature to 25 K and the other branch is that obtained by warming the specimen back to room temperature. T_c is considered to be at (or near) ~150 K, the inflection point of the warmup curve. PPC induce an additional electron density, $\Delta n_s = 3.4\times10^{11}$ cm^2, above the equilibrium value of n_s at 25 K, and the corresponding incremental change in the ionized electron density, ΔN_T can be calculated from the relation

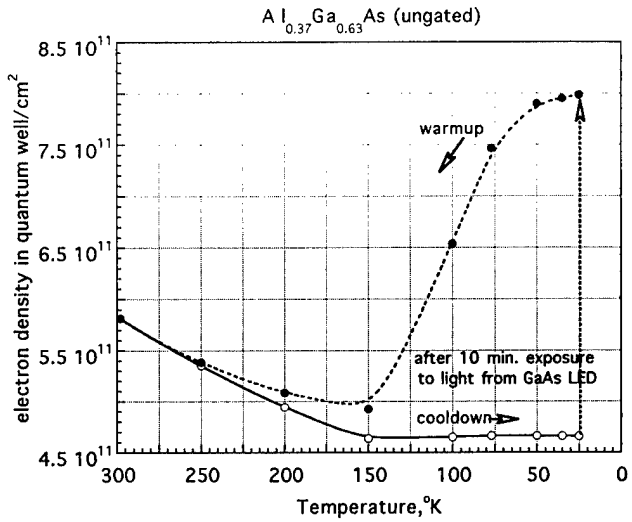


FIG. 6. Electron density in the quantum well of an ungated Hall bar (similar in other respects to that shown in Figs. 2–5), at $V_g=0$, as a function of temperature, in the normal and in the persistent photoconductive mode.

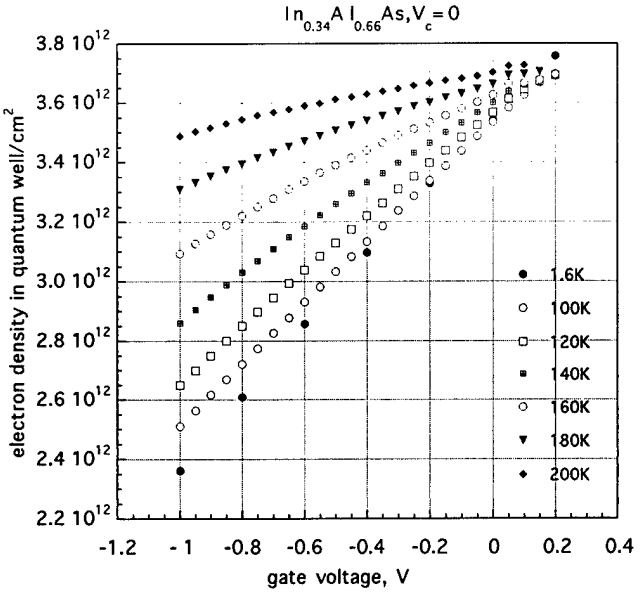


FIG. 7. Electron density measured as a function of gate voltage and temperature in the quantum well of an $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ heterostructure.

$$\Delta N_T=\frac{\Delta n_s}{(\partial n_s/\partial V_c)}\left(\frac{dN_T}{dV_c}\right). \tag{2}$$

At constant $V_g=0$, $(\partial n_s/\partial V_c)=-7.27\times10^{11}$ and $dN_T/dV_c=-9.40\times10^{11}$ in Fig. 3; consequently, $\Delta N_T=4.4\times10^{11}$ cm^2. PPC converts the *DX* centers into ionized shallow donors; hence $N_T(\mathrm{PPC})=N^+ + N^X$ and, because $\Delta N_T=N_T(\mathrm{PPC})-N_T=2N^X$, it follows that $N^X=2.2\times10^{11}$ cm^2, $N^+=1.74\times10^{12}$ cm^2, and $N_T(\mathrm{PPC})=1.96\times10^{12}$ cm^2.

This is also consistent with the grand canonical ensemble of Lazzouni and Sham:[16]

$$\frac{N^+-N^X}{N^++N^X}=\frac{1-4\exp[\beta(2E_{\mathrm{FB}}-E_S-E_{DX})]}{1+\exp[\beta(2E_{\mathrm{FB}}-E_S-E_{DX})]}, \tag{3}$$

where $\beta=1/kT$ and k is Boltzmann's constant, $E_S\sim10$ meV, is considered to be the energy of the shallow donor, E_{DX} is the equilibrium energy of the electrons on the *DX* centers; at 25 K the ratio: $(N^+-N^X)/(N^++N^X)=0.77$. Therefore, $E_{DX}=280$ meV which corresponds to 140 meV per electron. This value is in good agreement with $E_{DX}=0.14$ eV, that of Chadi and Chang[2] and $E_{DX}=0.13$ eV determined by Leuther *et al.*,[11] derived from Hall measurements made, at 4 K, on ungated Van der Pauw-type heterostructure.

III. $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ HETEROSTRUCTURES

$In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ lattice matched heterostructures were grown on GaAs substrates by solid source molecular beam epitaxy. Compositionally step-graded, $x=0.1$/step and 0.2 μm/step $In_xGa_{1-x}As$ buffer layers were used to relax the mismatch strain between these heterostructures and their substrates.[6] Each, nominally, 40-nm-thick $In_{0.34}Al_{0.66}As$ barrier layer is δ-doped (Si), adjacent to its interface with an undoped, 3-nm-thick, $In_{0.34}Al_{0.66}As$ spacer layer which separates it from a 50-nm-thick, undoped, $In_{0.36}Ga_{0.64}As$ layer. The latter contains the two-dimensional electron gas (2DEG), generated by modulation doping. The heterostructures were processed by means of photolithography, etching and lift-off procedures into the gated Hall bar test structures shown in Fig. 1. The fractional In concentrations were determined by means of a double crystal x-ray diffractometer with a resolution of 2 arcsec. The $In_{0.34}Al_{0.66}As$ layer is in the direct band gap range just above the direct-to-indirect transition and consequently the same methods can be used to derive its properties as those used for the above described $In_{0.37}Al_{0.66}As$-based heterostructures.

Figure 7 shows $n_s(V_g)$ as a function of temperature between 1.6 and 200 K of an $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ gated Hall bar structure. While $n_s(V_g=0)=3.54\times10^{12}$ cm^2 at 1.6 K, at other temperature $n_s(V_g=0)$ has a slight dispersion, unlike the temperature independence of this parameter found on $Al_{0.37}Ga_{0.63}As/GaAs$ heterostructures. Due to the onset of non-negligible gate current we were unable to obtain a strong tendency to saturation of $n_s(V_g)$. However, Fig. 8, which shows the dependence of $n_s(V_g)$ on V_c at 1.6 K, is qualitatively similar to that shown in Fig. 3; at $V_c=0, V_T=-3.0$ V; a $N_T(V_g=0)=3.37\times10^{12}$ cm^2 was determined by means of Eq. (1) with $\phi_b=1.2$ eV, $\Delta E_c=0.69$ eV, and with $C_a=1.53\times10^{-7}$ F/cm^2 essentially constant and independent of V_c. These parameters were also used for Poisson/Schrödinger calculations, similar to those shown in Fig. 4, with the result, shown in Fig. 9, which demonstrates that for $V_g=0$ the Fermi level in the barrier layer is $E_{\mathrm{FB}}=0.18$ eV (given the dispersion in the temperature dependence of $n_s(V_g=0)$ shown in Fig. 7, we estimate that $E_{\mathrm{FB}}=0.18\pm0.01$ eV).

Figure 10 shows the optically induced PPC, as a function of temperature, measured on an identical but ungated Hall bar at 25 K. The change in the electron density produced by the PPC is $\Delta n_s=4\times10^{11}$ cm^2 and by means of the data from

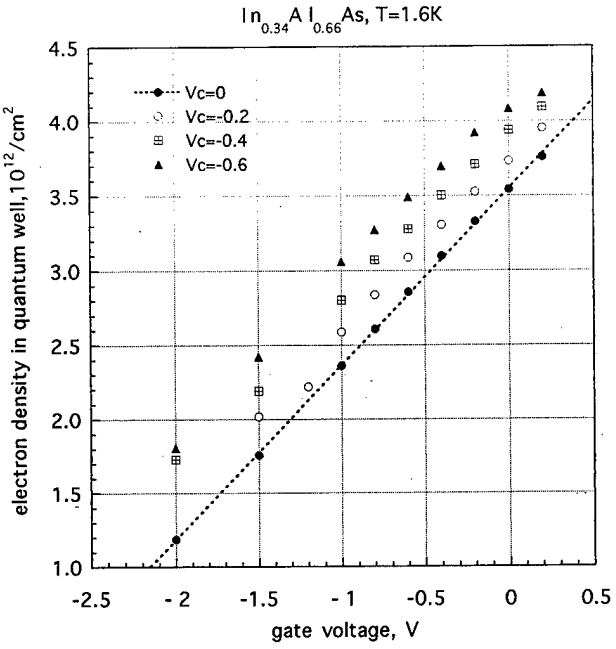


FIG. 8. Electron density, n_s, measured as a function of gate voltage, on the same quantum well shown in Fig. 7, as a function of the freeze-in potential, V_c at constant temperature of 1.6 K.

Fig. 8, $\partial n_s/\partial V_c = -9\times10^{11}$ cm^2 and $dN_T/dV_c = 1.18\times10^{12}$ cm^2. It follows from Eq. (2) that $N^X = 2\times10^{11}$ cm^2 and $N^+ + N^X = 3.8\times10^{12}$ cm^2. The application of the grand canonical ensemble of Eq. (3) to this heterostructure yields $E_{DX} = 0.18$ eV for a ratio $(N^+ - N^X)/(N^+ + N^X) = 0.88$.

The composition dependence of the fundamental band gap of $In_xAl_{1-x}As$, $E(\Gamma) = 3.13 - 3.55x + 0.84x^2$, can be obtained from an interpolation of the $E(\Gamma)$ of AlAs[17] and InAs.[18] The composition dependence of the $In_xAl_{1-x}As$ L band may be interpreted in terms of Adachi's data[19] for AlAs at 1.5 K, $E(L) = 2.35$ eV, and from the theoretical calculations of Cohen and Chelikowski[20] for InAs, $E(L) = 1.5$ eV, at 1.5 K. A linear interpolation between them is, therefore, $E(L) = 2.35 - 0.85x$ and the direct to indirect transition of $In_xAl_{1-x}As$[21] occurs at $x = 0.32$.The band edge of the X valley is uncertain; we consider it, tentatively,[11] to be of the form $E(X) = 2.23 - 0.47x$.

Figure 11 shows, in this context, the composition dependence of the Γ, L, and X bands relative to the valence band edge of $In_xAl_{1-x}As$. Also shown is $E_{DX}(x=0) = 2.11$ eV,

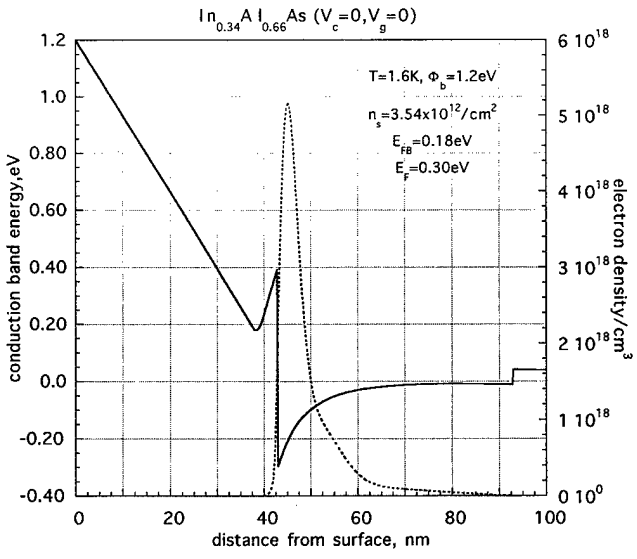


FIG. 9. Poisson/Schrödinger self-consistent simulation of the conduction band edge dependence on position within the same heterostructure, shown in Fig. 7, with n_s at $V_g = 0$; ΔE_c = conduction band offset between $In_{0.34}Al_{0.66}As$ and $In_{0.36}Ga_{0.64}As$. E_{FB} = Fermi level relative to conduction band minimum in barrier layer, E_F = Fermi level relative to conduction band minimum of quantum well. Dashed line shows position dependence of electron density in quantum well.

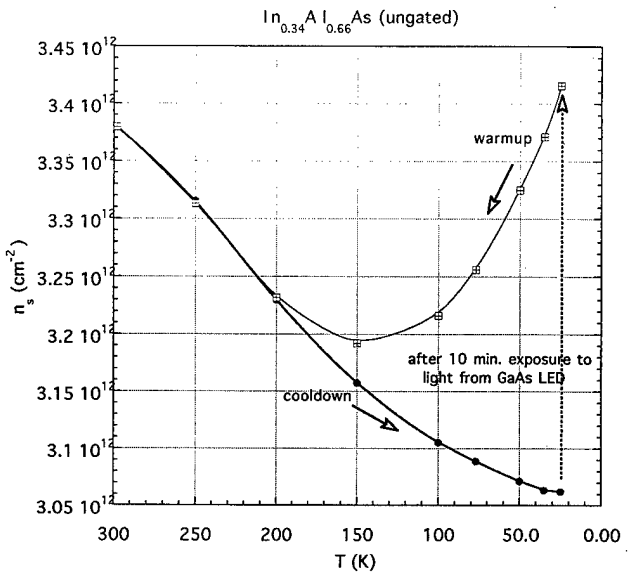


FIG. 10. Electron density in the quantum well of an ungated Hall bar (similar in all other respects to that shown in Figs. 7–9), at $V_g = 0$, as a function of temperature, in the normal and in the persistent photoconductive mode.

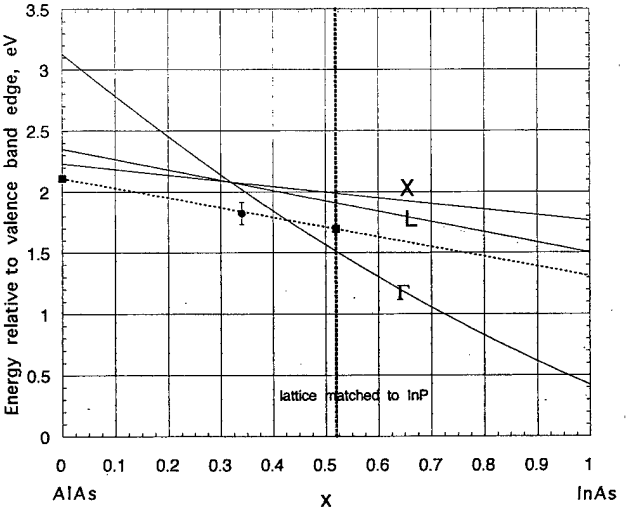


FIG. 11. *DX* center energy of $In_{0.34}Al_{0.66}As$ calculated from the data in Figs. 7–10 shown in context of the composition dependence and band edges of $In_xAl_{1-x}As$ as well as the data obtained by other (Refs. 11 and 22) and an assumed composition dependence of the *DX* centers in this alloy system.

determined by Leuther *et al.*,[11] and $E_{DX}(x=0.52)=1.69$ eV, which is 183 ± 10 meV above the conduction band edge of $In_{0.52}Al_{0.48}As$, determined by Fruth *et al.*,[22] using hydrostatic pressure-dependent galvanomagnetic measurements. Figure 11 also shows our own $E_{DX}(0.34)=1.84$ eV and a dashed line through these data suggests that $E_{DX}(x)$ is correlated but is not identical with $E_L(x)$; this might be expected from the considerations advanced by Chadi and Chang[3] that *DX* centers might be coupled to the average of the principal band edge minima of the entire Brillouin zone weighted by their densities of states rather than to a specific conduction band minimum.

IV. CONCLUSIONS

Experimental results, obtained from galvanomagnetic measurements, made as a function of temperature, on gated Hall bar test structures, indicate that the equilibrium Fermi level, $E_{FB}=0.14$ eV, in the barrier layer of a modulation doped $Al_{0.37}Ga_{0.63}As/GaAs$ heterostructure, cooled below T_c, in $V_g=0$, is consistent with the average energy of the three lowest lying *DX* centers of $Al_{0.37}Ga_{0.63}As$, $E_{DX}\sim0.146$ eV. This E_{FB} is consistent with the theoretically predicted energy of a single *DX* center and with measurements made on epitaxial layers and on heterojunctions using hydrostatic pressure-dependent and deep level transient spectroscopic measurements. In a similar manner the Fermi level in the barrier layer of $In_{0.34}Al_{0.66}As/In_{0.36}Ga_{0.64}As$ heterostructures, cooled below T_c, comes to rest in $V_g=0$ at $E_{DX}=0.18$ eV. This is consistent with $E_{DX}(x)$ represented as a linear interpolation between E_{DX}(AlAs) and $E_{DX}(In_{0.52}Al_{0.48}As)$, determined by others. The sublinear $n_s(V_g)$ and its saturation attributed to trapping of electrons on *DX* centers and on pinning of the Fermi level on the highest unoccupied *DX*0 level.

ACKNOWLEDGMENTS

The authors are grateful for the support of the Office of Naval Research in the course of this work. A portion of this paper was presented orally at UCSB in celebration of Professor Herbert Kroemer's 70th birthday.

[1]P. M. Mooney, J. Appl. Phys. **67**, R1 (1990).
[2]D. J. Chadi and K. J. Chang, Phys. Rev. Lett. **61**, 873 (1988).
[3]D. J. Chadi and K. J. Chang, Phys. Rev. B **39**, 10 063 (1989).
[4]T. Ishikawa, T. Yamamoto, K. Kondo, J. Komeno, and A. Shibatomi, Inst. Phys. Conf. Ser. **83**, 99 (1986).
[5]A. Y. Du, M. F. Li, T. C. Chong, and S. J. Chua, Appl. Phys. Lett. **66**, 1391 (1995).
[6]N. Theis, P. M. Mooney, and S. L. Wright, Phys. Rev. Lett. **60**, 361 (1988).
[7]T. Baba, M. Mizuta, T. Fujisawa, J. Yoshino, and H. Kukimoto, Jpn. J. Appl. Phys., Part 2 **28**, L891 (1989).
[8]E. Calleja, F. Garcia, A. Gomez, E. Munoz, P. M. Mooney, T. N. Morgan, and S. L. Wright, Appl. Phys. Lett. **56**, 934 (1990).
[9]P. M. Mooney, N. S. Caswell, and S. L. Wright, J. Appl. Phys. **62**, 4786 (1987).
[10]E. Buks, M. Heiblum, Y. Levinsohn, and H. Shtrikman, Semicond. Sci. Technol. **9**, 2031 (1994).
[11]A. Leuther, A. Forster, H. Luth, H. Holzbrecher, and U. Breuer, Semicond. Sci. Technol. **11**, 766 (1996).
[12]A. P. Young and H. H. Wieder, J. Vac. Sci. Technol. B **14**, 2944 (1996).
[13]H. Sari and H. H. Wieder, J. Appl. Phys. **85**, 3380 (1999).
[14]I-H. Tan, G. L. Snider, L. D. Chang, and E. L. Hu, J. Appl. Phys. **68**, 4071 (1990).
[15]P. M. Mooney, Semicond. Sci. Technol. **6**, B1 (1991).
[16]M. E. Lazzouni and L. J. Sham, Phys. Rev. B **48**, 8948 (1993).
[17]B. Monemar, Phys. Rev. B **8**, 5711 (1973).
[18]Landolt/Bornstein Series, *Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung (Springer, New York, 1987), Vol III/22a.
[19]S. Adachi, J. Appl. Phys. **58**, R1 (1985).
[20]M. L. Cohen and J. R. Chelikowski, *Structure and Optical Properties of Semiconductors*, 2nd ed. (Springer, New York, 1989).
[21]M. R. Lorenz *et al.*, Proc. 10th Internat. Conf. Phys. Semicond., 1970 (unpublished), p. 444.
[22]F. E. Fruh, J. M. Salese, M. Beck, D. K. Maude, U. Wilke, M. Rabary, J. C. Portal, and M. Ilegems, Solid State Commun. **89**, 323 (1994).

Behavior of a new ordered structural dopant source in InAs/(001) GaP heterostructures

V. Gopal,[a] E.-H. Chen, E. P. Kvam, and J. M. Woodall
Purdue University, W. Lafayette, Indiana 47907

(Received 19 January 1999; accepted 17 May 1999)

We report the characteristics of molecular-beam epitaxy grown InAs on highly lattice mismatched (001) GaP substrates. Strain relaxation in this system occurs at low thickness by the generation of a periodic two-dimensional square grid network of 90° misfit dislocations at the heterointerface. The very high interface dislocation density ($\sim 10^{13}$ intersections/cm^2) exerts a unique influence on the electronic properties of the system. An extended defect structure at the intersection of 90° misfit dislocations is proposed to act as an ordered structural donor source. Hall effect measurements indicate that this source is fully ionized with a constant sheet carrier concentration of 10^{13} cm^{-2}, irrespective of the InAs layer thickness, and exhibits no freeze out at low temperatures. We have also demonstrated that the electron mobility increases significantly with InAs layer thickness, reaching values in excess of 10 000 cm^2/V s in nominally undoped layers. The high threading dislocation density ($\sim 10^{10}$ cm^{-2}) in the InAs epilayers does not appear to have a deleterious effect on the transport properties of majority carrier electrons. © *1999 American Vacuum Society.* [S0734-211X(99)06904-8]

I. INTRODUCTION

Advances in the molecular-beam epitaxial (MBE) growth of compound semiconductors have resulted in the growth of a wide variety of III–V compounds and their alloys on the commercially available (001) oriented substrates such as GaAs, InP, and GaP. The growth only of lattice-matched epilayers, such as $Al_xGa_{1-x}As$ on GaAs, places a limit on the band gap range available for device design. While lattice-mismatched epitaxy removes this constraint, the consequent mismatch strain is accommodated by the generation of misfit dislocations (MD) if a critical layer thickness is exceeded. Much effort has been devoted to understanding the mechanisms of MD generation and minimizing their propagation into the epilayer as threading dislocations (TD).[1,2] However, when the lattice mismatch is high (>7%), MDs and TDs can no longer be avoided. In order to design devices using highly lattice-mismatched systems the electronic behavior of the dislocations must be understood.

Highly lattice-mismatched systems are well suited for a study of MDs for two reasons. First, the interface dislocation population is very high, so that the interface electronic properties tend to be dominated by those of the dislocations. Second, unlike in the case of lower mismatch where an irregularly spaced network of interfacial dislocations results, high mismatches favor the generation of a regular periodic network of 90° misfit dislocations oriented in the [110] and the [$\bar{1}$10] crystallographic directions. The periodicity imposed by such a "superlattice" dislocation microstructure results in interesting electronic behavior.

In the present study, we have investigated the direct growth of InAs on (001) GaP substrates. The lattice mismatch is 11%, the largest among all the III-arsenides and the III-phosphides. The difference in band gap is also large: InAs–0.36 eV, GaP–2.25 eV. $Al_xIn_{1-x}As$ layers with $x = 0.1$ and 0.2 grown on GaP, which have a higher band gap than InAs but not very significantly different lattice parameters, have also been studied.

II. EXPERIMENT

InAs/GaP heterostructures were grown by solid source MBE using a Varian Gen-II system. Commercially obtained GaP substrates were heated to 710 °C for oxide desorption under a P_2 overpressure. A 20 period superlattice consisting of 5 nm alternating layers of GaP/AlGaP was grown to prevent the outdiffusion of impurities from the substrate. This was followed by the growth of a 200 nm buffer layer of undoped GaP at 660 °C. The growth of phosphides was achieved by using a valved phosphorous cracker.[3,4] Reflection high-energy electron diffraction was used to monitor surface morphology, growth rate, and surface V–III incorporation ratio.

The substrate temperature was lowered to 350 °C for the growth of an undoped InAs epilayer. At higher temperatures (450–500 °C), an interfacial reaction between the InAs and GaP was reported to occur. The first few monolayers were grown under low V/III ratios to promote a smoother interface.[3] Finally, a 5 nm capping layer of undoped $Al_{0.2}In_{0.8}As$ was grown above the InAs layer to suppress carrier generation due to surface Fermi level pinning. Samples with varying thickness of InAs ranging from 3 nm to 2 μm were grown in this manner. Figure 1 is a schematic of the multilayer structure. The growth of $Al_xIn_{1-x}As$ on GaP was carried out similarly and these layers were capped by growing the final 5 monolayers with a higher Al fraction.

Transmission electron microscopy (TEM) was employed for defect microstructure investigations. Cross-sectional samples were prepared and examined in a JEOL 2000 FX microscope with an accelerating voltage of 200 kV to obtain

[a]Electronic mail: gopal@ecn.purdue.edu

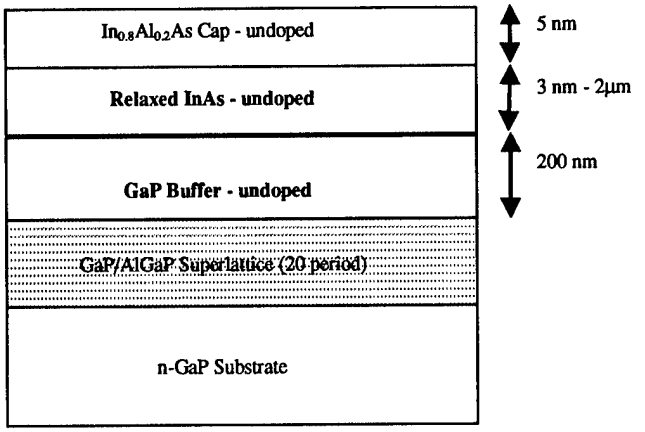


FIG. 1. Schematic of the multilayer structure grown by MBE.

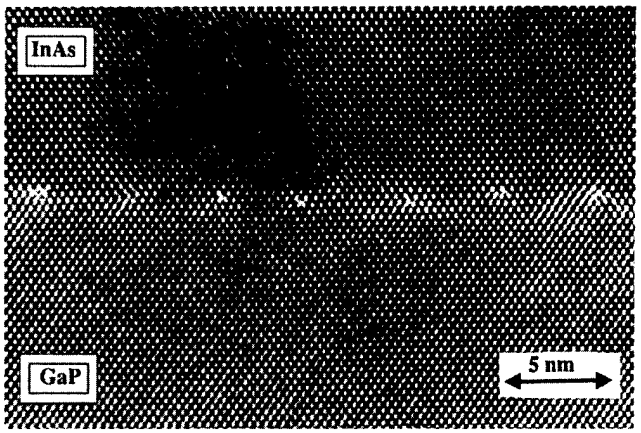


FIG. 2. HREM micrograph of InAs/GaP heterostructure. MDs appear as bright spots.

high resolution TEM (HREM) images of the heterointerface. Plan view samples were examined in a JEOL 200 CX microscope. 200 and 220 weak beam dark field images were obtained to study the threading dislocation microstructure in InAs epilayers.

Electrical properties were characterized using the Hall effect technique and electrochemical capacitance–voltage (ECV) depth carrier concentration profiling. Systematic Hall effect measurements were performed at both room temperature (300 K) and liquid N_2 temperature using the van der Pauw technique on InAs/GaP and $Al_xIn_{1-x}As$/GaP heterostructure samples. Variable temperature Hall effect data between 5 and 310 K were obtained for select samples. The majority carrier type, density, and mobility were determined from the data in the standard manner.

III. RESULTS AND DISCUSSION

A. Microstructure evolution

Strain relaxation during epitaxial growth occurs by the generation of a network of MDs in the interfacial (001) plane. Low mismatch strains (<2%) are accommodated predominantly by the glissile 60° interfacial dislocation segments formed during the glide of epithreading dislocations. Dislocation nucleation sources have been proposed to be active at the free surface of the growing film.[5] The resulting network consists of irregularly spaced long MD segments. At higher mismatches (2%–4%), the formation of short sessile 90° MD segments were observed by Kvam, Maher, and Humphreys,[6] who proposed a dislocation reaction between two complementary 60° dislocations to yield a 90° MD, e.g.,

$$\tfrac{1}{2}[01\bar{1}] \text{ on } (\bar{1}11)+\tfrac{1}{2}[\bar{1}01] \text{ on } (1\bar{1}1)=\tfrac{1}{2}[\bar{1}10] \text{ on } (001). \quad (1)$$

For very high mismatch (>7%), the MD microstructure consists of a regular, periodic network of 90° misfit dislocations. Direct introduction of sessile 90° dislocations at the edge of Ge islands growing on (001) Si (4% mismatch) has been reported by LeGoues *et al.*[7] Strain relaxation in the InAs/GaP system is expected to yield such a regular periodic MD network.

However, our HREM images provide evidence that strain relaxation in this system occurs through both direct introduction and by the glide of 60° dislocations, with the former being the dominant mechanism. Chang, Chin, and Woodall[3] reported that 2.5 nm InAs films grown on GaP were 85% strain relaxed with a dislocation spacing of 4.1±0.1 nm. They did not observe any 60° MDs in their HREM images and proposed that the 90° dislocations were directly incorporated into the growing InAs layer. As the layer thickness is increased the energy required for direct introduction of MDs becomes very high, and further strain relaxation must occur by the glide and reaction of 60° MDs to form 90° dislocation segments. Evidence for this mechanism is seen in Fig. 2, which is a HREM micrograph from a 15 nm InAs sample. A closely spaced pair of 60° dislocations is observed and a Burgers circuit around this pair yields a closure gap corresponding to that of a 90° edge dislocation. Table I shows the degree of strain relaxation for InAs layers of varying thickness grown on GaP. At about 30 nm strain relaxation is complete, with a mean dislocation spacing of ~3.85 nm.

Figure 3 is a 220 dark field micrograph of a 1 μm InAs epilayer. Threading dislocations are seen at a density of $\sim 10^{10}$ cm^{-2}. While this is a high density, it is comparable to those in ~2% mismatched SiGe layers grown on Si, where the MD population is significantly less dense. A simple calculation, assuming that each MD segment is associated with two TDs at its ends, yields an average MD segment length of 10 μm. This is far in excess of the size of InAs islands (~10–100 nm) during the initial stages of growth, and implies that a single MD propagates through several islands as

TABLE I. Degree of strain relaxation of InAs epilayers of differing thickness grown on GaP.

InAs layer thickness (nm)	Mean dislocation spacing (nm)	% strain relaxed
3.0	4.19	91
15	3.95	97.5
30	3.85	100

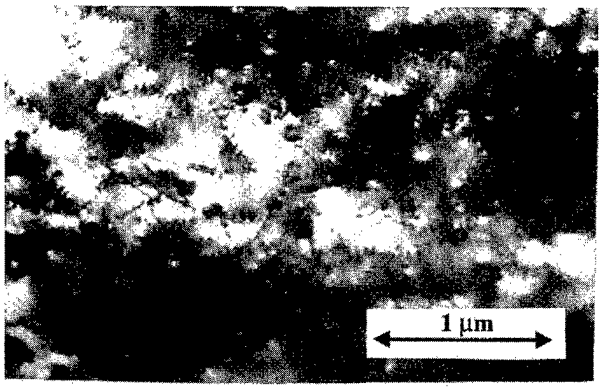


FIG. 3. Plan view TEM micrograph showing dense TD network in the InAs epilayer.

the islands coalesce without the dislocation threading into the epilayer. In summary, the dislocation microstructure consists of a regularly spaced network of (predominantly) 90° MDs with a few (<5%) closely spaced pairs of 60° dislocations. Isolated 60° MD segments have not been identified in HREM micrographs, but their occurrence cannot be ruled out.

The atomic level core structures of 90° MDs in zinc blende crystal structures have a sevenfold-fivefold ring structure, as shown in Fig. 4. Such dislocations do not have dangling bonds associated with them. We believe this to be the dominant MD core structure in the InAs/GaP system. On the other hand, 60° MDs can have dangling bonds at their cores. 90° MDs can also take on an eightfold-sixfold ring configuration with two dangling bonds per Burgers vector length at their core. The fraction of 90° MDs with this core structure is not known in InAs/GaP, but we believe it to be small.

Mostoller, Chisholm, and Kaplan proposed the occurrence of an 18 atom point defect structure—the "dreidl"—at the

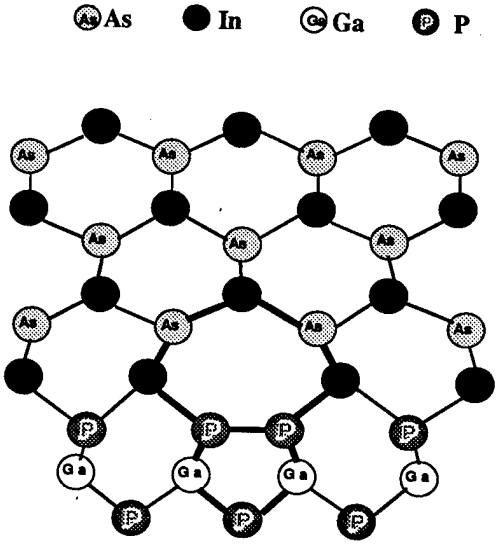


FIG. 4. Sevenfold-fivefold ring structure of a 90° MD.

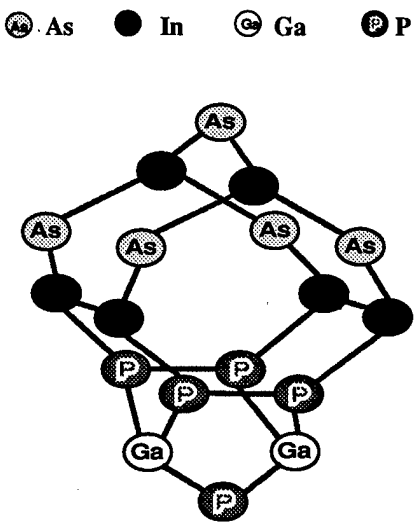


FIG. 5. 18 atom dreidl. Note the occurrence of Group III and Group V dipoles.

intersection of 90° MDs in Ge/(001) Si.[8] A similar defect structure would exist in the InAs/GaP system and is shown schematically in Fig. 5. Of particular interest is the occurrence of both III–III and V–V dipoles in each dreidl, which could contribute to the electrical activity of these defects.

B. Electronic behavior

Even though all the samples were nominally undoped, the sign of the Hall coefficient indicated *n*-type conductivity. This may be due to the large hole to electron effective mass ratio for InAs. Hence, donor-like defect levels will be shallower than acceptor levels.[4] Table II displays the sheet carrier concentration (N_s) and Hall mobility (μ) for the InAs/GaP samples tested at 300 and 77 K.[4,9] The sheet carrier concentration is very high— 10^{13} cm^{-2}—for the thin layer samples (Sample Nos. 1 through 5), and does not vary with epilayer thickness. This implies a high density of carriers localized at either the heterointerface or at the free surface. Since the InAs is capped off by an $Al_{0.2}In_{0.8}As$ layer, surface generation and accumulation was prevented. ECV profiling confirms that the carriers are localized at the InAs/GaP interface. Figure 6 is a semilog plot of carrier concentration versus depth for an ~1 μm InAs layer grown on GaP. Carrier accumulation is clearly seen on the low bandgap (InAs) side of the heterojunction.

There is also no freeze out, even at 5 K. Variable temperature Hall measurements were performed on Sample No. 4 as a representative of the thin layer samples, and the plot of N_s versus temperature is shown in Fig. 7. The sheet carrier concentration remains constant with temperature down to the lowest temperature, 5 K. This strongly indicates that the states contributing to the interfacial carriers lie above the conduction band edge of InAs and are degenerate. There are several possible sources for these carriers:

TABLE II. Sheet carrier density and electron mobility data for InAs/GaP heterostructures at 300 and 77 K.

Sample No.	InAs layer thickness (nm)	300 K Sheet density ($\times 10^{13}$ cm^{-2})	300 K Hall mobility (cm^2/V s)	77 K Sheet density ($\times 10^{13}$ cm^{-2})	77 K Hall mobility (cm^2/V s)
1	5	1.2	40	1.2	25
2	10	1.1	500	1.0	360
3	15	0.75	340	0.75	310
4[a]	20	1.0	460	0.8	430
5	30	1.0	800	0.9	730
6	250	1.0	4000	1.0	3 780
7[a]	500	1.2	5310	1.1	5 390
8	1000	1.5	7620	1.3	8 120
9	2000	2.0	9920	1.6	10 900

[a]5 K data.

(1) Atomic sites along 90° MDs: There are a total of 10^{14} sites along 90° MD lines per cm^2, but the measured sheet carrier concentration is an order of magnitude lower. Since the energy states responsible for these carriers do not display freeze out, it is unlikely that a fractional ionization event occurs.

(2) Atomic sites along 60° MDs: A small fraction (<5%), of closely spaced pairs of 60° MDs have been observed. Such dislocations can have dangling bonds at their cores that would be expected to display electrical activity. However, the sites along these dislocations would not be sufficient in number to account for 10^{13} carriers per cm^2.

(3) TDs: While TDs may be electrically active, they are not sufficiently dense ($\sim 10^{10}$ cm^{-2}) to account for the carriers in the thin layer samples. More importantly, TD contribution would result in a thickness dependence of N_s, which is not observed.

(4) Intrinsic band bending: InAs and GaP have large conduction and valence band offsets. However, calculations carried out using the band structure simulator ADEPT show that the intrinsic band bending would result in the Fermi level being pinned midgap in the InAs at the interface and, consequently, low carrier accumulation. A high density (10^{13} cm^{-2}) sheet of electronic charge at the interface requires the Fermi level to be pinned 0.2 eV above the conduction band edge of InAs at the interface.[9] This could be accomplished by a defect related donor state.

(5) 90° MD intersection sites: The dreidl structure (Fig. 5) might occur at the intersection of 1/2⟨110⟩ 90° MDs at the InAs/GaP interface, forming a square lattice with a periodicity of ~4 nm. These would appear with a density of $\sim 10^{13}$ cm^{-2}, a numeric match with the N_s values of Table II. Thus, the MD intersection sites may act as an array of ordered structural dopants. As mentioned previously, this structure in III–V compounds would contain strong internal dipoles, which may be the carrier generation source.

The carriers thus generated are confined to a two-dimensional electron gas near the interface and are scattered strongly by the dense MD network, as evidenced by the low mobility values in Table II. As the temperature is decreased,

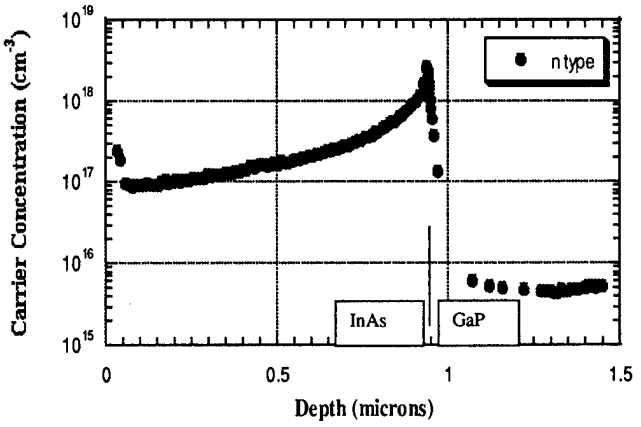


FIG. 6. ECV profile of InAs/GaP heterostructure showing carrier accumulation at the interface.

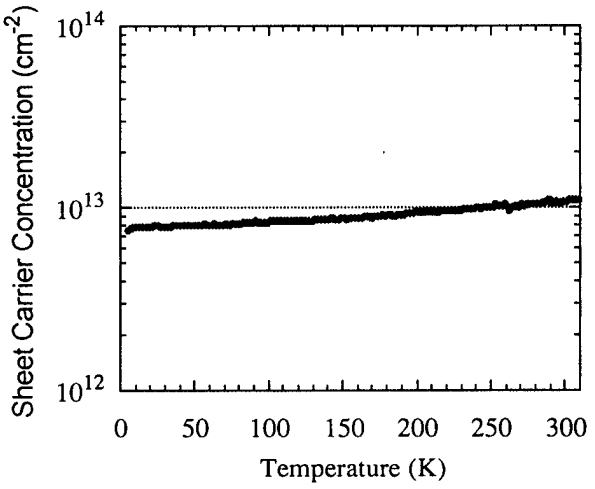


FIG. 7. Variation of sheet carrier concentration, N_s, with temperature. Note the absence of freeze out down to 5 K.

TABLE III. Comparative values of sheet carrier concentration and mobility for InAs and $Al_xIn_{1-x}As$ samples grown on GaP.

Epilayer thickness (nm)	Sample	300 K		77 K		5 K	
		N_s (cm^{-2})	μ (cm^2/V s)	N_s (cm^{-2})	μ (cm^2/V s)	N_s (cm^{-2})	μ (cm^2/V s)
20	InAs	1.0×10^{13}	460	0.8×10^{13}	430	0.75×10^{13}	470
20	$Al_{0.1}In_{0.9}As$	0.47×10^{13}	190	0.55×10^{13}	123	0.55×10^{13}	85
20	$Al_{0.2}In_{0.8}As$	5.0×10^{11}	380	7.4×10^{10}	290	...	...

the confinement would be stronger, i.e., the mobility would decrease with decreasing temperature which is consistent with our observations.[9]

Table III is a comparison of the sheet carrier concentration and mobility data for 20-nm-thick samples of InAs and $Al_xIn_{1-x}As$ ($x=0.1,0.2$). The addition of Al increases the band gap very significantly while retaining a large lattice mismatch with GaP, so the MD microstructure is expected to remain smaller. It is clear that the electronic properties differ considerably. N_s decreases with increasing Al content and carriers freeze out even at 77 K for the $Al_{0.2}In_{0.8}As$ sample. The energy states contributing to interfacial carriers are nondegenerate in this case and must lie within the band gap. Both the conduction band and the valence band offsets decrease upon addition of Al, although the relative change in each is unknown. A more quantitative analysis has not yet been performed. It can be speculated that the dislocation-related energy level remains constant with respect to the vacuum level, while the conduction band edge moves upward.

The thick layer InAs samples (Nos. 6–9 in Table II) display a thickness dependence of the sheet carrier concentration. The N_s versus t plot is linear with a zero thickness intercept of 9.0×10^{12} cm^{-2}, as seen in Fig. 8. The linearity is indicative of a constant bulk carrier concentration ($\sim5\times10^{16}$ cm^{-3}) from the InAs epilayer, in addition to the constant N_s contribution from the interface. These bulk carriers could be generated by states associated with threading dislocations. The comparative values at 77 K indicates that the bulk carriers in InAs are subject to a partial freeze out. There

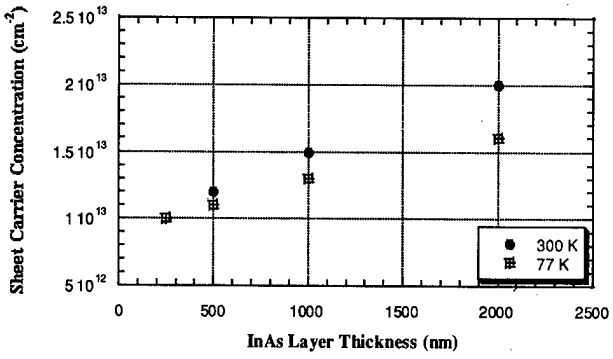


FIG. 8. Dependence of sheet carrier concentration on epilayer thickness for the thick InAs layer samples.

are thus two parallel channels or pathways for conduction:[4] (i) a high carrier concentration, low mobility interface channel, and (ii) a lower carrier density, high mobility bulk channel. As thickness is increased, the contribution of the bulk channel increases, as evidenced by the sharp rise in mobility. Using this two channel model Chen[4] has calculated that the electron mobility in the InAs epilayer (bulk channel) exceeds 20 000 cm^2/V s at 77 K—nearly 2/3 the theoretical electron mobility of InAs (33 000 cm^2/V s).

The dense TD network ($\sim10^{10}$ cm^{-2}) in the epilayer thus does not appear to scatter carriers strongly. This is in contrast to the strong scattering exhibited by the MD network. The InAs/GaP system is promising for high speed device applications despite having a high defect density.

IV. CONCLUSIONS

We have investigated the evolution of the dislocation microstructure during the MBE growth of InAs on GaP and the electronic behavior of the defects. MD introduction occurs by two distinct mechanisms and the resulting microstructure consists mainly of 90° edge MDs, but also a few closely spaced pairs of 60° dislocations. We have demonstrated the existence of a dense sheet of electron charge at the heterointerface, which is linked to the MD network. A defect structure proposed to occur at the intersection of 90° MDs is a likely source of the carriers and may act as an ordered structural dopant. Two parallel conducting pathways for lateral carrier transport in InAs epilayers were proposed—a low mobility interface pathway and a high mobility bulk pathway. High mobility in the bulk channel is achieved despite a high density of TDs.

ACKNOWLEDGMENTS

The authors thank Dr. A. L. Vasiliev for assistance in TEM. This research was funded by the National Science Foundation through Grant No. 9400415-DMR.

[1]E. P. Kvam and S. B. Samavedam, J. Minerals, Metals, Mater. Soc. **46**, 47 (1994).
[2]A. E. Romanov, W. Pompe, S. Mathis, G. E. Beltz, and J. S. Speck, J. Appl. Phys. **85**, 182 (1999).
[3]J. C. P. Chang, T. P. Chin, and J. M. Woodall, Appl. Phys. Lett. **69**, 981 (1996).
[4]E.-H. Chen, Ph.D. thesis, Purdue University, 1998.
[5]J. W. Matthews, S. Mader, and T. B. Light, J. Appl. Phys. **41**, 3800 (1970).

[6]E. P. Kvam, D. M. Maher, and C. J. Humphreys, J. Mater. Res. **5**, 1900 (1990).
[7]F. K. LeGoues, J. Tersoff, M. C. Reuter, M. Hammar, and R. Tromp, Appl. Phys. Lett. **67**, 2317 (1995).
[8]M. Mostoller, M. F. Chisholm, and T. Kaplan, Phys. Rev. Lett. **72**, 1494 (1994).
[9]V. Gopal, E. P. Kvam, T. P. Chin, and J. M. Woodall, Appl. Phys. Lett. **72**, 2319 (1998).

Nanoelectronic device applications of a chemically stable GaAs structure

D. B. Janes,[a] V. R. Kolagunta,[b] M. Batistuta, and B. L. Walsh
School of Electrical and Computer Engineering, Purdue University, W. Lafayette, Indiana 47907

R. P. Andres, Jia Liu, J. Dicke, J. Lauterbach, and T. Pletcher
School of Chemical Engineering, Purdue University, W. Lafayette, Indiana 47907

E. H. Chen, M. R. Melloch, E. L. Peckham, H. J. Ueng, and J. M. Woodall[c]
School of Electrical and Computer Engineering, and NSF MRSEC for Technology Enabling Heterostructure Materials, Purdue University, W. Lafayette, Indiana 47907

Takhee Lee and R. Reifenberger
Department of Physics, Purdue University, W. Lafayette, Indiana 47907

C. P. Kubiak and B. Kasibhatla
Department of Chemistry, University of California at San Diego, La Jolla, California

(Received 10 February 1999; accepted 3 May 1999)

We report on nanoelectronic device applications of a nonalloyed contact structure which utilizes a surface layer of low-temperature grown GaAs as a chemically stable surface. In contrast to typical *ex situ* ohmic contacts formed on *n*-type semiconductors such as GaAs, this approach can provide uniform contact interfaces which are essentially planar injectors, making them suitable as contacts to shallow devices with overall dimensions below 50 nm. Characterization of the native layers and surfaces coated with self-assembled monolayers of organic molecules provides a picture of the chemical and electronic stability of the layer structures. We have recently developed controlled nanostructures which incorporate metallic nanoclusters, a conjugated organic interface layer, and the chemically stable semiconductor surface layers. These studies indicate that stable nanocontacts (4 nm×4 nm) can be realized with specific contact resistances less than $1\times10^{-6}\,\Omega\,\mathrm{cm}^2$ and maximum current densities (1×10^6 A/cm^2) comparable to those observed in high quality large area contacts. The ability to form stable, low resistance interfaces between metallic nanoclusters and semiconductor device layers using *ex situ* processing allows chemical self-assembly techniques to be utilized to form interesting nanoscale semiconductor devices. This article will describe the surface and nanocontact characterization results, and will discuss device applications and novel techniques for patterning close-packed arrays of nanocontacts and for imaging the resulting structures. © *1999 American Vacuum Society.* [S0734-211X(99)05504-3]

I. INTRODUCTION

There have recently been numerous examples of prototype electronic devices having nanometer scale dimensions.[1–4] While the dimensions of some features within these devices are in the nanometer range, frequently the overall device dimension is orders of magnitude larger, particularly in structures employing semiconductor channels and source/drain regions. This size discrepancy is due to the need for ohmic contact structures which are 1 μm or greater in lateral extent and typically 100 nm or greater in depth. For example, in compound semiconductor devices based on GaAs, contacts such as alloyed Au/Ge/Ni on *n*-type layers are spatially nonuniform and also consume a significant surface layer in order to provide suitably low specific contact resistivity.[5] Thus, even though the active area of the device is in the nanometer scale, the contacts to the device still require areas $\sim10^2$–10^4 greater than the active device. For devices with vertical current flow, such as resonant tunneling diodes (RTDs), the use of alloyed contacts often dictates that the active area of the device is buried deep beneath the semiconductor surface. In contrast, an ohmic contact technology which can provide nanometer contact dimensions, both laterally and vertically, would allow the demonstration of the high circuit densities promised by nanometer scale device concepts. Contacts suitable for high density nanoelectronic devices must provide low contact resistance and must be spatially uniform at the nanometer length scale. This article describes a nonalloyed ohmic contact structure suitable for *n*-type GaAs and discusses developments aimed at nanoelectronic device and circuit applications.

II. OHMIC CONTACT STRUCTURE

The ohmic contact structure[6] employs a surface layer of low-temperature grown GaAs (LTG:GaAs), i.e., GaAs grown at a temperature of 250–300 °C by molecular beam epitaxy. Because of the low growth temperature, ~1–2% *excess* arsenic is incorporated during growth. This excess arsenic is responsible for several interesting electronic properties, including a short minority carrier lifetime and bulk Fermi level pinning (in as-grown material).[7] For as-grown material, which is used exclusively in the work reported in

[a]Electronic mail: janes@ecn.purdue.edu
[b]Permanent address: Motorola Inc., Austin, TX 78721.
[c]Permanent address: Yale University, New Haven, CT 06520.

this article, the excess arsenic is distributed in high concentration ($\sim 1.0\times 10^{20}\,\mathrm{cm}^{-3}$) of point defects, primarily as arsenic antisite defects.[7] These defects are observed as a band of states located approximately midgap in the GaAs.[8] Previous studies indicate that these states prevent the GaAs surface from rapidly oxidizing due to the relatively low concentration of minority carrier holes in the surface layer.[9,10] As a result, the presence of the gap states can be observed using scanning tunneling microscopy (STM) even following brief air exposure of the samples.[9]

The *ex situ*, nonalloyed ohmic contacts employing a LTG:GaAs surface layer can provide specific contact resistivity (ρ_c) below $1\times 10^{-6}\,\Omega\,\mathrm{cm}^2$.[6] Applications of this contact to shallow device layers and studies of the temperature stability have been reported.[11] These contacts may be appropriate for nanometer scale device applications since they would not suffer from the deep interface and spatial nonuniformity of Au/Ge/Ni contacts. In addition, this type of contact structure and the chemically stable LTG:GaAs surface layer are compatible with chemical self-assembly techniques, which can provide ordered organic monolayer films, one and two-dimensional arrays of nanoscale features, and other interesting structures.[12]

We have previously shown that high performance nano-contacts can be formed to GaAs device layers using this nonalloyed ohmic contact approach.[13] In that study, a controlled area nano-contact was formed by a single-crystal, 4-nm-diam Au nanocluster deposited on the surface. The semiconductor layer structure employs a thin (10 nm) layer of LTG:GaAs to facilitate a high quality, nano-contact to $n+$GaAs(100) layers grown at standard temperatures. Controlled mechanical tethering and electronic coupling between the Au cluster and the LTG:GaAs surface was achieved via a self-assembled monolayer (SAM) of p-xylene-α, α'-dithiol ($C_8H_{10}S_2$), also denoted as xylyl dithiol (XYL), which was formed on the (LTG:GaAs) surface before Au cluster deposition. Probing of the resulting nano-contact structure using STM current–voltage spectroscopy indicates that the specific contact resistance, $1\times 10^{-6}\,\Omega\,\mathrm{cm}^2$, and maximum current density, $1\times 10^{6}\,\mathrm{A/cm}^2$, of the nano-contact are both comparable to that achieved in high quality large-area ohmic contacts to n-type GaAs.[13]

III. PATTERNING TECHNIQUES

The current work involves definition of device structures in the GaAs material using XYL based etch masks and the development of shallow device technologies which can be compatible with such patterning techniques. For the monolayer based etch masks, a SAM of XYL is deposited in selected regions on the GaAs surface using an elastomeric stamp pad technique, as illustrated in Fig. 1. A stamp pad of poly(dimethylsiloxane) (PDMS) elastomer was made using a mold consisting of an oxidized silicon wafer in which the desired pattern had been defined by photolithography and transferred into the silicon dioxide using wet chemical etching. The GaAs samples were exposed to air for several days before preparation for stamped deposition. Following a HCl based oxide strip, the samples were transferred into a dry nitrogen atmosphere for the stamping. The elastomer stamp pad was "inked" by soaking it in a 1 mM solution of XYL in ethanol for 5 min. After the excess solvent had evaporated, the stamp pad was placed on the GaAs surface for a period of 18 h. Afterwards, the GaAs sample was thoroughly rinsed in ethanol to remove excess XYL.

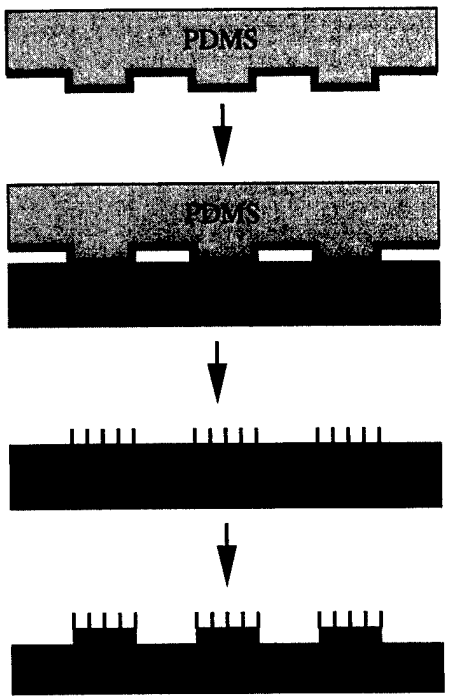


FIG. 1. Schematic diagram of the procedure used to deposit a monolayer of xylyl dithiol in specified regions on the LTG:GaAs surface. The procedure uses an elastomer stamp pad which has been lithographically patterned to provide the desired xylyl dithiol pattern. Illustrated steps are: (i) stamp pad is inked with xylyl dithiol, (ii) inked stamp pad placed on GaAs surface for 18 h, (iii) stamp pad is removed and sample is rinsed in ethanol to leave a patterned monolayer, and (iv) GaAs is etched using wet chemical etching.

Optical methods with a high surface sensitivity, such as ellipsometry, have been widely applied in surface science studies. However, it has not been until recently that ellipsomicroscopy for surface imaging (EMSI) has been developed and its submonolayer sensitivity has been exploited to obtain information about spatio-temporal pattern formation in heterogeneously catalyzed reactions.[14] EMSI is also well suited to study other ultrathin adsorbate layers and has the advantage that it can characterize the uniformity and stability of a thin (monolayer) coating as a function of position and time. With suitable optics, the EMSI technique has a spatial resolution of ~ 3 μm. For these reasons, EMSI was used to investigate the surface of representative samples of both the LTG:GaAs-capped and the xylyl dithiol coated LTG:GaAs wafers. Figure 2 shows an EMSI image of the patterned regions of XYL on LTG:GaAs. This image was obtained after the third step illustrated in Fig. 1, i.e., before etching of the GaAs layers. The contrast is therefore due to the presence of a SAM of XYL on the surface, with a thickness of approximately 1 nm. The stable optical properties observed in this study and in experiments on unpatterned layers of XYL on

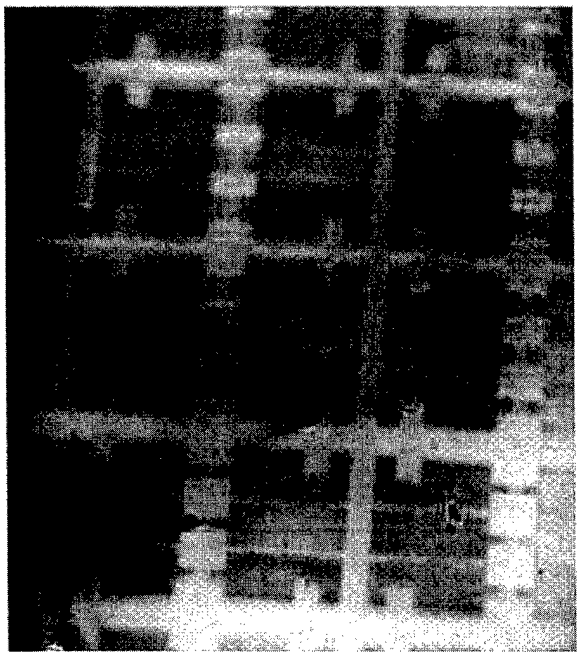


FIG. 2. Ellipsometric image of patterned regions of xylyl dithiol on LTG: GaAs, before GaAs etching. Light regions are coated with xylyl dithiol.

LTG:GaAs indicate that the XYL coated LTG:GaAs is a stable surface even under air exposure. It is believed that the sulphur to GaAs bond provides passivation comparable to that observed in studies involving elemental sulphur, with additional stability provided by both the characteristics of the LTG:GaAs and the organic tail of the XYL molecule.

The patterned monolayer of XYL was used as an etch mask for wet chemical etching of the GaAs layers. Figure 3 shows a scanning electron micrograph of a pattern etched into the LTG:GaAs capped device structure using a HCl: H_2O_2:H_2O based wet etch. In the semiconductor layer structure, a thin (5–10 nm) LTG:GaAs layer provides effective surface passivation. This layer prevents oxidation of underlying doped layers, as well as providing bulk Fermi level pinning. Because of the pinning effect, it is possible to

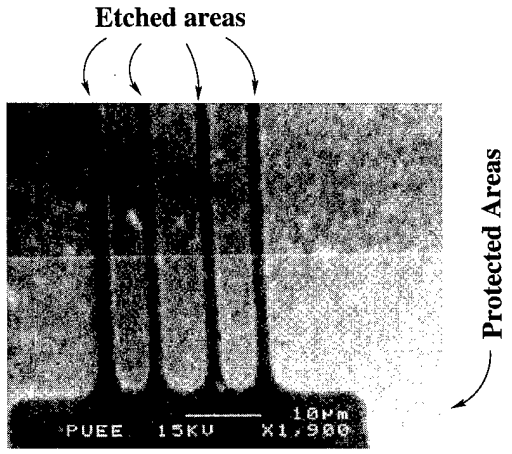


FIG. 3. Scanning electron micrograph of regions of GaAs patterned using wet chemical etching with stamped xylyl dithiol as etch mask.

achieve very high activated donor density ($\sim 1\times10^{20}/\text{cm}^3$) in the space charge regions immediately beneath the LTG: GaAs cap.[6] These characteristics enable the low-resistance nonalloyed ohmic contact, even for LTG:GaAs cap layers as thin as 2 nm. The removal of the LTG:GaAs cap is sufficient to destroy this passivation behavior and is therefore sufficient to pattern devices which use relatively thin doped (10–20 nm) layers beneath the cap. We have shown that GaAs based field-effect transistor layers can be effectively patterned by removal of approximately 3 nm of LTG:GaAs.

IV. SHALLOW DEVICES/ETCHING

To explore the possibility of making controlled heterostructures extremely close to the surface we have applied thin LTG:GaAs cap layers on thin $n++$ layers. For example, it has been possible to grow a 100-Å-thick Si-doped GaAs layer doped at $\sim10^{20}/\text{cm}^{-3}$ on an *i*-GaAs (unintentionally doped) buffer layer and capped by a 35 Å LTG: GaAs layer. To illustrate the effective surface passivation and the ability to isolate regions using a shallow etch, we have built transmission line method (TLM) structures[15] on this layer structure. The sheet resistance of this 135 Å contact structure measured from TLM patterns with *ex situ* Ti/Au contacts fabricated using conventional UV lithography was $R_{sh}\approx633\,\Omega/\text{square}$ with a contact resistance $\rho_c\approx7\times10^{-7}\,\Omega\,\text{cm}^2$. The variation across the sample was less than 15%. The measured sheet resistance and contact resistance indicate that the ohmic contacts to the layer are low resistance and that the thin LTG:GaAs cap layer had provided effective passivation of the heavily doped GaAs layer even in those regions not covered by metal. The measured sheet resistance is consistent with that predicted for a 10-nm-thick layer with activated donor density well above the bulk limit of $\approx5\times10^{18}\,\text{cm}^{-3}$ Given that the $n++$ layer is only 3.5 nm from the semiconductor surface, an unpassivated layer of this thickness would be expected to be largely (or completely) depleted by the surface Fermi level pinning, coupled with effects of surface oxidation. In that case, the low donor activation found in bulk GaAs doped with Si (as described earlier) would be expected in the remainder (un-oxidized portion) of the layer.

The next step is to demonstrate that controlled stripping of this LTG:GaAs cap layer causes the oxidation and/or repinning of the Fermi level of the underlying $n++$ layer wherever the LTG:GaAs cap has been removed, thus causing the electrical isolation of the various $n++$ islands whose surface is protected by the contact metal. It has to be noted that while the oxidation rate of the LTG:GaAs layer is very small compared to that for conventional *n*-GaAs, a thin protective oxide layer ($t_{ox}\approx30\,\text{Å}$) will form on the surface on exposure to the ambient. This oxidation process can be accelerated either by increased minority carrier generation in the LTG:GaAs induced by incident photons or by the use of a strong oxidizing ambient. Upon stripping of this thin oxide layer, the fresh surface begins oxidizing. Repeated formation followed by stripping of the thin oxide layer allows one to etch controllably thin layers of LTG:GaAs. Once the LTG:

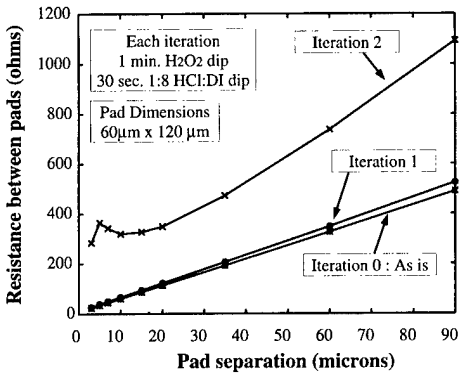


FIG. 4. Measured resistance vs pad separation for TLM structure with 60 μm×120 μm pads on device layer. Curves are shown for (i) unetched sample (as deposited), after one etch iteration of approximately 2.5 nm, and after two etch iterations, with a total etch depth of approximately 5 nm. The change in resistance indicates a dramatic change in the sheet resistance between the contacts.

GaAs cap has been etched, the dynamics of the surface Fermi level pinning causes the formation of a relatively thick oxide and a large depletion region at the surface. Due to this behavior, controlled chemical etching using oxidation followed by stripping of the oxide thus formed leads to etching of only the LTG:GaAs cap and can be used to isolate $n++$ islands as shown in Fig. 4. In this case, the sample surface was oxidized by exposing it to a hydrogen peroxide solution in steps of 60 s. Each oxidation iteration was followed by a stripping of the oxide using a 30 s dip in a dilute HCL:DI solution (1:8). The resistance between the TLM contact pads was measured after each such oxidation and strip "iteration." This oxidation and stripping iteration etches about 30 Å material at each instance. After three such iterations the contacts were completely isolated. The data plotted in Fig. 4 show the measured resistance between adjacent pads versus pad separation for both the "as-deposited" case and following one and two iterations of etching. The increase in resistance following controlled etching of the LTG:GaAs cap layer is associated with a decrease in the conductance of the $n++$ layer between two contacts. The measured resistance versus pad spacing relationship changes little after the first iteration, indicating that the passivation effect of the LTG:GaAs layer is still active. After the second iteration, the measured resistance increases dramatically, indicating that the contacts have been effectively isolated. This is as expected, since the entire LTG:GaAs layer has been removed at this point, so the passivation effect is gone. Since the y intercept, as well as the slope of the curve, changes dramatically it appears that the conductive layer between electrodes has been effectively pinched off. Although the proximity of metal electrodes likely plays a role in the exact nature of the resistance versus spacing relationship after etching, the general trend should be observed even if metal electrodes were not present. The results of this study lead us to conclude that the passivation characteristics of thin layers of LTG:GaAs make this structure well suited for shallow-etched nanoelectronic stuctures. This feature will likely be essential for the realization of high density circuits, since relatively gentle (and low-damage) etching techniques can be used to pattern and isolate devices.

The ability to isolate islands of $n++$ regions by stripping the LTG:GaAs is a novel way to fabricate both conventional devices and nanoelectronic devices. For example, we have been able to fabricate a metal semiconductor field effect transistor (MESFET) by combining the ability to make low resistance ohmic contacts using LTG:GaAs capped $n++$ layer and the modification of the conduction into the substrate upon stripping the LTG:GaAs cap. The MESFET structure has an active region comprised of a 1500 Å n-doped layer ($\approx 2\times10^{17}\,\mathrm{cm}^{-3}$) separated from the nonalloyed ohmic contact structure ($\approx 10^{20}\,\mathrm{cm}^{-3}$ capped by a 35-Å-thick LTG:GaAs) by a 50 Å n-doped ($\approx 1\times10^{16}\,\mathrm{cm}^{-3}$) layer. Ti/Au contacts using conventional UV lithography were used to define the source and drain contacts. Before the gate deposition the surface was treated with 3–5 iterations of the oxidation/strip iteration used in the study shown in Fig. 4. This would etch almost all of the top contact layer structure and the metal would be deposited on the lightly doped n-GaAs layer to form a Schottky contact. The resulting low-leakage MESFET indicates that we have been able to achieve good FET performance using a single structure to provide the source/drain contacts and the low-resistance channel access region. The chemical stability and surface electronic characteristics of the LTG:GaAs layer play essential roles in providing an effective capping layer. Thus we are able to exploit the passivation/surface pinning characteristics of the LTG:GaAs layer on thin $n++$ layers and modify the type of contact (i.e., either a ohmic or Schottky contact) to the underlying heterostructure to form ohmic contacts and gate regions, respectively.

For nanoelectronic device applications, these shallow etching techniques are especially interesting since they are compatible with self-assembly techniques, including the SAM based patterning described earlier in this article. SAM based resists can potentially provide nanometer scale patterns, but typically do not allow deep etching due to their limited resistance to wet chemical and dry etching techniques. However, it should be feasible to remove 10–20 nm of material using these materials as resists. In addition, self-assembly techniques can provide structures such as uniform islands and arrays of dots or lines, all at the nanoscale. The demonstration of suitable device and contact structures therefore opens up the possibility of realizing nanoelectronic circuits using relatively high throughput fabrication techniques.

V. CONCLUSIONS

In conclusion, we have described a nonalloyed contact structure which may be suitable for high density nanoelectronic device applications and discussed several experiments aimed at developing nanoscale device structures. A self-assembled monolayer of an organic tether molecule has been deposited selectively in regions defined by an elastomer stamp pad. The patterned SAM has been imaged using an ellipsometric imaging technique and has been used as an

effective mask for the wet chemical etching of the GaAs device layers. In addition, it has been shown that the thin layer of LTG:GaAs provides effective surface passivation for a FET-like device structure and that such devices can be patterned by shallow etching (~5 nm) of the layers. These demonstrations provide device approaches and fabrication techniques which can be integrated to develop a high density nanoelectronic device technology with high throughput fabrication processes.

ACKNOWLEDGMENTS

This work was partially supported by DARPA/Army Research Office under Grant No. DAAH04-96-1-0437, NSF MRSEC program under Grant No. 9400415-G-0144, and AFOSR Grant No. F49620-96-1-0234A. The authors would like to thank S. Datta and J. Gomez for many helpful discussions throughout the course of this work.

[1]D. L. Klein, R. Roth, A. K. L. Lim, A. P. Alivisatos, and P. L. McEuen, Nature (London) **389**, 699 (1997).
[2]A. N. Korotkov, in *Molecular Electronics*, edited by J. Jortner and M. Ratner (Blackwell, Oxford, 1997).
[3]D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner, Nature (London) **391**, 156 (1998).
[4]H-L. Lee, S-S. Park, D-I. Park, S-H. Hahm, J-H. Lee, and J-H. Lee, J. Vac. Sci. Technol. B **16**, 762 (1998).
[5]A. G. Baca, F. Ren, J. C. Zopler, R. D. Briggs, and S. J. Pearton, Thin Solid Films **308-309**, 599 (1997).
[6]M. P. Patkar, T. P. Chin, J. M. Woodall, M. S. Lundstrom, and M. R. Melloch, Appl. Phys. Lett. **66**, 1412 (1995).
[7]M. R. Melloch, J. M. Woodall, E. S. Harmon, N. Otsuka, F. H. Pollak, D. D. Nolte, R. M. Feenstra, and M. A. Lutz, Annu. Rev. Mater. Sci. **25**, 547 (1995).
[8]R. M. Feenstra, J. M. Woodall, and G. D. Pettit, Phys. Rev. Lett. **71**, 1176 (1993).
[9]S. Hong, D. B. Janes, D. McInturff, R. Reifenberger, and J. M. Woodall, Appl. Phys. Lett. **68**, 2258 (1996).
[10]T. B. Ng, D. B. Janes, D. McInturff, and J. M. Woodall, Appl. Phys. Lett. **69**, 3551 (1996).
[11]H. J. Ueng, V. R. Kolagunta, D. B. Janes, K. J. Webb, D. T. McInturff, and M. R. Melloch, Appl. Phys. Lett. **71**, 2496 (1997).
[12]R. P. Andres, J. D. Bielefeld, J. I. Henderson, D. B. Janes, V. R. Kolagunta, C. P. Kubiak, W. Mahoney, and R. G. Osifchin, Science **273**, 1690 (1996).
[13]T. Lee, J. Liu, D. B. Janes, V. R. Kolagunta, J. Dicke, R. P. Andres, J. Lauterbach, M. R. Melloch, D. McInturff, J. M. Woodall, and R. Reifenberger, Appl. Phys. Lett. **74**, 2869 (1999).
[14]G. Haas, T. D. Pletcher, G. Bonilla, T. A. Jachimowski, H. H. Rotermund, and J. Lauterbach, J. Vac. Sci. Technol. A **16**, 1117 (1998).
[15]H. H. Berger, Solid-State Electron. **15**, 145 (1972).

Role of As_4 in Ga diffusion on the GaAs(001)-(2×4) surface: A molecular beam epitaxy-scanning tunneling microscopy study

H. Yang, V. P. LaBella, D. W. Bullock, and P. M. Thibado[a)]
Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701

(Received 19 January 1999; accepted 5 May 1999)

The role of As_4 molecules in Ga diffusion on the GaAs(001)-(2×4) reconstructed surface has been studied using a combined molecular beam epitaxy and scanning tunneling microscopy multichamber facility. We deposited 10% of a plane of Ga atoms onto an otherwise pristine surface, while exposed to two separate As_4 beam equivalent pressures of 10^{-5} and 10^{-6} Torr. The higher As_4 flux resulted in the production of fewer and larger islands, indicating that increasing the As_4 flux increases the total interrogation area available to the Ga atoms before forming islands. © *1999 American Vacuum Society.* [S0734-211X(99)05904-1]

I. INTRODUCTION

The fabrication of III–V semiconductor structures is a major component of the rapidly growing wireless communications and optoelectronics industry.[1] Unlike silicon based devices, which are mainly fabricated by ion implantation,[2] III–V semiconductor structures must be fabricated by depositing one plane of atoms on top of another until the entire structure is formed. Consequently, III–V structure fabrication occurs solely at a surface. Naturally, a deeper understanding of the fundamental processes involved in growth, such as surface diffusion and island nucleation, may result in higher performance structures.[3] The most utilized surface for devices, and hence the most important surface for developing this understanding, is the GaAs(001)-(2×4) reconstructed surface.

To date, Ga surface diffusion on the GaAs(001) surface has been extensively investigated using reflection high-energy electron diffraction (RHEED).[4,5] These studies have played an important role in making predictions for the diffusion coefficient and the activation energy for the migration of Ga atoms. However, associating these values with the pure migration of Ga atoms can be ambiguous since the influence of surface defects, steps, and interactions between other adatoms cannot be taken into account. The pure migration of individual adatoms on elemental semiconductor surfaces has been successfully measured without these complications by depositing submonolayer coverages onto otherwise pristine surfaces and using scanning tunneling microscopy (STM) to quantify the outcome.[6] For binary compound semiconductors there is added complexity due to the unknown influence of the group V species, however we previously reported how to overcome these complications and successfully measured the Ga diffusion under normal growth conditions using STM.[7,8] In molecular beam epitaxy (MBE) growth of GaAs, the flux of the Ga determines the growth rate, while the flux of the arsenic is self-limiting. However, it is not known exactly how the As_4 flux effects the surface diffusion of the Ga atoms. In this article, we investigate the role of the arsenic molecules on the migration of the Ga adatoms.

a)Electronic mail: thibado@comp.uark.edu

II. EXPERIMENTAL PROCEDURE

Experiments were carried out in an ultrahigh vacuum (UHV) multichamber facility ($5-8\times10^{-11}$ Torr throughout) which contains a customized, commercial MBE (Riber 32P) chamber and a surface analysis chamber with a commercial STM (Omicron).[9] The MBE chamber was modified to include a highly accurate and fast optical substrate temperature measurement system (±2 °C from 0 to 700 °C and updated at 1 Hz) which utilizes the temperature dependence of the fundamental band gap.[10] Commercially available, "epiready," $n+$ (Si doped $10^{18}/cm^3$) GaAs(001)±0.1° substrates were loaded into the MBE system without any chemical cleaning. Once the surface oxide layer was removed, a 1.5 μm thick GaAs buffer layer was grown at ~1 μm/h as determined by RHEED.

The surface of the buffer layer was found to have several monolayers of roughness and a multitude of other atomic-scale defects. An algorithm was developed to remove this roughness and produce enormous terraces (~1 μm wide) essentially free of defects while cooling the sample quickly to room temperature. The algorithm consisted of first holding the substrate temperature at 600 °C under an As_4 beam equivalent pressure (BEP) of 10^{-6} Torr for 15 min to eliminate step bunching and produce large terraces. The substrate is then annealed at 570 °C under an As_4 flux of 8×10^{-7} Torr for another 15 min to eliminate adatom and vacancy islands on the terraces. Next, the sample temperature was set to drop (heater current set to zero) to 450 °C (~1.5 °C/s) while decrementing the As_4 flux in 10^{-7} Torr steps for every 10 °C drop in sample temperature (i.e., the As valve is closed once the sample temperature falls below 500 °C). Once the substrate temperature reaches 450 °C the power to the substrate heater was restored to maintain a constant temperature for 15 min. During this period the As_4 flux still inside the growth chamber was pumped out to minimize condensation onto the sample's surface.[11] Note, during the cool down from 570 to 500 °C, the As_4 flux plays a pivotal role in producing a high-quality surface. This is because above 500 °C arsenic sublimates from the GaAs surface and an impinging As_4 flux is required to replenish the lost arsenic. However, too much arsenic flux will drive the surface into supersaturation and

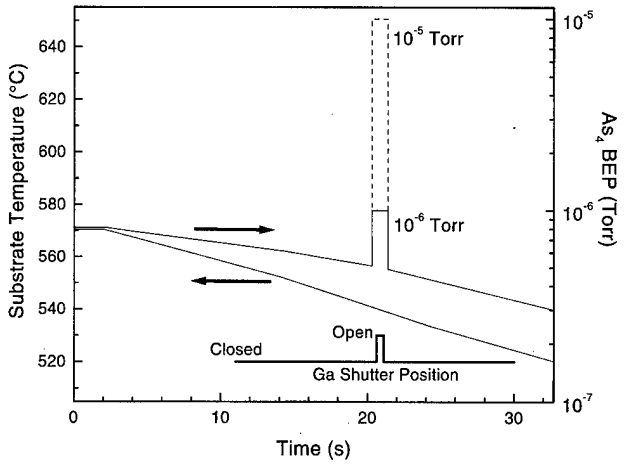


FIG. 1. Timing chart showing substrate temperature, Ga shutter position, and the As_4 BEP for a deposition scenario near 540 °C. Once the power to the substrate heater is turned off, the substrate temperature decreases at a rate of ~1.5 °C/s. Once the substrate reaches the desired temperature, the As_4 flux is increased to either 10^{-6} (solid line) or 10^{-5} (dashed line) Torr and the Ga shutter is opened for 0.5 s, resulting in a growth of 0.1 ML of GaAs.

the (2×4) surface reconstruction will be driven into the more arsenic rich $c(4\times4)$ surface reconstruction.

The above algorithm allows for an island-free surface with larger terraces to be produced. This algorithm was slightly modified to allow us to deposit 10% of a plane of Ga atoms, while simultaneously exposing the surface to two different As_4 BEPs. To illustrate how the growth of the submonolayer coverages was completed, a timing chart showing the substrate temperature, the Ga shutter position, and the As_4 BEP is shown in Fig. 1. Just prior to deposition, as stated above, the substrate is annealed at 570 °C while exposed to an As_4 BEP of 8×10^{-7} Torr. Next, the heater power to the sample is set to zero and the sample begins to cool at ~1.5 °C/s. Once the sample temperature has reached the desired value the As_4 BEP is increased to either 10^{-6} or 10^{-5} Torr, and the Ga shutter is opened for 0.5 s to deposit 10% of a plane of GaAs (i.e., the growth rate is set to 20% of a plane per second). Since all the Ga metal is deposited in 0.5 s, and the sample temperature is dropping at ~1.5 °C/s we can accurately determine the temperature. The As_4 BEP of 10^{-6} Torr and the Ga deposition rate of 20% of a plane per second yields an As_4/Ga flux ratio of ~15, thereby recreating typical GaAs growth conditions which use a ratio between 10 and 20. When the As_4 BEP is 10^{-5} Torr, the As_4/Ga flux ratio is ~150, which allows us to study the Ga diffusion under extremely high arsenic concentration conditions. Note, we could not study the Ga diffusion at lower arsenic fluxes, because, as explained earlier, the surface begins to degrade if the arsenic pressure is too low.

The Ga deposition on the (2×4) surface reconstruction was studied over as large a substrate temperature range as possible (570–460 °C) with two different arsenic fluxes. After deposition, the sample was transferred to the STM chamber through an all UHV transfer module. Once the sample and the STM reached thermal equilibrium with the STM imaging stage (~15 min), multiple images (6–12) were acquired from several, random, step-free regions that were either 200 nm by 200 nm or 100 nm by 100 nm in size to get an average island number density for each run.

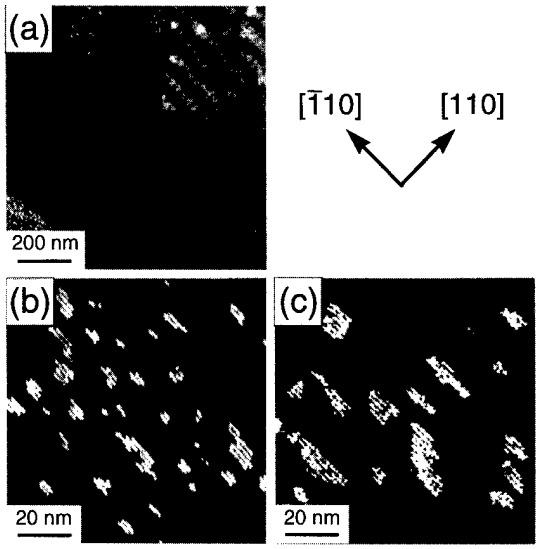


FIG. 2. Filled state STM images (~3.0 V, ~0.2 nA) of the GaAs(001)-(2×4) surface after deposition of 10% of a monolayer of GaAs onto a pristine surface. The deposition temperature is 540 °C for all the images: (a) 1 μm by 1 μm image, inset image shows an atomically well-ordered substrate; (b) 100 nm by 100 nm image of surface prepared with an As_4 BEP of 10^{-6} Torr, yielding 41 islands; (c) same conditions as in (b) but with an As_4 BEP of 10^{-5} Torr yielding 19 islands.

III. RESULTS AND DISCUSSION

Typical filled state STM images of the GaAs surface after the deposition of 10% of a monolayer are shown in Fig. 2. The routine observation of 1-μm-wide terraces allowed for

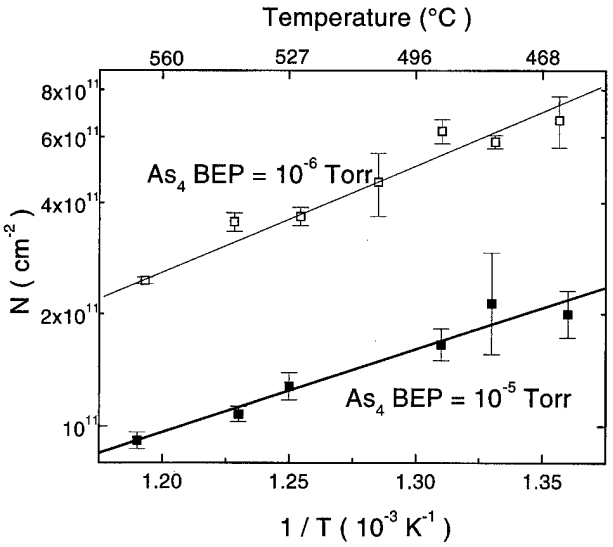


FIG. 3. Complete data set of island number density vs substrate temperature for two different As_4 BEP values. The error bars represent ±1 standard deviation as calculated from a set of images from each deposition temperature.

these and several other STM images to be acquired far from surface steps and shown in Fig. 2(a). The surfaces were well ordered and essentially free of defects down to the atomic scale as shown in the inset image within Fig. 2(a). Typical STM images for 10% of a plane of Ga atoms deposited on a surface at 540 °C using an As_4 BEP of 10^{-6} and 10^{-5} Torr are shown in Figs. 2(b) and 2(c), respectively. One striking feature is that there are less islands and the islands are larger in size when the As_4 BEP is higher. The increase in the island size is because the same amount of material is deposited for both As_4 BEPs, however a fewer number of islands were created.

The tendency of producing larger and hence fewer islands when the Ga atoms are deposited in a higher As_4 BEP environment persisted across the entire temperature range studied. For both As_4 BEPs, the average island number density is plotted on a logarithmic scale versus inverse temperature and shown in Fig. 3. These data points represent the average island number densities counted from all STM images acquired. One surprising feature is that the slopes for both 10^{-6} and 10^{-5} Torr As_4 BEPs are almost identical. This indicates that the activation energy barrier height for Ga diffusion is the same. However, the intercepts for these data sets are significantly different. The intercept change reflects the change in island number density between the two As_4 BEPs. Fewer islands under higher As_4 pressures indicates that the Ga atoms are able to interrogate a larger area of the surface than when exposed to lower As_4 pressures.

First-principles total-energy calculations of the activation energy for the pure migration of Ga atoms in an arsenic-free environment show that the diffusion is highly anisotropic.[12] The barrier height was found to be 1.5 eV along the [110] direction, while only 1.2 eV along the $[1\bar{1}0]$ direction. This difference results in Ga atoms being 100 times more likely to diffuse along the $[1\bar{1}0]$ direction than along the [110] direction at normal growth temperatures. Consequently, the Ga atoms in an arsenic-free environment are diffusing only in one dimension or along a line on the crystal's surface. Therefore, the Ga atoms will interrogate much fewer surface sites in a given unit time, than if they were able to diffuse freely in all directions.

We believe the low arsenic pressure data represent essentially the pure migration of Ga atoms on the surface with minimal influence due to arsenic. Under these conditions the Ga diffuses only in one dimension and hence can only interrogate the crystal surface along a line. When the arsenic pressure is high, the Ga atom diffusion is modified in such a way that the Ga atom can interrogate a larger area of the crystal surface but still experience the same energy barrier height (see Fig. 3). Thus, we believe the effect of the increase in the As_4 flux is to lower the activation energy barrier for hopping in the [110] direction closer to the barrier height in the $[1\bar{1}0]$ direction (1.2 eV). This results in the Ga atoms diffusing more isotropic in two dimensions, which allows them to interrogate a much greater area of the surface, while still yielding the same activation energy for hopping.

IV. CONCLUSION

The effect of As_4 flux on the diffusion of Ga atoms on the GaAs(001)-(2×4) reconstructed surface has been studied. It has been determined that the higher arsenic pressure leads to the production of larger and fewer islands when compared to those obtained at the lower As_4 BEP. These two observations suggest that the additional arsenic flux increases the mobility of the Ga atoms by possibly opening previously prohibited diffusion channels.

ACKNOWLEDGMENTS

This work is supported, in part, by ONR Grant No. N00014-97-1-1058 and NSF Grant No. 9733994.

[1]See, for example, Compd. Semicond. **4**, 26 (1998).
[2]M. I. Current, J. Vac. Sci. Technol. A **14**, 1115 (1996).
[3]Y. Nakamura, S. Koshiba, M. Tsuchiya, H. Kano, and H. Sakaki, Appl. Phys. Lett. **59**, 700 (1991).
[4]J. H. Neave, P. J. Dobson, B. A. Joyce, and J. Zhang, Appl. Phys. Lett. **47**, 100 (1985).
[5]T. Shitara, D. D. Vvedensky, M. R. Wilby, J. Zhang, J. H. Neave, and B. A. Joyce, Appl. Phys. Lett. **60**, 1504 (1992).
[6]Y. W. Mo, J. Kleiner, M. B. Webb, and M. G. Lagally, Phys. Rev. Lett. **66**, 1998 (1991).
[7]H. Yang, V. P. LaBella, D. W. Bullock, Z. Ding, J. B. Smathers, and P. M. Thibado, J. Cryst. Growth **201/202**, 88 (1999).
[8]H. Yang, V. P. LaBella, D. W. Bullock, W. G. Harter, and P. M. Thibado, Phys. Rev. Lett. (submitted).
[9]J. B. Smathers, D. W. Bullock, Z. Ding, G. J. Salamo, and P. M. Thibado, J. Vac. Sci. Technol. B **16**, 3112 (1998).
[10]P. M. Thibado, G. J. Salamo, and Y. Baharav, J. Vac. Sci. Technol. B **17**, 253 (1999).
[11]M. D. Johnson, J. Sudijono, A. W. Hunt, and B. G. Orr, Surf. Sci. **298**, 392 (1993).
[12]A. Kley, P. Ruggerone, and M. Scheffler, Phys. Rev. Lett. **79**, 5278 (1997).

Cross-sectional scanning tunneling microscopy of InAsSb/InAsP superlattices

S. L. Zuo and E. T. Yu[a)]
Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, California 92093-0407

A. A. Allerman and R. M. Biefeld
Sandia National Laboratories, Albuquerque, New Mexico 87185-0601

(Received 20 January 1999; accepted 20 May 1999)

Cross-sectional scanning tunneling microscopy (STM) has been used to characterize compositional structures in $InAs_{0.87}Sb_{0.13}/InAs_{0.73}P_{0.27}$ and $InAs_{0.83}Sb_{0.17}/InAs_{0.60}P_{0.40}$ strained-layer superlattice structures grown by metal-organic chemical vapor deposition. High-resolution STM images of the (110) cross section reveal compositional features within both the $InAs_xSb_{1-x}$ and $InAs_yP_{1-y}$ alloy layers oriented along the $[\bar{1}12]$ and $[1\bar{1}2]$ directions—the same as those in which features would be observed for CuPt–B type ordered alloys. Typically one variant dominates in a given area, although occasionally the coexistence of both variants is observed. Furthermore, such features in the alloy layers appear to be correlated across heterojunction interfaces in a manner that provides support for III–V alloy ordering models which suggest that compositional order can arise from strain-induced order near the surface of an epitaxially growing crystal. Finally, atomically resolved $(1\bar{1}0)$ images obtained from the $InAs_{0.87}Sb_{0.13}/InAs_{0.73}P_{0.27}$ sample reveal compositional features in the [112] and $[\bar{1}\bar{1}2]$ directions, i.e., those in which features would be observed for CuPt–A type ordering.
© *1999 American Vacuum Society.* [S0734-211X(99)08404-8]

I. INTRODUCTION

Mid- to long-wavelength infrared emitters and detectors are of great importance for applications such as chemical sensing, fiber communication, molecular spectroscopy, and infrared detection. $InAs_xSb_{1-x}$-based alloys and superlattices have shown great promise for mid- to long-wavelength infrared devices operating at wavelengths from 3 to 12 μm.[1,2] These materials are of particular interest because in narrow band gap materials, the threshold energy for Auger processes, which frequently suppress the radiating recombination,[3] is reduced and the wavelength at which emitters operate at room temperature is therefore typically limited to ~2.1–2.3 μm.[4] For $InAs_xSb_{1-x}$-based materials grown on GaSb or InAs substrates, the compressive strain in the $InAs_xSb_{1-x}$ layer will facilitate population inversion in laser structures[5] and will suppress Auger processes by decreasing the hole effective mass and increasing the Auger process threshold energy.[3] In addition, $InAs_xSb_{1-x}/InAs_yP_{1-y}$ heterostructures have been of considerable recent interest for such applications, because the enhanced valence-band offset improves the heavy-light hole splitting within the compressively strained $InAs_xSb_{1-x}$ layer, further suppressing Auger processes.[6] Crystal quality is also expected to be improved due to strain compensation.

For a variety of ternary and quaternary III–V alloys, features in compositional structure such as ordering, clustering, phase separation, and spontaneous compositional modulation have been observed to occur.[7–10] Such features may significantly influence material properties such as band gap,[11] interface quality, carrier mobility, or luminescence spectra.[12–15] Indeed, partial CuPt–B type ordering has been observed to occur, and to lead to a reduction in band gap, within $InAs_xSb_{1-x}$ alloys for Sb compositions as low as $x \approx 0.07-0.14$.[2,10] In $InAs_xSb_{1-x}$ alloys, ordering-induced band gap narrowing can extend the useful photoresponse to 8–12 μm, corresponding to band gaps as small as ~100 meV, much smaller than expected for random alloys.[16] The existence of a similar compositional structure, and its precise nature, within $InAs_xSb_{1-x}/InAs_yP_{1-y}$ superlattices, and especially within the $InAs_xSb_{1-x}$ layer, may significantly influence material and device properties. A detailed understanding of the atomic-scale compositional structure within such heterostructures is therefore essential for optimization of material properties and device performance.

In this article, we describe ultrahigh-vacuum cross-sectional scanning tunneling microscopy (STM) studies of pseudomorphic $InAs_xSb_{1-x}/InAs_yP_{1-y}$ strained-layer superlattices (SLS) grown on InAs (001) substrates. Details concerning the experiments and sample structures will be discussed in Sec. II. High-resolution STM images obtained from both cross-sections of $InAs_xSb_{1-x}/InAs_yP_{1-y}$ SLS samples revealing features in compositional structure with an orientation similar to that observed in both CuPt–B type and CuPt–A type ordering will be presented in Sec. III. Furthermore, compositional features with similar local strain relative to that in the surrounding material are observed to be correlated across the heterojunction interfaces, suggesting that strain plays a very important role in the determination of compositional structure in these materials. Section IV concludes the article.

[a)]Electronic mail: ety@ece.ucsd.edu

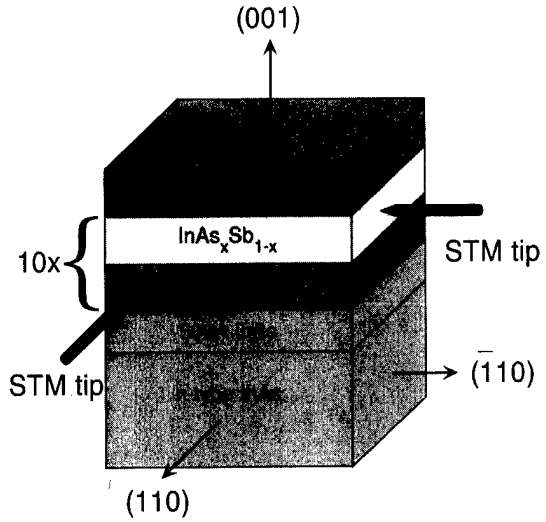


FIG. 1. Schematic diagram of the sample structure and STM geometry used in these studies. Samples consist of a 500 Å InAs buffer layer grown on an *n*-type InAs (001) substrate, followed by a ten-period $InAs_xSb_{1-x}$/$InAs_yP_{1-y}$ multiple-quantum-well structure capped by a final $InAs_yP_{1-y}$ layer.

II. EXPERIMENT

The $InAs_xSb_{1-x}$/$InAs_yP_{1-y}$ multiple-quantum-well (MQW) samples used in these studies were grown by metal-organic chemical vapor deposition (MOCVD). The growth temperature was 500 °C and the growth rate was 2.5 Å/s; additional details concerning the growth apparatus and procedures are described elsewhere.[6,17] Figure 1 shows a schematic illustration of the sample structure and STM geometry. For each sample, a 500 Å InAs buffer layer was grown initially on *n*-type InAs (001) substrates, followed by a ten-period $InAs_xSb_{1-x}$/$InAs_yP_{1-y}$ MQW structure, and finally another $InAs_yP_{1-y}$ layer. Here we present results obtained from two samples—a 65 Å $InAs_{0.87}Sb_{0.13}$/90 Å $InAs_{0.73}P_{0.27}$ MQW structure sample, and a 110 Å $InAs_{0.83}Sb_{0.17}$/112 Å $InAs_{0.6}P_{0.4}$ MQW structure sample. STM studies were performed on both (110) and (1$\bar{1}$0) cross-sectional surfaces exposed by *in situ* cleaving in an ultrahigh-vacuum (UHV) STM chamber at a pressure of $\sim 7\times10^{-11}$ Torr. Electrochemically etched W tips and commercially available Pt/Ir tips cleaned *in situ* by electron bombardment were used for these studies.

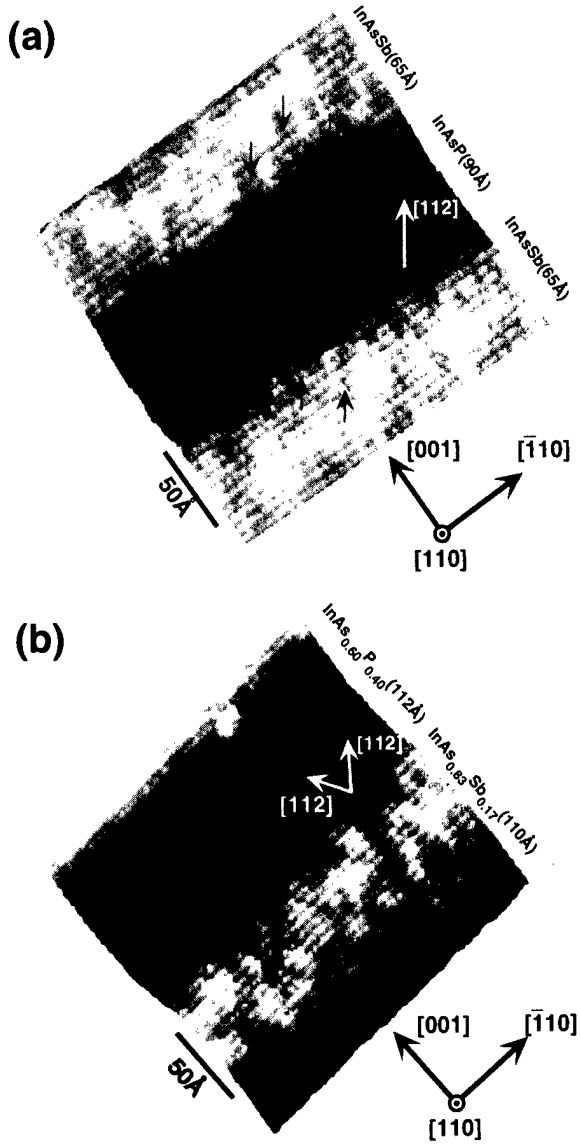


FIG. 2. (a) Three-dimensional rendering of a 210 Å×210 Å (110) constant-current STM image of the $InAs_{0.87}Sb_{0.13}$/$InAs_{0.73}P_{0.27}$ multiple-quantum-well structure, obtained at a sample bias voltage of −1.6 V and a tunneling current of 0.1 nA. [1̄12] compositional features correlated across heterojunction interfaces are marked by arrows. (b) Three-dimensional rendering of a 230 Å×230 Å constant-current (110) STM image of the $InAs_{0.83}Sb_{0.17}$/$InAs_{0.60}P_{0.40}$ multiple-quantum-well structure, obtained at a sample bias voltage of −1.8 V and a tunneling current of 0.1 nA. [1̄12] and [11̄2] compositional features correlated across heterojunction interfaces are marked by arrows.

III. RESULTS AND DISCUSSION

Figure 2(a) shows a three-dimensional rendering of a 210 Å×210 Å filled-state (110) constant-current STM image of the 65 Å $InAs_{0.87}Sb_{0.13}$/90 Å $InAs_{0.73}P_{0.27}$ multiple quantum well structure, obtained at a sample bias of −1.6 V and a tunneling current of 0.1 nA. The $InAs_xSb_{1-x}$ layer appears brighter (higher topographically), while the $InAs_yP_{1-y}$ layer appears darker (lower topographically), with a typical topographic contrast of ~2–2.5 Å between the two layers. The topographic contrast observed between the $InAs_xSb_{1-x}$ and $InAs_yP_{1-y}$ layers can arise from two sources. First, the valence-band edge of $InAs_{0.87}Sb_{0.13}$ is higher than that of $InAs_{0.73}P_{0.27}$, leading to a larger tip-sample separation, i.e., greater topographic height, in the $InAs_xSb_{1-x}$ layer in a filled-state constant current STM image. Second, the compressive and tensile strain in the $InAs_xSb_{1-x}$ and $InAs_yP_{1-y}$ layers, respectively, will lead to actual topographic variation on the cleaved surface. The compressively strained $InAs_xSb_{1-x}$ layer will protrude slightly, while the $InAs_yP_{1-y}$ layer, which is under tensile strain, will be slightly depressed. The morphological variations will also cause the $InAs_xSb_{1-x}$ layer to appear higher and the $InAs_yP_{1-y}$ layer

to appear lower topographically. A simple calculation based on elasticity theory[18,19] indicates that the amplitude of these undulations should not exceed a few tenths of an angstrom; the contrast between the $InAs_{0.87}Sb_{0.13}$ and the $InAs_{0.73}P_{0.27}$ layers therefore appears to be primarily electronically induced.

In addition to the overall contrast between different layers, detailed variations in contrast corresponding to local variations in electronic and, by extension, compositional structure are clearly visible. Specifically, compositional features aligned along the $\langle 112 \rangle$ directions in both the $InAs_xSb_{1-x}$ and $InAs_yP_{1-y}$ layers are clearly seen. In Fig. 2(a), compositional features along the $[\bar{1}12]$ direction in the (110) plane (perpendicular to the $[1\bar{1}1]$ direction in the crystal), are clearly visible in both the $InAs_{0.73}P_{0.27}$ and $InAs_{0.87}Sb_{0.13}$ alloy layers. These $[\bar{1}12]$ features are associated with compositional variations along the $[1\bar{1}1]$ direction in the crystal. In CuPt–B type ordering, the type most commonly observed in III–V zincblende semiconductors, the composition is periodic with a period of $2a/\sqrt{3}$ (where a is the cubic lattice constant) along the $[\bar{1}11]$ and $[1\bar{1}1]$ directions in the crystal. The features seen in Fig. 2(a) are therefore reminiscent of those associated with the $(1\bar{1}1)$ variant of CuPt–B type ordering. However, the compositional structure observed does not correspond to actual CuPt–B type ordering, for which the periodicity in the $[1\bar{1}1]$ direction would be only ~7 Å. The observation of compositional variations along only the $[1\bar{1}1]$ direction is most likely a local characteristic of the area imaged by STM, as previous studies of $InAs_xSb_{1-x}$ alloys have shown that CuPt–B type ordering can occur in these materials with local variations present in the relative intensities of the $(\bar{1}11)$ and $(1\bar{1}1)$ variants.[10]

Closer examination of the STM image reveals that compositional structure in the alloy layers appears to be correlated across heterojunction interfaces. Specifically, dark regions in the $InAs_yP_{1-y}$ layer in Fig. 2(a) appear to extend across the heterojunction interfaces in the $[\bar{1}12]$ direction and marge with dark regions in the $InAs_xSb_{1-x}$ alloy. In the $InAs_xSb_{1-x}$ alloy, increasing the Sb concentration will cause the valence-band edge to move higher in energy and the degree of compressive strain to increase, while in the $InAs_yP_{1-y}$ alloy, increasing the P concentration will cause the valence-band edge to move lower and the degree of tensile strain to increase. The areas of dark contrast in the $InAs_yP_{1-y}$ layer, therefore, correspond unambiguously to regions of P-rich composition, with greater tensile strain than the surrounding alloy, while those in the $InAs_xSb_{1-x}$ layer correspond to As-rich regions, with less compressive strain relative to the surrounding material. The alignment of P-rich regions in the $InAs_yP_{1-y}$ layer with As-rich regions in the $InAs_xSb_{1-x}$ layer observed in the image therefore corresponds to an alignment within $(1\bar{1}1)$ planes across heterojunction interfaces of regions with a smaller average lattice constant than the surrounding material.

Figure 2(b) shows a three-dimensional rendering of a 230 Å×230 Å filled-state (110) constant-current STM image of the 110 Å $InAs_{0.83}Sb_{0.17}$/112 Å $InAs_{0.60}P_{0.40}$ multiple-quantum-well structure, obtained at a sample bias of −1.8 V and a tunneling current of 0.1 nA. In this image, compositional features aligned along the $[\bar{1}12]$ and $[1\bar{1}2]$ directions are visible in both the $InAs_{0.83}Sb_{0.17}$ and $InAs_{0.60}P_{0.40}$ layers. Unlike the region shown in Fig. 2(a), where compositional variation is observed predominantly along the $[1\bar{1}1]$ direction, features corresponding to compositional variation in both the $[\bar{1}11]$ and $[1\bar{1}1]$ directions are observed to coexist in this image. On the basis of previous transmission electron diffraction studies of $InAs_xSb_{1-x}$ alloy ordering in which the relative intensities of CuPt–B type ordering variants were found to vary with location in the sample,[10] we believe that observations of compositional features along both the $[\bar{1}11]$ and $[1\bar{1}1]$ directions or of compositional features along only one of these directions in different images reflect local variations within a given sample, rather than systematic differences between samples.

Figure 2(b) also reveals the presence of correlations of compositional features across the heterojunction interfaces similar to those seen in Fig. 2(a). As with the previous sample, increasing the Sb concentration in the $InAs_xSb_{1-x}$ alloy will cause the valence-band edge to move higher in energy and the degree of compressive strain to increase, while increasing the P concentration in the $InAs_yP_{1-y}$ alloy will cause the valence-band edge to move lower and the degree of tensile strain to increase. Darker features in both the $InAs_xSb_{1-x}$ and $InAs_yP_{1-y}$ layers are therefore associated with an average lattice constant smaller than that in the surrounding material, while brighter features are associated with a larger average lattice constant. The observation, in both samples studied, of correlation across heterojunction interfaces of regions with similar strain relative to the surrounding material suggests that local strain is a significant factor in the determination of compositional structure in these materials. Theoretical models of ordering in III–V alloys have suggested that alloy ordering arises from strain-induced order in surface reconstructions present during growth that is subsequently incorporated into the structure of the epitaxially grown alloy.[20–22] Our observations described above are consistent with such models of ordering during epitaxial growth, and further suggests that, to some degree, such ordering can propagate across heterojunction interfaces.

We have also obtained STM images of the $(1\bar{1}0)$ cross-sectional surface for the 65 Å $InAs_{0.87}Sb_{0.13}$/90 Å $InAs_{0.73}P_{0.27}$ multiple-quantum-well sample. These images provide additional information about compositional structure within the group V sublattice. Figure 3(a) shows a three-dimensional rendering of an atomically resolved $(1\bar{1}0)$ filled-state STM image obtained at a sample bias of −1.65 V and a tunneling current of 0.1 nA. From this image we see that features oriented along the [112] and $[\bar{1}\bar{1}2]$ directions are present within the $InAs_xSb_{1-x}$ layer. Such features correspond to compositional variation along the $[11\bar{1}]$ and $[\bar{1}11]$ directions, respectively, the same as those in which compositional variations occur for CuPt–A type ordering, which is observed much less frequently than CuPt–B type ordering.[22,23]

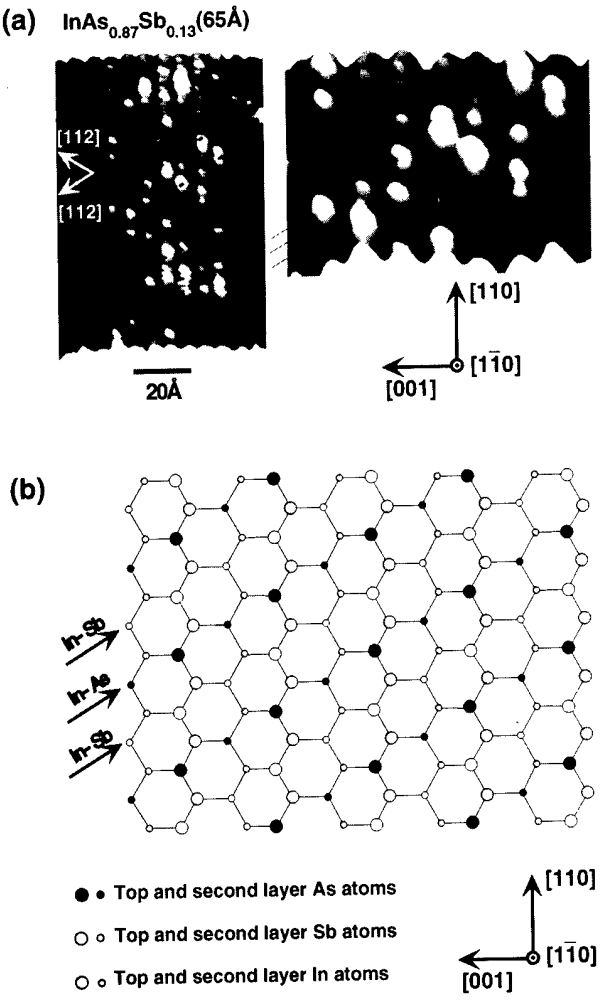


FIG. 3. (a) Three-dimensional rendering of a constant-current ($1\bar{1}0$) STM of an $InAs_{0.87}Sb_{0.13}$ layer within the $InAs_{0.87}Sb_{0.13}/InAs_{0.73}P_{0.27}$ superlattice, obtained at a sample bias voltage of -1.65 V and a tunneling current of 0.1 nA. Compositional features aligned in the [112] and [$\bar{1}$12] directions are delineated by dotted lines. The magnified view at right contains an area with a high degree of CuPt–A type ordering. (b) Schematic illustration of the (111) variant of perfect CuPt–A type ordering in $InAs_xSb_{1-x}$.

The resolution attained in this image is sufficient for more detailed investigations of compositional structure on the exposed ($1\bar{1}0$) surface, which reveal nanoscale regions within the alloy in which strong CuPt–A type ordering appears to occur. In the magnified view of the $InAs_xSb_{1-x}$ layer shown in Fig. 3(a), a compositional feature aligned in the [$\bar{1}$12] direction is indicated. The bright and dark atoms in the image correspond to Sb and As, respectively. Simple counting yields a composition of approximately $InAs_{0.77}Sb_{0.23}$ in this region of the alloy. In the magnified view, three arrows indicate the intersections of three adjacent (111) planes in the crystal with the exposed ($1\bar{1}0$) cross-sectional surface. Figure 3(b) shows the compositional structure that would correspond to ideal CuPt–A type ordering, with the planes corresponding to those indicated in Fig. 3(a) marked by arrows. In the uppermost plane indicated in Fig. 3(a), six of the seven group V atoms in the top layer are Sb, with one being undetermined. In the bottom plane indicated in Fig. 3(a), five of the seven group V atoms in the top layer are Sb. In the middle plane, only the rightmost [110] row of atoms in the top layer clearly contains Sb; the remainder appear to be As atoms. For perfect CuPt–A type ordering, the top and bottom planes would contain only Sb, and the middle plane only As. While the group V composition in the second-layer atomic rows cannot be determined unambiguously from this image, the compositional structure observed in the top-layer rows strongly suggests that the area described above possesses a high degree of the (111) variant of CuPt–A type ordering. Because the average Sb composition in the alloy is only ~13%, it is not unexpected that only small regions of highly ordered material would be observed.

IV. CONCLUSION

Cross-sectional STM has been used to investigate compositional structure in $InAs_{0.87}Sb_{0.13}/InAs_{0.73}P_{0.27}$ and $InAs_{0.83}Sb_{0.17}/InAs_{0.60}P_{0.40}$ strained-layer superlattice samples. Filled-state (110) cross-sectional images of both samples reveal compositional features within both $InAs_xSb_{1-x}$ and $InAs_yP_{1-y}$ layer with orientations that are the same as those present for CuPt–B type ordering; regions in which compositional variations in both the [$\bar{1}$11] and [$1\bar{1}$1] directions coexist as well as areas in which one orientation dominates are observed. Furthermore, clear evidence is found that certain compositional features are correlated across heterojunction interfaces within the superlattices. Detailed analysis reveals that these compositional features are associated with local strain, suggesting that strain plays a very important role in the determination of compositional structure within these materials. These observations lend support to theories of ordering in epitaxially grown material in which surface reconstructions present during growth induce near-surface compositional order, which is then incorporated into the epitaxially grown alloy. Finally, atomically resolved ($1\bar{1}0$) cross-sectional images of an $InAs_{0.87}Sb_{0.13}/InAs_{0.73}P_{0.27}$ multiple-quantum-well sample reveal nanoscale regions in the $InAs_xSb_{1-x}$ layer within which the compositional structures appears to possess a high degree of CuPt–A type ordering.

ACKNOWLEDGMENTS

Part of this work was supported by NSF (ECS 95-07986). E.T.Y. would like to acknowledge financial support from the Alfred P. Sloan Foundation.

[1]R. M. Biefeld, A. A. Allerman, and S. R. Kurtz, Mater. Sci. Eng., B **51**, 1 (1998).
[2]S. R. Kurtz, L. R. Dawson, R. M. Biefeld, D. M. Follstaedt, and B. L. Doyle, Phys. Rev. B **46**, 1909 (1992).
[3]S. R. Kurtz and R. M. Biefeld, Proceedings of Seventh International Conference on Narrow Gap Semiconductors, Santa Fe, NM, 8–12 Jan. 1995 (unpublished), p. 18.
[4]H. K. Choi and S. J. Eglash, Appl. Phys. Lett. **61**, 1154 (1992).

[5]E. Yablonovitch and E. O. Kane, J. Lightwave Technol. **6**, 1292 (1988).
[6]R. M. Biefeld, A. A. Allerman, S. R. Kurtz, and J. H. Burkhart, Compound Semiconductors 1997: Proceedings of the IEEE Twenty Fourth International Symposium on Compound Semiconductors, San Diego, CA, 8–11 Sept. 1997 (unpublished), p. 113.
[7]H. R. Jen, D. S. Cao, and G. B. Stringfellow, Appl. Phys. Lett. **54**, 1890 (1989).
[8]H. R. Jen, K. Y. Ma, and G. B. Stringfellow, Appl. Phys. Lett. **54**, 1154 (1989).
[9]Y.-E. Ihm, N. Otsuka, J. Klem, and H. Morkoç, Appl. Phys. Lett. **51**, 2013 (1987).
[10]D. M. Follstaedt, R. M. Biefeld, S. R. Kurtz, and K. C. Baucom, J. Electron. Mater. **24**, 819 (1995).
[11]K. A. Mäder and A. Zunger, Appl. Phys. Lett. **64**, 2882 (1994).
[12]G. B. Stringfellow and G. S. Chen, J. Vac. Sci. Technol. B **9**, 2182 (1991).
[13]A. Zunger and S. Mahajan, *Handbook on Semiconductors*, edited by S. Mahajan (North–Holland, The Netherlands, 1994), Vol. 3B, p. 1399.
[14]P. K. Bhattacharya and J. W. Ku, J. Appl. Phys. **58**, 1410 (1985).
[15]A. Gomyo, T. Suzuki, and S. Iijima, Phys. Rev. Lett. **60**, 2645 (1988).
[16]M. Y. Yen, R. People, K. W. Wecht, and A. Y. Cho, Appl. Phys. Lett. **52**, 489 (1988).
[17]R. M. Biefeld, K. C. Baucom, and S. R. Kurtz, J. Cryst. Growth **137**, 231 (1994).
[18]See, e.g., L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, London, 1959).
[19]H. Chen, R. M. Feenstra, R. S. Goldman, C. Silfvenius, and G. Landgren, Appl. Phys. Lett. **72**, 1727 (1998).
[20]S. Froyen and A. Zunger, Phys. Rev. Lett. **66**, 2132 (1991).
[21]S. Froyen and A. Zunger, Phys. Rev. B **53**, 4570 (1996).
[22]S. B. Zhang, S. Froyen, and A. Zunger, Appl. Phys. Lett. **67**, 3141 (1995).
[23]T. Suzuki, T. Ichihashi, and T. Nakayama, Appl. Phys. Lett. **73**, 2588 (1998).

Characterization of AlSb/InAs surfaces and resonant tunneling devices

B. Z. Nosho[a] and W. H. Weinberg
Center for Quantized Electronic Structures and Department of Chemical Engineering, University of California, Santa Barbara, California 93106

W. Barvosa-Carter, A. S. Bracker, R. Magno, B. R. Bennett, J. C. Culbertson, B. V. Shanabrook, and L. J. Whitman[b]
Naval Research Laboratory, Washington, DC 20375

(Received 12 February 1999; accepted 5 May 1999)

We have studied the evolution of AlSb-on-InAs(001) surfaces and interfaces grown by molecular-beam epitaxy using *in situ* scanning tunneling microscopy. We find that forming InSb-like interfacial bonds on an InAs(001)-(2×4) surface creates surface roughness because the surface In coverage inherent to the (2×4) reconstruction is insufficient to form a complete InSb(001)-(1×3)-like surface layer. This morphological roughness can be eliminated by depositing additional In to compensate for the different compositions of the reconstructions. We have also grown three different 5-monolayer-thick films of AlSb on the InSb-like interface to study the effect of growth conditions on the film surface morphology. The AlSb surface can be improved by either raising the growth temperature or by growing the film using migration-enhanced epitaxy. Finally, we present electrical characterization of InAs/AlSb/GaSb resonant interband tunneling devices fabricated with different growth procedures. The possible effects of various growth procedures on interfacial quality and device properties are discussed. © *1999 American Vacuum Society.* [S0734-211X(99)05404-9]

I. INTRODUCTION

The formation of high-quality interfaces between the nearly lattice matched "6.1 Å" family of III–V semiconductors, InAs, GaSb, and AlSb, is an important step in the development and fabrication of electronic and electro-optic devices from this material system. These materials have been used to make both type-I resonant tunneling diodes (RTDs) and type-II resonant interband tunneling diodes (RITDs) that show great promise for high-speed (terahertz) electronics.[1,2] Whereas the RTDs use InAs for both the cladding and quantum well material and AlSb for the tunneling barriers, the RITDs use GaSb instead as the quantum well material. In both cases the AlSb barrier layers are typically very thin, ≤5 monolayer (ML), and the transmission coefficient for electrons to tunnel through both barriers is strongly peaked for particular energies. Accordingly, the electrical properties of these structures may be sensitive to both submonolayer fluctuations in the barrier thickness and atomic-scale compositional variations across the interfaces.[1,3,4] The incorporation of these structures into reliable high-speed circuits therefore requires an understanding of how to create smooth and abrupt interfaces with minimal interfacial disorder.

Interfacial disorder is caused by both kinetics and thermodynamics, and can be characterized by two components. When discussing interfacial disorder, we generally distinguish between morphological *roughness*, i.e., nanoscale variations in the position of the interface, and *intermixing*, i.e., fluctuations in the chemical composition on the atomic scale. Roughness is typically associated with surface topography that has evolved during the growth. When epitaxial growth does not occur in a step-flow manner, but proceeds instead by the nucleation and growth of islands on the surface, islands present at the time an interface is formed will cause interfacial roughness. The second component, intermixing, may occur *during* growth due to local disorder or inter-atomic exchange, or *after* an interface is formed due to thermodynamically driven interdiffusion. Although in some cases compositional variations may be favorable, by helping to reduce strain at the interface, for example, more often the ill-defined interfaces resulting from such disorder are expected to degrade device performance.

In previous work, we employed *in situ* scanning tunneling microscopy (STM) to characterize the surfaces that evolve during the growth of AlSb/InAs RTD-like structures by molecular-beam epitaxy (MBE).[5] We found that exposing an InAs surface to a Sb_2 flux creates a bilevel surface with vacancy islands covering ~25% of the surface. Upon further deposition of AlSb and then InAs layers on top of such a bilevel surface, the surface roughness due to islanding increased with each successive layer. Here, we describe our efforts to improve these growth surfaces and interfaces by understanding the atomic-scale structures that arise during growth, and thereby develop growth procedures that minimize interfacial roughness (although not necessarily intermixing, which we are also presently in the process of characterizing). Furthermore, we have begun to examine how these new procedures affect the electrical characteristics of functional RITDs.

II. EXPERIMENT

The experiments were carried out in an interconnected, multichamber ultrahigh vacuum facility that includes a solid-

[a]Electronic mail: brettn@engineering.ucsb.edu
[b]Electronic mail: Lloyd.Whitman@nrl.navy.mil

source MBE chamber equipped with reflection high-energy electron diffraction (RHEED), and a surface analysis chamber equipped with a STM.[6] All samples for STM study were grown without intentional doping on InAs(001) substrates using "cracked" arsenic and antimony. Growth rates for InAs and AlSb were calibrated by RHEED intensity oscillations. Following oxide removal, ~0.5-μm-thick InAs buffer layers were grown with a 5:1 beam equivalent pressure ratio of As:In at 1 ML/s, with 30 s interrupts under As_2 every 90 s. The substrate temperature during the buffer layer growth was approximately equal to the congruent sublimation temperature of InAs, estimated to be 470 °C. We have previously shown that by gradually reducing the As_2 flux during a 10 min interrupt following growth of the buffer layer, an island-free InAs surface is produced with a well-ordered As-terminated (2×4) reconstruction.[5]

After completion of the InAs buffer layer, the substrate temperature was reduced and AlSb growth was initiated by forming InSb interfacial bonds via migration-enhanced epitaxy (MEE). A layer of In was first deposited on the clean, As-terminated InAs surface, followed by a brief Sb_2 exposure. For example, the following shutter sequence would be used at the end of the 10 min interrupt: close the As shutter, open the In shutter for 1 s (1 ML), then open the Sb_2 shutter for 2 s. Next, a 5-ML-thick AlSb film was deposited on the InSb-like interface. Three different growth procedures for this layer were examined, including different growth temperatures, as will be described later. In each case, after depositing the 5 ML AlSb film a 15 s interrupt under Sb_2 was performed, and then the sample was allowed to cool to room temperature. In a RTD or RITD structure, these AlSb surfaces would represent possible starting surfaces upon which the InAs or GaSb quantum well material would be deposited. Immediately following the completion of each growth, the samples were removed from the MBE chamber and transferred *in vacuo* to the surface analysis chamber (base pressure $<1\times10^{-10}$ Torr), as described previously.[5] All STM images shown here were acquired in constant-current mode with sample biases ranging from −1.2 to −3.3 V and tunneling currents between 30 pA and 0.7 nA.

As a first step to correlating device performance with growth procedures specifically developed using the results of our surface and interface characterization, we have fabricated several complete InAs/AlSb/GaSb RITD structures on both InAs and GaAs substrates. For each substrate type, two samples were grown: one using a "conventional" growth procedure, and the other implementing a number of changes in the growth intended to affect interfacial structure. All four RITD samples were grown at ~440 °C, and consist of a 9 nm (30 ML) GaSb quantum well between two 1.5 nm (5 ML) AlSb barriers. Silicon was used as a *n*-type dopant for all the doped layers of the structure. For the InAs substrates, a 1-μm-thick InAs buffer layer doped at 3×10^{18} cm^{-3} was grown first, followed by 30 nm of InAs doped at 1×10^{17} cm^{-3}, and then 12 nm of undoped InAs. The first AlSb barrier was grown on this undoped layer, with InSb-like interfacial bonds formed using MEE. For the conventional growth, 1 ML of In+2 s Sb_2 was used during the MEE. For the alternate growth, additional In was deposited (as will be explained later). At the interface between the second AlSb barrier and the adjoining InAs layer, InSb-like bonds were also formed using MEE (with 1 ML of In). The InAs immediately on top of the second AlSb barrier was undoped for the first 12 nm, doped at 1×10^{17} cm^{-3} for the next 30 nm, and then doped at 3×10^{18} cm^{-3} for the final 0.2 μm. During the alternate growth procedure, in addition to using additional In during formation of the first AlSb/InAs interface, 100 s interrupts were added after the growth of the first AlSb barrier layer, the GaSb quantum well, and the second AlSb barrier. Similar RITD samples were grown on the semi-insulating GaAs substrates; the only difference is that a 0.2-μm-thick AlSb buffer layer and 0.5 μm of undoped InAs were grown before the doped InAs buffer layer. After growth, standard photolithography procedures were used to fabricate an array of gold ohmic contacts that also served as an etch mask for the formation of mesa RITDs. Current density-versus-voltage spectra were then recorded at room temperature for devices from each sample.

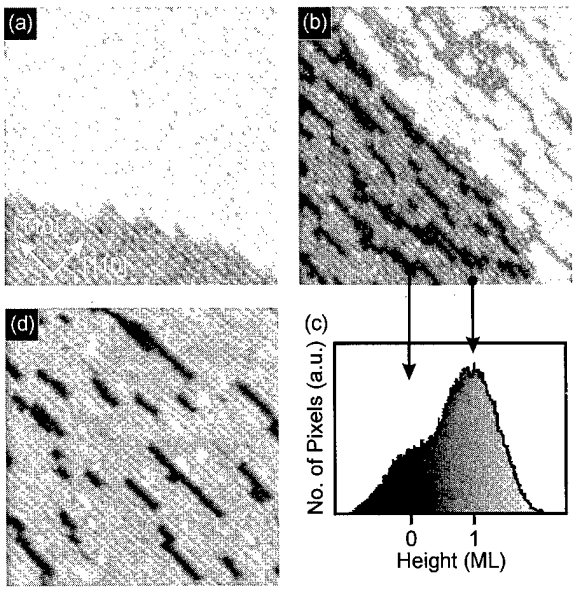


FIG. 1. Filled-state STM images (60 nm×60 nm) of (a) clean InAs(001)-(2×4); (b) InAs after depositing 1 ML of In and exposing to 2 s of Sb_2; and (d) InAs after depositing about 1.25 ML of In and exposing to 2 s of Sb_2. The height histogram for a single substrate terrace on the surface shown in (b) is displayed in (c).

III. RESULTS AND DISCUSSION

A typical STM image of the InAs buffer layer surface (after the 10 min interrupt) is shown in Fig. 1(a). The surface is nearly ideal, composed of atomically smooth terraces separated by monolayer-height (3 Å) steps. As previously reported, micron-scale images show that the terraces are ~0.5 μm wide with essentially no islands.[5] On the atomic scale, the surface exhibits a well-ordered (2×4) reconstruction, consistent with the sharp diffraction spots observed in the RHEED. Atomic resolution images (not shown) are simi-

lar in appearance to those previously reported for the As-terminated InAs and GaAs(001)-(2×4) surfaces, and are consistent with the generally accepted $\beta 2(2\times4)$ model for this reconstruction.[7] In this model, the surface III–V layer is nonstoichiometric, consisting of 1/2 ML of As on top of 3/4 ML of In. The 1/2 ML of As is in the form of dimers with each dimer bond oriented in the $[\bar{1}10]$ direction. Pairs of these dimers align in rows along the $[\bar{1}10]$ direction [the rows visible in Fig. 1(a)], separated by a row of single As dimers one InAs layer below.

The first interface of interest in a RITD structure is the AlSb-on-InAs interface, i.e., the interface directly before the first barrier layer. A key issue in the heteroepitaxial growth of antimonides on arsenides (or vice versa) is the type of interfacial bonds that are formed. For AlSb/InAs, the interfacial bonds can be either InSb-like or AlAs-like. For a number of reasons, including the general observation that InSb-like interfacial bonds lead to more abrupt interfaces,[8–11] we usually prepare our interfaces with this bond type. When this interfacial layer is created from the InAs starting surface by MEE (1 ML of In followed by 2 s of Sb_2), the RHEED pattern changes from the sharp (2×4) of clean InAs, to a streaky (1×3). Although we expect that the surface formed in this way should be terminated primarily with Sb, there may still be some remnant As present. [We will refer to this surface henceforth as the InSb(As) surface or interface.] An image of such a surface is displayed in Fig. 1(b). As discussed previously,[5] these surfaces exhibit a disordered (1×3)-like reconstruction with vacancy islands covering a quarter of the surface. The bifurcation of the surface into two levels is shown quantitatively in Fig. 1(c) by the height histogram for a single substrate terrace.

The (1×3)-like reconstruction of this InSb(As) surface is similar in appearance to that observed for other III–Sb materials. Based on our work[5] and previously reported results and models,[12–14] we believe the InSb(As) surface has a structure like the $(1\times3)/c(2\times6)$ reconstruction proposed for the clean III–Sb surfaces: a full plane (1 ML) of group III atoms covered by 1 2/3 ML of group V atoms (in this case possibly Sb and As). We have demonstrated elsewhere that the formation of vacancy islands is a direct result of the different compositions of the clean InAs (0.5 ML As/0.75 ML In) and Sb-terminated surfaces (~1.7 ML Sb+As/1 ML In).[15] When depositing 1 ML of In during MEE, one can think of 1/4 ML of this In "filling in" the In missing from the original (2×4)-reconstructed surface, while the other 3/4 ML remains as additional islands on the surface. Terminating this In with Sb, to complete the Sb-rich (1×3)-like reconstruction, leads to the observed 75%/25% bilevel surface morphology. Given this understanding of the vacancy island formation, and *additional* 0.25 ML of In can be added during the MEE procedure to compensate for the different compositions of the reconstructions. The much smoother morphology that results is shown in Fig. 1(d). For this growth the vacancy island coverage has been reduced to <7%. The primary limitation to getting a flat surface is depositing the precise amount of In which is needed.[15]

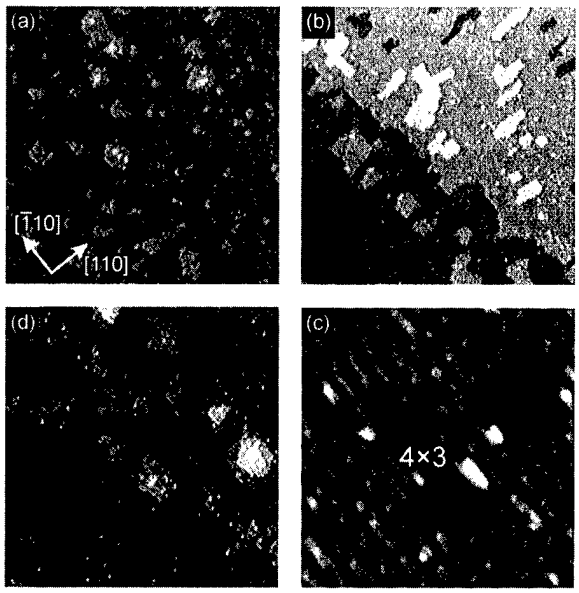


FIG. 2. Filled-state images of a 5-ML-thick AlSb film grown on a surface like that shown in Fig. 1(d). (a) Grown at 400 °C. (b) Grown at 470 °C. A higher resolution image of this surface, 17 nm×17 nm, is shown in (c). (d) Grown with layer-by-layer MEE at 400 °C. Images (a), (b), and (d) are 130 nm×130 nm, and have a gradient component added to the gray scale to accent the step edges.

We now proceed to examine the second interface—the one between the first AlSb barrier and the InAs or GaSb quantum well material. Three 5-ML-thick AlSb films were grown under different conditions, each starting on a flat, freshly prepared InSb(As) surface made using 1.25 ML of In, as just described. The first film was grown at 400 °C using a constant growth rate of 0.5 ML/s, producing the morphology shown in Fig. 2(a). The surface has a disordered (1×3)-like reconstruction and primarily consists of three different levels per substrate terrace, covering 21%, 67%, and 11% of each terrace (lowest to highest). It is apparent that the growth of the AlSb film is proceeding via island growth, not step flow, consistent with the observation of RHEED intensity oscillations during the growth. This growth mode is not surprising given that 400 °C is a relatively low temperature for AlSb growth.

To explore the effect of growth temperature on the AlSb film morphology, the second AlSb film was grown at the same growth rate as the previous sample, but at the InAs growth temperature of ~470 °C. As shown in Fig. 2(b), there is a dramatic change in the film surface morphology. The surface is still composed of three levels per terrace but is much smoother, with islands and vacancies comprising 17% and 3% of the surface, respectively. The growth is getting close to step flow due to the increased diffusion lengths at this higher temperature. The surface is also better ordered on the atomic scale, having relatively kink-free dimer rows [Fig. 2(c); compare with Fig. 4(c) in Ref. 5]. Based solely on its topography, this surface would appear to be a better starting point for the adjacent quantum well. However, it is important to recall that a variety of both kinetically and thermodynamically driven processes occur during epitaxy. For example,

the higher growth temperature may increase the amount of interfacial intermixing, or promote the formation of extended dislocations in a strained heterostructure. Thus, whereas the growth of AlSb at this higher temperature produces a much smoother growth surface, it ultimately may or may not lead to better device performance.

An interesting consequence of the increased atomic scale order on the AlSb film surface grown at higher temperature is the observation of a periodic "defect" that occurs in the surface Sb dimer rows within the supposed $c(2\times6)$ reconstruction. As shown in Fig. 2(c), in filled-state images it appears as if one Sb atom (half a dimer) is missing from every fourth dimer along the $[\bar{1}10]$ direction, giving the most well-ordered regions of the surface an overall (4×3) symmetry. Interestingly, hints of this structure are also apparent upon close inspection of previously published atomic-scale images of AlSb(001),[5,14] suggesting that this (4×3) structure may, in fact, be a low-energy reconstruction. Further investigations, both experimental and theoretical, are in progress with the goal of definitively determining the (4×3) structure and stability.

Because it may be advantageous to grow arsenide-antimonide heterostructures at relatively low temperatures (<450 °C), we have investigated whether the otherwise kinetically limited AlSb film morphology can be improved using layer-by-layer MEE growth. Specifically, the third of our 5-ML-thick AlSb films was grown by alternately depositing 1 ML of Al (2 s at 0.5 ML/s) and 5 s of Sb_2 repeated five times. Depositing the Al in the absence of a Sb_2 flux is expected to increase the diffusion length of Al adatoms, and the short interrupts under Sb_2 should help to coarsen any islands and thereby smooth the surface further. These effects are indeed apparent in the resulting surface morphology, shown in Fig. 2(d). Compared to the surface grown without the interrupts at the same temperature, Fig. 2(a), this surface has generally larger islands and fewer vacancies (i.e., the lower level on each terrace, whose area is reduced from 21% to 10%). Although the effectively lower growth rate and periodic interrupts improve the surface morphology, it should be noted that there might also be some problems created by this growth procedure. One problem is that the thickness of the layer could be more susceptible to flux transients or "bursts" that arise when a shutter is opened. Another issue is the longer growth times, during which the detrimental incorporation of dopant-like impurities into the active layers could become significant.

As a first step in exploring how growth procedures affect device characteristics, we have characterized full RITD structures fabricated using a number of procedures developed to reduce interfacial roughness. As described in Sec. II, one sample was grown on each substrate (GaAs and InAs) using conventional methods, i.e., the AlSb-on-InAs interface was formed by MEE using 1 ML of In. We will denote these samples "GaAs–A" and "InAs–A." A second pair of samples, "GaAs–B" and "InAs–B," was prepared using 1.25 ML of In (to produce a flatter AlSb-on-InAs interfacial layer), plus additional 100 s interrupts after the growth of each barrier layer and quantum well. Typical current–voltage spectra for InAs–A and InAs–B RITDs are displayed in Fig. 3. The spectra exhibit characteristic resonant tunneling behavior, with peak currents near ±130 mV and valley currents in the vicinity of ±350–400 mV. (The sharp features between the peaks and valleys are due to bistability effects.) Note that at positive bias the top of the mesa is positive with respect to the substrate side of the device. In general, we observed little difference in the negative current densities between sample sets for either substrate; however, the positive current densities are consistently lower for the samples with modified interfaces (the "B" samples). For example, the positive bias peak current density is $\sim2.5\times10^4$ A/cm^2 for InAs–A RITDs, and $\sim2.2\times10^4$ A/cm^2 for the InAs–B in Fig. 3 (values very similar to those previously reported for similar RITD structures).[16,17] Because both peak and valley currents decreased about the same amount, the peak-to-valley ratio, often used as a figure of merit for RITDs, did not change significantly. Considering the differences in the peak currents between the A and B samples relative to the variations in peak currents across both samples, the different interface treatments did not significantly affect the peak current densities or peak-to-valley ratios.

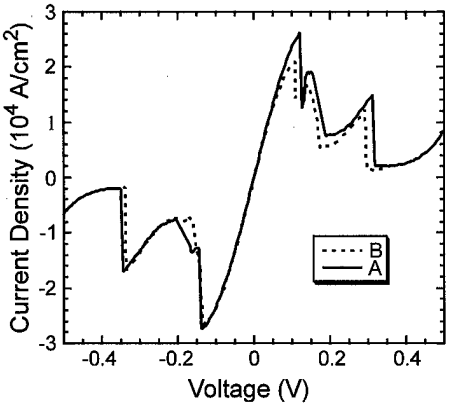


FIG. 3. Current density as a function of bias voltage acquired at room temperature for 3 μm diameter mesa RITDs grown on InAs substrates. The peak and valley currents are at approximately ±130 and ±300 mV, respectively, with positive bias corresponding to the mesa at a positive voltage with respect to the substrate. The sharp features between the peak and valley are due to bistability effects. Sample A was grown using conventional procedures, and sample B with a number of changes expected to reduce interfacial roughness.

The ratio of the peak current for positive bias to the peak current for negative bias, I_+/I_-, is an alternate characteristic of RITDs that provides a rough measure of the asymmetry in the structure along the growth direction. When the growth does not produce a structure with a mirror plane in the center of the GaSb well, $I_+/I_-\neq1$. Because different methods were used to prepare all the interfaces within the A and B samples, it is reasonable to expect that I_+/I_- will vary. The average values and standard deviations of I_+/I_- measured for devices at many locations on the samples, with mesa diameters ranging from 2 to 20 μm, are the following: InAs–A=0.95±0.06, InAs–B=0.78±0.03, GaAs–A=0.84

± 0.01, and GaAs–B $= 0.65 \pm 0.01$. On both substrates the ratio is consistently smaller for the samples with specially treated interfaces. This effect is mainly associated with a reduction in the size of the peak current at positive bias (see Fig. 3). However, because the growth procedures for all four interfaces within the RITD structure were changed from A to B, the decrease in peak current density is a convolution of the effects of the changes to all the interfaces. At this time, we do not know how all these growth modifications lead to the asymmetries in current–voltage measurements that we have observed. One possibility is that the extra In used at the first interface may be changing the effective band offset at the first barrier, thereby reducing the forward tunneling current. Alternately, the additional growth interrupts may somehow change the dominant scattering mechanisms at the interfaces or the overall tunneling probabilities. The greater problem that remains is to develop methods that will result in enhanced peak current densities.

IV. CONCLUSIONS

We have used *in situ* STM to study the surfaces and interfaces that evolve during the MBE growth of AlSb on InAs(001). We have discovered that the change in surface reconstruction that occurs across the III–Sb/III–As interface creates interfacial roughness. Specifically, when forming InSb-like interfacial bonds on an InAs(001)-(2×4) surface, vacancy islands emerge because the surface In coverage is insufficient to form a complete InSb(001)-(1×3)-like surface layer. This roughness can be greatly reduced by depositing the appropriate amount of additional In to compensate for the compositional differences between the reconstructions. We have also deposited 5-ML-thick AlSb films on the flatter InSb-like interfacial surface using three different growth conditions. The AlSb film surface morphology significantly improves when the AlSb is grown at higher temperature, as expected. We find it is also possible to reduce the island density at lower substrate temperatures by growing the AlSb film layer-by-layer using MEE. Atomic resolution images of the well-ordered AlSb surfaces resulting from higher temperature growth reveal periodic defects in the surface Sb dimer rows along the $[\bar{1}10]$ directions that give the surface reconstruction an overall (4×3) symmetry.

We have begun to implement some of the growth techniques aimed at reducing interfacial roughness into RITDs. Current–voltage spectra were acquired for RITDs that were fabricated using either 1 or 1.25 ML of In during the MEE at the AlSb-on-InAs interface. In addition, 100 s growth interrupts were added at the other interfaces. In general, the devices with specially treated interfaces had slightly different electrical characteristics (slightly lower current densities at positive bias). Although we have not yet correlated changes in device performance with particular changes in interfacial structure, it is clear that such an understanding can be achieved by closely coordinating surface and interface characterization with device fabrication and testing.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research, DARPA, QUEST (a NSF Science and Technology Center for Quantized Electronic Structures, Grant No. DMR 91-20007), and the W. M. Keck Foundation.

[1] J. R. Söderström, E. R. Brown, C. D. Parker, L. J. Mahoney, J. Y. Yao, T. G. Andersson, and T. C. McGill, Appl. Phys. Lett. **58**, 275 (1991).
[2] J. S. Scott, J. P. Kaminski, S. J. Allen, D. H. Chow, M. Lui, and T. Y. Liu, Surf. Sci. **305**, 389 (1994).
[3] P. Roblin, R. C. Potter, and A. Fathimulla, J. Appl. Phys. **79**, 2502 (1996).
[4] H. Kitabayashi, T. Waho, and M. Yamamoto, Appl. Phys. Lett. **71**, 512 (1997).
[5] B. Z. Nosho, W. H. Weinberg, J. J. Zinck, B. V. Shanabrook, B. R. Bennett, and L. J. Whitman, J. Vac. Sci. Technol. B **16**, 2381 (1998).
[6] L. J. Whitman, P. M. Thibado, F. Linker, and J. Patrin, J. Vac. Sci. Technol. B **14**, 1870 (1996).
[7] Q. Xue, T. Hashizume, and T. Sakurai, Prog. Surf. Sci. **56**, 1 (1998).
[8] B. Brar, J. Ibbetson, H. Kroemer, and J. H. English, Appl. Phys. Lett. **64**, 3392 (1994).
[9] J. Schmitz, J. Wagner, F. Fuchs, N. Herres, P. Koidl, and J. D. Ralston, J. Cryst. Growth **150**, 858 (1994).
[10] B. R. Bennett, B. V. Shanabrook, and E. R. Glaser, Appl. Phys. Lett. **65**, 598 (1994).
[11] B. R. Bennett, B. V. Shanabrook, E. R. Glaser, and R. J. Wagner, Mater. Res. Soc. Symp. Proc. **340**, 253 (1994).
[12] M. T. Sieger, T. Miller, and T.-C. Chiang, Phys. Rev. B **52**, 8256 (1995).
[13] U. Resch-Esser, N. Esser, B. Brar, and H. Kroemer, Phys. Rev. B **55**, 15401 (1997).
[14] P. M. Thibado, B. R. Bennett, B. V. Shanabrook, and L. J. Whitman, J. Cryst. Growth **175/176**, 317 (1997).
[15] B. Z. Nosho, W. H. Weinberg, W. Barvosa-Carter, B. R. Bennett, B. V. Shanabrook, and L. J. Whitman, Appl. Phys. Lett. **74**, 1704 (1999).
[16] D. H. Chow, H. L. Dunlap, W. Williamson III, S. Enquist, B. K. Gilbert, S. Subramaniam, P.-M. Lei, and G. H. Bernstein, IEEE Electron Device Lett. **17**, 69 (1996).
[17] J. N. Schulman, H. J. D. L. Santon, and D. H. Chow, IEEE Electron Device Lett. **16**, 220 (1996).

Analysis of buried (Al,Ga)As interfaces after molecular-beam epitaxy overgrowth

M. Wassermeier,[a] R. Hey, M. Höricke, E. Wiebicke, and H. Kostial
Paul-Drude-Institut für Festkörperelektronik, Hausvogteiplatz 5–7, D-10117 Berlin, Germany

(Received 19 January 1999; accepted 5 May 1999)

The buried interface (IF) formed by molecular-beam epitaxy overgrowth on differently prepared (Al,Ga)As templates has been analyzed by secondary ion mass spectrometry and capacitance/voltage profiling. For the regrowth of GaAs on GaAs the lowest IF contamination is achieved by a H-assisted oxide desorption at 450 °C. This process retains atomically flat surfaces and is highly efficient in reducing the carbon contamination and the interfacial carrier depletion compared to thermal oxide desorption at 580 °C. For the regrowth on $Al_xGa_{1-x}As$ templates, due to the high reactivity of Al, the carbon and oxygen contaminations increase with increasing Al-mole fraction. An efficient reduction of the IF contamination can be accomplished by reevaporating a sacrificial GaAs cap layer. Carrier deficits as low as 7×10^{11} cm^{-2} for the H-assisted oxide-desorbed GaAs homointerface and 7.5×10^{11} cm^{-2} for $Al_xGa_{1-x}As$ on $Al_{0.5}Ga_{0.5}As$ after reevaporating the sacrificial GaAs cap could be achieved. The carrier deficit can be compensated by a tailored doping profile around the IF. © *1999 American Vacuum Society.* [S0734-211X(99)05304-4]

I. INTRODUCTION

Epitaxy on patterned structures greatly enhances the flexibility in designing novel devices. An example are semiconductor lasers with a distributed feedback (DFB). These structures contain a periodic modulation of refractive index and/or gain of the material close to the optically active region. This may be realized by patterning a partial laser structure with a grating and completing the laser by a second epitaxial growth. Crucial for the operation of such a device is the cleanness of the buried interface (IF). Any residual IF contamination forms defect states that may act as nonradiative recombination centers and traps for electrons and holes resulting in a reduction of the performance of the device.

We have investigated overgrowth by molecular-beam epitaxy (MBE) on patterned templates in the $Al_xGa_{1-x}As$ system. This materials system is chosen because epitaxial growth and chemical processing are well known, making this a potential candidate for industrial production. For the fabrication of DFB lasers, however, it is a challenging system for the following reasons. The wavelength regime that may be accessed by this materials combination is below 870 nm, which for first order gratings, results in periods around 120 nm. These small periods can only be fabricated by *e*-beam lithography[1] or focused ion beam (FIB) implantation.[2,3] Especially the former technique requires special cleaning procedures due to the use of poly(methylmethacrylate) resist. Another difficulty arises due to the high reactivity of the Al-containing layers. Exposed to air extremely stable oxides form on the surface that cannot be desorbed thermally.

To cope with these difficulties we have investigated two strategies, namely H-assisted oxide desorption[4–7] and reevaporating a protective GaAs cap layer.[8,9] The former process results in a lower IF contamination compared to thermal oxide desorption on GaAs.[10] The latter process prevents the exposure of the Al-containing layers to air during processing and is fully compatible to patterning by FIB implantation.

II. EXPERIMENTAL DETAILS

Complete and partial laser structures as well as various test structures for IF characterization have been grown by solid source MBE. The silicon doping level in the test structures for capacitance/voltage (C/V) profiling is 2×10^{17} cm^{-3}.

H-assisted oxide desorption is carried out in a separate preparation chamber attached to the MBE system by exposing the surface at 450 °C to thermally activated atomic hydrogen. This process is monitored by reflection high-energy electron diffraction (RHEED) and stopped after 20–40 min, when a (2×4) or (1×1) reconstruction, depending on the Al-mole fraction, is observed. For Al-containing layers that have been exposed to air, an ozone cleaning step combined with a dip into H_2SO_4 prior to oxide desorption is found to further reduce the IF contamination.

To prevent the exposure of Al-containing layers to air, a protective GaAs cap layer is grown prior to the surface preparation. This cap layer is reevaporated after the H-assisted oxide desorption before the second epitaxy. It is performed at 680 °C for 20 min in the growth chamber and is monitored with RHEED along the [110] azimuth. As_4 is provided by a valved source, such that the RHEED pattern remains stable, i.e., the As pressure is raised as soon as the appearance of quarter RHEED streaks indicate a Ga termination of the surface. The Al-containing template serves as reevaporation stop. The reevaporation rate at these conditions is 10 nm/min. The reevaporation experiments included a variation of cap layer thickness, Al-mole fraction in the reevaporation stop as well as in the overgrown material, and the annealing time after reevaporation.

The regrowth starts with the deposition of three monolayers (ML) GaAs at 550 °C with 2 min growth interruption

[a]Electronic mail: matwas@pdi-berlin.de

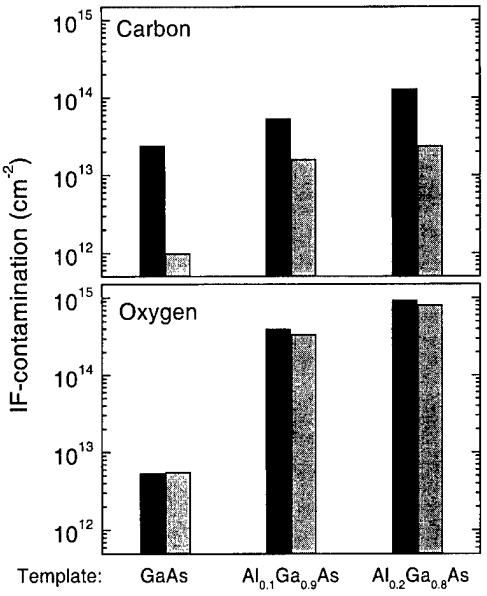


FIG. 1. Carbon and oxygen IF contamination determined by SIMS for the IF between regrown GaAs on GaAs (left two columns), on $Al_{0.1}Ga_{0.9}As$ (middle two columns), and on $Al_{0.2}Ga_{0.8}As$ (right two columns). For the homointerface, thermal oxide desorption (dark gray) is compared with H-assisted oxide desorption (light gray). For the heterointerfaces, an additional ozone treatment prior to the H-assisted oxide desorption (light gray) reduces the carbon IF contamination compared to plain H-assisted oxide desorption (dark gray).

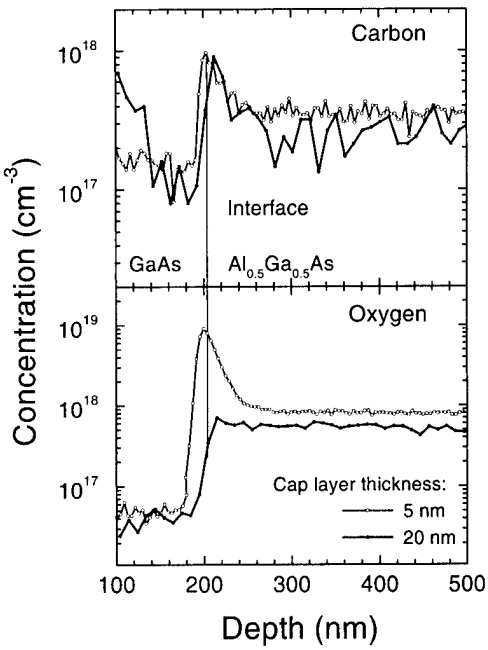


FIG. 2. Carbon and oxygen SIMS traces for the buried IF between regrown GaAs on $Al_{0.5}Ga_{0.5}As$. The $Al_{0.5}Ga_{0.5}As$ template was protected with a sacrificial GaAs cap layer that was reevaporated prior to regrowth. The higher oxygen contamination for the thinner cap indicates oxygen diffusion through the cap.

between each ML resulting in a (2×4) RHEED pattern. From this point the substrate temperature may be tuned to the appropriate value for the completion of the desired structure.

The IF contamination by carbon and oxygen was measured by secondary ion mass spectrometry (SIMS). The IF carrier depletion was characterized by *CV* profiling at 100 kHz. The surface topography after overgrowth was imaged by atomic force microscopy.

III. RESULTS AND DISCUSSION

A. H-assisted oxide desorption

For GaAs surfaces exposed to air, the H-assisted oxide desorption offers the following advantages with respect to thermal oxide desorption. On the one hand, it is a low-temperature process that preserves the initial surface topography, an important issue for patterned surfaces. On the other hand, it results in a reduction of the IF contamination. The oxygen and carbon IF contamination determined by SIMS are compared in Fig. 1. The H-assisted oxide desorption results in a reduction of the carbon contamination at the GaAs homointerface of more than an order of magnitude to a value below 1×10^{12} cm^{-2}.

The oxide that forms on Al-containing layers exposed to air is extremely stable. Due to the high reactivity of the Al, the residual carbon and oxygen IF contaminations increase with increasing Al-mole fraction even after H-assisted oxide desorption as shown in Fig. 1 for $Al_{0.1}Ga_{0.9}As$ and $Al_{0.2}Ga_{0.8}As$. An ozone treatment prior to the H-assisted oxide desorption reduces the carbon IF contamination; the oxygen contamination is, however, not affected. Although smooth overgrowth on Al-containing layers to a mirror-like finish and a good structural perfection judged by RHEED is obtained by this process, such high IF-contamination levels are detrimental for most device applications. The measured carrier deficit for all but one of the buried IFs lies above the value of 2×10^{12} cm^{-2}, the upper limit that may be determined with the employed bulk doping of 2×10^{17} cm^{-3}. Only the H-assisted oxide desorption on GaAs results in a lower carrier depletion of 0.7×10^{12} cm^{-2}.

B. Capping and reevaporation

In order to reduce the IF contamination in Al-containing layers to acceptable values, the exposure of Al-containing layers to air has to be avoided. This may be achieved by protecting the Al-containing layers with a sacrificial GaAs cap that is evaporated prior to the regrowth. This strategy is fully compatible to patterning by FIB implantation.

In the following we show that the IF contamination depends on the cap layer thickness, the Al-mole fraction of the template, and the overgrown material and on the time the surface has been exposed to air. The dependence of the IF contamination on the cap layer thickness is shown in the SIMS traces in Fig. 2 for the growth of GaAs on $Al_{0.5}Ga_{0.5}As$. The carbon contamination (1×10^{12} cm^{-2}) is more than an order of magnitude below the value obtained for the growth of GaAs on air-exposed $Al_{0.1}Ga_{0.9}As$ and independent on cap layer thickness. The oxygen contamination, reduced by several orders of magnitude with respect to

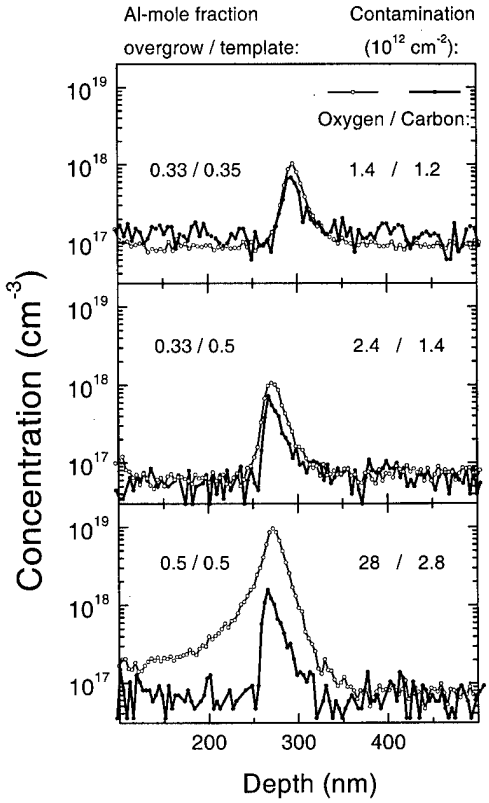


FIG. 3. Oxygen contamination for heterointerfaces formed by reevaporation of a sacrificial GaAs cap and regrowth of $Al_xGa_{1-x}As$ on $Al_xGa_{1-x}As$. The IF contamination depends on the Al-mole fraction x in the template (upper two pair of curves) and in the regrown material (lower two pair of curves).

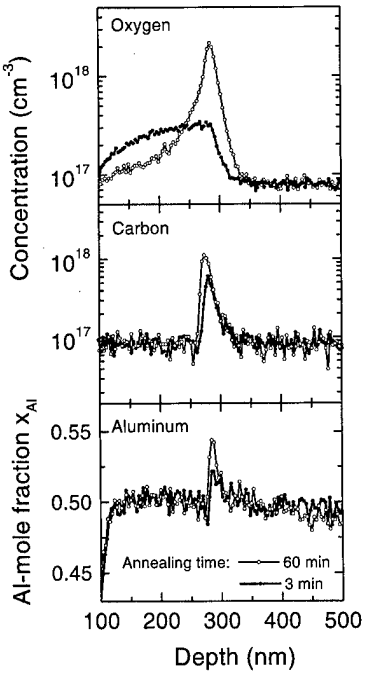


FIG. 4. Oxygen, carbon, and aluminum SIMS traces across the heterointerface formed by GaAs reevaporating, annealing, and regrowth of $Al_{0.5}Ga_{0.5}As$ on $Al_{0.5}Ga_{0.5}As$ for two annealing times of 3 and 60 min. The background level of the SIMS detection system is 10^{17} cm^{-2} for carbon and oxygen.

air-exposed layers, depends on the cap layer thickness and can be reduced to values below the detection limit for cap layer thicknesses above 20 nm. This behavior indicates a diffusion of oxygen through the GaAs cap layer during the residence time of the sample in air.

Another important parameter that affects the IF contamination is the Al-mole fraction in the template. On the one hand, it has to be high enough such that the template functions as reevaporation stop layer. An Al-mole fraction of only $x=0.2$ was found to result in a roughening of the IF due to spatially inhomogeneous Ga evaporation out of the $Al_{0.2}Ga_{0.8}As$ template. An Al-mole fraction above $x=0.35$, however, is sufficient to stop the reevaporation. On the other hand, we find that the oxygen IF contamination increases with increasing Al-mole fraction in the template. This behavior, that was already found for the air-exposed layers earlier, is shown in Fig. 3 for $Al_{0.33}Ga_{0.67}As$ grown on two different templates (the carbon contamination is not affected). Interestingly, the oxygen IF contamination also increases with increasing Al-mole fraction in the overgrown material as shown in Fig. 3. This may be explained by assuming a release of oxygen from the sample holder during the high-temperature reevaporation. This oxygen is gattered at the IF, a process that depends on both the Al-mole fraction of the template and the regrown material. The residual background pressure in the MBE chamber could be ruled out as a source for this contamination because a growth interruption of 1 h within an $Al_{0.5}Ga_{0.5}As$ layer did not show any oxygen accumulation. Further confirmation of the earlier explanation may be drawn from the dependence of the IF contamination on the annealing time after reevaporation. Annealing is essential for FIB implanted structures in order to heal out the structural defects introduced during FIB implantation. In Fig. 4 the carbon, oxygen, and aluminum traces of buried $Al_{0.5}Ga_{0.5}As$-on-$Al_{0.5}Ga_{0.5}As$ interfaces prepared with 3 and 60 min annealing times are compared. The oxygen and carbon contaminations at the interface are reduced for the shorter reevaporation time. The large amount of oxygen in the overgrown $Al_{0.5}Ga_{0.5}As$ layer supports the notion of a release of oxygen from the sample holder being incorporated either at the interface for the long annealing time or in the overgrown material for the short annealing time. Also interesting is the aluminum SIMS trace, that clearly shows an Al accumulation due to reevaporation of Ga from the $Al_{0.5}Ga_{0.5}As$ template.

The IF contaminations for the optimized structures, both GaAs homointerfaces and $Al_xGa_{1-x}As$ heterointerfaces lie in the low 10^{12} cm^{-2} range. At such low values the associated IF carrier depletion may be determined by CV profiling. For applications, in which a constant doping level is required, the residual carrier depletion may further be reduced by including a doping profile in order to compensate for the depletion. This is shown for the $Al_{0.33}Ga_{0.67}As$-on-$Al_{0.5}Ga_{0.5}As$ IF in Fig. 5. Template and regrown material are Si doped with 2×10^{17} cm^{-3}. The carrier depletion without additional doping (solid line+open symbols) is 1.4×10^{12} cm^{-2}. Including additional δ-doped layers to the template

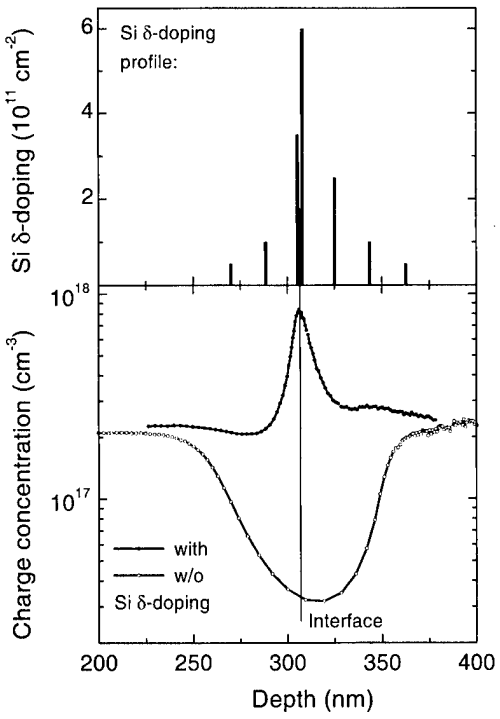


FIG. 5. *CV* profiles across the heterointerface formed by GaAs reevaporating and regrowth of $Al_{0.33}Ga_{0.67}As$ on $Al_{0.5}Ga_{0.5}As$. The residual carrier depletion of 7.5×10^{11} cm^{-2} is compensated by Si doping. The doping profile consists of seven δ-doped layers shown in the upper part of the figure.

and the regrown material as shown in the upper part of Fig. 5 results in a *CV* profile shown as solid line with filled symbols. With the known integrated additional Si doping of 1.5×10^{12} cm^{-2} and the determined carrier accumulation of 7.5×10^{11} cm^{-2} the carrier depletion due to IF contamination results to 7.5×10^{11} cm^{-2}. This lower value compared to the undoped template may be due to a different incorporation of defects on doped versus undoped surfaces. Another possibility could be the longer storage time at air of the undoped sample (34 days) compared to the doped sample (8 days).

An oxygen and carbon contamination of 2×10^{12} cm^{-2} seems to be the lower limit that may be achieved for the regrowth of Al-containing layers on an Al-containing template using the reevaporation strategy described earlier.

IV. CONCLUSIONS

The buried IF formed by MBE regrowth on differently prepared templates has been studied by SIMS and *CV* profiling. For the regrown IF on air-exposed surfaces we find that H-assisted oxide desorption considerably reduces the IF contamination compared to conventional thermal oxide desorption. For air-exposed Al-containing layers the IF contamination depends on the Al-mole fraction. An additional ozone treatment prior to the H-assisted oxide desorption allows to obtain smooth MBE overgrowth. Protecting Al-containing templates with a sacrificial GaAs cap that is reevaporated before overgrowth is a successful strategy to reduce the IF contamination in Al-containing layers. The GaAs cap thickness has to be at least 20 nm and the duration of exposure to air should be minimized to a few days. The Al-mole fraction in the template should be above $x=0.35$ to serve as a reevaporation stop layer. An accumulation of Al at the IF indicates an evaporation of Ga from the template for long annealing times after reevaporation. The obtained IF contaminations for an optimized procedure lie in the low 10^{12} cm^{-2} regime, which is an exceptionally low value for Al-containing layers comparable to values obtained for the GaAs homointerface. The associated carrier depletion may be compensated by δ doping. Also for the reevaporation, the IF contamination increases with increasing Al-mole fraction in the template and the overgrown material. This behavior indicates that the IF contamination may further be reduced by optimizing the initial processing steps after reevaporation.

ACKNOWLEDGMENT

The authors would like to acknowledge financial support by the European Community within the ESPRIT NANOLASE Project No. 22497.

[1]K. H. Wang, A. Pecher, E. Höfling, and A. Forchel, J. Vac. Sci. Technol. B **15**, 2829 (1997).
[2]H. König, J. P. Reithmaier, A. Forchel, J. L. Gentner, and L. Goldstein, Appl. Phys. Lett. **73**, 2703 (1998).
[3]H. König, N. Mais, E. Höfling, J. P. Reithmaier, A. Forchel, H. Müssig, and H. Brugger, J. Vac. Sci. Technol. B **16**, 2562 (1998).
[4]T. Sugaya and M. Kawabe, Jpn. J. Appl. Phys., Part 2 **30**, L402 (1991).
[5]Z. Lu, M. T. Schmidt, and R. M. Osgood, J. Vac. Sci. Technol. A **9**, 1040 (1991).
[6]K. D. Choquette, M. Hong, S. N. G. Chu, H. S. Luftman, J. P. Mannaerts, R. C. Wetzel, and R. S. Freund, Appl. Phys. Lett. **62**, 735 (1993).
[7]T. Niwa, Furukata, and T. Maeda, J. Cryst. Growth **175/176**, 441 (1997).
[8]T. Kojima, N. J. Kawai, T. Nakagawa, K. Ohta, T. Sakamoto, and M. Kawashima, Appl. Phys. Lett. **47**, 286 (1985).
[9]K. Ohta, Surf. Sci. **298**, 415 (1993).
[10]R. Hey, M. Wassermeier, M. Höricke, E. Wiebicke, and H. Kostial, J. Cryst. Growth (to be published).

Mechanistic studies of silicon oxidation

M. K. Weldon,[a] K. T. Queeney, Y. J. Chabal, B. B. Stefanov, and K. Raghavachari
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

(Received 17 January 1999; accepted 30 April 1999)

The microscopic mechanism of the formation of ultrathin oxides on Si(100) has been investigated using a combination of infrared spectroscopy and *ab initio* quantum chemical cluster calculations. The 0→2 monolayer oxide films are grown sequentially from the "bottom-up" using repeated water exposures and annealing cycles, with the partial pressure of water ranging from 10^{-10} to 10 Torr. The resultant films were then compared to the equivalent thicknesses of thermal and native oxide films. In this way, we obtain unprecedented insight into the essential chemical structures formed during the initial oxidation and subsequent layer growth of these technologically relevant films. © *1999 American Vacuum Society.* [S0734-211X(99)02804-8]

I. INTRODUCTION

Gate oxide thicknesses of ~20 Å are required for future generations of metal–oxide–semiconductor devices, in order to meet the scaling requirements of Moore's Law. In this ultrathin film limit, the interfacial layer comprises at least 15% of the total gate oxide thickness, so that understanding and controlling the interfacial composition is clearly of paramount importance. Despite the critical need for detailed characterization of the microscopic structure of the Si/SiO_2 interface, remarkably little is known even after decades of intensive scientific effort. This dearth of information is due in large part to the absence of a definitive spectroscopic probe with the requisite sensitivity. Infrared (IR) spectroscopy is almost uniquely suited in this regard as it is highly sensitive (down to fractions of a monolayer of oxide), selective (able to distinguish different Si–O bonding environments), nondestructive, and can be used under ambient conditions as a process control tool. There are, however, a number of factors that have prevented widespread utilization of IR spectroscopy for this purpose: first, commercially available instruments have only recently reached the point where a simple one-pass external transmission (ET) geometry affords (sub)monolayer sensitivity with a few minutes of scan time. This advance is significant since the ET geometry allows access to the entire 500–4000 cm^{-1} spectral range of interest, which is not the case for the multiple internal reflection (MIR) geometry that has typically been utilized for monolayer studies.[1,2] Specifically, the MIR geometry cannot be used to probe the Si–O phonon modes, due to the pronounced Si multiphonon absorption that arises when using the long optical pathlengths intrinsic to the MIR approach, and which effectively prohibits study of vibrations occurring below 1300 cm^{-1}.

The second barrier to routine IR spectroscopic analysis of ultrathin oxides is the fact that the Si–O phonon bands are broad and relatively featureless, due to the large variation in Si–O–Si bond angles that naturally occurs in SiO_2. This makes the interpretation of spectral changes fraught with ambiguity, without recourse either to *ab initio* methods of analysis or to detailed modeling using the known optical constants of silica and other pertinent species that may be present. Recently, the first reports of *ab initio* calculations of a silica slab using gradient corrected density functional theory have emerged.[3,4] The resultant theoretical spectra were found to be in excellent agreement with experimental data typical of thicker films, although much work now remains to be done to understand the origin of the unique spectral changes observed in the thin film limit. Alternative approaches to the analysis of the IR spectra of thin oxide films based on the effective medium theory have also been employed in recent years.[5,6] However, the absence of comprehensive data for the **n** and **k** values for different films of known stoichiometry or composition similarly limits the current utility of this approach.

Importantly, neither optical modeling nor the *ab initio* slab calculations described earlier can reveal the microscopic details of the *mechanism* of the oxide film growth since they rely solely on analysis of the final film composition. Therefore, we have focused on the growth of ultrathin oxide films from the dilute limit starting from submonolayer coverages of H_2O on Si(100)–(2×1), through completion of the first monolayer, to development of a true three-dimensional thin film.

Water exposure of clean Si(100)–(2×1) represents a particularly well-defined initial system to study silicon oxidation. After an early suggestion that molecular adsorption predominated at a low temperature,[7] electron energy-loss spectra[8] and IR[1,9] studies conclusively showed that H_2O principally undergoes dissociative adsorption into H and OH (with unit probability), a finding that was later confirmed by scanning tunneling microscopy (STM) studies.[10,11] A minor amount of molecular water physisorbs at defects,[11] but these molecules desorb upon annealing and therefore do not participate in the oxidation process.

The structure of the dissociated water is also well-characterized. Using single domain Si(100) surfaces, electron stimulated ion angular desorption measurements showed enhanced desorption perpendicular to the dimer axis.[12,13] This asymmetry was interpreted as evidence for a tilt of the OH bond out of the Si–Si dimer plane, since H^+ desorption

[a]Electronic mail: marcus@bell-labs.com

comes primarily from OH rather than SiH, suggesting that the OH species may be stabilized at low coverages by the interaction between the O lone pair electrons and the Si dangling bond on the neighboring dimer. Recently, however, first principles calculations have shown that the lowest energy state was indeed that with OH perpendicular to the dimer axis, even without dimer–dimer interaction.[9,14,15] These latter studies were a clear demonstration of the potential power of a combined experimental and theoretical approach, since the *ab initio* quantum cluster calculations provided a detailed understanding of the IR spectrum of the water-exposed Si(100)–(2×1) surface.

Recently, similar combined IR studies and *ab initio* calculations have provided evidence for inter-dimer coupling. Advances in both Fourier-transform infrared (FTIR) spectrometers and the increase in processing power have made it possible to resolve the fine structure of the water-exposed Si(100)–(2×1) surface and definitively ascribe it to (OH-mediated) coupling between neighboring dimers.[16] The ability to now probe the extended coupling among dimer units suggested by this work is very exciting and is a central theme of this article.

In this work, we have again employed a combined experimental (IR spectroscopy) and theoretical approach (*ab initio* cluster calculations), to delineate the myriad discrete and coupled oxide structures that form during film growth. These results are then compared with those obtained for ultrathin thermal and native (chemical) oxides in order to make contact with the actual films used in silicon technology. Specifically, we are able to show that oxidation occurs via insertion of O into the dimer bond of the reconstructed (2×1) Si(100) surface, which facilitates subsequent insertion into the Si backbonds. As the temperature is raised to ~500 °C, rapid agglomeration of oxygen occurs to form multiply oxidized, 3–5 O containing hydrogenated dimers. Upon annealing to 600 °C, dehydrogenation of these structures occurs, resulting in the formation of novel epoxide linkages comprised of SiOSi triangular rings. The epoxide is then the stable unit structure up to the desorption temperature of these films. Subsequent repeated water dosing and annealing cycles leads to a complete monolayer of agglomerated epoxide structures which is stable with respect to further water exposures up to partial pressures of several Torr. As the pressure is brought up to 1 atm, a rapid transformation of the film occurs via distinct oxide structures which are intermediates to the formation of the corner-sharing SiO_2 tetrahedra that comprise the bulk-like stoichiometric oxide film. The final film formed possesses a true transverse optical (TO) and longitudinal optical (LO) pair of modes that closely resembles that observed for a thin (4–6 Å) thermal or native oxide film. Consequently, these studies provide a detailed insight into the microscopic mechanism of silicon oxidation and suggest that the goal of decades of studies of Si oxidation may now be within reach.

II. EXPERIMENTAL METHODS

IR data were acquired under dry N_2 purge conditions using either a Nicolet 860 FTIR equipped with a broadband MCT–B detector or a Nicolet 760 FTIR with a DTGS detector. The ultrahigh vacuum (UHV) experiments were performed in a stainless steel chamber (base pressure $<2\times10^{-10}$ Torr), equipped with a double directed doser and variable-control leak valve to allow simultaneous, precise exposures of the front and back surfaces with minimal background pressure rise. A water dose of 1×10^{-9} Torr for 20 s was found to be sufficient to produce the saturated (0.5 ML of oxygen) water-covered Si(100)–(2×1) surface. UHV dosing was performed with the sample held at −50 °C (as measured by a pair of chromel/alumel thermocouples clipped to the top and bottom surfaces of the sample). The high-pressure (10^{-6}–10 Torr) water exposures were performed by isolating the pumping stages then back-filling the chamber with water vapor through the variable control leak valve.

The sample was heated using a four-point electrical contact fed through a feedback circuit that allowed heating over the −50–1000 °C temperature range with a precision of ±0.2 °C. IR spectra were typically acquired after the sample temperature had stabilized at −50 ±0.2 °C, for the initial low-pressure water exposures, or at ambient temperature (~23–25 °C) for doses in the 10^{-6}–10 Torr pressure range. The epoxide-covered surface that forms at elevated temperatures was produced using a multiple stage annealing protocol, wherein the sample temperature was incrementally raised to 600 °C, pausing for 2 min at several intermediate temperatures (i.e., 375, 450, and 500 °C). This protocol was found to facilitate oxygen agglomeration rather than desorption of gaseous Si–O. The sample spectra were referenced to the clean surface spectrum (recorded immediately prior to water exposure) for the former dose regime or, for the latter high-pressure doses, to a Si(100)–(2×1):H terminated reference that had previously been acquired with the sample at ambient temperature. The low temperature data acquisition and precise temperature control minimized the contribution of Si substrate phonons to the observed spectra.

The Si(100) sample (cut from a low resistivity, 2–10 Ω cm, double-side polished float zone wafer) was cleaned using a modified RCA clean prior to mounting in vacuum. The system was baked at 160 °C for 48 h, then the sample flashed to 1000 °C for 1 min to remove the protective native (chemical) oxide formed during the cleaning step. This protocol has been shown to minimize recontamination of the sample during pump-down and bake-out of the system.

The thermal oxide studied was grown at 800 °C using O_2 in a conventional oxidation furnace. The initial 31 Å film (thickness determined by ellipsometry) was given a brief (1 min) dip in a 1:1:4 SC-1 solution at 80 °C in order to remove surface contaminants, then individual pieces were thinned using a dilute HF (0.05% HF) solution. The etch time was varied for each piece depending on the desired final thickness. This etching procedure has been shown to uniformly thin the film, without significant surface roughening, so that it faithfully reproduces the true composition of the film at each thickness.[5]

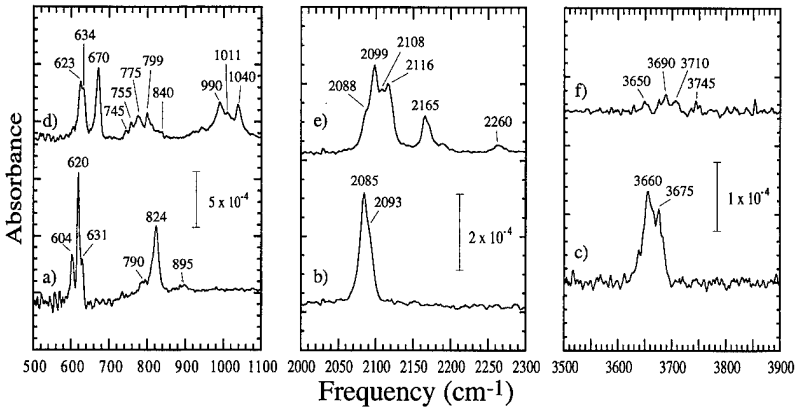


FIG. 1. IR spectra of the H_2O-exposed Si(100)–(2 $\times$1) surface; (a)–(c) −50 °C and, (d)–(f) after annealing to 400 °C.

III. THEORETICAL METHODS

As in previous theoretical studies of Si surfaces,[9,15,17,18] a cluster containing nine silicon atoms was used to represent the local structure of a single dimer on the reconstructed Si(100) surface. This cluster can be viewed as representing a unit cell of the surface and the three underlying layers of the dimerized surface, where the dangling bonds resulting from truncation of Si–Si bonds are terminated by H atoms in order to avoid artifacts due to excess spin or charge. For interactions involving two adjacent dimers in the same dimer row (representing oxygen agglomerated species, *vide infra*), a cluster containing 15 silicon atoms was used. Such a cluster has been used previously to represent hydrogen bonding interactions between neighboring dimers.[16] The largest cluster considered in this work, representing two neighboring hydrogenated oxygen-agglomerated dimers, was $Si_{15}O_{10}H_{20}$.

Additional boundary constraints are then imposed on the cluster models to avoid unphysical relaxations that are not possible on the real surface. As in previous work, third and fourth layer atoms in the extended solid are not allowed to relax but are held fixed at ideal diamond lattice-like positions. In contrast, the atoms of the first and second layers and the adsorbate atoms are allowed to relax fully without any constraints.

The geometry that minimizes the total energy, subject to the aforementioned constraints, is then obtained for each cluster via gradient corrected density functional calculations using the B3LYP method.[19] This combines the gradient corrected three-parameter exchange functional due to Becke[20] with the Lee–Yang–Parr correlation functional.[21] We use the polarized 6–31 G** basis set on the first and second layer silicon atoms as well as for all the adsorbed hydrogen and oxygen atoms. The third and fourth layer silicons and the hydrogens representing truncated Si–Si bonds are treated with the double-zeta 6–31 G basis set. The resulting overall basis set is denoted 6–31 G(**). This method is particularly accurate in reproducing frequency shifts due to substituent effects in related chemical species. In order to obtain absolute frequency values, the calculated frequencies are calibrated within each spectral region by simple scaling, as described elsewhere.[6,17]

IV. RESULTS

A. Initial oxidation

The spectrum of the initial water-exposed Si(100)–(2 $\times$1) surface is shown in Figs. 1(a)–1(c). As discussed earlier, this spectrum has previously been well-characterized and shown to consist of a mixture of isolated (uncoupled to any other dimer unit) and coupled dimers, wherein the coupling is mediated by hydroxyl groups on the same end of neighboring dimers within a dimer row.[16] This determination was facilitated by the extended spectral range that can now be probed with high sensitivity using the simple transmission geometry (400–4000 cm^{-1}); the observation of multiple Si–H bending modes at 604, 620, and 631 cm^{-1}, in combination with the Si–O modes at 790 and 824 cm^{-1} [Fig. 1(a)] provides clear evidence for the coupled normal modes that result from this inter-dimer interaction.[16] The presence of two distinct O–H stretching modes, one down-shifted by $\sim$15 cm^{-1} with respect to the other [Fig. 1(c)] is also indicative of a hydrogen bonding interaction (the "donor" O–H bond of a H-bonded pair is known to exhibit a lower stretching vibration).[22] Importantly, all the preceding observations are in excellent agreement with the results of *ab initio* cluster calculations of the relevant energetics and vibrational frequencies.[16] The observation of a limited-range inter-dimer coupling is also consistent with previous STM studies of this system, which showed short chains of OH groups aligned along the same side of a dimer row.[10]

Profound changes are apparent in the IR spectrum obtained after annealing this surface to 400 °C [Figs. 1(d)–1(f)], indicating that decomposition of the surface hydroxyl groups has occurred and the constituent O and H atoms have incorporated into the Si surface. The O–H stretching mode at 3660 cm^{-1} is significantly (90%) diminished [Fig. 1(f)], and five distinct Si–H stretching bands are apparent at 2088, 2099, 2108, 2116, and 2165 cm^{-1} [Fig. 1(e)]. Notably, the total integrated Si–H absorbance has also increased by a factor of 2.0$\pm$0.3 relative to that at 20 °C. These data strongly suggest that large-scale decomposition of the surface hydroxyl groups has occurred after annealing to this temperature, and that the constituent hydrogen atoms are in-

corporated as surface hydride species. The modes at 2088 and 2099 cm^{-1} have previously been assigned[17] to the asymmetric and stretching modes of the coupled monohydride species, H–Si–Si–H, based on the observed polarization and the agreement with the literature values for the pure hydrogen-terminated Si(100)–(2×1) surface.[2] The 2108 and 2116 cm^{-1} features were then assigned to the analogous modes of a coupled monohydride species with oxygen inserted into the dimer bond, i.e., H–Si–O–Si–H.[17] The observation of the (weak) feature at 2165 cm^{-1} following the loss of the surface hydroxyl species results from the next insertion product, namely H–Si–O–Si(O)–H. However, the real evidence for O insertion is obtained by analysis of the low frequency region. Specifically, by comparison of the experimental frequencies and polarizations with those derived from *ab initio* cluster calculations of oxygen-containing dimers, the three new features observed at 990, 1011, and 1040 cm^{-1} [Fig. 1(d)] can be assigned[17] to the asymmetric Si–O–Si stretching vibrations of (i) a single oxygen inserted into the surface dimer bond and (ii) two oxygen atoms incorporated into a single dimer, one inserted into the dimer and the other into a silicon backbond.

Consideration of the calculated energetics reveals that insertion of oxygen into the dimer bond is favored relative to backbond insertion[23] (due to the strained nature of the long Si–Si dimer bond) and gives rise to an Si–O mode at 990 cm^{-1}, and associated Si–H bending and stretching modes at 670/740 and 2111/2115 cm^{-1},[24] in excellent agreement with the experimentally observed modes at 670, 745, 2108, and 2116 cm^{-1} [Figs. 1(d) and 1(e)]. Furthermore, calculations also show that the insertion of a second oxygen into these Si backbonds is greatly facilitated by the former reaction (due to the "flexibility" that the first Si–O–Si bond introduces into the structure), with the resultant structure giving rise to the modes at 1011, 1040 [symmetric/antisymmetric Si–O–Si(O) modes], and 2165 cm^{-1} [Si–H stretch of the O–Si(O)–H subunit][17] as well as bending modes in the 740–840 cm^{-1} frequency range.[24] Once again, this is precisely in accordance with our experimental observations. In fact, the backbond insertion becomes sufficiently facile after dimer insertion that coalescence of two H–Si–O–Si–H structures to form one H–Si–O–Si(O)–H and one H–Si–Si–H structure is *thermoneutral*.[23] Consequently, an equilibrium between these three structures is expected, just as observed experimentally.

Notably, calculations of the energetics of higher order oxidation products reveal that there is a strong tendency for O to subsequently agglomerate on the Si(100)–2×1 surface,[23] with the 3- and 5-oxygen agglomerated structures predicted to be particularly stable. However, the question arises as to how these findings are modified upon desorption of the passivating hydrogen atoms (which is known to commence above 500 °C).

We have thus investigated the high-temperature annealing behavior of the H_2O:Si(100)–2×1 system.[18] Briefly, we find that lateral agglomeration of the surface oxygen continues to occur, producing 5-oxygen containing hydrogenated

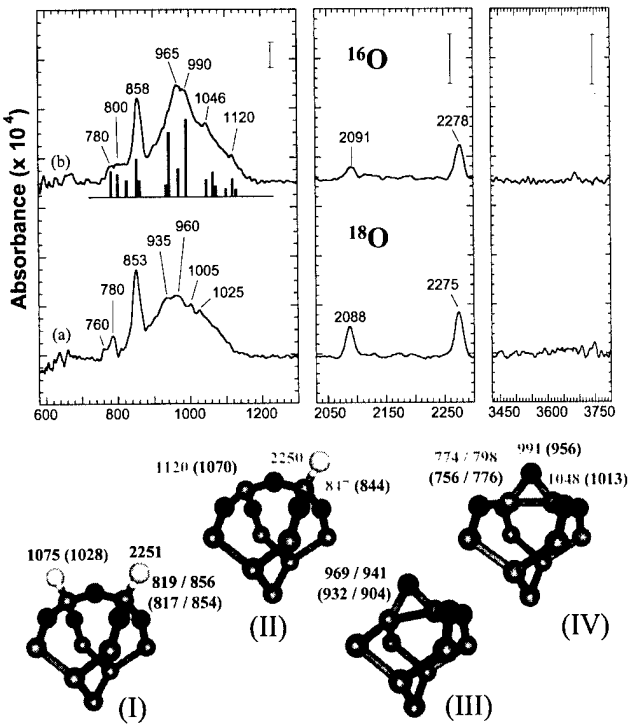


FIG. 2. IR spectra of the Si(100)–(2×1) surface exposed to water at −50 °C and annealed to 600 °C. Data obtained using; (a) $H_2{}^{18}O$ and (b) $H_2{}^{16}O$. Lower panel: Cluster structures to which the observed modes are assigned together with calculated frequencies for ^{16}O (^{18}O frequencies in parentheses) (I) Hydrogenated 5 oxygen-containing dimer, H–(O_2)Si–O–Si(O_2)–H; (II) Partially dehydrogenated 5 oxygen dimer structure, *(O_2)Si–O–Si(O_2)–H; (III) 3 oxygen epoxide; (IV) 5 oxygen epoxide. In all cases Si atoms are depicted as gray spheres, O atoms are dark gray and H atoms are the small, light gray spheres.

dimer units, denoted H–(O_2)Si–O–Si(O_2)–H and *(O_2)Si–O–Si(O_2)–H [Fig. 2, structures I and II], as evidenced by the presence of asymmetric stretching modes at 1120 cm^{-1} and significantly shifted Si–H bending and stretching modes at 858 and 2278 cm^{-1}, respectively, associated with the –O–Si(O_2)–H structural unit [Fig. 2(b)]. However, 90% of the passivating H desorbs on annealing to 600 °C, so that the pronounced growth in Si–O intensity below 1000 cm^{-1} that occurs over this temperature range clearly cannot be assigned to hydrogenated oxide structures. The (broad) Si–O stretching band exhibits discrete features at 965 and 990 cm^{-1} together with a weaker third absorption band at 1046 cm^{-1}, all of which show anomalously small isotopic shifts on ^{18}O substitution [Fig. 2(a)]: ~35 cm^{-1} compared to ~50 cm^{-1} for the hydrogenated oxide structures. Furthermore, two weak features are also observed at 780/800 cm^{-1} for $H_2{}^{16}O$ and 760/780 cm^{-1} for $H_2{}^{18}O$ (Fig. 2), clearly suggesting assignment to another Si–O mode, despite the unusually small isotopic shift. The puzzling nature of these modes is again resolved by *ab initio* density functional cluster calculations which predict that a unique "epoxide" structure, comprised of a three-membered SiOSi ring [Fig. 2, structures (III) and (IV)], is the favored product upon dehydro-genation,[18] in the event of oxygen agglomeration. Specifically, while the familiar dangling bond-containing

product is favored for 1- and 2-oxygen containing dimer units, the epoxide is favored in the limit of higher oxidation, due to the flexibility introduced by the Si–O–Si backbond linkages. Once again, the agreement between theoretically predicted and experimentally observed vibrational frequencies is remarkable: the 5-oxygen epoxide is predicted[18] to have Si–O modes at 774/798 (symmetric) and 991/1048 cm^{-1} (antisymmetric), respectively, in excellent agreement with the experimentally observed bands at 780–800 and 990/1046 cm^{-1}. The calculated isotopic shifts of 18–22 and ~35 cm^{-1} for the symmetric and antisymmetric modes, respectively, are also in good agreement with the observed ~20 and 30 cm^{-1} shifts for films grown using ^{18}O. The remaining peak at 965 cm^{-1} (935 cm^{-1} for ^{18}O) is then attributed to the 3-oxygen epoxide analog for which the calculated frequency is 969 cm^{-1} with a 37 cm^{-1} isotopic shift.

The origin of both the low Si–O asymmetric stretching frequencies and the small isotopic shifts can also be clarified: the former arise from the unusually long Si–O bonds of the SiÔSi ring (1.71 Å compared to 1.64–1.68 Å normally observed) and the latter from the substantial component of Si–Si motion associated with these modes.

The preceding discussion typifies the level of understanding that is now possible of chemical phenomena on semiconductor surfaces for relatively homogeneous overlayers at low temperatures, but the real interest must still lie in describing the subsequent processes that lead to genuine surface oxidation. In the following we will describe recent experiments that use the same combined experimental and theoretical approach to uncover the rich tapestry of reactions that occur as the surface is further oxidized.

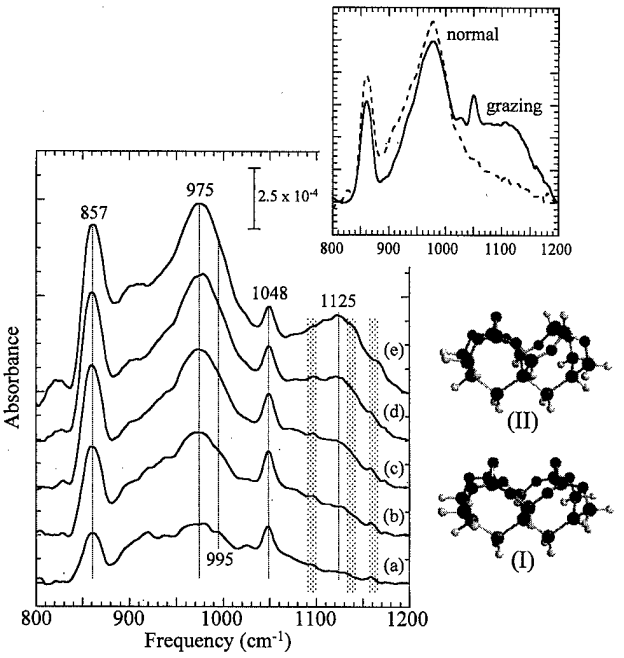


FIG. 3. IR spectra of the Si–O region acquired after each of 5 consecutive doses of H_2O followed by annealing to 600 °C each time. In each case, the same exposure of H_2O ($\sim 1\times10^{-9}$ Torr for 20 s) was used to saturate the remaining bare Si dimers. All spectra are ratioed to the initial (clean) Si(100)–(2×1) surface. The shaded regions highlight the weak additional features at ~1100, 1140, and 1165 cm^{-1}. Inset: spectra obtained with the IR beam normal, and at grazing incidence (60°), to the surface. The two structures pictured represent two different coupled species: two coupled 5-oxygen epoxides (I) and one 5-oxygen epoxide coupled to a hydrogenated agglomerate (II). Si atoms are depicted as gray spheres, O atoms are (smaller) dark gray, and H atoms are the (smallest) light gray spheres.

B. Completion of the first monolayer: Epoxide agglomeration

Repeated H_2O dosing and annealing cycles create a surface of extended, agglomerated oxygen structures. After epoxide formation via heating the first 0.5ML H_2O exposure to 625 °C, the resulting surface was redosed with H_2O and heated to the same temperature (termed 1 "cycle"). This process was repeated until no further H_2O could be dissociatively adsorbed at −50 °C, based on the absence of any of the characteristic vibrational features after prolonged H_2O exposure at $\sim 1-3\times10^{-9}$ Torr.

Figure 3 shows the IR spectra acquired after each subsequent exposure cycle. The IR spectrum obtained after the initial H_2O exposure cycle [Fig. 3(a)] shows a mixture of epoxide and hydrogenated, oxygen-agglomerated species, as discussed earlier. Growth of the 1048 cm^{-1} peak after the second cycle shows the formation of additional 5-oxygen epoxide species, although this peak does not grow in intensity after the second cycle, indicating no net growth of the *isolated* epoxide species after this step.

Three principal features are otherwise apparent in the spectra of the multiply-exposed surface: a mode at 857 cm^{-1} that can be unambiguously attributed to a Si–H deformation mode, $\delta(O_3Si–H)$, and two broad Si–O features centered at 975 and 1125 cm^{-1}, polarized parallel and perpendicular to the surface plane, respectively (Fig. 3, inset). Interestingly, this polarization of the low and high frequency vibrational bands is identical to that required for the TO and LO modes of an extended oxide network, suggesting that the 975 and 1125 cm^{-1} features may be the microscopic precursors to the genuine TO/LO mode pair of the resultant extended film.

The growth of the 857 cm^{-1} $\delta(O_3Si–H)$ peak indicates that the addition of a (small) number of hydrogenated oxide structures occurs throughout the oxidation process. The fact that the highest frequency mode associated with such a hydrogenated, 5-oxygen species is an 1120 cm^{-1} mode would, on first inspection, seem to be consistent with the concomitant growth of a feature at ~1125 cm^{-1} as the dose and anneal cycles are repeated [Figs. 3(b) and 3(c)]. However, two observations preclude such an assignment: (i) the growth of the 1125 cm^{-1} mode does not scale with that of the 857 cm^{-1} feature and (ii) the pronounced tail to the high frequency absorption that exists out to 1170 cm^{-1} cannot be attributed to *any* isolated oxidized dimer structure by reference to the calculated frequencies for all such possibilities.

Now, before proceeding to interpret the spectra obtained after multiple exposure cycles, it is important to realize that the dimer structure of the Si(100)–(2×1) surface is retained throughout. The persistence of the 1048 cm^{-1} peak in all spectra demonstrates that the initial 5-oxygen epoxides maintain their isolated dimer-based structures. Furthermore,

each subsequent H_2O dose, while decreasing in total amount due to the reduction in the number of "vacant" (dangling bond-terminated) dimers, occurs by the same single-dimer dissociation pathway as the initial H_2O dose, based on the observation of the absorption modes characteristic of the HO–Si–Si–H dimer structure (data not shown). Thus, in the consideration of possible surface species that explain the observed spectra, we are necessarily confined to dimer-based oxidized structures.

This inability of single-dimer oxidized structures to explain the spectrum observed in the high oxygen coverage limit [Fig. 3(e)], despite the persistence of the surface dimer structure, suggests that *coupling* of neighboring dimers to yield more extended, agglomerated structures must be considered. In order to address this possibility, we have performed quantum cluster calculations of such coupled dimer units. Three types of dimer pairings were considered,[25] all joining two oxygen-agglomerated dimers within the same dimer row: two coupled epoxides [Fig. 3(I)], one epoxide coupled to one hydrogenated agglomerate [Fig. 3(II)] and two hydrogenated agglomerates (not shown). The latter coupled species will not be considered further here, since all experimental evidence suggests that this species is present in negligibly small amounts since the total number of hydrogenated dimers is estimated to be less than 30% of the surface by integration of the ν(Si–H) peak at ~2275 cm^{-1}. Thus, the probability of finding two neighboring hydrogenated dimers in the same row is <5%. Both of the remaining two-dimer species possesses a pair of relatively intense Si–O modes polarized parallel and perpendicular to the surface: at 961 and 1103 cm^{-1} for species 3 I and 979 and 1142 cm^{-1} for species 3 II, respectively.

By inspection of the relevant experimental spectrum [Fig. 3(e)], we find that both the persistence of a parallel band between 950–1000 cm^{-1} and the appearance of a perpendicular mode in the 1100–1170 cm^{-1} region are consistent with the theoretically predicted trends for coalescence of the oxidized dimer units. The intermediate frequency of the experimentally observed modes compared to the theoretical predictions for the 2 coupled dimer species is then attributed to the co-existence of such double-dimer species and higher-order coupled dimer species on this surface. This assertion is supported by the observation of various discrete (weak) features superimposed on the perpendicular absorption band. Specifically, the feature at ~1100 cm^{-1} is consistent with the existence of a minor amount of the two-dimer epoxide species (3 I) which, given the theoretical finding that the perpendicular mode shifts to higher frequency with increasing agglomeration, suggests that the 1125 cm^{-1} peak may arise from the three-dimer epoxide analog. By analogy, the ~1140 and 1160–1170 cm^{-1} shoulders would then arise from species 3 II and the three-dimer analog, respectively.

It is interesting to note that the spectrum in Fig. 3(e) is somewhat reminiscent of an oxide, given the LO- and TO-like polarization behavior of the 1130 and 980 cm^{-1} modes, although these features must be understood in the context of an oxidized (2×1) surface rather than a more conventional amorphous, oxide-like film. The surface shown in Fig. 3(e), after five cycles of H_2O dosing at −50 °C and annealing to 600 °C, is unreactive to further H_2O adsorption under UHV conditions. Therefore, in order to make contact between this well-controlled surface and that typical for a truly three-dimensional silicon-oxide film, we have exposed the agglomerated surface to water at increasingly high pressures (up to ~8 Torr) at room temperature, collecting IR spectra continuously during this process, as described in the following section.

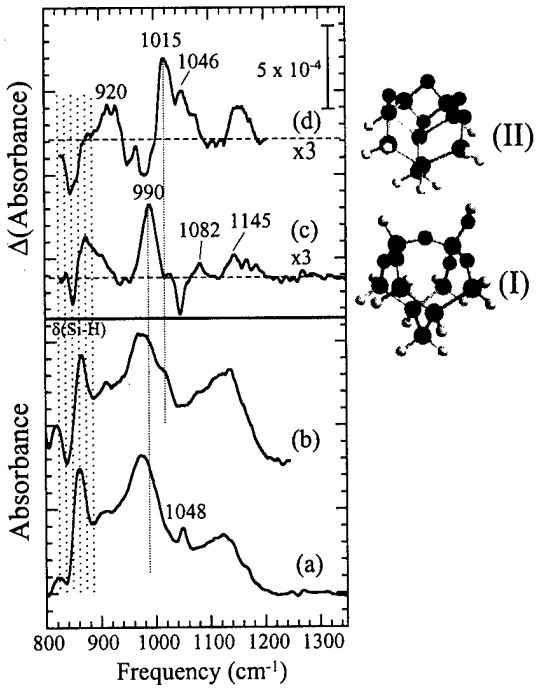


FIG. 4. IR spectra showing the development of the epoxide surface over ten orders of magnitude in water exposure. (a) The same spectrum as in Fig. 3(e), (b) this surface after water exposures up to 8 Torr at room temperature; (c) and (d) are difference spectra showing the changes between ~mTorr and 1 Torr exposures (c) and 1 and 8 Torr exposures (d), with the scales expanded for clarity. The structures denoted I and II are two of the cluster models used to evaluate the proposed intermediate structures. Different spheres depict Si/O/H atoms as described in Fig. 3.

C. Water-induced reaction of the agglomerated surface

Figure 4 shows the initial transformation of this surface over 10 orders of magnitude in H_2O partial pressure, from the agglomerated epoxide surface of Fig. 3(e) [shown again for reference in Fig. 4(a)] to the spectrum in Fig. 4(b), which was obtained after several days of water exposure at room temperature. While the overall change in the number of Si–O bonds detected is small—only ~7% as estimated by the integrated areas of the spectra in Figs. 4(a) and 4(b)—the transformations observed represent important steps in the growth of the first layer of oxide. The main peak at ~975 cm^{-1} in Fig. 4(a) is unshifted in Fig. 4(b), but a distinct shoulder at 1015 cm^{-1} is apparent after extended water exposure. The higher frequency, perpendicularly polarized peak at ~1125 cm^{-1} in Fig. 4(a) has grown in intensity

and shifted to ~1130 cm^{-1} in Fig. 4(b), whereas the sharp peak at 1048 cm^{-1} that corresponds to isolated epoxide species is absent in the latter spectrum.

By examining difference spectra obtained at select points during this extended high-pressure water exposure, we can pinpoint the more discrete changes in surface composition which lend some insight into possible mechanisms for oxide growth from the initial well-defined epoxide-covered surface. Two such difference spectra are displayed in Figs. 4(c) and 4(d) on an expanded scale. A spectrum acquired after exposure to ~1 Torr H_2O, ratioed to a spectrum acquired during mTorr H_2O exposure [Fig. 4(c)], demonstrates quite clearly that isolated epoxide species are consumed during this exposure range, since the 1048 cm^{-1} peak is completely depleted. Concomitant growth of discrete features at 990, 1082, and 1145 cm^{-1} suggests that these features arise from the reaction product of H_2O and 5-oxygen epoxide; growth of ν(OH) features around 3740 cm^{-1} (data not shown) indicates that the chemistry taking place is hydroxyl-mediated. By reference to results of cluster calculations, we find that one energetically favorable reaction ($\Delta E = -4.3$ eV) entails breaking the epoxide bond to adsorb H and OH on a 5-oxygen inserted dimer [Fig. 4(I)]. Since such a species would give rise to vibrational modes in good agreement with the three experimental peaks (calculated at 964, 1083, and 1149 cm^{-1}), we propose that the difference spectrum displayed in Fig. 4(c) arises from formation of such an intermediate.

The changes occurring as H_2O partial pressure is increased from 1 to ~8 Torr, shown in Fig. 4(d) indicate the growth of yet another intermediate species. As the dominant feature in this difference spectrum, at 1015 cm^{-1}, persists during the formation of a true oxide film (see later), it is likely that the intermediate giving rise to this peak involves the decomposition of Si–OH species to form siloxane (Si–O–Si) bonds. Indeed, we find that one such structure [which is closely related to the OH–containing intermediate proposed in Fig. 4(I)] gives rise not only to a calculated vibrational mode at 1024 cm^{-1}, but also modes at 911 and 1041 cm^{-1}, corresponding closely to the experimentally observed features at 920 and 1046 cm^{-1} in Fig. 4(d). This structure, shown in Fig. 4(II), contains a two-membered Si–O–Si ring on a fully oxygen-inserted dimer. While this structure is somewhat strained by the presence of this small ring, it does provide a possible link between hydroxylated species and the extended Si–O tetrahedral network which forms in the next step. It is particularly interesting to note that the two top silicon atoms in this structure are tetrahedrally coordinated to oxygen; addition of one more oxygen atom to this structure to break one of the shared Si–O bonds would then naturally result in the formation of two corner-sharing tetrahedra, the structural unit of amorphous SiO_2. Such two-membered Si–O–Si rings have been observed in fibrous silica, although they are typically found to be energetically unfavorable (+1.2 eV) relative to relaxed SiO_2 tetrahedra in a periodic structure.[26] It is reasonable, however, to posit that the observation of such a "kinetic" (metastable) product may be facilitated by the utilization of low temperatures (20 °C) in our extended oxide growth regime.

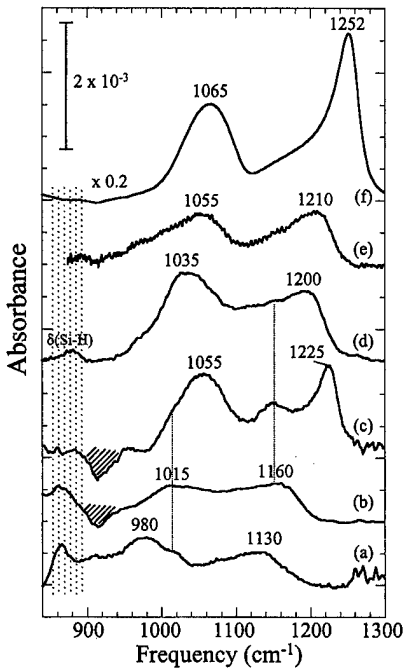


FIG. 5. IR spectra showing the final water-induced oxidation of the epoxide surface. (a) The same spectrum as in Fig. 4(b), obtained after exposure to ~8 Torr partial pressure of H_2O; (b) after condensation of water on this surface; (c) after approximately 3 weeks of controlled exposure to a saturated humid environment; (d) IR spectrum of a chemical oxide prepared by the modified RCA protocol, annealed to 200 °C in UHV; (e) and (f) are spectra of commercially grown thermal oxides; the 30 Å oxide in (f) was thinned in dilute HF to yield the 4.6 Å oxide in (e). The negative absorbance feature at ~905 cm^{-1} arises from the dihydride scissors mode of the hydrogen-terminated reference.

As a final step in water-induced oxidation, we condensed H_2O on the surface in Fig. 4(d) by raising the partial H_2O pressure to its dew point. The resulting surface, now exposed to atmospheric pressure of air, is shown in Fig. 5(b). As is evident by comparison to the spectrum from Fig. 4(b), repeated in Fig. 5(a), the lower frequency, parallel-polarized mode at 975 cm^{-1} has shifted upward to ~1015 cm^{-1}, coincident with the intermediate shoulder frequency identified in Fig. 4(d). At the same time, the higher frequency mode has shifted further upward to 1160 cm^{-1}. In addition, there is still significant intensity between these two maxima, suggesting that the epoxide-derived species giving rise to intensity at 1130 cm^{-1} in Fig. 5(a) persist in the film of Fig. 5(b). It is also important to note that there is virtually no change in the hydrogen content of the film during this water condensation step, as judged by integration of the Si–H modes in the IR spectrum, demonstrating that the spectral changes observed arise from growth in the Si–O network rather than from the introduction of new hydrogenated oxide defect structures.

D. Comparison with chemical and thermal oxides

Exposure of the surface of Fig. 5(b) to condensed water in a controlled humid environment causes further oxide growth

over a period of weeks. The spectrum in Fig. 5(c) shows the oxide film grown after approximately 3 weeks of such exposure. Remarkably, the dominant features of this spectrum are a SiO_2 TO/LO pair with frequencies indicative of a thermal-quality oxide film, 1055 and 1225 cm^{-1}. This is even more remarkable upon comparison to a chemical oxide film of similar thickness, such as the modified RCA spectrum displayed in Fig. 5(d). Even though the oxide grown from the epoxided surface is *thinner* than this chemical oxide, as judged by total Si–O intensity, the TO and LO frequencies of the epoxide-derived oxide are closer to those of a thermal oxide and thus indicative of a less defective film. Both a thick [Fig. 5(f), ~30 Å] and thinned [Fig. 5(e), ~4.6 Å] thermal oxide spectrum are also presented for comparison; it is clear that the final film produced in this work is more reminiscent of high quality stoichiometric SiO_2 than the 4.6 Å thinned thermal oxide, as exemplified by the position and linewidth of the respective LO modes.

Another important feature of the epoxide-derived oxide is the presence of a distinct feature at 1160 cm^{-1} [Fig. 5(c)], which can be traced back to the high-frequency peak of the previous spectrum [Fig. 5(b)] of the first stage of ambient oxide formation on this surface. Clearly, then, the intermediate/defect structures present in this initial surface leave an imprint on the subsequent oxide film, so that "thermalization" (i.e., attainment of an equilibrium state) is not quite complete. Further annealing studies are currently under way to investigate this latter transformation.

V. CONCLUSIONS

In summary, we have shown how unprecedented insight can now be obtained into the microscopic processes that occur during Si oxidation. We have delineated mechanistic pathways for thin film growth from isolated dimer units to continuous ultrathin oxide films. However, much remains to be explored before the generality of our findings can truly be assessed. To this end, we are currently studying the oxidation of hydrogen-terminated surfaces under well-controlled conditions, since these are the surfaces most commonly used as a starting point for growth of ultrathin gate oxides. In addition, from a theoretical standpoint, we are continuing to explore the myriad possible intermediate structures that may form as kinetic products during the latter stages of the oxidation process outlined earlier. The existing body of work should serve to stimulate further studies of these and other systems of interest such as silicon nitrides and oxynitrides which will form the next generation gate dielectric.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the invaluable technical assistance of E. E. Chaban and S. B. Christman, without whose involvement this work would not have been possible. A. B. Gurevich, J. P. Chang, R. L. Opila, and J. Sapjeta are also acknowledged for their involvement in the various experimental parts of this work.

[1]Y. J. Chabal and S. B. Christman, Phys. Rev. B **29**, 6974 (1984).
[2]Y. J. Chabal and K. Raghavachari, Phys. Rev. Lett. **53**, 282 (1984).
[3]J. Sarnthein, A. Pasquarello and R. Car, Science **275**, 1925 (1997).
[4]A. Pasquarello and R. Car, Phys. Rev. Lett. **79**, 1766 (1997).
[5]M. K. Weldon, K. T. Queeney, J. P. Chang, R. L. Opila, and Y. J. Chabal, Appl. Phys. Lett. (in press).
[6]T. Ohwaki, M. Takeda, and Y. Takai, Jpn. J. Appl. Phys., Part 1 **36**, 5507 (1997).
[7]See for example, P. A. Thiel and T. E. Madey, Surf. Sci. Rep. **7**, 211 (1987), and references therein.
[8]H. Ibach, H. Wagner, and B. D, Solid State Commun. **42**, 457 (1982).
[9]L. M. Struck, *et al.*, Surf. Sci. **380**, 444–454 (1997).
[10]M. Chander, Y. Z. Li, J. C. Patrin, and J. H. Weaver, Phys. Rev. B **48**, 2493 (1992).
[11]L. Andersohn and U. Kohler, Surf. Sci. **284**, 77 (1993).
[12]C. U. S. Larsson, A. L. Johnson, A. Flodstrom, and T. E. Madey, J. Vac. Sci. Technol. A **5**, 842 (1987).
[13]Q. Gao, Z. Dohnalek, C. C. Cheng, W. J. Choyke, and J. T. Yates, Jr., Surf. Sci. **312**, 261 (1994).
[14]K. Raghavachari, Y. J. Chabal, and L. M. Struck, Chem. Phys. Lett. **252**, 230 (1996).
[15]R. Konecny and D. J. Doren, J. Chem. Phys. **106**, 2426 (1997).
[16]A. B. Gurevich, B. B. Stefanov, M. K. Weldon, Y. J. Chabal, and K. Raghavachari, Phys. Rev. B **58**, R13434 (1998).
[17]M. K. Weldon, B. B. Stefanov, K. Raghavachari, and Y. J. Chabal, Phys. Rev. Lett. **79**, 2851 (1997).
[18]B. B. Stefanov, A. B. Gurevich, M. K. Weldon, Y. J. Chabal, and K. Raghavachari, Phys. Rev. Lett. **81**, 3901 (1998).
[19]M. Frisch, *Gaussian 94* (Gaussian Inc., Pittsburgh, PA, 1995).
[20]A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[21]C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[22]R. Frochtenicht, M. Kaloudis, M. Koch, and F. Huisken, J. Chem. Phys. **105**, 6128 (1996).
[23]B. B. Stefanov and K. Raghavachari, Surf. Sci. **389**, L1159 (1997).
[24]M. K. Weldon, A. B. Gurevich, Y. J. Chabal, B. B. Stefanov, and K. Raghavachari, J. Chem. Phys. (in press).
[25]While it is impossible to rule out the growth of 3-oxygen agglomerated structures during extended surface agglomeration, we confine our discussion to the coupling of the (dominant) 5-oxygen containing species, since these species can account for all the major IR features.
[26]D. R. Hamann, Phys. Rev. B **55**, 14784 (1997).

Constraint theory and defect densities at (nanometer SiO_2-based dielectric)/Si interfaces

J. C. Phillips[a)]
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974-0636

(Received 19 January 1999; accepted 3 May 1999)

Native SiO_2 passivating films grown thermally on flat Si(100) surfaces exhibit many ideal properties that justify regarding them as well-defined (albeit noncrystalline) phases with a continuous and almost defect-free interface with the Si substrate in spite of a large difference in density. Global constraint theory explains this remarkable situation without recourse to elaborate mathematical or geometrical local models, and it correctly predicts the scaling dependence of defect densities in films doped with N and H with essentially no adjustable parameters. © *1999 American Vacuum Society.* [S0734-211X(99)05204-X]

I. CONTINUITY OF NATIVE OXIDES

Many materials oxidize, but most oxides are polycrystalline, with discontinuities and high defect densities at grain boundaries which provide easy paths for catastrophic dielectric breakdown. Thermally grown SiO_{2-x}/Si(100) films are continuous and can have very low densities of point defects and chargeable traps. Because of the technological importance of SiO_2 passivating films, there is a large literature[1] on chemical structures of both thin (nanometer SiO_2)/Si interfaces and thicker interfaces up to ~100 nm. Substantial ordering of the amorphous SiO_2 has been observed by electron and x-ray diffraction[1,2] and ion channeling,[1,3] which indicates coherence lengths in the amorphous SiO_2 of order 10 nm or more. For ultrathin (a few nanometers) films on atomically flat Si(100) surfaces this means that the film can be regarded as a nearly homogeneous transition layer or interfacial phase[4] with its own characteristic properties; the defect density of the latter depends on oxygen diffusion kinetics and step densities (surface roughness)[5] of the Si substrate, but otherwise should exhibit systematic trends in chemical properties, much like bulk network glasses. In particular, it may be possible to find simple scaling relations which predict chemical trends in minimal or optimized defect densities, the property of greatest technological importance.

Most theoretical discussions of native Si/SiO_2 films have utilized local geometrical (ball-and-stick) models containing a few atoms to explain differences in properties of films prepared in somewhat different ways.[6] In general, such local models do not explain the crucial fundamental features of the films, their continuity, quasi-epitaxy which is virtually stress-free[7] in spite of a 40% density difference between c-Si and g-SiO_2, and with optimal preparation, defect densities[2] below 10^{-5}. The reason the geometrical models are inadequate is that they do not specify how the local geometry is related to the network as a whole, in other words, they do not specify the boundary conditions. To see how serious this problem is, suppose a continually branching (no rings) Bethe lattice model is attached to a small molecule. The resulting structure has an effective dimensionality D which $\rightarrow \infty$ as the boundary expands. This is *not* a space-filling model, in other words, it is basically unrelated to the continuity condition which is the *sine qua non* of Si/SiO_2.

One can construct[8] space-filling models of Si/SiO_2 by starting from the high-temperature crystalline phases (cristobalite and tridymite) of SiO_2 and mapping these epitaxially onto Si(100). The energetics of forming a defect (such as an O vacancy) can then be modeled with empirical force fields, but the crystalline character of the initial SiO_2 makes it necessary to neglect Si–Si bonding, in other words, the network character is lost. In general, such efforts founder on large cancellations between repulsive and attractive energies, and that appears to be the case here as well.

Physically the film grows without reference to a crystalline topology, and the Si–Si bonds appear naturally as a result of incomplete oxidation during growth with a minimal *long-range* strain field. The author has found only one previous theoretical model which is close in spirit to the present global model. This is the interface model constructed by hand and relaxed by computer of Pantelides and Long.[9] Their most interesting result was that the elastic distortion energy was minimized by an abrupt interface. Of course, the largest term in the total energy is that associated with the heat of formation of Si–O bonds, but this is fixed by the number of such bonds, so that the elastic strain energy is the relevant factor for determining the equilibrium configuration. The physical configuration is expected to be close to the equilibrium one, but it is modified by growth kinetics, as discussed later.

One can understand the growth of such a continuous, nearly defect-free film if one assumes that the growth takes place in a nearly stress-free way, compatible with nearly thermal equilibrium energetics; in other words, one treats the film as if it were almost a bulk glass without ever using a crystalline starting point to fill space. The main difference from the bulk glass is that oxygen must diffuse to the SiO_{2-x}/Si(100) interface, and for a nonzero growth rate this means $x>0$ in the film, while $x=0$ in the bulk glass as normally prepared.

a)Retired; electronic mail: jcphillips@lucent.com

II. CONSTRAINT THEORY

All the foregoing features of Si/SiO_2 can be understood by utilizing the global topological methods of constraint theory,[8] which has largely evolved in the context of chalcogenide, rather than oxide, network glasses, because chemical trends in the former can be studied more easily, including departures from stoichiometry The abruptness of interfaces[1] suggests that the defect density of thermally grown or optimally annealed films may be a characteristic function of the network chemistry and structure (charge transfer and bonding or space-filling topology). Experience with good glass formers (such as SiO_2 and chalcogenide glasses such as As_2Se_3) has shown that, so long as only single bonds are present, charge transfer plays a minor part in determining structure, with the major part being played by the network stress which arises for a given space-filling bonding topology. This space-filling characteristic is the microscopic mathematical analog of film continuity.

Constraint theory[10,11] (originally developed by Maxwell to describe mechanical scaffolds) has proved to be a remarkably accurate description of network stress and its consequences; it has even identified network stiffness phase transitions[11–13] which have properties strikingly analogous to those of fully equilibrated disordered electronic metal-insulator phase transitions. In any network the internal stress is extensive in mechanical equilibrium, much like the internal electric field in the disordered electronic case. The theory is based on the idea that all the bonding forces (stretching, bending, etc.) in the network can be arranged in a hierarchy, from strong to weak. The constraining effects of these forces are a linear function of the average coordination number N_{av}. If both bond-bending and stretching forces are present [some chalcogenide alloy glasses, such as (Si,Se) or (Ge,Se)], the optimal average coordination number N_{av}^* which matches constraints to degrees of freedom is 2.4 [$(Si,Ge)Se_4$)], but for SiO_2-based alloys, $N_{av}^*=2.67$ [stoichiometric SiO_2] because the bond-bending forces at the O atoms are too weak to function as constraints at the growth temperature.[10,13]

III. INTERMEDIATE PHASE (REVERSIBILITY WINDOW)

One of the important recent accomplishments of constraint theory is that it has identified not just one, but actually two,[13,14] stiffness transitions, at $N=N_{av}^*$ and at $N=N_{av}^{**}$. The difference $\Delta N_{av}^*/N_{av}^*$ between these two values is normally small, and experimentally it appears to depend on the presence of short-range order (molecular clusters, often with crystalline topologies). The details of these topologies are unimportant from a space-filling viewpoint; instead, what matters is that these topologies define characteristic subspace dimensionalities d^*. One would expect $\Delta N_{av}^*/N_{av}^*$ to decrease with increasing subspace dimensionality d^*, as d^* approaches d. Experimentally, this appears to be the case, as $\Delta N_{av}^*/N_{av}^*$ decreases from 6% for (Si,Se) [dominated by the chain subspace morphology of c-$SiSe_2$, $d^*=1$] to 2% for (Ge,Se) alloys [dominated by the layer subspace morphology of c-$\beta GeSe_2$, $d^*=2$] to, presumably, ~1% for SiO_2-based alloys (subspace morphology of the high temperature crystalline phases cristobalite or tridymite, $d^*=3$). (Note that this subspace device incorporates all the essential knowledge provided by the equilibrium crystal structure and phase diagram *without* imposing a crystalline topology on the amorphous film.) Apparently $N_{av}^*=2.67$ is associated with the hydrostatic stiffness transition, while $N_{av}^{**}=2.70$ is associated with a shear stiffness transition.[13]

Imagine now that we follow the growth of an $SiO_{2-x}/Si(100)$ nm film. Starting from $x=2$, where the network of the film is crystalline and severely overconstrained, it gradually becomes amorphous. The reaction $Si+O_2 \rightarrow SiO_2$ is exothermic and kinetically limited by diffusion of oxygen to the Si interface. Some internal stress builds up because of misfit between the amorphous SiO_2 film and the crystalline Si substrate, and this may limit film growth on processing time scales. However, when x decreases to ~0.01, $N_{av}=N_{av}^{**}=2.70$, and Young's modulus becomes nearly zero, so that the misfit stress can be relaxed at constant volume by volume-conserving uniaxial deformations normal to the interface. At this point the diffusion coefficient increases rapidly, as it always does in a liquid (which also has hydrostatic but not Young's modulus stiffness) compared to a crystal. Film growth will take place rapidly and will be limited by time, temperature, and oxygen supplied. This means that $N_{av}=N_{av}^{**}=2.70$ is a natural kinetic condition which can determine the composition of an ultrathin film. Attempts to oxidize x to 0 in the ultrathin film will merely grow a thicker film.

In the window $N_{av}^*<N_{av}<N_{av}^{**}$ the network behaves much like an incompressible liquid, which will not support a large shear stress, and it has a glass transition from the supercooled liquid to the glass which is nearly thermally reversible.[14] In a supercooled liquid containing network fragments one can associate irreversibility with fragment entanglement or knot formation (compare polymer chains). This solves the problem that puzzled Mott,[7] namely how SiO_2 can grow epitaxially (no defects) on Si, in spite of the large density difference. It also explains how annealing stepped Si/SiO_2 interfaces anneals away the steps[15] through the instability of SiO_2 against shear.

IV. DEFECT CONCENTRATIONS IN ULTRATHIN FILMS

We can now extend this discussion of internal stress to predict defect concentrations c as a function of N_{av} and N_{av}^{**} for optimized processing conditions in ultrathin films. The basic idea is borrowed from the theory[16] of stress relief at crystalline heterointerfaces, which occurs via the formation of misfit dislocations (which is also kinetically limited). Stress accumulates harmonically, that is to say, quadratically, as a function of interfacial misfit, which in the crystalline case is measured by the lattice constant difference. There are no "lattice constants" for an amorphous network, but the configuration coordinate N_{av} can (and judging from many experiments,[14,17] does) play an equivalent role.

More specifically, for $N_{av} > N_{av}^{**}$ (an overconstrained network) in SiO_2-based films which include N and H, the bond-bending constraints of the N may be weak, as for O, and there are no bending constraints from one-fold coordinated H. Then all the stretching constraints are stronger than the Si bending constrains, which cannot all be satisfied and so can be termed marginal. In other words, strain energy accumulated among these constraints, with the average angle θ_{ij} between two Si bonds being distorted from the ideal local value θ_{ij}^{*} by an amount

$$\delta\theta \propto \delta N_{av}^{*} = N_{av} - N_{av}^{**} . \quad (1)$$

The total strain energy is proportional[16] to $(\delta\theta)^2$, as is the defect density.

With the increase of strain energy defects can be created, and the cost in energy per defect (such as a dangling bond) is constant. As in the case of misfit dislocations,[18] the accumulation of strain energy in the network is interrupted (or relaxed) by the presence of defects. Here these are point, rather than line or half-plane, defects, but similar reasoning applies. In fact, at crystalline semiconductor heterointerfaces, it has been observed that, for a given lattice constant misfit, the density of misfit dislocations can be reduced by the introduction of impurities which in the partially disordered environment of the interface also relieve internal strain.[18] Thus, the concentration of network defects should scale as $(N_{av} - N_{av}^{**})^2$, in good agreement with experiment.[19]

Genuine columnar epitaxial SiO_2/Si has been identified[2] in synchrotron diffraction experiments; the crystalline material apparently fills ~1% of the sample volume. According to constraint theory, this material should not be viewed as having grain boundary interfaces with amorphous SiO_2, but rather as simply losing planar orientation relative to the Si substrate. The fact that the observed interfacial lateral coherence lengths vary from 130 to 400 A (30–100 lattice constants) as the film thickness varies from 100 to 1000 A confirms that Young's modulus of the SiO network is almost zero. The actual interfacial boundary conditions that generate this kind of columnar growth are not clear, and may involve a different kind of extended defect which is absent from ultrathin films.

V. RELATED EXPERIMENTS

Recent work[20] on atomically flat Si(100) surfaces prepared by oxide removal with a Si flux showed a relatively high density y of monatomic steps (~10%). The density of Si–Si bonds ($\sim 1\% = x \sim y^2$) predicted by our model, which assumes that thermal oxide growth occurs at the composition of the stiffness threshold for shear (Young's modulus normal mode), is much smaller. This means that either optimized oxide thermal growth erases most of the steps, or that optimized oxidized steps do not contain Si–Si bonds. Perhaps these bonds occur at optimized oxidized kinks in the steps, and the density of these kinks z is relaxed in ultrathin (<30 nm) oxide films to the value expected from constraint theory. Incidentally, there is some indirect evidence from Fourier transform infrared-attenuated total reflection spectroscopy that such relaxation is characteristic of ultrathin films.[21]

ACKNOWLEDGMENT

This article would not have been written without encouragement and stimulation from Professor G. Lucovsky.[19]

[1]T. Hattori, Crit. Rev. Solid State Mater. Sci. **20**, 339 (1995).
[2]A. Munkholm, S. Brennan, F. Comin, and L. Ortega, Phys. Rev. Lett. **75**, 4254 (1995).
[3]Y. P. Kim, S. K. Choi, H. Y. Kim, and D. W. Moon, Appl. Phys. Lett. **71**, 3504 (1997).
[4]G. Renaud, P. H. Fuoss, A. Ourmazd, J. Bevk, B. S. Freer, and P. O. Hahn, Appl. Phys. Lett. **58**, 1044 (1991).
[5]C. Zhao, Y. Z. Hu, T. Labayen, L. Lai, and E. A. Irene, J. Vac. Sci. Technol. A **16**, 57 (1998).
[6]N. F. Mott, S. Rigo, F. Rochet, and A. M. Stoneham, Philos. Mag. B **60**, 189 (1989).
[7]N. F. Mott, Philos. Mag. B **55**, 117 (1987).
[8]S. Carniato, G. Boureau, and J. H. Harding, Philos. Mag. A **75**, 1435 (1997).
[9]S. T. Pantelides and M. Long, in *The Physics of SiO_2 and Its Interfaces*, edited by S. T. Pantelides (Pergamon, New York, 1978).
[10]J. C. Phillips, J. Non-Cryst. Solids **34**, 153 (1979); **43**, 37 (1981).
[11]J. C. Phillips, Solid State Commun. **47**, 203 (1981).
[12]H. He and M. F. Thorpe, Phys. Rev. Lett. **54**, 2107 (1985).
[13]J. C. Phillips, in *Rigidity Theory and Applications*, edited by M. F. Thorpe and P. Duxbury (Michigan State University Press, MI, 1999).
[14]D. Selvanathan, W. J. Bresser, P. Boolchand, and B. Goodman, Solid State Commun. (in press).
[15]E. P. EerNisse, Appl. Phys. Lett. **30**, 290 (1977).
[16]J. H. Van der Merwe, J. Appl. Phys. **34**, 123 (1963).
[17]P. Boolchand, X. Feng, D. Selvanathan, and W. J. Bresser, in *Rigidity Theory and Applications*, edited by M. F. Thorpe and P. Duxbury (Michigan State University Press, MI, 1999).
[18]L. Zou, Z. G. Wang, D. Z. Sun, T. W. Fan, X. F. Liu, and J. W. Zhang, Appl. Phys. Lett. **72**, 845 (1998). As implanted in SiGe reduces misfit dislocation densities by precipitating GeAs.
[19]G. Lucovsky, Y. Wu, H. Niimi, V. Misra, and J. C. Phillips, J. Vac. Sci Technol. B (these proceedings).
[20]G. D. Wilk, Y. Wei, H. Edwards, and R. M. Wallace, Appl. Phys. Lett. **70**, 2288 (1997).
[21]K. Inanga, T. Nakahara, T. Furukawa, and K. Ono, Appl. Surf. Sci. **100–101**, 421 (1996).

Bonding constraint-induced defect formation at Si-dielectric interfaces and internal interfaces in dual-layer gate dielectrics

G. Lucovsky[a)]
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

Y. Wu
Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

H. Niimi
Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

V. Misra
Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

J. C. Phillips
Lucent Bell Laboratories, Murray Hill, New Jersey 07974

(Received 19 January 1999; accepted 30 March 1999)

As aggressive scaling of integrated circuits continues into the next century, insulators with dielectric constants higher than SiO_2 with different local bonding arrangements will be required to increase gate dielectric capacitance in field effect transistor devices. An important issue in semiconductor device physics is determining whether differences between the bonding at (i) Si–SiO_2 interfaces and (ii) interfaces between crystalline Si and alternative gate dielectric materials will result in increased densities of electrically active defects at the alternative dielectric interfaces, thereby limiting targeted levels of performance and reliability. In particular, it is important to understand from a chemical bonding perspective why Si–SiO_2 interfaces display both low defect densities and high reliability, while other interfaces such as Si–Si_3N_4 with similar bonding chemistry, display defect densities that are at least two orders of magnitude higher. Building on previously established criteria for formation of low defect density glasses and thin films, constraint theory is extended to crystalline Si-dielectric interfaces that go beyond Si–SiO_2 through development of a model that is based on the average bonding coordination at these interfaces. This approach identifies quantitative bonding criteria that distinguish between device-quality and highly defective interfaces. This extension of constraint theory is validated by its application to interfaces between Si and stacked silicon oxide/nitride dielectrics which demonstrates that as in bulk glasses and thin films an average coordination, $N_{av}>3$ yields increasingly defective interfaces. Finally, the universality of this application of constraint theory is demonstrated by showing that defect densities scale with overcoordination in the same way in thin films and at interfaces. © *1999 American Vacuum Society.*
[S0734-211X(99)03904-9]

I. INTRODUCTION

As source-drain distances in field effect transistors (FETs) are scaled to <150 nm and ultimately to ~50 nm, the oxide equivalent thickness, $t_{ox\text{-}eq}$ of gate dielectrics must be reduced initially to <2 nm, and then to <1 mn. Values of $t_{ox\text{-}eq}$ are calculated from the experimentally determined gate capacitance in the accumulation region, C_{acc}, by assuming that the dielectric film or composite multilayer stack is equivalent to an SiO_2 film with a static dielectric constant, k_0, of 3.8, so that

$$t_{ox\text{-}eq}=k_0\epsilon_0 A/C_{acc}, \tag{1}$$

where ϵ_0 is the permittivity of free space, and A is the area of the capacitor. C_{acc} can be obtained directly from experimental $C-V$ data for *thicker* films, $t_{ox\text{-}eq}>2.5-3$ nm, at accumulation voltages >~3 V, or from a fit to capacitance–voltage data for *ultrathin* dielectrics when $t_{ox\text{-}eq}<2.5$ nm.

At $t_{ox\text{-}eq}<1.5$ nm direct electron tunneling through SiO_2 films at ~1 V exceeds 1 A/cm^2, a value at which circuit performance and reliability are degraded. One approach to reduce direct tunneling current is to replace SiO_2 by physically thicker films with higher dielectric constants. The first step is this replacement process has been to substitute Si_3N_4 for SiO_2 so that the film thickness can increased by ~2, the ratio of the respective dielectric constants. Successful substitution requires (i) that the Si_3N_4 bulk films have fixed charge and trap densities that are comparable to SiO_2, i.e., $<1\times10^{11}$ cm^{-2}, (ii) that the interfaces between Si and Si_3N_4 have fixed charge and trap densities that are comparable to Si–SiO_2, i.e., $<5\times10^{10}$ cm^{-2}, and (iii) that changes in other material and interface properties, such as tunneling masses and barriers do not offset reductions in tunneling current anticipated from increased physical thickness. Recent experi-

a)Electronic mail: gerry_lucovsky@ncsu.edu

mental studies have demonstrated that reductions in electron tunneling masses in Si_3N_4 thin films, and in conduction band offset energies between Si_3N_4 and Si compensate almost exactly for increases in dielectric constant that allows the use of physically thicker SiO_2 films.[1] These aspects of Si_3N_4/SiO_2 substitutions are addressed in more detail elsewhere, and suggest that composite SiO_2/Si_3N_4 gate dielectric films may not satisfy many of the targeted goals for aggressive scaling required in advanced the most FET devices.[2]

With the exception of issues relating to tunneling currents, it has been demonstrated that bulk Si_3N_4 films prepared by remote plasma-enhanced chemical vapor deposition (RPECVD) at 300 °C, and annealed at 900 °C for at least 30 s in an inert, nonoxidizing ambient have bulk properties comparable to thermally grown SiO_2 when incorporated into capacitors with stacked oxide/nitride/oxide (ONO) dielectrics with $Si–SiO_2$ or nitrided $Si–SiO_2$ interfaces,[3,4] or in FETs with stacked oxide/nitride (ON) dielectrics with $Si–SiO_2$ or nitrided $Si–SiO_2$ interfaces.[5–7] In contrast to nitrides prepared by other techniques, such as rapid thermal CVD (RTCVD),[8] or jet vapor deposition (JVD),[9] which generally require an oxidizing anneal after film deposition to form oxynitride alloys with low bulk defect densities, the annealed RPECVD nitrides derive their low bulk defect densities from incorporation of ~10–15 at. % hydrogen.[4,7] The as-deposited RPECVD films have approximately 20–30 at. % bonded hydrogen, mostly in *near neighbor* in Si–H and SiN–H arrangements,[10–12] and after the 900 °C anneal, the remaining 10–15 at. % is bonded in *isolated* SiN–H arrangements. Detailed studies of the specific hydrogen bonding arrangements with nearest-neighbor Si–H and/or SiN–H groups, that readily evolve H_2 molecules upon annealing, are discussed in detail elsewhere.[11,12] This article focuses on issues relative to $Si–Si_3N_4$ interface bonding and defects.

It is shown that a direct substitution of Si_3N_4 for SiO_2 in FETs is limited by charged defects at the $Si–Si_3N_4$ interface that will be shown to result from constraints imposed by overcoordination in the spirit of Refs. 13–15. In particular, this article extends constraint theory, originally developed for bulk oxide and chalcogenide bulk glasses by Phillips[13,14] and thin films by Lucovsky and Phillips[15] and thereby provides a quantitative approach to explaining the significant differences in defect properties between $Si–SiO_2$ and $Si–Si_3N_4$ interfaces. Another paper at this conference by Misra and co-workers,[7] presents electrical measurements on both *p*-channel (hole conducting) and *n*-channel (electron conducting) FETs, and concludes that increases in the density of donor-like defects in the lower half of the gap *p*-channel FETs render these devices totally dysfunctional. This article presents complimentary studies of *p*-channel FETs and capacitors with $Si–Si_3N_4$ interfaces with essentially the same result. The results reported in this paper and in Ref. 7 further demonstrate that insertion of ~0.6 nm (or about two molecular layers) of SiO_2 between the Si substrate and the Si_3N_4 film deposited by RPECVD at 300 °C and annealed in an inert ambient for at least 30 s at 900 °C restores interface devices properties to essentially the same

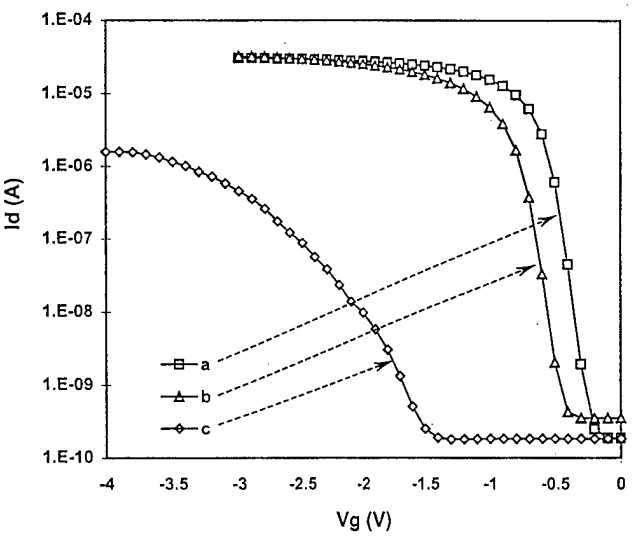


FIG. 1. Drive current–gate voltage ($I_d–V_g$) characteristics for PMOSFETs with $t_{ox\text{-}eq}$~2 nm: (a) a 1.5 nm oxide separating a 1.0 nm nitride from the Si substrate, (b) a 0.6 nm oxide separating a 2.4 nm nitride from the Si substrate, and (c) a 4 nm nitride layer. The threshold voltage shift between (a) and (b) is due in part to substrate doping differences (0.16 V) and in part to positive charge at the oxide-nitride interface (0.04 eV).

level as in devices with thermally grown $Si–SiO_2$ interfaces. This dramatic change in interfacial defect properties will also be explained in the context of the extension of constraint theory.

Section II presents experimental results for $Si–Si_3N_4$ interfaces and Sec. III discusses constraint theory as proposed originally for bulk glasses[13,14] and thin films.[15] The application of constraint theory to thin films is used to develop quantitative relationships between departures from ideal average bonding coordination and defect concentrations, which serve as a basis for the development of a quantitative approach to defects semiconductor-dielectric interfaces and internal dielectric interfaces. Section IV extends constraint theory to interfaces and Sec. V applies this extended theory to the experimental results of Sec. II. The model is further generalized in Sec. V to interfaces between Si and the so-called *alternative high-K dielectrics* such as TiO_2 (Ref. 16) and Ta_2O_5 (Ref. 17) that are being considered as replacements for SiO_2 in aggressively scaled FET devices. In Sec. V, the theory is applied to internal interfaces in stacked dielectrics, such as the oxide/nitride interfaces of Refs. 1–6. Finally, in Sec. VI, the universality of quantitative defect scaling with overcoordination in thin films and at interfaces is demonstrated.

II. EXPERIMENTAL RESULTS FOR $Si–Si_3N_4$ INTERFACES

Figure 1 displays current–voltage, $I–V$, curves for *p*-channel FETs for different gate dielectrics with $t_{ox\text{-}eq}$ ~2 nm: (i) a 4 nm RPECVD nitride, (ii) a 0.6 nm plasma-oxide with a 2.4 nm RPECVD nitride, and (iii) a 1.5 nm thermal oxide with a 1.0 nm RPECVD nitride. The substrates are Si(100), doped to $5\times10^{17}\,cm^{-3}$ for (i) and (iii) 1.1

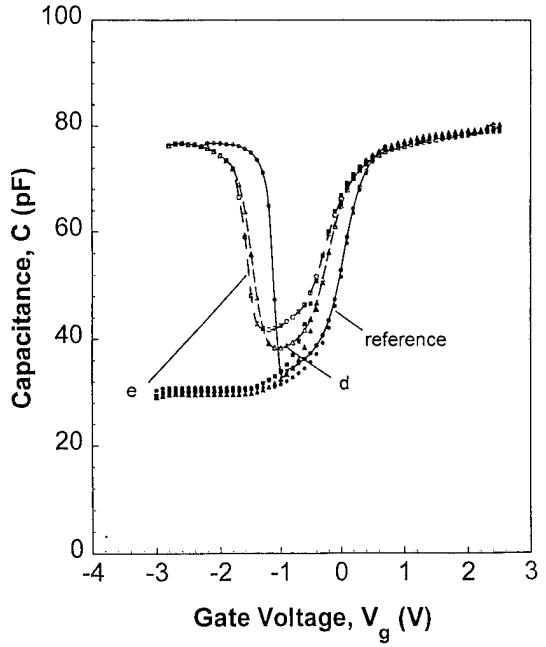


FIG. 2. $C-V$ characteristics demonstrate shift in flatband voltage due to positive charge, and increased separation between high frequency and quasi-static plots due to interface trapping accompanying direct deposition of thin nitride films onto Si. Each of these capacitors has $t_{ox\text{-}eq}\sim 2$ nm: (i) a reference oxide and (ii) two stacked NO structures with the nitride layer in contact with the Si substrate. The nitride layer thickness is 0.4 nm for (d), and 0.8 mn for (e).

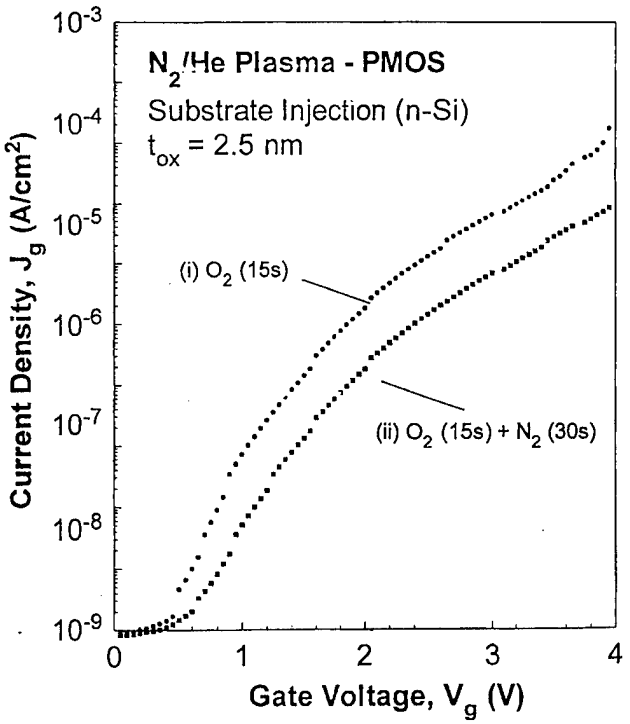


FIG. 3. $J-V$ characteristics for capacitors prepared on lightly doped n-type Si(100) with n^+ polycrystalline Si gate electrodes for $t_{ox\text{-}eq}\sim 2.5$ nm. The upper curve is for a device with an RPECVD SiO_2 layer and a non-nitrided $Si-SiO_2$ interface, and the lower curve is for a device with an RPECVD SiO_2 layer and a monolayer nitrided $Si-SiO_2$ interface.

$\times 10^{18}\,cm^{-3}$ for (ii). $I-V$ traces for (i) and (iii) display excellent turn-on behavior and the same current drive, with differences in threshold voltage between (ii) and (iii) derived primarily from differences in substrate doping noted above ($\sim$0.160 V out of the 0.200 V difference). In contrast, for the FET with the 4 nm nitride: (i) threshold voltage is shifted to negative voltages by >1 V, (ii) turn-on is soft, and (iii) channel drive current is reduced by $\sim$50. Figure 2 displays capacitance–voltage, $C-V$, characteristics for p-type metal–oxide semiconductor (PMOS) devices with $t_{ox\text{-}eq}\sim 4.3$ nm: one with a plasma-oxide, and two with stacked dielectrics with RPECVD nitride interface layers of 0.4 and 0.8 nm, respectively. Shifts in threshold (and flatband) voltage relative to $Si-SiO_2$ indicate increased fixed charge for devices with nitride interfaces: $\Delta q_f = C_{ox}\Delta V_{th}(\text{or}\,\Delta V_{fb})$. Qualitatively similar results have been obtained for n-type metal–oxide–semiconductor (NMOS) devices with nitride layer interfaces.[7] Figure 3 displays current density–voltage, $J-V$, plots for a capacitor fabricated on lightly doped n-type substrates and n^+ polycrystalline Si gate electrodes for an oxide thickness of 2.5 nm as determined from analysis of $C-V$ data.[18] There are two types of devices shown in Fig. 3: (i) the device exhibiting the higher current was prepared by remote plasma processes that resulted in a non-nitrided interface, whereas (ii) the device exhibiting reduced tunneling current was prepared by remote plasma process that resulted in a monolayer nitrided interface.[18] The $C-V$ data establish that the flatband voltages are determined by the substrate and polycrystalline Si doping and not changed by interface nitridation. For this pair of devices with the same effective thickness, the tunneling current is reduced by the monolayer interface nitridation, indicating that insertion of a monolayer of nitride at the $Si-SiO_2$ interface is beneficial. This means that monolayer interface nitridation is *qualitatively different* from having one or more molecular layers of nitride present the interface with the crystalline Si substrate.

Summarizing the data in Figs. 1–3 and data presented in other papers,[19,20] the $I-V$ and $C-V$ data establish that: (i) that deposition of Si_3N_4 films by RPECVD at Si-dielectric interfaces degrades interface properties, (ii) that interface properties can be effectively restored by interposing an ultrathin ($\sim$0.6 nm) layer of plasma-grown SiO_2 between the Si substrate and the RPECVD Si_3N_4 film, and finally, (iii) monolayer interface nitridation either during remote plasma oxidation,[19] or after remote plasma oxidation[18] resulted in improvements in interface properties, rather than degradation.[7,20]

III. CONSTRAINT THEORY FOR BULK GLASSES AND THIN FILMS

Phillips showed that fully bonded, strain-free three-dimensional continuous random networks (CRNs) can be formed at compound and alloy chemical compositions by counting average bond constraints associated with applicable valence force fields, C_{av}, and then matching constraints to the dimensionality of the network structure.[13–15] For three-dimensional melt-quenched glasses these fields are simply

bond-stretching and bond-bending valence forces. The quantitative relationships between (i) local bonding coordination and geometry, and (ii) constraints are presented in the next paragraph. Under the condition of exactly matching constraints to network dimensionality, the average coordination number/atom, N_{av}, is 2.4, corresponding to the well-known glass-formers arsenic trisulfide and arsenic triselenide, As_2S_3 and As_2Se_3, respectively. Good glass formation has also been obtained in the Ge–S and Ge–Se systems, even at the compound compositions (GeS_2 and $GeSe_2$), where $N_{av}=2.67$, and $C_{av}=3.67$, exceeding the optimum value of 3. Recent studies have shown ideal-glass formation in Ge–S and Ge–Se alloy systems occurs at a chalcogenide-rich composition corresponding to $N_{av}\sim 2.45$ with a value of C_{av} close to 3.[21] For the SiO_x ($x\leq 2$) system, O-atom bond-bending forces are ineffective at fictive (or quench) temperatures above the effective viscoelastic relaxation temperature of ~1000 °C, so that the applicable forces in bulk-quenched glasses are bond-stretching forces on both atoms, and bond bending forces only on the Si atoms. $N_{av}=2.67$ at the composition of SiO_2, and the number of constraints per atom, neglecting the O-atom bond-bending forces, is also $C_{av}=3$. This explains the ease of glass formation by quenching from melts exceeding the viscoelastic relaxation temperature.[13–15]

As indicated above, the calculation of C_{av} is based on counting constraints associated with valence forces. The number of bond-stretching constraints/atom is $m/2$, where m is the coordination number, and for three-dimensionally coordinated atoms, the number of bond-bending constraints is $2m-3$.[13] Applied to As_2S_3, the number of bond-stretching constraints is equals 6, and the number of bond-bending constraints is equals 9, so that the total number of constraints is 15, and $C_{av}=3$. Applied to SiO_2 glasses that are quenched from temperatures exceeding the viscoelastic relaxation temperature, the total number of bond-stretching constraints is 4, and the number of bond-bending constraints is associated with Si atoms is 5, so that $C_{av}=3$, as well.

The bonding of N in Si–N and Ge–N systems is planar or effectively two dimensional with the N atom bonded to three Si or Ge atoms in an NX_3 geometry, $X=$ Si or Ge.[22] This bonding arrangement is stabilized by $p\pi$-$d\pi$ interactions in which the nonbonding pair of the N atom in p_z orbital perpendicular the Si(Ge–N) bonding x-y plane backdonates to unoccupied antibonding Si or Ge orbitals that have a d^3p symmetry. As a consequence of this planar bonding of N atoms in Si–N alloys, the number of N-atom bond-bending constraints is reduced to $m-1$, so that for Si_3N_4, $C_{av}=5.0$, and $N_{av}=3.43$. However, this value of C_{av} is significantly higher than the network dimensionality so that glass formation from the melt does not occur.

The optimum value of $C_{av}=3$ applies to ideal melt-quenched networks, and this criterion must be modified for vapor-deposited films prepared at temperatures well below the glass transition temperature. These films are generally not strain free, and contain voids as well.[14] For nonhydrogenated a-Si $N_{av}=4$ and $C_{av}=7$, and hydrogen-free a-Si films produced by sputtering generally contain high concentrations of dangling bonds, e.g., the concentration of Si dangling bonds detected by electron spin resonance is $\sim 10^{18}\,cm^{-3}$ after optimized thermal annealing.[23] The hydrogenated films also have internal voids which serve to reduce strain, thereby reducing the effective number of dangling bond defects below a value that is simply determined by C_{av}.[15] N_{av} and C_{av} can be reduced further by hydrogenation leading to significant reductions in the dangling bond density. For example, for an a-Si:H alloy with 10–15 at. % H, $N_{av}=3.5$, and $C_{av}=5.7$. These alloys display Si atom dangling bond densities of order $5\times 10^{16}\,cm^{-3}$, which is low enough for promoting good semiconductor properties as for example high substitutional doping efficiencies, thereby allowing the use of these doped a-Si:H films in photovolatic cells and thin film transistors (TFTs).[24]

TABLE I. Calculated values of N_{av} and N_{av} for melt-quenched and deposited dielectrics.

Material	Atomic coordination and (dimensionality) bulk-quenched glasses		N_{av}	C_{av}
$As_2S(Se)_3$	As: 3(3)	S(Se):2(3)	2.40	3.00
$GeS(Se)_2$	Ge: 4(3)	S(Se):2(3)	2.67	3.67
$GeS(Se)_{3.5}$	Ge: 4(3)	S(Se):2(3)	2.44	3.10
SiO_2	Si:4(2)O: 2[a]		2.67	3.00
	Plasma-deposited films			
SiO_2	Si: 4(3)O: 2(3)		2.67	3.67
Si_3N_4	Si: 4(3)N: 3(2)		3.43	5.00
a-Si:N:H[b]	Si: 4(3), N: 3(2), H: 1(1)		2.68	3.58
a-Si	Si: 4(3)		4.00	7.00
a-Si:H[c]	Si: 4(3)H: 1(1)		3.50	5.70

[a]Neglects bond-bending constraints for O atoms.
[b][N]=0.42, [Si]=0.28, and [H]=0.30.
[c][Si]=0.8, [H]=0.2.

Table I includes calculated values of N_{av} and C_{av} for ideal melt-quenched glasses, and for thin films. As shown in Fig. 4, increases in C_{av} correlate linearly with increases in N_{av} so that either descriptive can be applied to bulk-quenched glasses or plasma-deposited amorphous materials and as shall be demonstrated in the next section, to crystalline Si-dielectric interfaces and internal dielectric interfaces as well. In the remainder of this article, the N_{av} metric will be used in discussing scaling of defects.

As discussed above, if both bond-bending and stretching forces are present, the optimal average coordination number, N^*_{av}, which matches constraints to degrees of freedom is 2.4 as in $As_2S(Se)_3$, however, for SiO_2 bulk glasses, $N^*_{av}=2.67$ is optimal because bond-bending forces at O atoms are too weak to function as significant constraints at growth or annealing temperatures.[15] For overconstrained networks such as in thin film Si_3N_4 ($N_{av}=3.43$), Si-atom stretching constraints are stronger than bending constraints, so that strain energy accumulates along the bending constraints. The average Si–N–Si bond angle θ_{ij} is distorted from the ideal local value θ^*_{ij} by an amount

$$\delta\theta\propto\delta N^*_{av}=N_{av}-N^*_{av}. \quad (2)$$

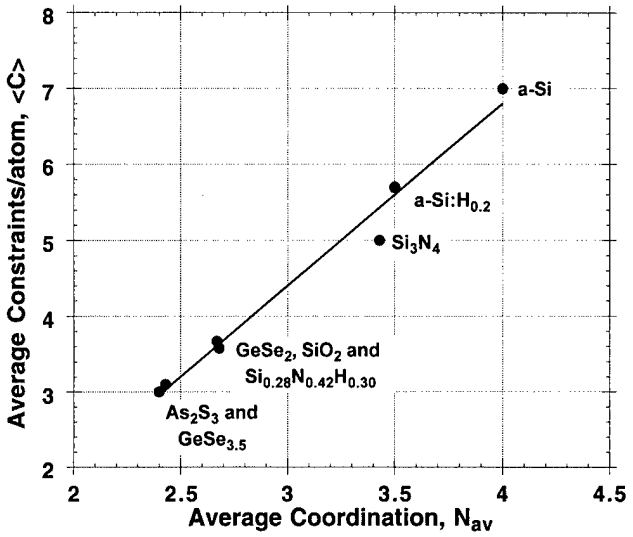


FIG. 4. Bonding constraints per atom, C_{av} vs average number of bonds per atom, N_{av}, for bulk-quenched glasses and deposited thin films (see Table I). The line in this figure is a least-squares fit the data points, excluding the value for bulk-quenched SiO_2.

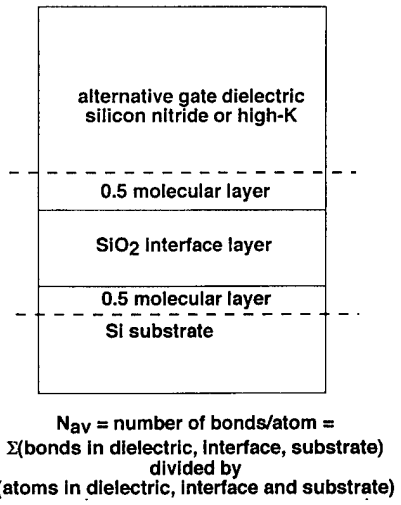


FIG. 5. Schematic representation of interface constituents used in the calculation of N_{av} for Si-dielectric interfaces.

Since total strain energy is proportional to $(\delta\theta)^2$,[25] it is then expected that defect creation, e.g., dangling Si or N bonds, will be proportional to $\{N_{av}-N_{av}^*\}^2$. Similar considerations apply to other thin film materials as well. Experiments have shown that $N_{av}\sim 3$ represents a reasonable demarcation criterion between low defect density ($\sim 10^{16}\,cm^{-3}$), and increasingly defective materials.[15] Equation (2) is also consistent with the structure of noncrystalline solids in which the statistical distribution of bond angles is much larger than the statistical spread in bond lengths, δr.

IV. EXTENSION OF CONSTRAINT THEORY TO Si-DIELECTRIC INTERFACES

Before applying constraint theory to Si-dielectric interfaces, it is important understand other factors than can play a role in interface bonding and the associated electronic structure. There are at least four factors related to bonding that can promote formation of interfacial defects; these are (i) differences between electronic charge required for bonding and for balancing the nuclear charge of the constituent atoms as first identified by Harrison and co-workers for group IV–group III–V heterojunctions such as Ge–GaAs (heterovalent as opposed to isovalent bonding),[26] (ii) interfacial dipoles produced by charge transfer between the Si substrate and the gate dielectric material,[27,28] (iii) molar volume differences between the Si substrate and gate dielectric that produce interfacial strain, and generally require interfacial dangling bond formation, and finally, (iv) overconstrained bonding due to differences between the average number of bonds per atom on each side of the interface.[29] Going down this list in order; (i) is not a consideration since both SiO_2 and Si_3N_4 can bond to a Si substrate by forming isovalent bonds between substrate Si atoms and either O or N atoms of the respective dielectrics. Second, it has been shown that interfacial charge transfer dipoles in (ii) result in relatively small differences in band offset energies at the Si–SiO_2 and Si–Si_3N_4 interfaces, and that charge transfer is smaller at the nitride interfaces that at oxide interfaces. Therefore it is not likely that these dipoles play a determinant role in defect or defect precursor formation since the defect densities at the Si–Si_3N_4 interfaces are larger than those at the Si–SiO_2 interfaces. Interfacial strain as in (iii) has been shown to be an important factor in interface quality. It has been shown that for thermally grown oxides on Si in a thickness range that extends to at least 100 nm, the residual density of interface traps, D_{it}, scales linearly with integrated strain in the oxide layer.[30] Since the molar volume mismatch between Si_3N_4 and Si is reduced with respect to that of SiO_2 to Si, residual interface strain cannot be the driver for the increased defect concentrations at Si–Si_3N_4 interfaces. The remainder of this section focus on the final interface consideration in (iv) dealing with bonding constraints.

The abruptness of Si–SiO_2 interfaces suggests that the defect density of thermally grown oxides and optimally annealed deposited oxides is a characteristic function of their bonding chemistry and structure. Experience with good glass-formers such as SiO_2 and As_2Se_3 has shown that as long as only single bonds are present charge transfer plays a minor part in determining structure. The major factor is the network stress which arises for a given space-filling bonding topology, and this factor is expected to extend to interfaces as well.

Extension of constraint theory to Si-dielectric-interfaces considers three interfacial contributions to N_{av}: (i) the Si substrate represented by one-half a Si atom, (ii) an ultrathin oxide or nitride interfacial layer (0.3–0.6 nm), and (iii) the bulk dielectric by one-half a molecular layer (see Fig. 5). Table II includes calculations of N_{av} for representative Si-dielectric interfaces. When a demarcation level $N_{av}\sim 3$ is applied between device-quality and highly defective interfaces, these calculations are in excellent agreement with experiment (see Figs. 1 and 2). For example, the model con-

TABLE II. Average bonding coordination at Si-dielectric interfaces.

Material system	Average coordination (N_{av})	Electrical quality
Si–SiO_2 (1.5 molecular layers)	2.8	excellent
Si–Si_3N_4 (1.5 molecular layers)	3.5	very poor (Fig. 1)
Si–{SiO_2}(t)–Si_3N_4	t=0.5 nm: 3.1	excellent (Fig. 1)
t=oxide layer thickness	t=1.5 nm: 2.9	excellent (Fig. 1)
Si–{Si_3N_4}(t)–SiO_2	t=0.4 nm: 3.3	poor (Fig. 2)
t=oxide layer thickness	t=0.8 nm: 3.4	poor (Fig. 2)
Si–N–SiO_2 {1 monolayer}	2.8	excellent (Ref. 18)
Si–$(SiO_2)_{0.9}\{Si_3N_4)_{0.1}$	10 at. % N: 2.9	very good (Ref. 31)
Si–$(SiO_2)_{0.7}\{Si_3N_4)_{0.3}$	26 at. % N: 3.1	poor (Ref. 31)
Si–TiO_2}[a] (1.5 molecular layers)	4.0	unreported
Si–Ta_2O_5}[b] (1.5 molecular layers)	3.5	unreported
Si–Al_2O_3}[c] (1.5 molecular layers)	3.6	unreported

[a]Average coordination: [Ti]=6, [O]=3.0 (rutile/anatase bonding).
[b]Average coordination: [Ta]=6, [O]=2.4 (Ref. 32).
[c]Average coordination: Al=[4.5], [O]=3.0 (3:1 ratio of tetrahedral to octahedral sites, see Ref. 33).

firms that Si–SiO_2 interfaces are expected to display excellent interface properties ($N_{av}\sim 2.8$), whereas Si–Si_3N_4 interfaces are not ($N_{av}\sim 3.5$). Equally important, the calculations demonstrate that interposition of ultrathin SiO_2 layers between Si and Si_3N_4 results in values of $N_{av}\sim 3$ consistent with very good electrical properties as in Fig. 1, whereas interposition of ultra thin Si_3N_4 layers between Si and SiO_2 results in $N_{av}>3$ and degraded electrical performance as in Fig. 2 Figure 6 is based on the data of Figs. 1 and 2, and demonstrates that defect scaling for bulk films, Eq. (1), also holds at interfaces. The defect densities, in this case increases fixed positive charge, Δq_f,

$$\Delta q_f = C_{ox}\Delta V_{th}(\text{or } \Delta V_{fb}), \tag{3}$$

scale with shifts of the flatband or threshold voltages (ΔV_{fb} or ΔV_{th}), from values anticipated from considerations of the substrate doping and gate electrode material; is the oxide capacitance, $C_{ox}\sim C_{acc}$.

The model can also been applied to interfaces between Si and (i) silicon oxynitride alloys and (ii) alternative high-K dielectrics. For oxynitride alloys a maximum interfacial N incorporation for low defects of ~15 at. % as indicated by the value of N_{av} in Table II; is consistent with results of Vogel *et al.*[31] The model calculations for Ta_2O_5, TiO_2, and Al_2O_3 explain the necessity for ultrathin SiO_2 layers between the Si substrate and these high-K oxides (see, for example, Ref. 17). The model thereby establishes important limitations for extension for gate dielectric interfaces other than Si–SiO_2. Specifically: (i) Si_3N_4 cannot be substituted for SiO_2 at Si substrates; and (ii) substitution of more highly coordinated high-K dielectrics such as Ta_2O_5, etc., requires SiO_2, or monolayer nitrided SiO_2 interfaces, thus establishing a limitation on the extent to which $t_{ox\text{-}eq}$ can be reduced below 1 nm.

V. EXTENSION OF CONSTRAINT THEORY TO INTERNAL DIELECTRIC INTERFACES

As noted above, the integration of high-K alternative gate dielectrics into aggressively scaled devices for ultrahigh density integrated circuits will require compliant interface layers to bridge reduce bond constraint interfacial defects. It is

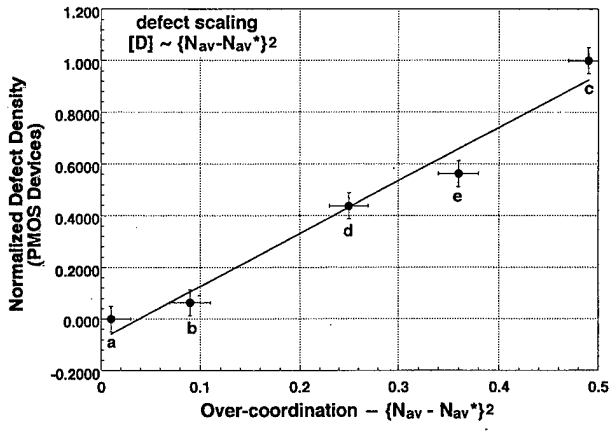


FIG. 6. Plot of normalized defect density as a function of $\{N_{av}-N^*_{av}\}^2$, where N^*_{av} is the value for an ideal Si–SiO_2 interface. Data points a, b, and c are from Fig. 1, and d and e from Fig. 2, and the solid line is a linear regression analysis of the data points.

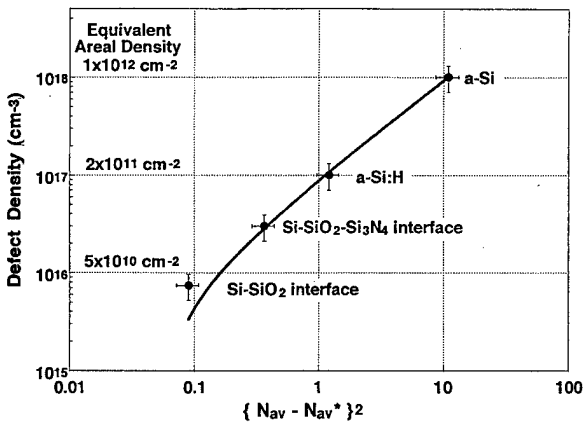


FIG. 7. Log-log plot of volume defect density vs $\{N_{av}-N^*_{av}\}^2$, where N^*_{av} =2.4 for ideal three-dimensional network structures. The solid line is a linear regression analysis of the data points.

therefore necessary to consider bond-constraint induced defects at the internal dielectric layer and the high-K materials. This has been studied in ONO stacks, where the density of positively charged defects at ON interfaces, well separated from the Si–SiO_2 interface (~5 nm), was found to be $\sim 3 \times 10^{11}\,cm^{-2}$ for stacked dielectrics deposited by RPECVD at 300 °C, and annealed in inert ambients at 900 °C.[2] Similar defect densities have been reported[20] for internal ON interfaces even when the SiO_2 layer thickness is reduced to <1 nm.

A two component model similar to the one developed earlier in this article for the Si-dielectric interface structure has been used as a basis for this calculation. In this model, N_{av} is simply the average of the N_{av} values of the interface constituents, so for SiO_2–Si_3N_4 internal interfaces, $N_{av}\sim 3$, consistent with the interfacial defect concentrations in the low $10^{11}\,cm^{-2}$ regime. Since N_{av} values are also ~3.5 for other candidate high-K materials such as Al_2O_3 and Ta_2O_5, similar interfacial defect densities are anticipated. This in agreement with experimental results reported for stacked SiO_2–Ta_2O_5 dielectrics.[17] The effect of these interfacial defects on performance and reliability has yet to be fully evaluated, but it is expected to a factor due to the large change in the longitudinal electric field at internal dielectric interfaces between SiO_2 and alternative high-K dielectric materials. The electric field in the SiO_2 will be the same as that calculated for a homogeneous oxide with a thickness equal to $t_{ox\text{-}eq}$, whereas the field in the high-K dielectric is reduced by the dielectric constant ratio.

VI. SUMMARY

It has been shown how constraint theory originally developed for bulk glasses[13,14] and thin films,[15] can be extended to (i) interfaces between crystalline Si and noncrystalline gate dielectric materials such as SiO_2, Si_3N_4, etc., and (ii) internal interfaces in stacked gate dielectrics such as SiO_2–Si_3N_4, SiO_2–Ta_2O_5. For the thin film materials, and the two interfacial materials systems, $N_{av}\sim 3$ represents a demarcation between device-quality and increasingly defective materials and interfaces. This is illustrated in Fig. 7, which is a log-log plot of defect density versus $\{N_{av}-N_{av}^*\}^2$. The data in this plot include different thin film materials and dielectric interfaces. By including both thin film a-Si and a-Si:H, the general trend in Fig. 7 establishes the universality of the approach to interfacial bonding constraints developed in this article.

ACKNOWLEDGMENTS

Research at NC State University is supported in part by the ONR, the AFOSR, the NSF, the SRC, and SEMATECH.

[1]H. Y. Yang, H. Niimi, and G. Lucovsky, J. Appl. Phys. **83**, 2327 (1998).
[2]H. Y. Yang and G. Lucovsky, IEEE Electron Device Lett. (submitted).
[3]Y. Ma, T. Yasuda, and G. Lucovsky, J. Vac. Sci. Technol. A **11**, 952 (1993); Y. Ma, T. Yasuda, S. Habermehl, and G. Lucovsky, J. Vac. Sci. Technol. B **11**, 1533 (1993); Yi Ma, T. Yasuda, and G. Lucovsky, Appl. Phys. Lett. **64**, 2226 (1994).
[4]S. V. Hattangady, H. Niimi, and G. Lucovsky, J. Vac. Sci. Technol. A **14**, 3017 (1996).
[5]C. G. Parker, G. Lucovsky, and J. R. Hauser, IEEE Electron Device Lett. **19**, 106 (1998).
[6]Y. Wu and G. Lucovsky, IEEE Electron Device Lett. **19**, 367 (1998).
[7]V. Misra *et al.*, IEEE Trans. Electron Device (submitted).
[8]B. Y. Kim, H. F. Luan, and D. L. Kwong, Tech. Dig. Int. Electron Devices Meet. 463 (1997).
[9]Y. Shi, X. Wang, and T. P. Ma, IEEE Electron Device Lett. **19**, 388 (1998).
[10]Z. Lu, M. J. Williams, P. F. Santos-Filho, and G. Lucovsky, J. Vac. Sci. Technol. A **13**, 607 (1995).
[11]G. Lucovsky, Z. Jing, P. Santos-Filho, G. Stevens, and A. Banerjee, J. Non-Cryst. Solids **198–200**, 19 (1996).
[12]P. Santos-Filho, G. Stevens, G. Lucovsky, T. Cull, P. Fedders, P. Leopold, and R. Norberg, J. Non-Cryst. Solids **198–200**, 77 (1996).
[13]J. C. Phillips, J. Non-Cryst. Solids **34**, 153 (1979).
[14]J. C. Phillips, J. Non-Cryst. Solids **47**, 203 (1983).
[15]G. Lucovsky and J. C. Phillips, J. Non-Cryst. Solids **227**, 1221 (1998).
[16]S. Campbell, D. Gilmer, X. Wang, M. Hsieh, H. Kim, W. Gladfelter, and J. Yan, IEEE Trans. Electron Devices **44**, 104 (1997); X. Guo, T. P. Ma, T. Tamagawa and B. L. Halpern, Tech. Dig. Int. Electron Devices Meet. 377 (1998).
[17]A. Chatterjee *et al.*, Tech. Dig. Int. Electron Devices Meet. 777 (1998).
[18]H. Niimi, H. Y. Yang, and G. Lucovsky, AIP Conf. Proc. **449**, 273 (1998).
[19]D. R. Lee, G. Lucovsky, M. R. Denker, and C. Magee, J. Vac. Sci. Technol. A **13**, 1671 (1995).
[20]Y Wu and G. Lucovsky, IEEE Trans. Electron Devices (submitted).
[21]X. Feng, W. J. Bresser, and P. Boolchand, Phys. Rev. Lett. **78**, 4422 (1997).
[22]J. E. Huheey, *Inorganic Chemistry* (Harper and Row, New York, 1978), Chap. 17.
[23]M. Stutzmann, *Properties of Amorphous Silicon*, 2nd ed. (INSPEC, London, 1989).
[24]J. C. Phillips, in *Rigidity Theory and Applications*, edited by M. F. Thorpe and P. Duxbury (Michigan State University Press, East Lansing, 1999) (to be published).
[25]J. H. Van der Merwe, J. Appl. Phys. **34**, 123 (1963).
[26]S. Y. Ren and J. D. Dow, J. Electron. Mater. **26**, 341 (1996).
[27]H. Z. Massoud, Mater. Res. Soc. Symp. Proc. **105**, 265 (1988); H. Z. Massoud, J. Appl. Phys. **63**, 2000 (1988).
[28]G. Lucovsky and H. Z. Massoud, J. Vac. Sci. Technol. B **16**, 2191 (1998).
[29]G. Lucovsky, Y. Wu, H. Niimi, V. Misra, and J. C. Phillips, Appl. Phys. Lett. **74**, 2005 (1999).
[30]C. H. Bjorkman, J. T. Fitch, and G. Lucovsky, Appl. Phys. Lett. **56**, 1983 (1990).
[31]E. Vogel, P. McLarty, and J. Wortman, IEEE Trans. Electron Devices **43**, 753 (1996).
[32]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.
[33]G. Lucovsky, A. Rozaj-Brvar, and R. F. Davis, in *The Structure of Non-Crystalline Materials 1982*, edited by P. H. Gaskell, J. M. Parker, and E. A. Davis (Taylor and Francis, London, 1983), p. 193.

Suppression of boron transport out of p^+ polycrystalline silicon at polycrystalline silicon dielectric interfaces

Y. Wu
Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

H. Niimi
Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

H. Yang
Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695-8202

G. Lucovsky[a)]
Departments of Electrical and Computer Engineering, Materials Science and Engineering, and Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

R. B. Fair
Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708-0291

(Received 19 January 1999; accepted 7 May 1999)

The transport of B atoms out of p^+ polycrystalline silicon (poly-Si) gate electrodes through SiO_2 gate oxides to the Si–SiO_2 interface during dopant activation anneals degrades performance and reliability of hole-conducting (p-channel) field effect transistors. This article studies the suppression of B atom transport by using remote plasma processing to form ultrathin Si_3N_4 and silicon oxynitride diffusion barrier layers between p^+ poly-Si gate electrodes and SiO_2 gate dielectrics. Suppression of B atom transport has been monitored through electrical measurements, demonstrating that ~0.8 nm of Si_3N_4, equivalent to a N areal density of $\sim 4.5\times 10^{15}$ atoms cm^{-2}, is sufficient to *effectively* suppress B out diffusion during aggressive anneals of ~1 min at 1000 °C. The suppression and transport mechanisms in nitrides, oxides, and oxynitrides have been studied by varying the N atom areal density by alloying. Quantum chemistry calculations suggest that B transport occurs through the formation of donor-acceptor pair bonds between B^+ ions and nonbonding electron pairs on oxygen atoms with the transport process requiring a connected O atom *percolation* pathway. Donor-acceptor pair bonds with B^+ ions are also formed with N atoms in nitrides and oxynitride alloys, but with a binding energy more than 1.5 eV higher than B^+ ion O-atom bonds so that nitrides and oxynitride alloys effectively block B diffusion through the formation of a deep trapping site. © *1999 American Vacuum Society.* [S0734-211X(99)05104-5]

I. INTRODUCTION

Aggressive scaling of complementary metal–oxide–semiconductor gate oxide thickness to below 2.0 nm for sub-250 nm devices is necessary to maintain drive currents and improve short channel behavior in field effect transistors (FETs). For p-channel devices with p^+ poly-Si gate electrodes, B transport through ultrathin gate oxides becomes an important issue for gate dielectric integrity. The out diffusion of B and subsequent transport occur during the high temperature anneals ($T_{an}>950$ °C) required to activate B dopant atoms implanted into poly-Si gate electrodes (Ref. 1, and references therein). The B diffusion process involves transport across the p^+ poly-Si–SiO_2 interface and through the thin SiO_2 gate oxides to the Si–SiO_2 interface. An effective way to suppress B atom transport out of the p^+ poly-Si gate electrodes is to interpose a dielectric barrier layer such as Si_3N_4, or an oxynitride alloy $[(SiO_2)_x(Si_3N_4)_{1-x}]$ between the SiO_2 dielectric and the poly-Si gate electrode.[2–4] In contrast, the formation of oxynitride layers at the Si–SiO_2 interface leads to decreased reliability, since B can diffuse through the oxide layer and pile up at the Si–SiO_2 interface.[5]

The diffusion of B through SiO_2 has been studied experimentally and modeled (Ref. 1, and references therein). In addition, it has been shown by x-ray photoelectron spectroscopy (XPS) that B–O bonds are present in oxide films into which B has diffused.[6] Analysis of experimental B diffusion data yields an activation energy for diffusion of ~3.6 eV. It was proposed in Ref. 1 that the B transport pathways are through peroxy (Si–O–O–Si) bonding defects. In the so-called *peroxy linkage defect* (PLD) controlled diffusion model, B atoms insert into peroxy defect sites which provide the transport pathway. The insertion reaction is given by

$$B+Si–O–O–Si\rightarrow Si–O–B–O–Si, \quad (1)$$

where the Si–O–O–Si is the peroxy bonding defect. For transport above the visco-elastic temperature of SiO_2 of ~950–1000 °C, bond breaking and network reorganization can readily occur; however, B diffusion has been studied at

a)Corresponding author; electronic mail: gerry_lucovsky@ncsu.edu

temperatures well below 950 °C, with no change in the activation energy or prefactor terms in the diffusion constant down to temperatures as low as 800 °C. Additionally there is no independent structural or spectroscopic evidence for the required concentrations of PLDs in thermally grown, or plasma deposited and annealed SiO_2, or $(SiO_2)_x(Si_3N_4)_{1-x}$ alloy films. The experimental data presented in this article relating to B transport through oxynitride alloys suggests that there is an alternative microscopic process that involves the motion of B^+ ions by hopping between near-neighbor O atoms, and the blocking of boron transport by bonding of B^+ ions at N atoms that act as deep traps. In this respect, the diffusion process model proposed in this article parallels charged carrier transport via a trap controlled drift mobility in the presence of deep trapping.[7] There is one major difference, in electrical conduction process, an applied voltage is necessary, whereas in the diffusion process analog, the driving force is an atomic concentration gradient. This article quantifies the effectiveness of both nitride and oxynitride barrier layers prepared by remote plasma-assisted processing.

Remote plasma processing and device fabrication are discussed in Sec. II. Suppression of B atom transport has been quantified in metal oxide semiconductor (MOS) devices prepared by two different processing approaches: (i) by deposition of ultrathin nitride barrier layers by remote plasma enhanced chemical vapor deposition (PECVD) at 300 °C,[8] and (ii) by a 300 °C plasma-assisted top surface oxide nitridation.[9] Both of these processes are followed by annealing at 900 °C prior to poly-Si deposition, B implantation and dopant activation. The extent to which B atom transport is suppressed has been studied in Sec. III by monitoring shifts in electrical properties using devices with oxide dielectrics as a control.[10] These studies demonstrated that ~0.8 nm of nitride was sufficient to *completely* suppress B atom out diffusion out of p^+ poly-Si gate electrodes as monitored electrically. To gain additional insights into the B transport process, experiments were performed in which B transport was studied as a function of the oxide to nitride ratio in silicon oxynitride alloys. In one series of experiments discussed in Sec. IV, the thickness of the oxynitride layer was fixed at ~0.8 nm, and the alloy fraction of silicon dioxide set at 30 and 70 at. %, reducing the areal density of N atoms. In a second series of experiments the areal density of N atoms was fixed at the same level as in an 0.8 nm nitride film, and the thickness was increased by oxynitride alloy formation. In Sec. V, a quantitative model for the B transport based on a quantum chemistry model is presented. The quantum chemistry calculations are consistent with B transport through SiO_2 and the oxynitride alloys occurring through the formation of donor-acceptor pair bonds between B^+ ions and nonbonding electrons pairs of the oxygen atoms. This transport mechanism requires a connected *percolation* pathway of oxygen atoms. The model calculations further demonstrate B^+ ion transport is blocked by N atoms in the nitride films and oxynitride alloys through the formation of donor-acceptor pair bonds between B^+ ions and the nonbond electron pair on a nitrogen atom. The binding energy of these B^+–N atom bonds is more than 1.5 eV greater than that of the B^+–O atom bond, so that B^+ ions are trapped at N atom sites, effectively blocking transport at the annealing temperatures.

II. DEVICE FABRICATION BY REMOTE PLASMA PROCESSING

Two different types of device structures were used in these studies, (i) *p*-channel FETs and MOS capacitors with dual layer oxide/nitride gate dielectrics in which nitride layers were formed by remote plasma deposition,[8,11] and the oxide layers were either thermally grown, or formed by remote plasma-assisted oxidation,[12] and (ii) MOS capacitors formed by remote plasma-assisted oxidation, nitridation, and oxide deposition processes in which the nitride blocking layers were formed by converting the top surface of the deposited oxide to an ultrathin nitride by a plasma-assisted nitridation process.[9,13]

A. Deposition of nitride films

Two different remote plasma enhanced chemical vapor deposition (RPECVD) processes have been used for deposition of bulk nitride films: one process used NH_3 as the N-atom source gas,[11] and the other used N_2.[8] Both processes were performed at a process pressure of 300 mTorr and a temperature of 300 °C and resulted in films with as-deposited bonded hydrogen concentrations, [H]~20–30 at. %. Following a 900 °C rapid thermal annealing (RTA) in an inert, nonoxidizing ambient, the bonded hydrogen was reduced by about a factor of two to about 10–15 at. %.[8,11,14–16] Initial evaluation of these nitride films was performed in devices with tri-layer dielectrics in oxide–nitride–oxide stacks.[11] Following the 900 °C anneal, the electrical performance of devices with these stacks, including interfacial and bulk defects, were equivalent to those of devices with thermally grown oxides, and the reliability was improved with respect to devices with thermally grown oxides.[11] Since other studies of *n*-channel FETs with stacked oxide/nitride gate dielectrics indicated no differences in performance or reliability for nitride films prepared with NH_3 and N_2 source gases, the nitride films of this article utilized the N_2 source gas process in order to reduce their initial bonded H content.[8,10]

A secondary mass spectrometry (SIMS) depth profile analysis was performed on a dual layer nitride/oxide dielectric in which ~0.8 nm of Si_3N_4 was deposited on to 4.0 nm of thermally grown SiO_2 (see Fig. 1). The trailing of the nitrogen signal into the oxide is a result of the SIMS analysis method. After annealing at 900 °C for 30 s, a distinct nitrogen peak appears at the oxide/silicon interface, showing that N-atoms diffuse into the oxide, and pile up at the oxide/silicon interface during the anneal, so that the annealed devices have a "NON" stacked structure. The concentration of N at the Si-dielectric interface is of the order of 2–5 at. %, and is considerably less than has been achieved by the plasma nitridation process of Ref. 12 (see Sec. III B and Fig. 5).

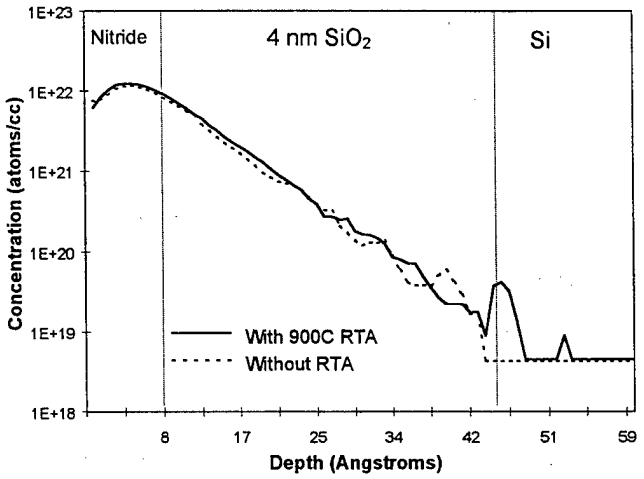


FIG. 1. Nitrogen depth profile determined by SIMS analysis for a nitride/oxide (~0.8 nm/4.0 nm) dual layer gate dielectric before and after a 900 °C post-deposition RTA.

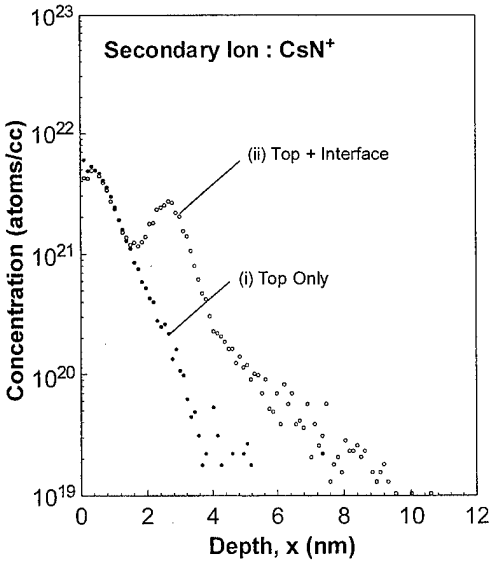


FIG. 2. Nitrogen depth profile determined by SIMS analysis comparing top plasma-assisted top surface nitridation, and the combination of plasma-assisted top surface and interface nitridation.

B. Top surface nitridation

Top surface plasma-assisted nitridations were performed with substrates at 300 °C on thermally grown and plasma-deposited SiO_2 films in a thickness range from 2.5 to 10 nm.[9,13] A remote He/N_2 discharge was initiated at a process pressure of 100 mTorr (0.1 Torr) which allows the plasma *after glow* to penetrate into the processing chamber. Following nitridation, RTA at 900 °C was performed on-line in a multichamber system at a pressure of 0.5 Torr in an Ar ambient for a time of approximately 40 s. On-line Auger electron spectroscopy was performed in the analysis chamber at intermediate stages of processing. Angle resolved XPS was performed *ex situ* using an Al $K\alpha$ x-ray source (1.486 keV) in both normal (detector at ~80° with respect to sample surface) and glancing-angle modes (detector at ~20° with respect to sample surface). SIMS analyses were performed *ex situ* using Cs^+ ions, and detecting $CsSi^+$, CsN^+, and CsO^+ species for analysis.[9,13] These experiments demonstrated that nitrogen atoms were bonded at the top surface of the oxide film, effectively converting that surface region to either a nitrogen rich oxynitride alloy or an ultrathin nitride layer (see Fig. 2). Test devices were prepared using a 20 min plasma exposure time which formed a Si_3N_4 layer ~0.8 nm thick on top of the underlying oxide film.[17]

C. Fabrication of *p*-channel FETs and capacitors

p-channel FETs with p^+-poly-Si gate electrodes were fabricated on *n*-type Si(100) substrates doped with phosphorus to 5×10^{17} cm^{-3}. Some wafers were implanted with phosphorous to increase the channel doping to 1.1×10^{18} cm^{-3}. Device processing was by standard lithography and patterning techniques. After growing a field oxide and creating wells for device fabrication, sacrificial thermal oxides were grown at 800 °C, annealed at 900 °C, and removed by rinsing in dilute HF prior to growth of gate oxide layers. Bottom oxides were grown by either (i) thermal oxidation in oxygen with 4.5% HCl at 800 °C for thicknesses of 1.5–4.0 nm, or (ii) remote plasma oxidation in N_2O at 300 °C to an approximate 0.6 nm thickness. This was followed by RPECVD nitride depositions in a range from ~0.2 to ~2.4 nm.[18] Boron was implanted into poly-Si gate electrodes deposited by conventional CVD, and into the source and drain regions of the device structure at energy of 20 MeV to a level of $5\times10^{15}/cm^2$. This was followed by the deposition of 200 nm of a low temperature CVD oxide to prevent the boron out-diffusion during dopant activation anneals. The anneals ranged from 950 to 1000 °C for 1 min. After deposition and patterning of Al for gate, source and drain contacts, a conventional 30 min post metallization anneal in forming gas (N_2/H_2) at 400 °C was performed. The equivalent oxide electrical thickness was determined by capacitance–voltage ($C-V$) measurements with devices biased into accumulation using quantum mechanical corrections.[19] *p*-channel FETs with thermal oxide gate dielectrics were fabricated as control devices.

For the top surface nitridation studies, MOS capacitors were formed on lightly doped *n*-type substrates ($n\sim5\times10^{15}$ cm^{-3}) using a combination of remote plasma-assisted oxidation, deposition, and nitridation to form two types of test structures.[13] Both devices had top surface nitride layers prepared by the process outlined earlier in Sec. II B, and plasma-deposited oxide layers ~2.5 nm thick. However, the $Si-SiO_2$ interfaces were formed differently: (i) in one group of devices, the interface was formed by remote plasma-assisted oxidation using excited species from a remote O_2/He plasma, and (ii) in the second group of devices, the oxidation process of (i) was followed by an interface nitridation process that utilized species from a remote N_2/He plasma.[13] The first interface formation process produced a non-nitrided interface and ~0.6 nm of SiO_2 which served as a substrate for film deposition by RPECVD using remote

plasma excited O_2/He and downstream injected SiH_4. The second process also produced ~0.6 nm of SiO_2, but following the 90 s nitridation process, had ~1 monolayer (or equivalently $\sim 7-8\times10^{14}$ N-atoms/cm^2) incorporated at the interface between the Si and the ultrathin SiO_2 layer. Capacitors were fabricated after top surface nitridation and p^+ poly-Si gate electrodes were applied following the same procedures as used for the p-channel FETs.

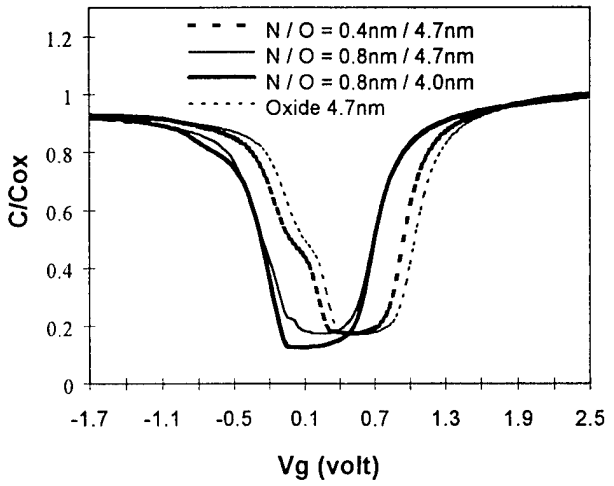
III. ELECTRICAL MEASUREMENTS AND DATA ANALYSIS

A. Dual layer *p*-channel devices

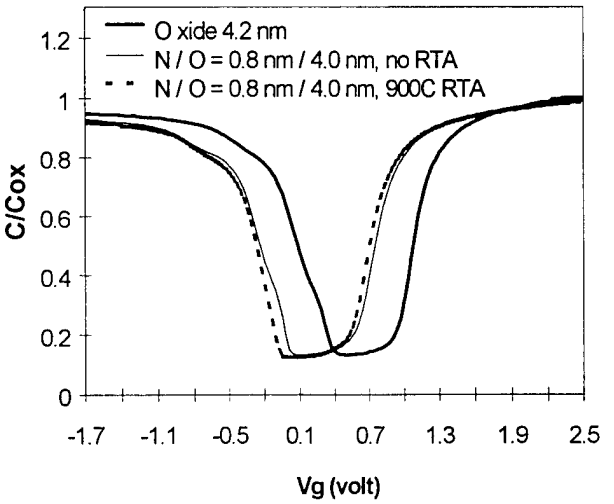
$C-V$ curves for p-channel devices with a 0.8 nm ultrathin nitride layer on top of thermally grown oxides indicated a significant suppression of B-atom penetration to the Si–SiO_2 interface compared to p-channel devices with thermally grown oxide dielectrics.[20] $C-V$ measurements were made on capacitors formed using the same process as described earlier for the p-channel FETs. Figure 3(a) contains the superposition of normalized quasistatic $C-V$ curves for capacitors with control oxides, and NON stacked dielectrics in which the deposited nitride layers are ~0.4 or ~0.8 nm thick. Oxide thickness for these devices ranged from ~4.0 to ~4.7 nm. The $C-V$ curve for control oxide is shifted to a more positive voltage by approximately +0.5 V with respect to the capacitor with 0.8 nm top nitride. Based on the calculated value of flatband voltage as determined from the substrate and gate electrode doping, the large positive shift of the reference oxide flat band voltage indicates significant B-penetration to the substrate for that device.[21] The flat band voltage shift of capacitor with 0.4 nm top nitride film is intermediate, indicating that B-atom transport can be controlled for a fixed thermal budget by simply changing the top nitride layer thickness.

In addition, the B-atom blocking capability is improved by performing the 900 °C RTA prior to the deposition of the polycrystalline Si gate electrode as indicated by the $C-V$ data in Fig. 3(b). Annealing of the RPECVD nitride film at 900 °C for 30 s prior to the poly-Si deposition, implantation, and activation anneal, retards the diffusion of B through the top nitride during the activation anneal, presumably by reducing the amount of available hydrogen to assist in the diffusion process.[1] As displayed in Fig. 3(b), the $C-V$ curve for stacked dielectric film without 900 °C annealing shows a small shift to positive voltage with respect to film with the post-deposition RTA. This indicates a small amount of B has migrated through the nonannealed nitride film to the substrate. Additionally, this $C-V$ curve exhibits some distortion at the onset of the inversion region, presumably due to modification of the channel region potential by B-atom compensation.

Locating a nitride film on top of oxide as a diffusion barrier provides other advantages in p-channel devices. Previous studies have shown that N-atom concentration peaks at the Si–SiO_2 interface for dielectrics grown by oxidation in NO or N_2O. Since B-atom transport is stopped at the



(a)



(b)

FIG. 3. (a) Normalized quasistatic $C-V$ curve for thermal oxide, and 0.4 and 0.8 nm top nitride deposited onto thermal oxides. Curves are shifted due to boron penetration through thin gate material. The dopant activation annealing is 1000 °C for 60 s. (b) Normalized quasistatic $C-V$ curve for oxide and N/O stack with and without post-deposition treatment. The RTA condition is at 900 °C for 30 s in the He ambient, and the dopant activation annealing is 1000 °C for 60 s.

Si–SiO_2 interface in these devices, this approach allows an accumulation of B atoms in the bulk oxide, degrading the dielectric reliability compared to oxides grown in O_2 without nitrogen incorporation.[22] By preventing B from diffusing into the bulk oxide layer, devices with NON dielectrics display improved reliability compared to devices with oxide grown in O_2. Figure 4 shows Weibull plots for gate dielectrics with t_{ox-eq} ~3.0 nm under a substrate injection stress of 500 mA cm^{-2}. The charge to breakdown, Q_{bd}, is improved by about an order of magnitude in the devices with the NON dielectrics.

It has been shown earlier that a 0.8 nm top nitride is effective in stopping boron penetration even with a thermal

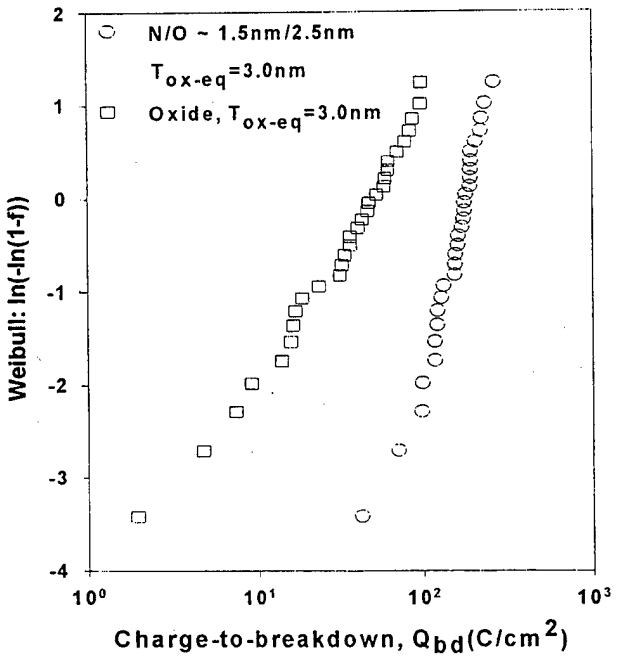


FIG. 4. Weibull plots of Q_{bd} for capacitors with 3.0 nm ON and oxide dielectrics under constant substrate injection current stressing at a current density of 500 mA cm^{-2}.

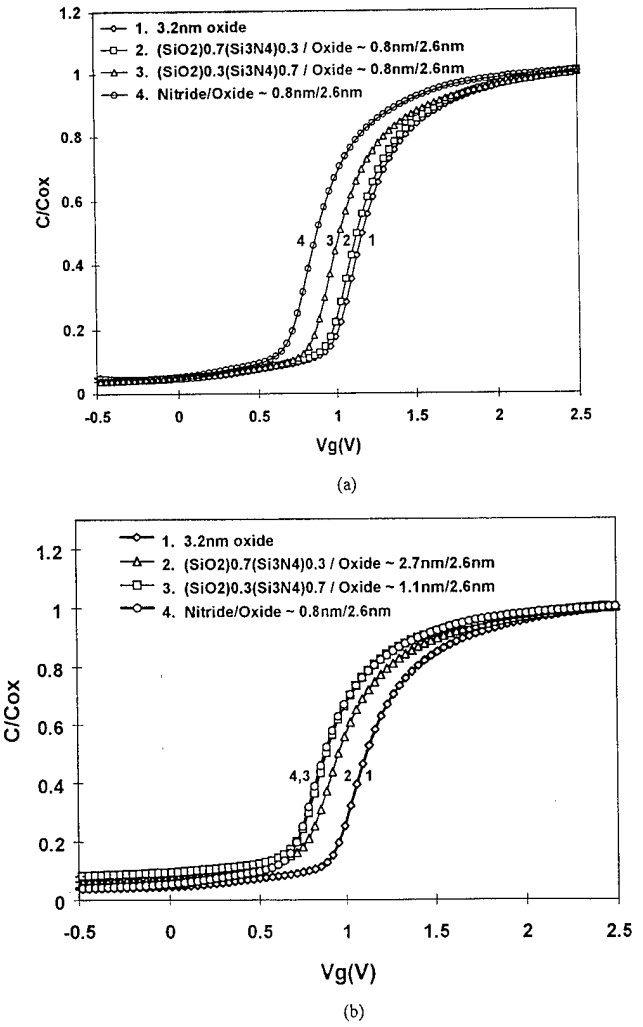


FIG. 5. Boron penetration as monitored by changes in flatband voltage for oxynitride alloys, $(SiO_2)_x(Si_3N_4)_{1-x}$ ($x=0$, 0.3, 0.7, and 1), deposited onto a 2.5 nm oxide. The alloys were formed by a RPECVD technique described in Ref. 8. In (a) the thickness of the oxynitride alloy is fixed at ~0.8 nm and the nitrogen concentration was reduced by alloying. In (b) the areal density of nitrogen atoms is fixed, and the barrier layer thickness is increased proportionally.

budget as high as 1000 °C for 4 min.[20] However, as shown in Fig. 3(a), some B-atom penetration is observed when the top nitride thickness is reduced to 0.4 nm. It is interesting to note that 0.8 nm of nitride is equivalent to an areal density of $\sim 4.5\times 10^{15}$ N atoms/cm^{-2}. If one assumes that B transport is impeded by N atoms, then an estimate of the cross section for a transport blocking reaction such as chemical capture as discussed later can be obtained by using the Si–N bond length, d_{Si-N} of 0.175 nm to estimate the size of the chemical trapping site. Taking the radius of this trap to be equal to $0.5d_{Si-N}$, the product of $\pi(0.5d_{Si-N})^2$ and the areal density is ~1.1, and is reduced proportionally for the 0.4 nm film. It is difficult to explain this thickness dependence of B-atom transport through these very thin nitride layers by conventional diffusion theory (Ref. 1, and references therein). To gain additional insights in the B transport process, B penetration to the Si–SiO_2 has been studied by changes in flat band voltage for devices with different types of oxynitride alloys deposited onto a 2.5 nm oxide. The $(SiO_2)_x(Si_3N_4)_{1-x}$ alloys were formed by a RPECVD technique described in Ref. 8. The SiO_2 fractions, x, in the $(SiO_2)_x(Si_3N_4)_{1-x}$ alloys were 0, 0.3, 0.7, and 1. In the first set of experiments, the thickness of the oxynitride alloy films was fixed at ~0.8 nm and the N concentration was reduced by increasing x. In the second set of experiments, the areal density of nitrogen atoms is fixed at $\sim 4.5\times 10^{15}$ cm^{-3} (as in a 0.8 nm Si_3N_4 film) and the oxynitride film thickness was increased proportionally as the SiO_2 fraction was increased. Figure 5(a) shows the effectiveness of the boron diffusion barrier formed by 0.8 nm of the $(SiO_2)_x(Si_3N_4)_{1-x}$ alloys. The extent of B-atom penetration is compared by studying the flat band voltage shifts from the theoretical value determined by the work function difference between gate and substrate Si. In the Fig. 6(a), the device with 0.8 nm top nitride films shows essentially no flat band voltage shift (to ±0.05 V) confirming the results shown in Figs. 3 and 4. By reducing the nitrogen areal density by substituting $(SiO_2)_x(Si_3N_4)_{1-x}$ alloys for the nitride layer, and keeping the physical thickness at 0.8 nm, the flat band voltage is shifted to more positive voltages indicating increased B-atom penetration. Similar results apply for the second set of experiments in which the nitrogen areal density is fixed by increasing the thickness of the $(SiO_2)_x(Si_3N_4)_{1-x}$ alloy films [see Table I and Fig. 5(b)]. Reduced B-atom penetration through the second set of films supports a model in which boron transport proceeds via a percolation-like process involving the connectivity of the oxygen atom pathways through the oxynitride alloys. This model is qualitatively dif-

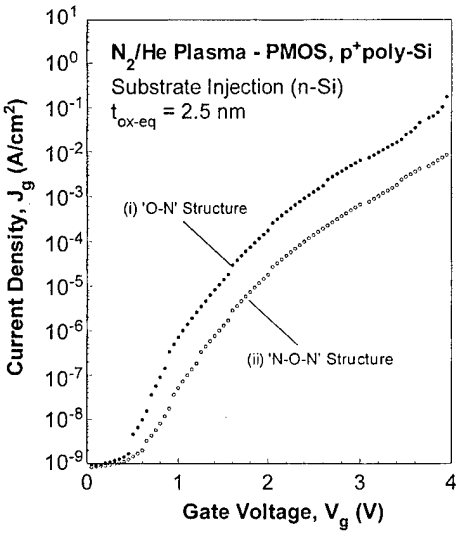


FIG. 6. $J-V$ characteristics for capacitors prepared on lightly doped n-type Si(100) with p^+ polycrystalline Si gate electrodes for $t_{ox-eq} \sim 2.5$ nm. The upper curve is for a device with an RPECVD SiO_2 layer and a non-nitrided $Si-SiO_2$ interface, and the lower curve is for a device with a RPECVD SiO_2 layer and a monolayer nitrided $Si-SiO_2$ interface.

ferent that the model for B diffusion, as proposed in Ref. 1, and will be discussed in more detail later.

B. $C-V$ current density–voltage ($J-V$) studies on capacitors

Electrical characterizations have also been performed on the capacitors subjected to top surface nitridation in order to demonstrate the equivalence of this approach for forming a barrier to B diffusion to the direct deposition of ~0.8 nm of Si_3N_4. In this case, tunneling $J-V$ data was used to demonstrate the effectiveness of top surface nitridation in suppressing B-atom transport out of p^+ poly-Si gate electrodes. Figure 6 shows $J-V$ data for the two different types of capacitors discussed above: (i) one with top surface nitridation and a non-nitrided interface, and (ii) the second with top surface nitridation and a monolayer nitrided interface. Flat band voltages for these devices, as determined by $C-V$ analysis, were the same to ± 0.05 V, indicating that B atoms have not been transported to the $Si-SiO_2$ interface in either of the devices. The reduction in tunneling current in the device with nitrided interface is similar to what has been shown found for devices with nitrided interfaces and either Al or n^+ poly-Si gate electrodes where B diffusion is not an issue.[13] Tunneling current in devices with interface nitridation is typically reduced by a factor of 7–10 due to decreases in effective surface roughness.[23]

TABLE I. Boron atom transport through oxynitride alloys as monitored electrically by flatband voltage shifts.

Gate dielectric	Flatband voltage, V_{fb}	Change in V_{fb}
N/O stack: 0.8 nm/2.6 nm	0.84 V	0.0 V
Reference 3.2 nm oxide	1.21 V	+0.37 V
(i) Fixed oxide (2.6 nm), and oxynitride (0.8 nm) alloy thickness		
Oxynitride alloys		
30% SiO2:0.8 nm	1.02 V	+0.18 V
70% SiO2:0.8 nm	1.15 V	+0.31 V
(ii) Fixed oxide thickness (2.6 nm), and variable oxynitride thickness		
Oxynitride alloys		
30% SiO2:1.1 nm	0.86 V	+0.02 V
70% SiO2:2.7 nm	1.09 V	+0.25 V

IV. MODEL FOR NITRIDE BARRIER LAYERS

The experimental results of Sec. III have established the following: (i) B penetration out of p^+ poly-Si gate electrodes following aggressive dopant activation anneals (~1 min at 950–1000 °C) of implanted B poly-Si is suppressed below levels of detection by electrical measurements when Si_3N_4 layers, prepared either by remote plasma CVD of Si_3N_4 or top surface nitridation of SiO_2 are about 0.8 nm thick, (ii) B penetration has been detected when the Si_3N_4 layer thickness is reduced to ~0.4 nm, and (iii) B penetration is occurs when oxynitride films $[(SiO_2)_x(Si_3N_4)_{1-x}]$ are substituted for the layers, even if the area density of N atoms is held constant by increasing the thickness of the oxynitride film in direct proportion to the SiO_2 fraction.

The combination of these results suggest that B transport is hindered at the atomic level when N atoms are substituted for O atoms in oxynitride alloys, and that B transport is by a percolation-like process that involves jumps between near, or nearest-neighbor O atoms in which B–O bonds are formed in accord with the results of XPS studies.[6] The model developed below proceeds in two stages: (i) quantum chemistry calculations do determine how B atoms and ions interact with O and N atoms in SiO_2, $(SiO_2)_x(Si_3N_4)_{1-x}$ and Si_3N_4, and (ii) transport calculations based on the connectivity of O pathways in $(SiO_2)_x(Si_3N_4)_{1-x}$ alloys.

A. Quantum chemistry calculations

Based on the electronic structure of B and two-fold coordinated O atoms in Si–O–Si bonding arrangements in SiO_2 and the pseudobinary oxynitrides alloys, a quantum chemistry study was made to determine if donor-acceptor pairs would form between B atoms and/or ions and the two-fold coordinated O atoms of the SiO_2 network. This bonding does not require a defect bonding environment and does not involve bond-breaking reactions and network restructuring. Figure 7(a) indicates the molecular model used to study donor-acceptor bond formation through attachment of B ions to the nonbonding pairs on the O atoms of the Si–O–Si groups. The calculations of this article are similar to those used to study defect metastability associated with H atoms at $Si-SiO_2$ interfaces,[24] and in hydrogenated amorphous silicon.[25] In particular, as indicated in the figure captions for Fig. 7 the Si atoms of the these clusters have been terminated Si–H groups. These calculations indicate that binding energy for attachment of neutral B atoms to the O nonbonding pair is very low, ~0.4 eV, but that the binding energy for attachment of B^+ ions to the O nonbonding pair is significantly higher, ~1.85 eV; the respective bond lengths are ~0.2 and

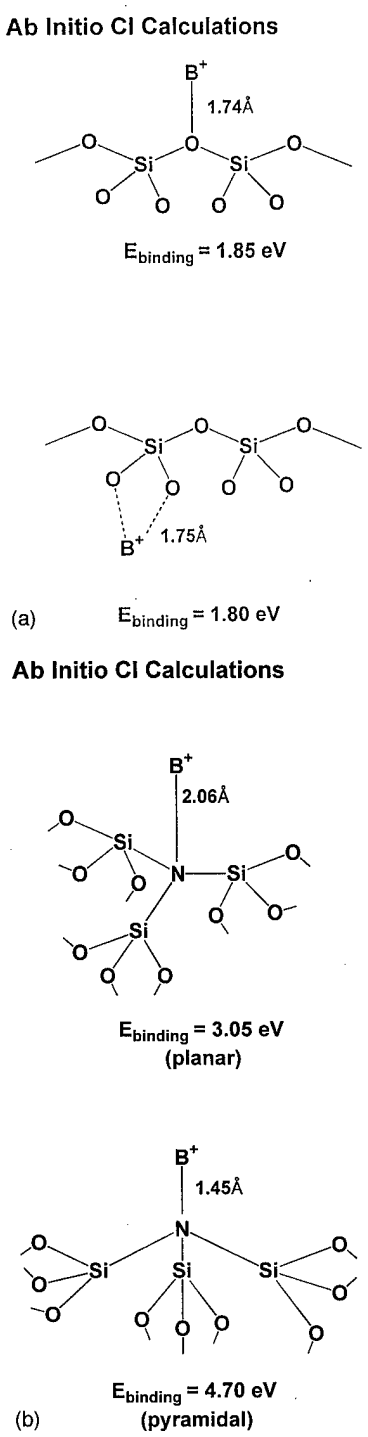


FIG. 7. Schematic representation of molecular structures used for quantum chemistry calculations: (a) for B^+ to Si–O–Si and (b) for B^+ to 3Si–N. (b) Local bonding geometry of the N atoms. The termination of these clusters is by H atoms.

0.175 nm. This calculation demonstrates that B atoms would not be impeded by attachment to O atoms, whereas B^+ ions would. Since the activation energy for B diffusion is high, ~3.6 eV, it can either involve bond breaking and restructing as in the PLD model of Ref. 1, or donor-acceptor bond formation involving ionic species as introduced in this article. The experiments on the relative blocking power of different silicon oxynitride alloys favor a site percolation model with near-neighbor jumps involving O atoms, and deep trapping or blocking involving N atoms. This approach is supported by the quantum chemistry calculations presented earlier, and those introduced later for B^+–N bonding. Therefore, the next part of this discussion focuses on (i) ion transport in SiO_2, and (ii) differences between the bonding of B^+ ions to nonbonding pairs on O and N atoms of SiO_2, $(SiO_2)_x(Si_3N_4)_{1-x}$ alloys and Si_3N_4.

While there is no direct experimental evidence that boron diffusion proceeds by ionic rather than neutral species, there is experimental evidence for formation of positive H^+ ions (protons) at Si–SiO_2 interfaces and their transport through SiO_2 films.[26] For example, studies of proton motion in SiO_2, where the source of H resides in poly-Si films deposited onto the SiO_2 surface, indicates that mobile proton motion is detected at temperatures >450 °C, and is the dominant H transport mechanism at temperatures >600 °C.[26,27] The transport of H^+ through SiO_2 may involve the formation of three-fold coordinated oxonium sites where the H^+ ion forms a donor acceptor pair bond through utilization of the nonbonding orbitals of the two-fold coordinated O atoms of Si–O–Si groups.[28] Since the ionization energy of H is greater than that for B, 13.6 eV as opposed to 8.3 eV,[29] it is reasonable to assume that B^+ ions can be generated at the poly-Si interface by a reaction similar to the one that generates H^+ ions. If this were the case, then the ionization energy of B atoms at the Si–SiO_2 interface would be reduced at least by the energy difference between the conduction band of Si and vacuum, ~4.2 eV. For this model to apply, the measured activation for diffusion[1] would have two contributions, (i) one associated with the ion formation process, and (ii) a second associated with the microscopic transport process. This separation into two terms is consistent with the activation energies for As and P diffusion in SiO_2 (see Ref. 1). In particular, the increased activation energies of As and P diffusion in SiO_2 with respect to B are consistent with their increased ionization, I, energies with respect to B (I_P=8.3 eV), I_P=10.5 eV, and I_{As}=9.8 eV.[29] However, the activation energy for the diffusion constant for Ga in SiO_2 is greater than that of B, but I_{Ga}=6.0 eV and is less than I_B so that the extension of this approach to other diffusing species such as Ga has issues yet to be resolved.

Quantum chemistry calculations have been applied to bonding between B^+ ions and N atoms bonded to three Si atoms (i) in a planar bonding environment like that found in Si_3N_4 and the oxynitride alloys,[29] and also (ii) in a more probable tetrahedral bonding environment as found in ammonium ions. In this regard, the bonding of N atoms in both crystalline and amorphous Si_3N_4 is in a planar configuration,[30,31] whereas the bonding of four-fold coordinated positively charged nitrogen is in either a tetrahedral or distorted tetrahedral configuration.[29] The calculation for the bonding of B^+ in a planar configuration is included to emphasize the increased stability of the tetrahedral (or distorted tetrahedral) arrangement. The neutral B to N binding energy in a planar configuration is too low (<0.5 eV) to provide a

TABLE II. Quantum chemistry calculations for interactions between B, P, and As atoms and ions and O and N atoms in Si–O–Si and 3Si–N bonding groups, respectively as in SiO_2, $(SiO_2)_x(Si_3N_4)_{1-x}$ alloys and Si_3N_4.

(i) B atom and ion bonding to O atoms in Si–O–Si groups [see Fig. 7(a)]		
Bonding group	Bond energy, eV	Bond length, nm
$B\cdots OSi_2$	<0.5	>0.2
$B^+\cdots OSi_2$	1.9	0.174
(ii) B atom and ion bonding to N atoms in 3Si–N groups [see Fig. 7(b)]		
Bonding group	Bond energy, eV	Bond length, nm
$B\cdots NSi_3$	<0.5	>0.2
Planar $B^+\cdots NSi_3$	3.0	0.21
Tetrahedral $B^+\cdots NSi_3$	4.7	0.15

chemical block, consistent with the results for B to O bonding where the proposed hopping transport mechanism can only apply if B^+ ions are the diffusing species. The results of the calculations for the B^+ ions are summarized in Table II, where it is shown that the binding energy of the B^+–N bond is more than 1.5 eV greater than that of the B^+O bond if the local geometry is planar, and more than 2 eV greater if the geometry is tetrahedral as in ammonium ions. Due to this difference in binding energies between B^+–N and B^+–O sites, the B^+–N sites can act as deep traps in the B diffusion process, and the B^+–O centers can act as jump sites with a significantly increased release rate determined by the difference in their respective binding energies. For a diffusion temperature of 1000 °C (1273 K), this is a factor of $\sim 5 \times 10^4$ for the planar B^+–N geometry, and more than 5×10^6 for the tetrahedral site expected for an ammonium ion-like center.

B. Percolation pathway calculations

It is now demonstrated that the experimental data on the blocking power of nitride layers, and the transport through the different oxynitride alloys is in good agreement with the predictions of a transport model based on site percolation. The model is based on a statistical analysis of bonding in the oxynitride alloys. Figure 8 presents the results of a calculation of the fractional concentration of the five possible local bonding groups in the oxynitride alloys. The calculation assumes that the alloy is pseudobinary (or chemically ordered) in character containing only Si–N and Si–O bonds. The Si atoms are four-fold coordinated in tetrahedral arrangements, the O and N atoms, are respectively two- and three-fold coordinated as in noncrystalline SiO_2[31] and Si_3N_4;[30,31] in particular the neutral N atoms are in a planar bonding geometry. As pointed out in Ref. 30, the bonding of neutral N atoms with Si atoms in amorphous materials is in planar bonding geometry just as in the crystalline phases. Infrared studies indicate that this same bonding geometry also occurs in silicon oxynitride alloys.[32]

Based on an assumption that B atom transport takes place through a connected array of oxygen atoms, the B transport proceeds through O pathways involving Si atoms bonded to two, three, or four oxygen atoms, and B transport is blocked

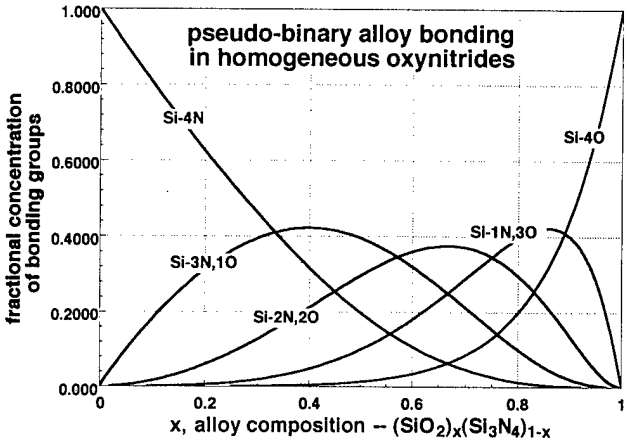


FIG. 8. Calculated fractional concentrations of Si–4N, Si–3N, 1O, Si–2N, 2O, Si–1N, 3O, and Si–4O tetrahedral groups as a function of x in silicon oxynitride alloys $[(SiO_2)_x(Si_3N_4)_{1-x}]$ using a statistical model.

by Si atoms bonded to either three or four N atoms. This definition of transport via the O atoms makes an *arbitrary assumption* that the transport is by a hopping process that involves jumps between O atoms bonded to the same Si atom, and that the probability of such a jump between available O atom sites is greater than the probability of trapping on a N atom bonded to the same Si atom. Obviously, these assumptions require additional critical study. Moving ahead, and based on these assumptions, Fig. 9 displays the fractional concentrations of these O *pathways*, and N *blockers*, as determined from Fig. 8. To determine the extent to which this type of model explains the results of the experiments described earlier, the shift in the flat band voltage, ΔV_{fb}, with respect to the value for complete suppression of B transport as determined from devices in which the Si_3N_4 thickness is 0.8 nm or more, is plotted in Fig. 10 as a function of the product of the blocking power of Si–N bonding groups as defined in Fig. 9 times the nitride or oxynitride

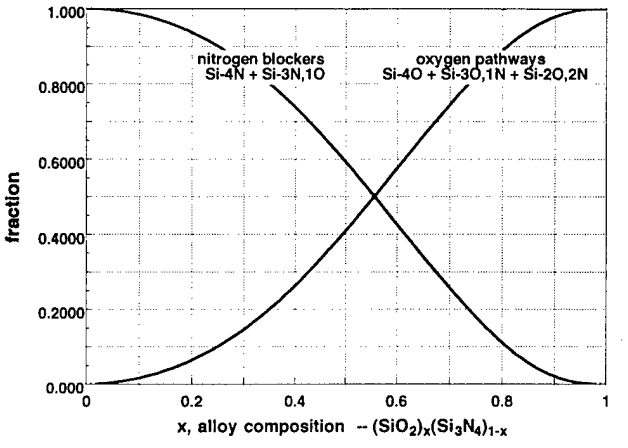


FIG. 9. Calculated fractions of nitrogen blocking groups, Si–4N, Si–3N, 1O, and oxygen pathway groups, Si–4O, Si–3O, 1N, Si–2O, 2N, as a function of x in silicon oxynitride alloys $[(SiO_2)_x(Si_3N_4)_{1-x}]$ using the statistical model calculations of Fig. 8.

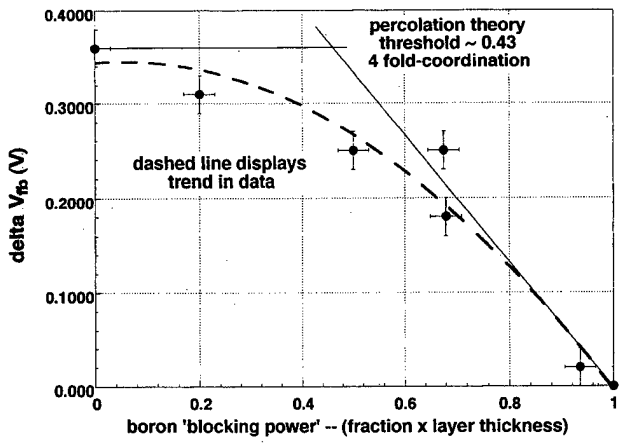


FIG. 10. Change in flat band voltage, as a function of boron "blocking power" defined as the fraction of nitrogen blocking groups (from the model calculation in Fig. 8) times the layer thickness (see Table I). The dashed line is quadratic fit to the data which is included to emphasize the trend in the data. The solid lines indicate the threshold behavior that is anticipated on the basis of the percolation theory as applied to films in which all dimensions are greater than the atomic scale dimensions associated with the transport process.

thickness (see Table I). The dashed line is included in Fig. 10 to emphasize the trend in these data. The solid lines indicate the character of the threshold behavior that is anticipated on the basis of percolation theory as applied to films in which all dimensions are greater than the atomic scale dimensions associated with the transport process.[33] The thickness dimensions of the films in this study do not meet the assumptions of percolation theory so that only qualitative agreement is to be expected. Note that the percolation threshold predicted by theory is ~0.43[33] and is approximately at the intersection of the two straight lines in Fig. 10 that reflect the qualitative dependence of ΔV_{fb} on boron blocking power.

V. DISCUSSION

The experimental studies of the suppression of B transport by nitride and oxynitride films prepared by RPECVD and subjected to post deposition annealing suggest that B atom transport takes place via percolation or connected pathways is nearest-neighbor O atoms. Quantum chemistry calculations suggest that the diffusing species is a B^+ ion rather than a neutral B atom. The measured activation energy for transport is 3.6 eV, and is greater than the binding energy of B^+ ions to O atoms through an interaction that forms a donor-acceptor pair bond. The ionization of B atoms that out-diffuse from the ion implanted poly-Si layer is assumed to take place at the poly-Si–SiO_2 interface paralleling the formation of H^+ atoms or protons, and contributes to the measured activation energy.

It will be necessary to determine if this model can explain enhanced B atom diffusion that takes place when either F or H is bonded in the SiO_2 network. Quantum chemistry calculations performed on Si–O–F alloys[34] indicated increases in the Si–O–Si bond angle, which affects the spatial orientation and relative *s*- and *p*-character of the nonbonding pairs on the O atoms. For example as the Si–O–Si bond angle increases, the nonbonding pairs become more *p*-like in character. Quantum chemistry calculations are in progress to determine whether this will increase or decrease bond energies. Finally, the diffusion constants for As and P in SiO_2 are similar to those for B,[1] with the activation being proportional to the As–O and P–O bond energies. Preliminary studies of dative bond formation between As^+ and P^+ ions and nonbonding pairs on O atoms and N atoms has been completed, with results parallel to those obtained for B^+ ion. For example, the binding energies of As^+ and P^+ ions to nonbonding pairs on O atoms of Si–O–Si groups are respectively 1.5 and 1.6 eV, whereas the binding energies of As^+ and P^+ ions to nonbonding pairs on N atoms of tetrahedral N–Si_3 groups are increased, respectively, to 4.4 and 4.8 eV. The similar diffusion constants for As and P in SiO_2, and the effective blocking of diffusion for As and P in oxynitride alloys supports the model proposed in this article for the transport of B^+ ions through SiO_2, and the blocking of these ions in nitride and oxynitride films.

ACKNOWLEDGMENTS

The Research at NC State University is supported in part by ONR, AFOSR, NSF, SRC, and SEMATECH.

[1]R. B. Fair, J. Electrochem. Soc. **144**, 708 (1997).
[2]G. Hu and R. Bruce, IEEE Trans. Electron Devices **32**, 584 (1985).
[3]J. Pfiester and F. Baker, IEEE Electron Device Lett. **11**, 247 (1990).
[4]T. Mogami, L. Johansson, I. Sakai, and M. Fukuma, Tech. Dig. Int. Electron Devices Meet. 533 (1991).
[5]D. Wristers, L. Han, and D. Kwong, Appl. Phys. Lett. **68**, 2094 (1996).
[6]K. A. Ellis, Ph.D. thesis, Cornell University, Department of Physics, Ithaca, NY, August 1998, p. 189; K. Kawagishi, M. Susa, T. Maruyama, and K. Nagata, J. Electrochem. Soc. **144**, 3270 (1997).
[7]M. D. Tabak, Phys. Rev. B **2**, 2104 (1970).
[8]S. V. Hattangady, H. Niimi, and G. Lucovsky, J. Vac. Sci. Technol. A **14**, 3017 (1996).
[9]S. V. Hattangady, H. Niimi, and G. Lucovsky, Appl. Phys. Lett. **66**, 3495 (1995).
[10]Y.-D. Wu and G. Lucovsky, IEEE Electron Device Lett. **19**, 367 (1998); C. G. Parker, G. Lucovsky, and J. R. Hauser, *ibid.* **19**, 106 (1998).
[11]Y. Ma, T. Yasuda, and G. Lucovsky, J. Vac. Sci. Technol. A **11**, 952 (1993); Y. Ma, T. Yasuda, S. Habermehl, and G. Lucovsky, J. Vac. Sci. Technol. B **11**, 1533 (1993); Y. Ma, T. Yasuda, and G. Lucovsky, Appl. Phys. Lett. **64**, 2226 (1994).
[12]G. Lucovsky, H. Niimi, K. Koh, D. R. Lee, and Z. Jing, in *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface*, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (Electrochemical Society, Pennington, NJ, 1996), p. 441.
[13]H. Niimi, H. Y. Yang, and G. Lucovsky, AIP Conf. Proc. **449**, 273 (1998).
[14]Z. Lu, M. J. Williams, P. F. Santos-Filho, and G. Lucovsky, J. Vac. Sci. Technol. A **13**, 607 (1995).
[15]G. Lucovsky, Z. Jing, P. Santos-Filho, G. Stevens, and A. Banerjee, J. Non-Cryst. Solids **198–200**, 19 (1996); P. Santos-Filho, G. Stevens, G. Lucovsky, T. Cull, P. Fedders, P. Leopold, and R. Norberg, *ibid.* **198–200**, 77 (1996).
[16]V. Misra *et al.*, IEEE Trans Electron Devices (submitted).
[17]H. Niimi, Ph.D. dissertation, North Carolina State University, Raleigh, NC (submitted).
[18]Y. Wu and G. Lucovsky, IEEE Trans Electron Devices (submitted).
[19]S. Lo, D. Buchanan, Y. Taur, L. Han, and E. Wu, Symp. VLSI Tech. (1997), p. 149.
[20]Y. Wu and G. Lucovsky, IEEE International Reliability Physics Symposium (1998), p. 70 (unpublished).

[21]L. Han, D. Wristers, J. Yan, M. Bhat, and D. Kwong, IEEE Electron Device Lett. **16**, 319 (1995).
[22]R. Fair, IEEE Electron Device Lett. **18**, 244 (1997).
[23]H. Y. Yang and G. Lucovsky (unpublished).
[24]Z. Jing, G. Lucovsky, and J. L. Whitten, J. Vac. Sci. Technol. B **13**, 1613 (1995).
[25]H. Yang and G. Lucovsky, J. Non-Cryst. Solids **227–230**, 281 (1998).
[26]K. Vanheusden, W. L. Warren, R. A. B. Devine, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, P. S. Winokur, and Z. L. Lemnios, Nature (London) **386**, 587 (1997).
[27]R. E. Stahlbush, R. K. Lawrence, and H. L. Hughes, IEEE Trans. Nucl. Sci. **45**, 2398 (1998).
[28]Z. Jing, G. Lucovsky, and J. L. Whitten, J. Vac. Sci. Technol. B **13**, 1613 (1995).
[29]J. E. Huheey, *Inorganic Chemistry* (Harper and Row, New York, 1978), Chap. 3.
[30]G. Lucovsky, J. Yang, S. S. Chao, J. E. Tyler, and W. Czubatyj, Phys. Rev. B **28**, 3234 (1983).
[31]S. S. Chao, J. E. Tyler, Y. Takagi, P. G. Pai, G. Lucovsky, S. Y. Lin, C. K. Wong, and M. J. Mantini, J. Vac. Sci. Technol. A **4**, 1574 (1986).
[32]D. V. Tsu, G. Lucovsky, M. J. Mantini, and S. S. Chao, J. Vac. Sci. Technol. A **5**, 1998 (1987).
[33]R. Zallen, *The Physics of Amorphous Solids* (Wiley, New York, 1983), Chap. 4.
[34]G. Lucovsky and H. Yang, J. Vac. Sci. Technol. A **15**, 836 (1997).

Energy-dependent conduction band mass of SiO_2 determined by ballistic electron emission microscopy

R. Ludeke[a)]
IBM T. J. Watson Research Center, Yorktown Heights, New York 10598

Andreas Schenk
Integrated Systems Laboratory, Swiss Federal Institute of Technology, Gloriastrasse 35, CH-8092 Zürich, Switzerland

(Received 19 January 1999; accepted 5 May 1999)

Quantum interference oscillations in ballistic electron emission microscopy (BEEM) spectra were observed for metal–oxide–semiconductor structures with 23 and 30 Å SiO_2 interlayers. Maxima in the transmission coefficients, obtained from solutions of the one-dimensional Schrödinger equation that included image force corrections, could be matched to the spectral maxima provided that the effective electron mass m_{ox}, an adjustable parameter, was increased at each of the consecutive higher energy maxima. The resulting energy dependence or dispersion of $m_{ox}(E)$ showed a dependence on the oxide thickness. The 23 and 30 Å oxides exhibit initial (zero kinetic energy) m_{ox} values of 0.52 m_0 and 0.45 m_0, respectively, that disperse upward with energy by $\approx$0.3 m_0 over a 0–2.5 eV range in kinetic energies. The range of m_{ox} values observed is substantially lower than the average m_{ox} values deduced from quantum interference in Fowler–Nordheim tunneling experiments. The origin of these differences are discussed, and it is argued that BEEM is an inherently simpler and less error prone technique to evaluate m_{ox}. © *1999 American Vacuum Society.* [S0734-211X(99)05004-0]

I. INTRODUCTION

The effective mass m_{ox} of the conduction band electrons in amorphous SiO_2 is a relevant parameter in the prediction of hot electron effects in metal–oxide–semiconductor (MOS) structures, as well as in performance simulations of field effect transistors.[1] Nevertheless, its value has been contentious over decades, a problem derived from its relagation to a fitting parameter that is used to force agreement of frequently over simplified transport equations with experiment. Moreover, few attempts have been made to assess the consequences of the physical constraints imposed in the data interpretation and their resulting impact on m_{ox}. As a result the reported values for m_{ox} ranged from a low of 0.3 m_0,[2] where m_0 is the free electron mass, to estimates as high as 0.85 m_0.[3,4] An underlying but never states premise in all measurements reported thus far is the assumption of an energy independent or dispersionless mass. This assumption, as we shall see shortly, is a consequence of the inability to extract a dispersive m_{ox} with conventional transport techniques. The current–voltage ($I-V$) measurements give instead values of m_{ox} averaged over an energy range determined by the experiment. In this work we present an approach based on quantum interference effects of electrons injected directly into the conduction band of SiO_2 from which it is possible to derive the energy dispersion of m_{ox}.

The most commonly employed $I-V$ transport measurement of MOS capacitor structures is Fowler–Nordheim (FN) tunneling, in which a sufficiently high potential V_{ox} is applied across the oxide to field-inject electrons via tunneling from near the Fermi level of the gate and thence into the conduction band of the SiO_2. This is schematically shown in the top illustration of Fig. 1. The electron, represented by a plane wave, tunnels through the trapezoidal barrier represented by the solid line and emerges in the conduction band of the SiO_2. There it continues to gain energy due to field acceleration as it approaches the SiO_2–Si interface. The current density is given by the well known FN expression obtained in the Wentzel–Kramers–Brillouin (WKB) approximation,[5] $J_{FN} \propto \exp[-4(2\ m_{ox})^{1/2}\Phi_B^{3/2} d_{ox}/3\hbar V_{ox}]$, where Φ_B is the barrier height, d_{ox} the oxide thickness, and V_{ox} the net oxide potential. Experimental $I-V$ curves are generally fitted to this simple relationship, with the assumption that both Φ_B and V_{ox} are known, but not m_{ox}, which thereby is treated as a fitting parameter. This expression for J_{FN} ignores screening effects arising from the presence of the interfaces, an intentional omission based on results for thicker oxides.[6] For ultrathin oxides (<10 nm) these effects, expressed through the classical concept of image force lowering of the barrier and illustrated by the dashed line in Fig. 1(a), cannot be ignored.[7,8] The inclusion of image force effects requires a numerical solution for J_{FN}, whether represented in terms of a complete Schrödinger equation, or as an approximation, such as WKB.[7] The mass deduced from fits is that of a particle tunneling through the barrier, with a value that should represent the effective mass of the imaginary branch of the band structure. It is expected to be different from the real conduction band mass, although at the band edge their values should be nearly equal. However, the electron is constantly changing its energy from the point of injection to its emergence in the conduction band (a change exceeding 3 eV), yet it is assumed that its mass is either constant or that the deduced mass represents an energy av-

a)Electronic mail: ludeke@watson.ibm.com

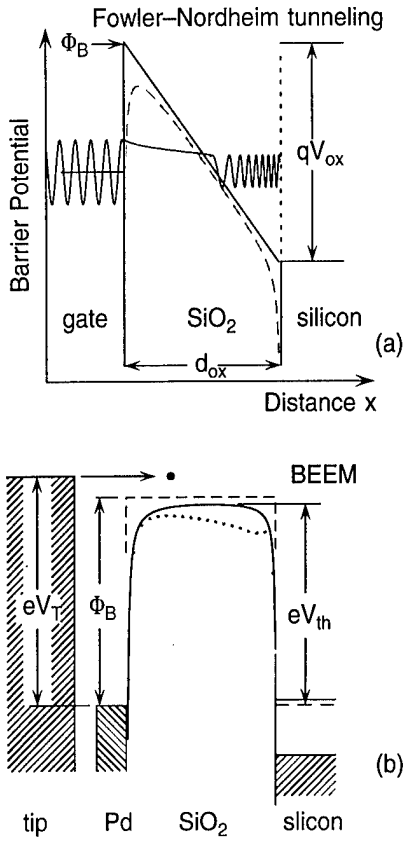


FIG. 1. (a) Energy diagram for FN tunneling is shown at the top under the usual assumption of a trapezoidal barrier that neglects image force effects. The influence of the latter is shown by the dashed curve (b) A corresponding diagram for a BEEM experiment. The barrier potential is shown with (solid) and without (dashed) image force corrections. Φ_B is the potential difference between the Fermi level of the metal and the SiO_2 conduction band minimum. eV_{th} represents the barrier height with the inclusion of image force lowering and corresponds to the threshold energy for current detection measured in a BEEM spectrum. V_T is the STM bias and corresponds to the kinetic energy of the electrons injected into the metal gate. The dotted potential represents the consequence of positive oxide charge near the SiO_2–Si interface.

eraged value. The same can be said of its mass in the conduction band, as the electron is accelerated towards the SiO_2–Si interface. The mass, the real m_{ox}, can also be obtained in a FN experiment from weak oscillations in J_{FN}. Such oscillatory structure arising from interference of the electron wave function in the "cavity" between the conduction band edge and the SiO_2–Si interface was predicted by Gundlach,[9] and first observed by Maserjian and Petersson,[3,10] and subsequently by others.[4,11–14] The experimental reality of a changing electron energy and its complications, together with the inherently weak oscillatory structure modulating a large background current, makes the technique unsuitable for measuring mass changes with electron energy. Band structure calculations indicate that for crystalline quartz considerable deviations from parabolicity occur within 1–2 eV of the lower band edge, with additional bands starting to contribute to the density of states near these energies.[15–17] However, the conduction bands are often inadequately treated in such calculations and it is therefore difficult to extract an accurate dispersion of m_{ox}. Theoretical estimates of m_{ox} for α-quartz range from m 0.3[16] to $0.5\ m_0$.[17] Even if accurate values of m_{ox} and their dispersions were calculated for the polyphases of crystalline quartz, a direct correspondence to amorphous quartz would still be speculative, although it has been argued that on the short range scale of a few nanometers the distorted amorphous phase still exhibits many band-like properties akin to crystalline SiO_2.[1]

Quantum interference oscillations have recently been observed in the $I-V$ spectrum of electrons injected directly into the conduction band of SiO_2.[18] The technique, ballistic electron emission microscopy or BEEM, uses the tip of a scanning tunneling microscope (STM) to inject electrons into the thin metal gate of a MOS structure, whence they proceed ballistically to enter the SiO_2 and subsequently the Si substrate. They emerge from the Si as a collector current that is modulated by the interference phenomenon in the oxide. For over-the-barrier transmission, maxima in the transmission probability for a rectangular barrier (i.e., in the absence of an internal oxide potential V_{ox}) occur at the following energies:[9]

$$E=(n\pi\hbar/d_{ox})^2/2m_{ox} \quad \text{with} \quad n=1,2,3..., \tag{1}$$

from which m_{ox} can be deduced by matching theoretical maxima to those obtained experimentally. However, the presence of an internal field plus the inclusion of image force effects requires that the equations be solved numerically. Using such an approach, a value of $m_{ox}=(0.63\pm0.09)m_0$ was obtained for a 2.8 mm oxide.[18] Subsequent improvements in the data quality have now necessitated the inclusion of an energy dependent mass to match experiment with theory, which is the topic of this article.

II. EXPERIMENTAL DETAILS

A. Ballistic electron emission microscopy/ spectroscopy

BEEM, as mentioned earlier, is a STM based microscopy that differs from conventional STM by the presence of a thin metal layer deposited on a substrate, in the present case a SiO_2–Si sample. The only purpose of the metal layer is to provide a reference electrode relative to which the STM tip is biased with a potential V_T. The electrons injected into the metal layer by the STM tip thus have an energy of eV_T. The thickness of the metal should be comparable or preferably less than the electron mean free path in the metal, so that the electrons can traverse the layer ballistically. If their energy is larger than the potential barrier posed by the oxide—in the present case the potential Φ_B representing the difference between the Fermi level in the metal and the bottom of the SiO_2 conduction band—some electrons will be injected into the conduction band, and after traversing the oxide and entering the Si, will emerge as a collector current I_c from the substrate. An energy diagram for a BEEM experiment at zero applied bias is shown at the bottom of Fig. 1. The probability for the electron to reach the Si is dependent on the overlap of the conduction band density of states at the interfaces, the transmission probability $T_{ox}(E)$ across the oxide, as well as transmission probabilities of a quantum mechanical origin

(e.g., transverse momentum conservation). These issues have been discussed elsewhere.[1,7,19] Although for oxide films thicker than about 4 nm T_{ox} is strongly dependent on electron-phonon scattering,[19,20] for the thinner oxides used here this scattering mechanism is not dominant, a conclusion supported by the observation of interference phenomena. Consequently, scattering will be ignored in our analysis of T_{ox}.

In the present application BEEM is exclusively used in the spectroscopy mode, in which 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 potential Φ_B. I_c becomes finite once V_T exceed Φ_B, or more precisely, Φ_B modified by the image potential and any oxide potential that affects the net barrier height. The consequence of the image force is included in the potential shown in Fig. 1(b) (solid line), the dashed line representing the bare potential. The STM is generally operated at a constant tunneling current I_T. An important experimental consideration is tip drift, which should be negligible over the acquisition time for a spectrum, which is typically ~1.5 min. We waited until tip drifts were below 1 Å/min before attempting the acquisition of spectra.

B. Sample preparation

Device-grade amorphous SiO_2 layers were thermally grown in dry oxygen at 800 °C. The substrates were 125 mm diameter Si(100) wafers, boron doped in the low 10^{16} cm^{-3} range. No additional treatments were performed after the oxidation. The thickness d_{ox} of the oxides studied in this work were 23 and 30 Å. The thickness was determined with an ellipsometer, and represents an average of over 50 measurements over the wafer. All measurements were within 1 Å of the averaged value. The thickness of the oxides was also obtained from capacitance–voltage ($C-V$) measurements, using 500-Å-thick W dots deposited *ex situ* by chemical vapor deposition. Their values were 1 Å less than those obtained ellipsometrically for oxides of thicknesses in the 20–40 Å range. We chose the ellipsometric values for the data analysis, as the $C-V$ data reduction routines included corrections that were unnecessary for metal gates and which resulted in a small (~1 Å) underestimate of the thickness.

Approximately 8×15 mm^2 samples were cleaved from the wafers in a dry box. An ohmic contact was made by scraping a small droplet of a Ga–In alloy into the backside of the samples. They were then introduced into an ultrahigh vacuum (UHV) preparation chamber, where each sample, prior to metallization, was annealed separately near 250 °C for 10 h to remove water and other volatile surface contaminants. Arrays of metal dots, 0.2 mm in diameter, were deposited by evaporating the metal through a shadow mask. The samples were cooled to a temperature of ~30 K in order to minimize surface diffusion and thereby achieve pinhole-free layers at the lowest possible coverages. For Pd, which was used for the 23 Å oxide, full coverage was achieved in the 30–40 Å range. The resulting morphology, shown in Fig. 2(a), consists of nodules typically 80 Å in diameter that protruded <10 Å above the valleys. Tungsten was used for the

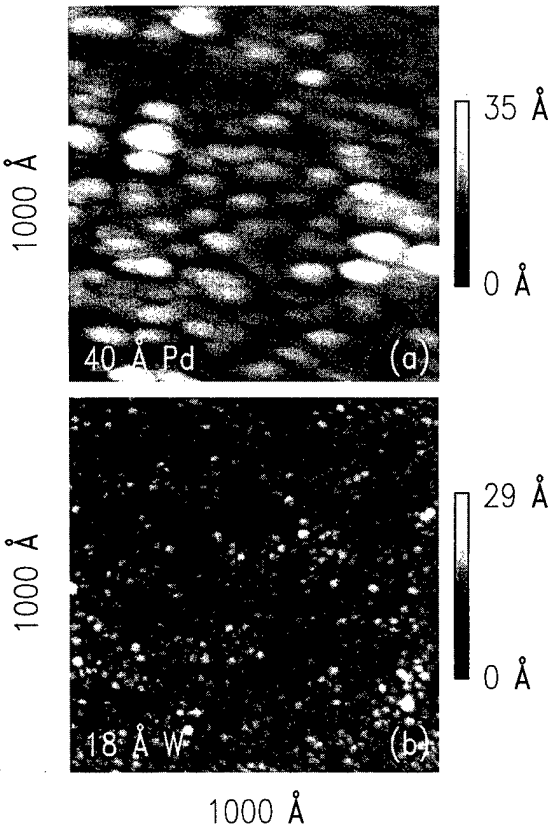


FIG. 2. 1000×1000 $Å^2$ topographic images for metal on SiO_2: (a) 40 Å Pd film and (b) 18 Å W film. Both films were deposited with the SiO_2/Si substrate near 30 K.

30 Å oxide sample. It was evaporated from a low voltage and low power (<1800 V, ~200 W) electron beam evaporator to minimize oxide damage, as well as to maintain a low chamber pressure during metallization (~10^{-8} Torr of mostly H_2). Pinhole-free films were obtained in the 15–18 Å range. The morphology of a W layer is shown in Fig. 2(b). Its nodular texture, which is substantially finer than that of Pd, exhibits a rather homogeneous distribution of grain sizes in the 15–20 Å range. The small size of the W grains serves to refocus on the importance of a low STM tip drift during the acquisition of a spectrum, as the tip position should be kept well within the area of the grain. Otherwise the curvature of the grain will cause undesirable changes in electron injection angle as the tip drifts near the grain boundary. It has been known for some time that the injection angle plays a critical role in the transmission.[21] The finished sample was allowed to warm up to room temperature and was subsequently transferred under UHV into the STM chamber. A reference electrode at the STMs ground potential, needed to bias the tip, was carefully positioned onto a selected metal dot by means of three orthogonally mounted Inchworms™. After the STM tip reached tunneling the drift was checked repeatedly until it decayed to the desired level, usually in a few hours. Once stabilized, large lateral movement of the tip were avoided as well to minimize tip creep.

BEEM spectra were usually taken on previously unmeasured areas of the sample to avoid trap generation and charg-

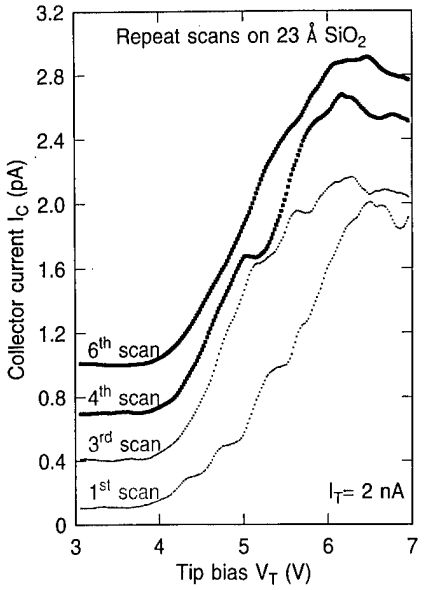


FIG. 3. BEEM spectra taken on the same location of a Pd/23 Å SiO_2/p-Si(100) MOS structure. The first scan was taken on a spot that had not been previously exposed to electrons of sufficient energy to inject electrons into the SiO_2. The quantum interference oscillations are substantially changed after repeat scans, being nearly obliterated after about 6 scans. The spectra show few additional changes beyond the sixth scan. The spectra are displaced vertically for clarity. I_T=2 nA.

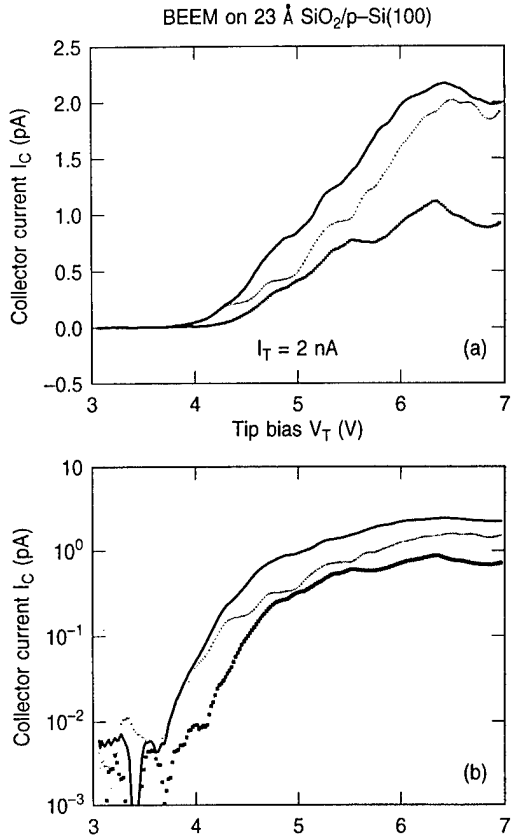


FIG. 4. BEEM spectra for a 40 Å Pd/23 Å SiO_2/p-Si(100) MOS structure. The bottom two curves are individual spectra, top curves are for an average of four spectra. Each spectrum was taken on a previously unexposed part of the sample. (a) Linear plot that enhances the interference structure at the higher energies. (b) Logarithmic plot of the same data as (a), which enhances the structure in the threshold region. I_T=2 nA.

ing from prior hot electron injections.[22] Any oxide charge affects V_{ox} locally, thereby altering the interference structure, which renders a determination of m_{ox} nearly impossible.[18] Adjacent measurement points were separated by at least 250 Å. In order to achieve flat band conditions (i.e., $V_{ox}=0$), a +0.3 V bias was applied to the Si for the W-"gated" 30 Å oxide structure. In the absence of an external bias, the Fermi level of the W is commensurate with the midgap energy of the Si at the SiO_2–Si interface.[23] In contrast, the Pd contacts to oxides grown on p-Si(100) did not require a bias to achieve flatband conditions, as our best estimate suggests $V_{ox}\approx 0$ in the absence of a bias.[22] This fortuitous situation avoids biasing the thin 23 Å oxide layer. Even biases of ~0.1 V would have resulted in large (direct) tunneling currents for the metal dot size used here, which would have saturated the operational amplifier.

III. INTERFACE PHENOMENA: SPECTROSCOPIC RESULTS

A. 23 Å oxide

BEEM spectra for a 40 Å Pd/23 Å SiO_2/p-Si(100) MOS structure are shown in Fig. 3. The STM tunneling current was set at 2 nA. An oscillatory component in the collector current is clearly discernible in the spectrum labeled first scan, which corresponds to the first scan on a virgin part of the sample surface. In this set of spectra we repeated the scans several times at the same location. The second scan already shows structure shifted to slightly different energies. The fourth scan is altogether different, and in general, shows weaker structure, a tendency that continues with subsequent scans, but does not change appreciably beyond the sixth scan shown in Fig. 3. We have observed this behavior on every occasion we performed repeated scans at a previously unexposed location of the surface. We generally observe a moderate increase in the collector current after the first and subsequent scans in the energy regions just above threshold. The increase is attributed to the generation of positive charge after the oxide layer has been electrically stressed with electrons of kinetic energy exceeding 2 eV ($V_T>|6|$ V).[18,22] The positive charge near the Si interface lowers locally the barrier height, as depicted by the dotted barrier profile in Fig. 1(b). Noise in the spectra, which increases noticeably for $V_T>5$ V, is an indication that charging and discharging events occur at trap sites generated by the hot electrons.[22] Thus, it is important to realize that only the first spectrum may be representative of a charge free oxide region. Preexisting traps in the oxide may also charge up and distort, or more likely obliterate the interference structure due to the inhomogeneous nature of the local fields. About 25% of the spectra show interference structure, but substantially fewer show relatively noise free characteristics over the whole spectral range from threshold, near 4 V, to the upper limit of 7 V.

Figure 4(a) depicts a couple of nearly identical spectra (the two lower curves) as well as an average over four spec-

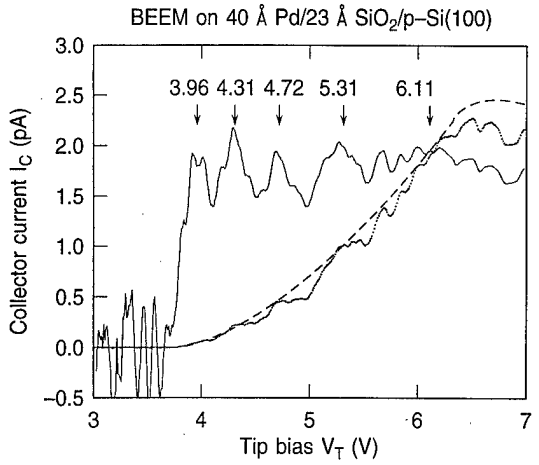


FIG. 5. Methodology for extracting interference maxima from BEEM spectra for a 23 Å oxide: a power law curve (dashed curve) is tangentially fitted to the experimental curve near the maxima. Their ratio simulates the transmission probability function, from which the indicated peak positions are easily obtained.

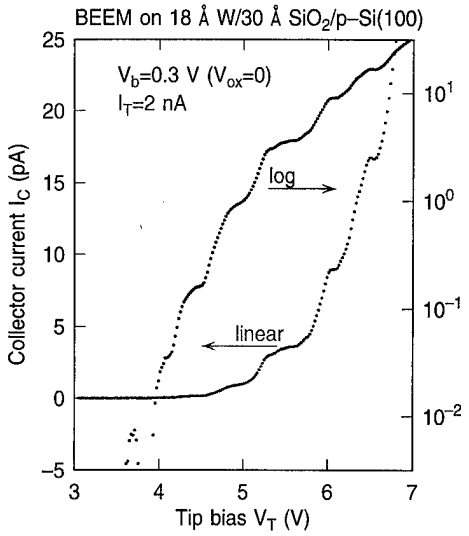


FIG. 6. Linear and logarithmic plot of a BEEM spectrum for a 18 Å W/30 Å SiO_2/p-Si(100) MOS sample. The logarithmic scale is shown on the right ordinate. I_T=2 nA.

tra (upper curve). Whereas the oscillation are clearly discernible above 4.5 V in the linear plot, the structure below this value are more readily seen in a logarithmic plot, shown in Fig. 4(b) for the same spectra. To extract the energy location of the interference maxima, we have used a simple approach that assumes $I_c(V_T)=\overline{I_c}T_{ox}$, where $\overline{I_c}$ is the collector current in the absence of interference. It can be simulated by fitting a smooth polynomial expression to the maxima of the experimental spectrum, as shown by the dashed line in Fig. 5. An "experimental" T_{ox} is then obtained by numerical division of I_c by $\overline{I_c}$ as depicted in Fig. 5. The peak positions of the interference maxima are readily extracted from this curve with an accuracy of ±0.02 V. Their values are shown above the maxima, and will be used in Sec. V to determine the effective mass.

B. 30 Å oxide

The tungsten metallization of a MOS structure on p-type Si results in an increase in the effective barrier height that reaches a maximum at the SiO_2–Si interface. In order to achieve flat band conditions it is therefore necessary to apply a positive oxide bias of 0.3 V to the substrate.[23] The resulting direct tunneling current of ~20 pA did not affect the BEEM spectrum and could be readily subtracted from the data. An exceptionally clean spectrum is shown in Fig. 6. As with the 23 Å oxide layer, about a quarter of the spectra showed interference oscillations, but only a fraction of these exhibited essentially identical features, which we interpreted to represent the characteristics of a charge free oxide. The plot of this spectrum on a logarithmic scale, also depicted in Fig. 5, again emphasizes the oscillatory structure immediately above the threshold of ~3.6 V. This value was obtained from computer-aided fits to the threshold region of many spectra.[19] The obvious difference between the spectra for the two different oxide thicknesses is that the 30 Å oxide exhibits two more maxima over the same energy interval. This is due to the increase in d_{ox} and can be readily understood by inspecting Eq. (1) for the ideal square barrier case. Taking the differential of Eq. (1), one gets

$$\Delta E/E=-2\Delta d_{ox}/d_{ox}. \tag{2}$$

Thus, an increase in d_{ox} results in a decrease in the energy separation of the maxima.

IV. TRANSMISSION PROBABILITIES AND FITTING PROCEDURE

The incorporation of screening effects and a more realistic treatment of the interfaces requires a numerical solution of the one-dimensional Schrödinger equation to calculate the energy dependent transmission probability T_{ox}. Only an outline will be presented here, as details can be found elsewhere.[18] At the outset, we assume a parabolic dispersion of $E(k)$ in the oxide, with an effective mass m_{ox} as parameter. Consequently, the momentum takes the form

$$k(x)=\sqrt{2m_{ox}/\hbar^2}\sqrt{E-[\Phi_B+eF_{ox}x+E_{im}(x)]}, \tag{3}$$

where Φ_B is the barrier height, E_{im} is the image potential that includes the effect of all images in the two electrodes,[7] and F_{ox} is the oxide field. The singularity of the classic image potential at the interfaces was "removed" by extending both the band edges in the semiconductor and the Fermi level in the metal until they intersect the image potential. We use the approach of Ando and Itoh for an arbitrary potential barrier by segmenting the barrier [such as depicted in Fig. 1(b)] into N equal intervals with coordinate x_1 marking the 1th segment.[24] With continuity of both wave function and quantum-mechanical current density as boundary conditions at each interface, $T_{ox}(E)$ is then expressed as

$$T_{ox}(E)=\frac{m_0}{m_{N+1}}\frac{k_{N+1}}{k_0}\frac{|\det M|^2}{|M_{22}|^2}, \tag{4}$$

where M is a (2×2) product matrix $M=\Pi_{l=0}^{N} M_l$ with transfer matrices $M_l(x_l)$ being only functions of momentum k_l and effective mass m_l in the lth segment.[24] With the metal-oxide interface at x_0 and the oxide-silicon interface at x_N, $m_0=m_M$ denotes the effective mass of the conduction electrons in the metal electrode and $m_N=m_{Si}$ the corresponding mass in silicon. For all other l we have $m_l=m_{ox}$. The following parameters were used in the simulations: (a) for both oxides: $N=30$, $m_{Si}=0.19m_0$, $m_M=m_0$, $\epsilon_{Si}=11.7$ and $\epsilon_{ox}=2.13$; (b) for the 23 Å oxide with Pd gates: $d_{ox}=23$ Å, $\Phi_B=4.01$ eV, $F_{ox}=V_{ox}/d=0$ and 0.07143 V/nm (potential peak at the oxide-metal boundary).[8] The finite value of the field corresponds to a $V_{ox}\approx 0.2$ V, and serves to assess the shifts in m_{ox} due to uncertainties in V_{ox}. (c) For the 30 Å with W gates: $d_{ox}=30$ Å, $\Phi_B=3.77$ eV, $F_{ox}=0$. The value of Φ_B agrees well with the 3.7 eV value for a 76 Å oxide obtained from $I-V$ measurements (FM) on W-gated capacitor structures.[23]

$T_{ox}(E)$ is a rapidly rising function for $E>eV_{th}$ that exhibits an oscillatory structure with maxima near unity,[9,18,24] and need not be reproduced here. The initial observations of interference phenomena in BEEM were characterized by relatively noisy spectra that exhibited only 3–4 clearly distinguishable peaks, whose positions could be fitted theoretically with a single valued m_{ox}.[19] A similar approach for the present multipeaked interference structure failed, however, and we had to resort to an energy dependent mass to match theory to experiment. The following procedure was used to match the theoretical interference maxima to the experimental data, an example of which is shown in Fig. 5 for the 23 Å oxide. With the parameters listed earlier, and m_{ox} as a variable, T_{ox} was numerically evaluated until agreement was obtained between the first theoretical and experimental peaks, thus yielding $m_{ox}(E_1)$. Then m_{ox} was increased until agreement was reached between the second theoretical and experimental peaks, yielding $m_{ox}(E_2)$. The same was done for all higher lying peaks, giving values $m_{ox}(E_i)$ at peak energies E_i by the best fit between theoretical and measured T_{ox}. $m_{ox}(E_i)$ defines a mass dispersion curve for the discrete peak energies E_i. Interpolation of $m_{ox}(E_i)$ between the peak energies results in a smooth mass dispersion $m_{ox}(E)$ with the property that the theoretical interference maxima match the measured interference peaks.

V. MASS DISPERSIONS AND DISCUSSION

For the indicated peak positions of the 23 Å oxide in Fig. 5 we have calculated $m_{ox}(E)$ under various assumptions related to uncertainties in the parameters. The results are shown in Fig. 7 plotted as a function of V_T. The solid curve represents the most likely dispersion using the best estimates for the values of the parameters, as given in Sec. IV. The symbols represent the calculated values, while the smooth curves are spline fits through the data points (for clarity we have omitted the data points in some curves). The dashed curve represents $m_{ox}(E)$ calculated from maxima obtained from the four averaged spectra shown in Fig. 4. As discussed earlier, uncertainties in the thickness have the largest effects

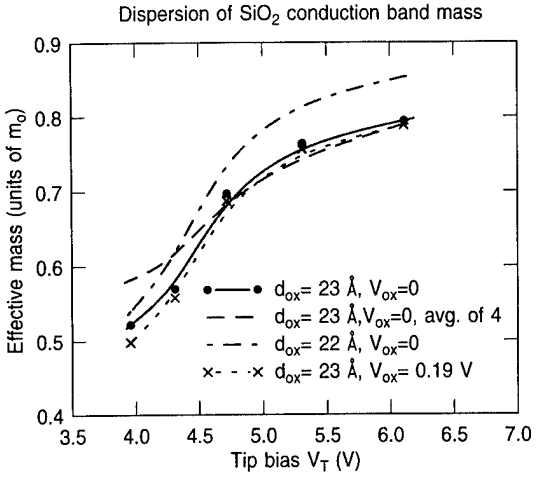


FIG. 7. Conduction band mass dispersions for a 23 Å oxide. The various curves depict $m_{ox}(E)$ calculated for variations in parameters for which the mass is particularly sensitive. The solid curve depicts the dispersion calculated with the best estimates for d_{ox}, Φ_B, and V_{ox}.

on m_{ox}. A change of 1 Å results in a rather large change in $m_{ox}(E)$, particularly at the higher energies, as shown by the dot-dashed curve in Fig. 7, which depicts the results for a decrease of 1 Å in the oxide to 22 Å. On the other hand, uncertainties in V_{ox} have a much smaller effect on the mass, as shown in Fig. 7 by the small change generated when V_{ox} is changed from 0 to 0.19 V (dotted curve). This value is outside an estimated uncertainty in V_{ox} of ±0.1 V.

The calculated mass dispersion based on the data of Fig. 6 for the 30 Å oxide is shown in Fig. 8 by the lower solid curve. The dotted curve represents the same data, but assumes a decrease in thickness to 29 Å, as was done for the 23 Å, oxide. The mass dispersion of the latter is again shown in the figure by the upper solid curve. The barrier height or threshold energy has been subtracted to obtain the kinetic

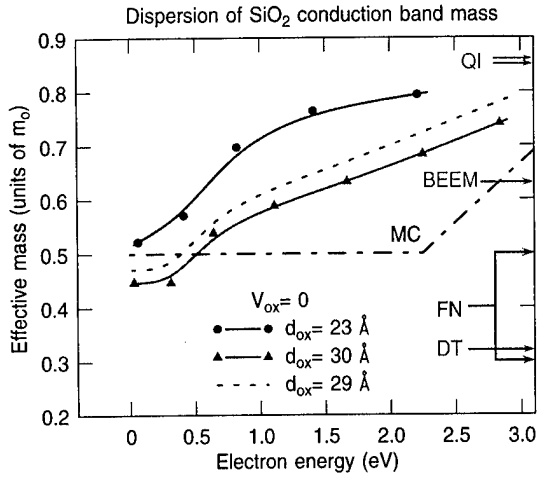


FIG. 8. Energy dispersions of the conduction band mass of SiO_2 for a 30 Å oxide (lower solid curve) and for a 23 Å oxide (upper solid curve). The dotted curve represents a reduction of 1 Å of the 30 Å oxide to $d_{ox}=29$ Å, which shifts the dispersion upwards. Some previously reported values of m_{ox} are indicated along the right ordinate.

energy in the conduction band of the SiO_2. Because of the curvature in the barrier profile due to image effects, as shown in Fig. 1(b), the electron has a position dependent kinetic energy in the oxide. The abscissa in Fig. 8, due to the subtraction of V_{th}, thus corresponds to the minimum kinetic energy experienced by the electrons, with $m_{ox}(E)$ representing a weighted average over a range of energies. However, we observed previously that the average value for m_{ox} is only about 3% smaller than the mass calculated by excluding image force effects.[18] The latter scenario, for $V_{ox}=0$, leads to a position independent kinetic energy. Thus, $m_{ox}(E)$ in Fig. 8 can be interpreted as representing to within a few percent the mass at the actual value of the kinetic energy.

It is clearly evident that an appreciable dispersion is observed in m_{ox} over the energy range to 3 eV. The expectation that m_{ox} disperses is not new,[1,25] but had not previously been demonstrated experimentally. A "dispersion" marked MC in Fig. 8 has been used in Monte Carlo calculations of a variety of hot electron transport phenomena in SiO_2.[19,25–27] A relevant question at this point concerns the disparity between the two curves for different oxide thickness: is there a real difference or can experimental uncertainties account for the difference? Because of the quadratic dependence of m_{ox} on d_{ox} an error in d_{ox} is doubled for m_{ox}, as is readily ascertained by differentiating Eq. (1) and obtaining $\Delta m_{ox}/m_{ox}=-2\Delta d_{ox}/d_{ox}$. This dominance of an uncertainty in m_{ox} outweighs that for the other parameters, as discussed earlier. The relative accuracy between the 23 and 30 Å oxides is within 1 Å, while their absolute thickness is estimated to have an uncertainty of ± 1 Å. Judging from the $C-V$ determined thicknesses that are about 1 Å smaller, the uncertainty should be weighted towards a smaller value of d_{ox}. However, assuming such an error, as represented by the dotted line for the 29 Å oxide in Fig. 8, it is thus unlikely that the difference in the two dispersions can be attributed to uncertainties in the values of d_{ox}. A second source of errors can arise from uncertainties in the determination of the interference maxima, as well as from small deviations between comparable spectra. However, comparing a single spectrum with an average of four spectra, shown as the dashed curve in Fig. 7, clearly shows that this possible source of error cannot account for the large differences observed for the different thickness. We therefore conclude that the 23 Å oxide exhibits a different, faster rising dispersive behavior than the 30 Å oxide. This difference may arise from band structure changes that result from an increased confinement in the direction normal to the film for the 23 Å oxide. Nevertheless, both curves show an initial slow change from a value near $0.5\ m_0$ that implies, particularly for the 30 Å oxide, the presence of a parabolic conduction band. This observation is in agreement with the parabolic behavior at the bottom of the conduction band predicted for crystalline SiO_2.[15–17]

Also indicated on the right margin of Fig. 8 are a number of values for m_{ox} reported in the literature. The range for FN determined values is shown for illustrative purposes only, as these values represent tunneling masses that are only indirectly related to the real conduction band mass. The two values near $0.85\ m_0$, marked QI, were obtained from quantum interference oscillations in FN experiments and were reported by Maserjian[10] and by Zafar *et al.*[4] The position labeled BEEM is from our earlier results for the 28 Å SiO_2 sample.[18] It's value of $0.63\ m_0$, based on a single mass fit of the interference structure, thus represents an average over the energy interval, and is in excellent agreement with an average m_{ox} for the 30 Å oxide. The question now arises as to why the QI results from the FN experiments are substantially larger than the range of values obtained with BEEM. Before addressing this point, it is instructive to briefly discuss the differences in the techniques and the critical parameters that enter in the determination of m_{ox}. Aside from BEEM being a local and FN a broad area probe, the crucial difference lies in the nearly monoenergetic nature of BEEM injected electrons[28] that maintain a relatively constant kinetic energy in the oxide, as opposed to the continuously changing energies of FN injected electrons. In order to observe interference phenomena the field has to be changed in FN measurements, which changes the injection conditions, i.e., cavity length, image force effects and current, with the consequence that the interference signal represents an energy averaged, relatively weak component modulating an exponentially increasing background current. Whereas the deduction of m_{ox} by BEEM depends crucially only on one parameter, namely the thickness d_{ox}, and to a substantially lesser extent on Φ_B and V_{ox},[18] the precise knowledge of all three parameters is critical for the FN determination of m_{ox}.[4] The latter depends, as for over-the-barrier injection in BEEM [Eq. (1)] on the inverse square of d_{ox}, but more significantly, on the inverse cube of Φ_B.[4] The effective barrier height is affected by the image force lowering that is significantly more pronounced at high fields (Schottky effect[8,29]) assuming, of course, that its zero field value is accurately known to begin with, itself a questionable assumption. For BEEM determined masses it was estimated that at relatively low fields of $\sim 6.8\times 10^5$ V/cm, the increase in the average mass due to image force neglect is about 3%.[18] The increases should be substantially larger in the FN experiments, for which the fields are 20-fold higher. Another contribution to experimental uncertainty is oxide charge, which affects the net oxide field, a problem that we assiduously avoided in BEEM through selective choice of local injection sites. The presence of positive oxide charge was reported,[3,10] and constitutes a nearly unavoidable problem in low level stressing of SiO_2.[30] Its effect further lowers the barrier height, as depicted by the dashed potential profile in Fig. 1(b).

VI. CONCLUDING REMARKS

We have further demonstrated here the versatility and power of the highly localized hot electron capabilities of BEEM. The relatively simple physical concept and formulation of the monochromatic over-the-barrier electron current was used here to obtain for the first time the energy dispersion of the effective mass of conduction band electrons in SiO_2. A thickness dependence was also observed, with the thinner 23 Å oxide exhibiting a larger mass than the 30 Å

oxide, although the net changes in dispersion were comparable over the same energy range. Since this difference was outside the estimated uncertainties in the thicknesses of the oxides, a primary source for error in determining m_{ox}, it was intimated that the thickness of the 23 Å oxide was sufficiently small to affect deviations from bulk-like features of the band structure. The validity of this conjecture could be assessed through an appropriate band structure calculation. The observed range of dispersion from ~0.5 to ~0.8 m_0 is substantially larger than the tunneling masses deduced from direct tunneling and FN experiments, although an agreement exists at the low kinetic energy value. However, these masses should not be confused, as frequently done in the literature. Due to experimental constraints of FN measurements, the electron masses are at best averaged values that are obtained from fits of data to simple analytic expressions that neglect several relevant physical phenomena. It is hoped that modeling sophistication will overcome this problem, and that at one point in the future it will be feasible as well to determine the dispersion of the tunneling mass.

ACKNOWLEDGMENTS

The authors acknowledge many fruitful discussions with Ed Cartier. Special thanks go to both Doug Buchanan, for the $C-V$ thickness measurements, and to Chris D'Emic for growing the oxide samples, as well as for the ellipsometric thickness measurements.

[1]M. V. Fischetti, S. E. Laux, and E. Crabbé, J. Appl. Phys. **78**, 1058 (1995), and references therein.
[2]T. Yoshida, D. Imafuku, J. L. Alay, S. Miyazaki, and M. Hirose, Jpn. J. Appl. Phys., Part 2 **34**, L903 (1995).
[3]J. Maserjian and G. P. Petersson, Appl. Phys. Lett. **25**, 50 (1974).
[4]S. Zafar, K. Conrad, Q. Liu, E. A. Irene, G. Hames, R. Kuehn, and J. J. Wortman, Appl. Phys. Lett. **67**, 1031 (1995).
[5]M. Lenzlinger and E. H. Snow, J. Appl. Phys. **40**, 278 (1969).
[6]Z. A. Weinberg and A. Hartstein, Solid State Commun. **20**, 179 (1976).
[7]A. Schenk and G. Heiser, J. Appl. Phys. **81**, 7900 (1997).
[8]H. J. Wen, R. Ludeke, D. M. Newns, and S.H. Low, J. Vac. Sci. Technol. A **15**, 784 (1997).
[9]K. H. Gundlach, Solid-State Electron. **9**, 949 (1966).
[10]J. Maserjian, J. Vac. Sci. Technol. **11**, 996 (1974).
[11]M. V. Fischetti, D. J. DiMaria, L. Dori, J. Batey, E. Tierney, and J. Stasiak, Phys. Rev. B **35**, 4044 (1987).
[12]J. C Poler, K. K. McKay, and E. A. Irene, J. Vac. Sci. Technol. B **12**, 88 (1994).
[13]S. Zafar, Q. Liu, and E. A. Irene, J. Vac. Sci. Technol. A **13**, 47 (1995).
[14]H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguru, S. Nakamura, M. Saito, and H. Iwai, IEEE Trans. Electron Devices **45**, 1233 (1996).
[15]P. M. Schneider and W. B. Fowler, Phys. Rev. Lett. **36**, 425 (1976); Phys. Rev. B **18**, 7122 (1978).
[16]J. R. Chelikowsky and M. Schlüter, Phys. Rev. B **15**, 4020 (1977).
[17]Y.-N. Xu and W. Y. Ching, Phys. Rev. B **44**, 11048 (1991).
[18]H. J. Wen, R. Ludeke, and A. Schenk, J. Vac. Sci. Technol. B **16**, 2296 (1998); R. Ludeke, H. J. Wen, and A. Schenk, Appl. Phys. Lett. **73**, 1221 (1998).
[19]R. Ludeke, A. Bauer, and E. Cartier, J. Vac. Sci. Technol. B **13**, 1830 (1995); Appl. Phys. Lett. **66**, 730 (1995).
[20]R. Ludeke, H. J. Wen, and E. Cartier, J. Vac. Sci. Technol. B **14**, 2855 (1996).
[21]R. Ludeke and M. Prietsch, J. Vac. Sci. Technol. A **9**, 885 (1991).
[22]H. J. Wen and R. Ludeke, J. Vac. Sci. Technol. A **16**, 1735 (1998).
[23]D. A. Buchanan, F. R. McFeely, and J. J. Yurkas, Appl. Phys. Lett. **73**, 1676 (1998).
[24]Y. Ando and T. Itoh, J. Appl. Phys. **61**, 1497 (1987).
[25]M. V. Fischetti, D. J. DiMaria, L. Dori, J. Baty, E. Tierney, and J. Stasiak, Phys. Rev. B **35**, 4404 (1987).
[26]E. Cartier and F. R. McFeely, Phys. Rev. B **44**, 10689 (1991).
[27]D. Arnold, E. Cartier, and D. J. DiMaria, Phys. Rev. B **49**, 10278 (1994).
[28]N. D. Lang, A. Yacoby, and Y. Imry, Phys. Rev. Lett. **63**, 1499 (1989); N. D. Lang (unpublished).
[29]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981).
[30]D. A. Buchanan, D. J. DiMaria, C.-A. Chang, and Y. Taur, Appl. Phys. Lett. **65**, 1820 (1994).

Band offsets for ultrathin SiO_2 and Si_3N_4 films on Si(111) and Si(100) from photoemission spectroscopy

J. W. Keister[a)]
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202 and National Research Council Postdoctoral Associate at the Army Research Office, Research Triangle Park, North Carolina 27709-2211

J. E. Rowe
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202, Physics and Astronomy Department and Laboratory for Surface Modification, Rutgers University, Piscataway, New Jersey 08855-0849, and Physics Division, Army Research Office, Research Triangle Park, North Carolina 27709-2211

J. J. Kolodziej
Physics and Astronomy Department and Laboratory for Surface Modification, Rutgers University, Piscataway, New Jersey 08855-0849

H. Niimi
Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

T. E. Madey
Physics and Astronomy Department and Laboratory for Surface Modification, Rutgers University, Piscataway, New Jersey 08855-0899

G. Lucovsky
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 26 March 1999; accepted 27 May 1999)

High resolution soft x-ray photoelectron spectroscopy with synchrotron radiation is used to study the interfaces of SiO_2/Si(111), SiO_2/Si(100), Si(111)/Si_3N_4, and SiO_2/Si_3N_4 for device-quality ultrathin gate oxides and nitrides. The thin oxides and nitrides were grown by remote plasma deposition at a temperature of 300 °C. Aftergrowth samples were further processed by rapid thermal annealing for 30 s at various temperatures from 700 to 950 °C. The Si(111)/Si_3N_4 samples were air exposed and formed a thin ~6 Å SiO_2 layer with a Si($2p$) core-level shift of 3.9 eV, thus allowing us to study both the Si(111)/Si_3N_4 and SiO_2/Si_3N_4 interfaces with a single type of sample. We obtain band offsets of 4.54±0.06 eV for SiO_2/Si(111) and 4.35±0.06 eV for SiO_2/Si(100) with film thicknesses in the range 8–12 Å. The Si(111)/Si_3N_4 nitrides show 1.78±0.09 eV valence-band offset for 15–21 Å films. This value agrees using the additivity relationship with our independent photoemission measurements of the nitride–oxide valence-band offset of 2.66±0.14 eV. However, we measure a substantially larger SiO_2/Si_3N_4 ΔE_V value of 3.05 eV for thicker (~60 Å) films, and this indicates substantial differences in core-hole screening for films of different thickness due to additional silicon substrate screening in the thinner (15–21 Å) films. © *1999 American Vacuum Society.* [S0734-211X(99)08904-0]

I. INTRODUCTION

Interfaces of SiO_2/Si(111) and SiO_2/Si(100) have been extremely well studied by photoemission spectroscopy[1–5] but little attention has been given to measurements of the valence-band offsets for ultrathin oxide systems. A crucial issue that continues to inhibit understanding of spectroscopic measurements is sample preparation at the device-grade level of processing such that interface details can be usefully compared to other measurements. In the present study we use ultrathin oxides and nitrides grown on Si(111) and on Si(100) that achieve these interface conditions. Since gate oxide thickness used in current devices has continued to decrease in thickness with corresponding improvements in oxide growth, postgrowth processing and device properties, we decided to reinvestigate the issue of the band offsets for interfaces of SiO_2/Si(111) and SiO_2/Si(100) using current state-of-the-art methods of gate oxide growth. Standard interface capacitance and other electrical measurements done on the same wafer as the electron spectroscopic experiments independently characterized our device-grade samples.[6–8]

II. EXPERIMENTAL DETAILS

A. Film preparation methods

Native oxide layers on Si(111) substrates were removed by etching in 40 wt % NH_4F for 4 min and then rinsing for 20 s with deionized water. The Si(100) wafers were treated with 1 wt % HF and also rinsed in de-ionized water for 20 s. These steps produced H-terminated Si surfaces. After this

[a)]Electronic mail: jkeister@unity.nesu.edu

step, the wafers were loaded into the vacuum chamber. The films were then grown by one of three methods.

1. Method 1

Ultrathin SiO_2 samples were made using remote plasma-enhanced (RPE) oxidation. Wafers of Si(111) or Si(100) were heated to 300 °C and exposed to a flow of excited oxygen formed in a remote He/O_2 rf plasma. The He and O_2 were flowed at 200 and 20 sccm ($cm^3\,min^{-1}$), respectively, at a total chamber pressure of 0.3 Torr, with the 13.56 MHz rf plasma power fixed at 30 W. On-line (i.e., *in situ*) Auger electron spectroscopy (AES) measurements characterized the growth rate as following a power law SiO_2 film thickness dependence on time: $t_{ox} \approx 7t^{0.28}$ (where t_{ox} is the SiO_2 thickness in Å and t is plasma exposure time in minutes).[9] The SiO_2 films reported in this study [measured by soft x-ray photoelectron spectroscopy (SXPS) to be in the range 9–22 Å] were made using exposure times of approximately 5 s to nearly 10 min.

2. Method 2

Remote plasma-enhanced chemical vapor deposition (RPECVD) of SiO_2 to produce thicker (~80 Å) films. These were produced by flowing dilute silane downstream from the oxygen plasma. After passing the plasma excitation, excited O_2 flowed through a rf ring of 2% SiH_4 in helium. The wafers were found to grow at ~34–40 Å per minute at a wafer temperature of 300 °C. The chamber pressure was maintained at 0.3 Torr during growth. Flow rates were: 200 sccm He, 20 sccm O_2, and 10 sccm of the 2% SiH_4 mixture.

3. Method 3

RPECVD was also used to produce (hydrogen-rich) Si_3N_4 films ranging in thickness from 10 to 80 Å. Excited N_2 from the He plasma was flowed through a ring of SiH_4. The wafers were observed to grow at a rate of ~6–8 Å/min at a wafer temperature of 300 °C. The chamber pressure was maintained at 0.2 Torr during growth. Flow rates were: 200 sccm He, 60 sccm N_2, 10 sccm of the 2% SiH_4 mixture. When these samples were exposed to air, reaction with atmospheric water produced a native silicon oxide overlayer (~5–10 Å thickness), presumably by the substitution reaction:

$$-NH-+H_2O \rightarrow -O-+NH_3. \quad (1)$$

After growth, some of the samples were subjected to *ex situ* rapid thermal annealing (RTA) using optical heating in an Ar atmosphere. In many cases, a single oxidized wafer was broken *ex situ* and fragments were annealed at various temperatures in the range 600–950 °C by RTA (for 30 s) in Ar, providing a means of comparing "as grown" and annealed samples under the same oxidation conditions.

The films were transferred under air to the ultrahigh vacuum XPS chamber of NSLS-U4A. Once in vacuum, the samples were annealed at ~500 °C by electron-beam heating of the metal sample holder. This removes contamination due to weakly bonded adsorbates picked up in air. The thick nitride films were further treated with neon ion etching (5 min at 500 eV beam energy, 3×10^{-5} Torr chamber pressure) to remove some of the SiO_2 overlayer. Then, a second ~600 °C anneal was performed after the etching in order to reequilibrate the film somewhat.

B. Soft x-ray photoemission spectroscopy (SXPS) measurements

The SXPS configuration at the U4A beamline of the National Synchrotron Light Source (NSLS) includes a 6 m toroidal grating monochromator which produces a photon beam with $\leq$0.2 eV resolution at photon energies ($h\nu$) of 10–200 eV.[9,10] At 130 photon energy this system has ~0.15 eV total resolution, photons, and electrons. The photoelectron kinetic energy (KE_e) was measured with a 100 mm hemispherical analyzer fixed at 45° to the photon beam axis. All the spectra presented here were obtained with the sample surface facing the analyzer at the normal emission geometry (α=90° take-off angle). The room temperature SXPS spectra were collected with the sample holder grounded, and the electron analyzer was used in fixed pass energy mode with a resolution of ~0.1 eV.

The Fermi energy E_F was used as the reference energy and was measured using a metal sample attached to the same sample holder. The E_F threshold appeared at a kinetic energy of 4.6 eV less than the photon energy. Valence-band offsets were measured within each spectrum individually and were not affected by any small offset bias applied to the sample. The sample SiO_2 film thickness was estimated from the Si($2p$) SXPS SiO_2 and substrate peak intensity ratio, as described elsewhere,[11] and were consistent with independent *in situ* (on-line) AES and ellipsometric measurements within experimental uncertainties.

III. RESULTS

A. Valence-band offsets for thin SiO_2 films on Si(111) and Si(100)

Because of the surface sensitivity of SXPS, the Si substrate signal can be seen only for thin films. Thus, the Si–SiO_2 valence-band offset ΔE_V was measured using thin RPE oxidation SiO_2 films of 10±2 Å thickness. This film thickness was measured from the core-level peak intensity ratio between Si and SiO_2 Si($2p$) peaks [binding energy (BE, (Si) ~99.3 eV)] which are distinct due to the ~4 eV chemical shift difference in BE (SiO_2) ~104 eV. As shown in Fig. 1, the valence band spectra for such films show the same features despite being obtained using different photon energies. The differences can be explained as: (i) Auger transitions which have fixed electron energies, or (ii) intensity changes due to energy-dependent electron escape depths. These spectra are dominated by a signal attributable to oxygen atoms in the SiO_2 film. The O($2s$) peak at binding energy ($h\nu$−KE)~26 eV can clearly be distinguished from the O($2p$) peak at 7 eV (with additional components at ~11 and 14 eV). The onset of the O($2p$) signal at BE ~4 eV corresponds to the valence band maximum (E_V) of the SiO_2 film.

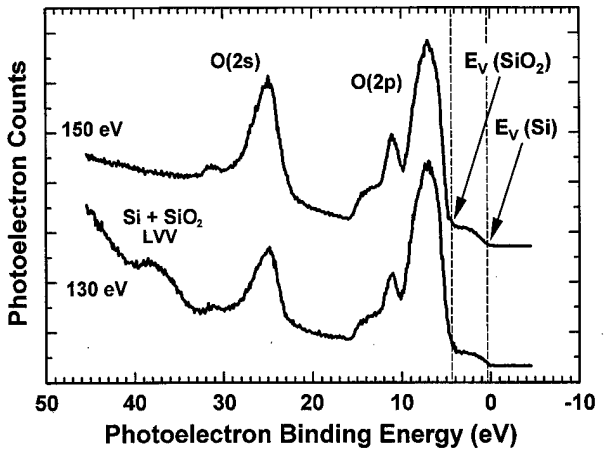


FIG. 1. Typical valence band scans for a thin (~15 Å) SiO_2 film grown on Si(111) substrate. Shown are spectra acquired at 130 and 150 eV photon energy. The higher energy shows more substrate signal (BE <4 eV) due to increased electron escape depth. The lower energy scan shows greater scattered electron background as well as silicon Auger signal near 90 eV electron kinetic energy. The small peak near 31 eV binding energy is Ta($4f$) XPS signal from the tantalum wire which was used to hold the silicon wafer fragment to the sample holder.

The signal at lower binding energy results from the Si substrate, which has an onset BE of nearly zero as shown in Fig. 1.

The difference between these two onsets corresponds directly to the valence-band offset of the interface. That is, $\Delta E_V = E_V(\mathrm{Si}) - E_V(\mathrm{SiO_2})$. Thus, we have measured ΔE_V by fitting the low-binding-energy part of the spectrum with two broadened step functions. In Fig. 2 is shown the edge region of an SXPS spectra obtained at 130 eV photon energy for a SiO_2 film of 10 Å thickness grown on Si(111). We have found that the spectrum in this region is well modeled by a pair of Gaussian-broadened Fermi functions (the Fermi width is virtually negligible in this case compared with the Gaussian component), with a parabolic peak added to match the signal found experimentally just below the Si band edge.

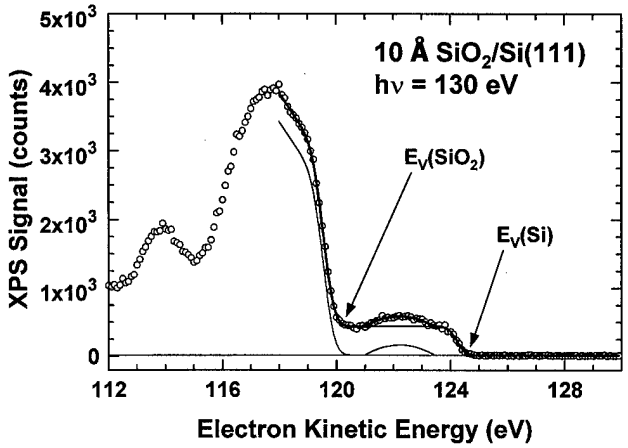


FIG. 2. Closeup view of the silicon and silicon dioxide valence-band edges measured for a 10 Å SiO_2 film grown on Si(111) substrate using a photon energy of 130 eV. The fitting functions are described in the text.

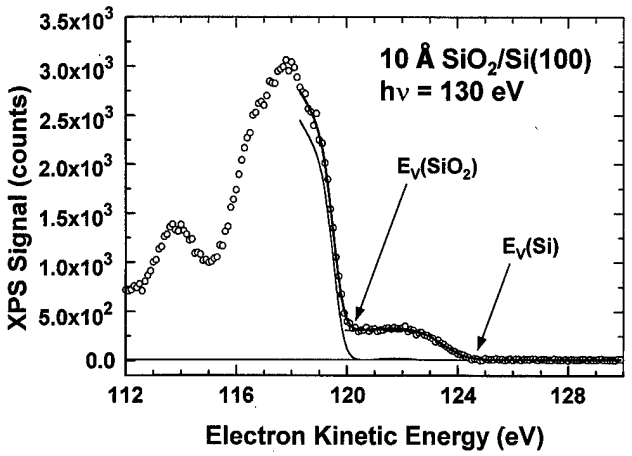


FIG. 3. Closeup view of the silicon and silicon dioxide valence-band edges measured for a 10 Å SiO_2 film grown on Si(100) substrate using a photon energy of 130 eV. The fitting functions are described in the text.

Fits such as these, made using samples of similar thickness and various postoxidation anneal temperatures, were used to derive an average value for ΔE_V of the thin SiO_2/Si(111) interface of 4.54±0.06 eV. ΔE_V was found to be quite constant independent of annealing. Similarly, we have fit the valence-band edge spectra for thin SiO_2 films on the Si(100) substrate. These spectra are similar (see Fig. 3), but with two distinct differences (i) the parabolic peak is much less for the Si substrate, and (ii) onset for Si bulk is much broader, at nearly twice the width of the peaks encountered so far (~1.5 eV compared to 0.7 eV). Again, the ΔE_V values obtained appear independent of annealing treatment, but show a distinct substrate-orientation dependence as pointed out by Alay and co-workers.[12] For ~10 Å SiO_2/Si(100) we find $\Delta E_V = 4.35 \pm 0.06$ eV, which is 0.19±0.08 eV lower than for Si(111). Quantitatively these results compare favorably with previous work (see Table I). However, the present work is done with better resolution and device-proven processes, so we believe our work to be superior.

This direct measure of the valence-band offset can be compared with an indirect approach as well.[13] In particular, this ΔE_V value can be used in conjunction with $(E_V - E_{CL})$ values (the electron energy difference between Si($2p$) core level peak and valence band edge) measured for thin and thick SiO_2 films, in order to derive a Si–SiO_2 core level shift comparable to known values.[1,2,11] The estimated ΔE_{CL} value is given by:

TABLE I. Compared band offset measurements (values are given in eV).

Value	Si(100)	Si(111)	ΔSi(100)–Si(111)
ΔE_V	4.3[a]	4.5[b]	
(eV)	4.43 wet[c]	4.36 (dry)[c]	0.13[c]
	4.49 dry[c]		
	4.54±0.06 eV[d]	4.35±0.06[d]	0.19±0.08[d]

[a]Reference 2.
[b]Reference 1.
[c]Reference 12.
[d]Present work.

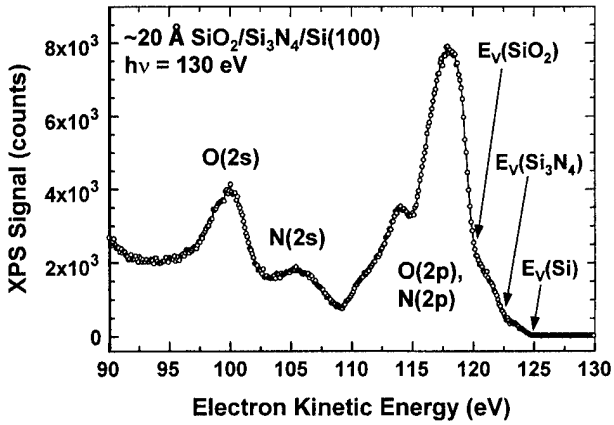


FIG. 4. Valence-band scan measured at 130 eV photon energy for a ~20 Å Si_3N_4 film deposited on Si(100) using the RPECVD method. A silicon dioxide overlayer resulting from atmospheric exposure is the dominant feature. In addition, the Si_3N_4 valence-band edge appears between those of silicon substrate and SiO_2. The N(2*s*) peak at ~106 eV kinetic energy is evident, and resembles the O(2*s*) peak at ~100 eV.

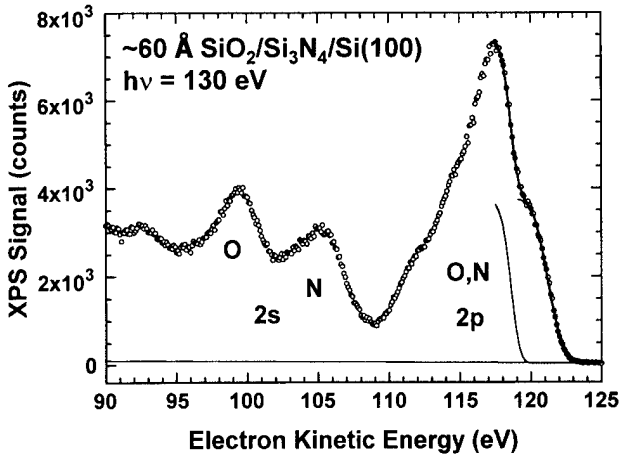


FIG. 5. Valence-band scan measured at 130 eV photon energy for a ~60 Å Si_3N_4 film deposited on Si(100) using the RPECVD method. This film is quite thick compared to the electron escape depth (~5–10 Å) so that no substrate signal is seen. The large valence-band offset for SiO_2/Si_3N_4 measured from this spectrum differs from that measured for thinner films, in part due to photoelectron hole screening by silicon substrate.

$$\Delta E_{CL} = \Delta E_V(\mathrm{Si/SiO_2}) + (E_V - E_{CL})(\mathrm{SiO_2}) - (E_V - E_{CL})(\mathrm{Si}). \quad (2)$$

We can average the Si(111) and Si(100) results for ΔE_V to give 4.44 eV, and obtain an average value of $(E_V - E_{CL})$(Si) of 98.52 eV by using data for the same thin (~10 Å) SiO_2 film samples. For the $(E_V - E_{CL})(SiO_2)$ value we turn to a thick film of SiO_2 (~80 Å) which yields 98.45 eV. The derived ΔE_{CL} value of 4.37 eV compares very well with a ΔE_{CL} value of 4.36 eV estimated for a 80 Å film using the core-hole screening approach.[11]

B. Valence-band offsets for thin Si_3N_4 films

We have also measured valence-band offsets for silicon nitride films deposited on Si(111) and Si(100) substrates. An example of such valence-band data for <20 Å Si_3N_4 films on Si(100) is shown in Fig. 4. The hydrogen-rich silicon nitride (made by RPECVD, using N_2^* and SiH_4) is known to react readily with atmospheric water to produce a surface layer of silicon (di) oxide. Thus, the spectra show three distinct offsets: (i) Si substrate, (ii) Si_3N_4, and (iii) SiO_2. We have used these data to measure ΔE_V for the Si_3N_4/Si interface. Unfortunately, the overlapping spectra in this case makes the Si_3N_4/Si offset ΔE_V less clear. In addition, the Si_3N_4 onset was not sharp enough to measure for many films thinner than ~16 Å. From the best of such data, we arrive at an average value for ΔE_V (Si_3N_4/Si) of 1.78±0.09 eV for an average total film thickness (oxide and nitride) of 18±3 Å. We could not find any statistically significant differences between samples of difference substrate orientation or annealing treatment.

We can estimate the Si_3N_4/SiO_2 valence-band offset from these data as well. Again, using a pair of broadened step functions, we have fit the nitride and oxide onsets from the spectra and found an average Si_3N_4/SiO_2 valence-band offset of 2.66±0.14 eV. It is comforting to find that the ΔE_V values measured independently for Si–SiO_2, Si–Si_3N_4, and SiO_2–Si_3N_4 follow a simple additivity rule: i.e., $\Delta E_V(\mathrm{Si–SiO_2}) = \Delta E_V(\mathrm{Si–Si_3N_4}) + \Delta E_V(\mathrm{Si_3N_4–SiO_2})$. In this case, adding the two nitride-relative offsets 2.66 and 1.78 eV yields 4.44 eV which is precisely the average of the ΔE_V values for the SiO_2 films on the two crystal faces (4.45 ±0.09 eV). However, when we measure the SiO_2–Si_3N_4 ΔE_V value from spectra of thicker films (~60 Å), we find it much higher (see Fig. 5). The ΔE_V value of 3.06 eV measured from the spectra in this figure may be higher due to the screening effect difference for different oxide overlayer thickness, as discussed below.

IV. DISCUSSION AND CONCLUSIONS

The valence-band offsets for silicon/silicon oxide and silicon/silicon nitride have been measured with good accuracy using soft x-ray photoemission. These values are critical in the design of oxide-semiconductor devices. In agreement with previous workers, we have found a difference in valence-band offset between the Si(100) and Si(111) crystal interfaces with silicon dioxide. This can be correlated with the number and type of other species at the interface, such as dangling bond defects or hydrogen.[14] This interpretation would suggest that the Si(111) interfaces have less interface H, assuming no dangling bonds. A second explanation for the ΔE_V difference is the orientation of the interface bond dipoles relative to the surface normal. The interface bond dipole is more aligned with the surface normal for the (111) orientation and less aligned (more glancing) for the (100) orientation. If the contribution of the interface bond dipole to the valence-band offset (total interface dipole) follows a cosine rule, this would also explain the lower ΔE_V for Si(100)/SiO_2 relative to Si(111)/SiO_2.

The silicon nitride valence-band edge has also been measured relative to those of Si bulk and SiO_2. These values appear to be less sensitive to the underlying Si crystal orien-

tation. The Si_3N_4–SiO_2 ΔE_V has been found to be somewhat film thickness dependent, increasing by ~0.4 eV between 18 and 60 Å film thicknesses. This can be explained as a result of core-hole screening by image charge in the higher k silicon substrate. In other words, the high dielectric constant ϵ_{opt} mismatch ($\Delta\epsilon_{opt}=12-2=10$) between silicon and the SiO_2 substrate is likely to be responsible for this shift. The mismatch is about half as great in the case of screening by the Si substrate of silicon nitride ($\Delta\epsilon_{opt}=12-7=5$). This screening effect causes lower binding energies for thin films and higher binding energies in thicker films.

The results presented here can be used to derive values for the conduction-band offsets as well, using the known E_g values. For example, the known SiO_2 gap energy of 8.95 eV[15] implies a Si–SiO_2 ΔE_C of 3.38 eV (using the average ΔE_V of 4.44 eV and Si bandgap of 1.1 eV).

ACKNOWLEDGMENTS

This work was supported, in part, by DOE-Office of Basic Energy Sciences, by NSF, by NRC, by the Office of Naval Research, and the Army Research Office.

[1]F. J. Grunthaner and P. J. Grunthaner, Mater. Sci. Rep. **1**, 65 (1986).
[2]F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B **38**, 6084 (1988).
[3]M. M. Banaszak-Holl, S. Lee, and F. R. McFeely, Appl. Phys. Lett. **65**, 1097 (1994).
[4]K. Z. Zhang, M. M. Banaszak-Holl, J. E. Bender, S. Lee, and F. R. McFeely, Phys. Rev. B **54**, 7686 (1996).
[5]F. R. McFeely, K. Z. Zhang, M. M. Banaszak-Holl, S. Lee, and J. E. Bender, J. Vac. Sci. Technol. B **14**, 2824 (1996).
[6]H. Niimi and G. Lucovsky, Surf. Coat. Technol. **98**, 1529 (1998).
[7]H. Niimi, K. Koh, and G. Lucovsky, Proc. Electrochem. Soc. **12**, 623 (1996).
[8]G. Lucovsky, H. Niimi, Y. Wu, C. R. Parker, and J. R. Hauser, J. Vac. Sci. Technol. A **16**, 1721 (1998).
[9]P. Thiry, P. A. Bennett, S. D. Kevan, W. A. Royer, E. E. Chaban, J E. Rowe, and N. V. Smith, Nucl. Instrum. Method Phys. Res. **222**, 85 (1984).
[10]G. K. Wertheim, J. E. Rowe, D. M. Riffe, and N. V. Smith, AIP Conf. Proc. **215**, 259 (1990).
[11]J. W. Keister, J. E. Rowe J. J. Kolodziej, H. Niimi, H.-S. Tao, T. E. Madey, and G. Lucovsky J. Vac. Sci. Technol. A (submitted).
[12]J. L. Alay and M. Hirose, J. Appl. Phys. **81**, 1606 (1997).
[13]E. T. Yu, M. C. Phillips, J. O. McCaldin, and T. C. McGill, J. Vac. Sci. Technol. B **9**, 2233 (1991).
[14]P. Perfetti, C. Quaresima, C. Coluzza, C. Fortunato, and G. Margaritondo, Phys. Rev. Lett. **57**, 2065 (1986).
[15]S. Miyazaki, H. Nishimura, M. Fukuda, L. Ley, and J. Ristein, Appl. Surf. Sci. **113/114**, 585 (1997).

Interfacial properties of ultrathin pure silicon nitride formed by remote plasma enhanced chemical vapor deposition

V. Misra,[a] H. Lazar, Z. Wang, and Y. Wu
Departments of Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695

H. Niimi
Departments of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695

G. Lucovsky
Departments of Electrical and Computer Engineering, Materials Science Engineering and Physics, North Carolina State University, Raleigh, North Carolina 27695

J. J. Wortman and J. R. Hauser
Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695

(Received 3 May 1999; accepted 5 May 1999)

This article addresses the electrical properties of interfaces between *n*- and *p*-type Si and remote plasma-deposited Si_3N_4, which are of interest in aggressively scaled advanced CMOSFETs. The nitride films of this article display excellent electrical properties when implemented into stacked oxide/nitride dielectrics in both NMOSFETs and PMOSFETs with oxide, or nitrided oxide interfaces. The same nitride layers deposited directly onto clean Si surfaces display degraded electrical properties with respect to devices with oxide, or nitrided oxide interfaces. PMOS interfaces are significantly more degraded than *n*-type metal–oxide semiconductors interfaces indicating a relatively high density of donor-like interface traps that inhibit channel formation.
© *1999 American Vacuum Society.* [S0734-211X(99)04904-5]

I. INTRODUCTION

Gate dielectric scaling requires new materials with dielectric constants (K) higher than SiO_2 to provide the increased capacitance without compromising gate leakage current.[1,2] One such material is Si_3N_4 that has approximately twice the dielectric constant of SiO_2, and additionally is effective in blocking diffusion of dopants such as boron. Si_3N_4 has also been proposed as an interface layer for high K transition metals oxides since it has a higher K than SiO_2 and is also an excellent diffusion barrier. To date, the interfacial and bulk properties of Si_3N_4 have not been encouraging, and only a few processes have shown promising results.[3–5] One such process is jet vapor deposition (JVD) in which the current transport mechanism in the dielectric films is predominantly electron tunneling similar to what has been observed in SiO_2.[4] However these films contain a relatively high concentration of O (~20 at %) which reduces the dielectric constant, and additionally, the devices have $Si–SiO_2$ interfaces. Recent studies at NCSU have shown that nitride films formed by RPECVD, and subjected to a postdeposition anneal in an inert ambient at 900 °C produce good quality films with relatively low concentrations of both O (<2 at %) and H (<14%).[3,6] When used in SiO_2/Si_3N_4 stacks (O/N), in which a thin oxide layer separates the nitride film from the substrate, these films display excellent electrical properties in both NMOS[7] and PMOS devices.[4] The goal of this work is to extend the study of the RPECVD nitride as the gate dielectric, and in particular to determine the electrical properties of $Si–Si_3N_4$ interfaces in both NMOS and PMOS devices.

II. EXPERIMENTAL PROCEDURES

Silicon nitride films were formed in a cluster tool in which the entire gate stack is completed before exposure to atmosphere. Prior to gate stack formation, a 200 nm field oxide was grown on 100 mm substrates followed by patterning of the active areas. A 10 nm sacrificial oxide was then grown and removed with 2% HF solution immediately before insertion into the cluster tool. Nitride layers were formed by RPECVD at 400 W, 300 mTorr and 300 °C using $SiH_4+NH_3(N_2)$.[3,7,8] Single layer nitrides varying in thickness from 1.5 to 3.0 nm were formed on both *n*- and *p*-type substrates. In addition to forming single layer nitride films, nitride/oxide stacks were also formed to study the properties of ultrathin nitride at the interface while using conventional characterization techniques. After deposition, films were annealed at 900 °C in vacuum to reduce the hydrogen content and promote formation of additional Si–N bond.[3,6] Gate dielectrics were then capped by rapid thermally deposited polysilicon in the cluster tool system. Control oxides of varying thickness were also prepared by thermal oxidation. Surface channel NMOS and PMOS devices were formed by conventional techniques, and a forming gas ($10\%H_2$ in N_2) anneal at 400 °C was performed after metallization.

[a]Electronic mail: vmisra@eos.ncsu.edu

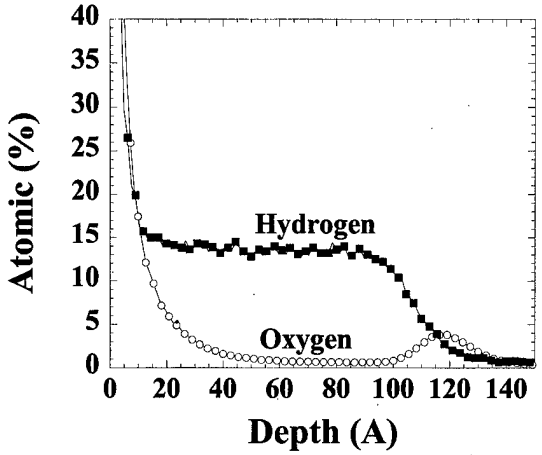


FIG. 1. Secondary ion mass spectroscopy depth profile of RPECVD nitride.

III. RESULTS

Secondary ion mass spectroscopy analysis of RPECVD nitride films is shown in Fig. 1. The hydrogen content is less than 15% throughout the film. The oxygen concentration is also very low (<3%) with a slight increase at the interface. The low hydrogen and oxygen content makes this film very different from the nitrides deposited via LPCVD techniques and is expected to exhibit different electrical characteristics. The slight increase of oxygen at the interface may be related to surface contamination from being exposed to the cleanroom before being loaded in the RPECVD chamber. In order to determine the dielectric constant, the physical thickness was measured by transmission electron microscopy and compared to the electrical oxide equivalent thickness. This gave a dielectric constant of 7.3.

I_{ds} vs V_g measurements were taken on NMOS and PMOS devices. The thickness were extracted from capacitance–voltage data using a least-square fit method.[9] The equivalent oxide thicknesses of the nitride and control oxide were 2.1 and 2.3 nm, respectively. As shown in Fig. 2(a), the data for NMOSFETs with nitride dielectrics had a lower V_t compared to the control devices agreeing with what has been observed for nitrided oxides; this is attributed to the fixed positive charge.[10–12] The nitride devices also displayed an ~50% degraded peak mobility, but an enhanced high-field mobility. The behavior of reduced peak degradation and enhanced

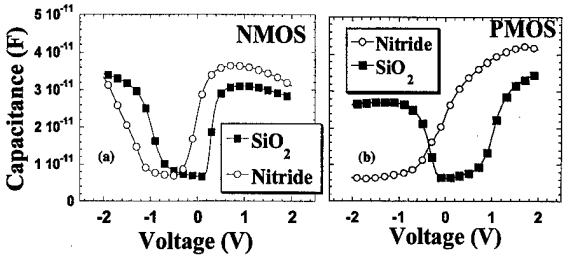


FIG. 3. Capacitance–voltage on $50\times50\ \mu m^2$ (a) NMOSFETs and (b) PMOSFETs taken at 1 MHz. The equivalent oxide thickness for nitride and SiO_2 are 2.1 and 2.3 nm, respectively.

high field mobility are consistent what has been observed in heavily nitrided oxides.[10,12] The PMOS data with nitride dielectric, shown in Fig. 2(b), display markedly different characteristics compared to NMOS devices. The channel conduction is significantly reduced, and there is a very large V_t shift, ~1.5 V. The degraded I_{ds} and large V_t shift make these PMOS devices effectively nonfunctional. To better understand the cause of the degraded PMOS behavior, capacitance–voltage (C–V) measurements were performed on transistors to obtain both accumulation and inversion characteristics.

As shown in Fig. 3(b), the C–V measurements on PMOSFET devices did not indicate the presence of an inversion layer which is consistent with the severely degraded drive currents obtained from the IDVG measurements. However, normal characteristics were observed in accumulation suggesting that only holes in the channel were being affected. To see if this behavior was independent of the substrate conductivity, C–V measurements were also made on NMOSFETs. As shown in Fig. 3(a), the NMOS devices showed degraded characteristics near the accumulation region, whereas the inversion C–V curves were normal. Although the degradation in the NMOS devices was lower than in the PMOS, these results demonstrated that the nitride dielectric was affecting channel holes significantly more than channel electrons.

To determine the effect of the interface nitride thickness on the C–V characteristics, the nitride/oxide stacks were also measured on PMOSFETs. As shown in Fig. 4, a nitride thickness of >5 Å can drastically influence the inversion characteristics. This implies that the interfacial nitride dras-

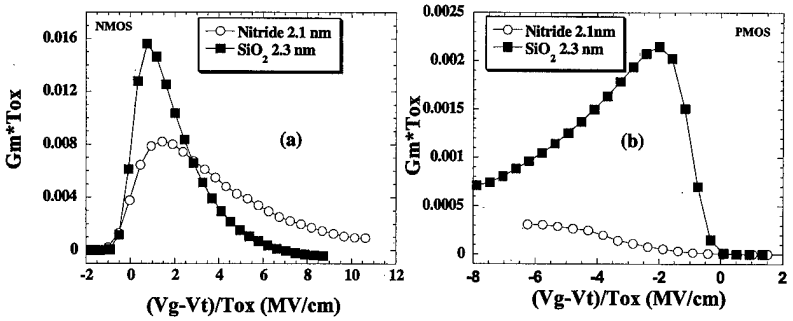


FIG. 2. $G_m{*}T_{ox}$ vs. oxide field for silicon nitride and SiO_2 for 0.6 μm channel length (a) NMOSFETs and (b) PMOSFETs. The oxide field was calculated as $(V_g - V_t)T_{ox}$.

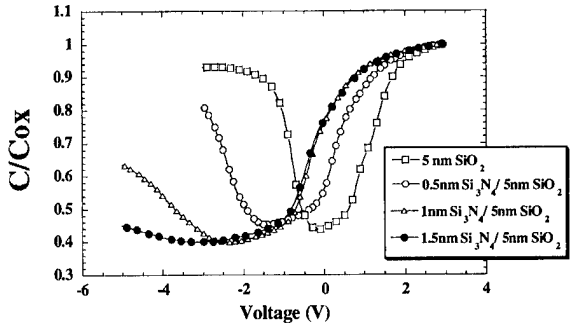


FIG. 4. Capacitance–voltage on 50×50 PMOSFET with varying nitride/oxide stack layers taken at 1 MHz.

tically modulates the hole carriers in the channel region.

Gate tunneling current was measured both in accumulation and inversion for both NMOS and PMOS devices (area=$10\times3\ \mu m^2$) and is shown in Figs. 5(a) and 5(b). The voltage drop across the oxide was obtained using V_g-V_t in inversion and V_g-V_{fb} in accumulation. The nitride shows slightly lower tunneling current than SiO_2 in the positive region for both NMOS [Fig. 5(a)] and PMOS [Fig. 5(b)] devices. However under negative bias conditions, the gate current through the nitride layers is excessive compared to SiO_2 even though the nitride layers are physically thicker.

In an effort to understand the mechanisms responsible for the degraded hole behavior, a carrier separation technique was used to separately measure the hole and electron currents.[13] A positive gate bias was applied to the NMOS devices, and a negative gate bias was applied to the PMOS, with the source, drain, and substrate grounded. The gate bias creates a channel and as carriers tunnel out of the channel into the gate dielectric they are replaced by carriers from the source/drain regions. Measurement of gate, channel, and substrate current provides a measure of the electron and hole current components. The results for SiO_2 for both PMOS and NMOS are shown in Figs. 6(a) and 6(b). For both transistors, the gate current is nearly equal to the channel current for all voltages implying that the majority of the carriers that tunnel to the gate come from the channel, i.e., electron current dominates in the NMOS and hole current dominates in the PMOS. The substrate current is a very small percentage

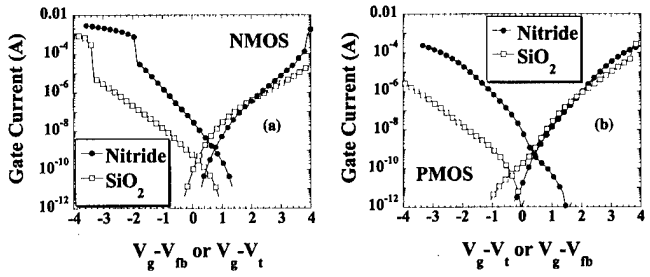


FIG. 5. Gate current measurements for (a) NMOSFET and (b) PMOSFET in inversion and accumulation. The equivalent oxide thickness for nitride and SiO_2 are 2.1 and 2.3 nm, respectively. The voltage across the oxide is taken as V_g-V_t in inversion and V_g-V_{fb} in accumulation.

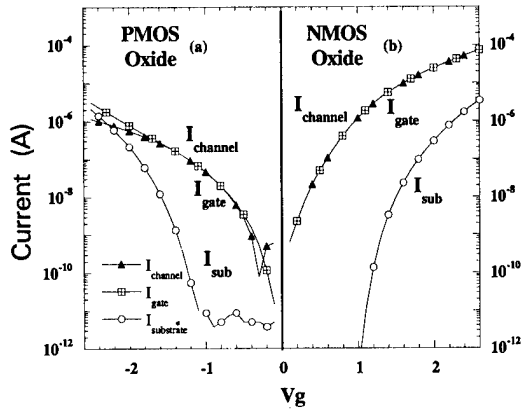


FIG. 6. Carrier separation measurements for silicon nitride on (a) PMOSFET and (b) NMOSFET devices.

of the gate current. The measurements for Si_3N_4 for PMOS and NMOS are shown in Figs. 7(a) and 7(b). The NMOS behavior is similar to the control oxide, i.e., the gate current is dominated by source/drain channel current suggesting that electron current is the dominating process in this regime. However, the PMOSFETs with nitride dielectric display markedly different behavior than their SiO_2 counterparts. First, at smaller gate biases the drain current is very low and is not equal to the gate current suggesting that the channel holes are not participating in the carrier conduction in this regime. Second, the substrate current is very high in this regime and equals the gate current. This implies that the electron conduction from the gate is the dominating process. It should also be noted that after V_g exceeds approximately -1.5 V the channel current starts increasing and becoming equal to the gate current.

IV. DISCUSSION

The proposed explanation for the above results is as follows. The nitride dielectric is believed to have a large density of interface traps below the midgap. These traps remain oc-

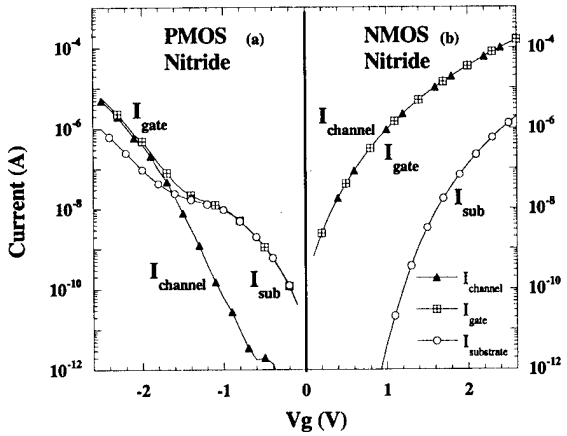


FIG. 7. Carrier separation measurements for silicon nitride on (a) PMOSFET and (b) NMOSFET devices.

cupied and neutral if the Fermi level lies above them as in the case of NMOS inversion. However, as the Fermi level starts moving down towards the Si valence band edge, as in PMOS inversion, these traps get depopulated and become positively charged. The high density of these positive traps in the nitride layers can screen the gate charge until all interface traps become depopulated. This is believed to occur at approximately -1.5 V, which is the same V_t value extracted from the I_d-V_g curves [see Fig. 2(b)]. After this voltage is reached, a surface potential is allowed to drop in the channel resulting in the onset of inversion. This is believed to be the mechanism responsible for the degraded PMOS characteristics in inversion.

As mentioned, the high density of traps below the midgap can effectively screen the gate charge preventing any voltage drop in the semiconductor. This screening process can also explain the high gate tunneling current observed with nitride dielectrics in the negative regime. A much higher electric field is created across the nitride where $V_{\text{nitride}}=V_{\text{applied}}$ as compared to the control oxide where $V_{\text{oxide}}=V_{\text{applied}}-V_{\text{substrate}}$. This higher field assists in large conduction current through the dielectric.

Recently, reports of bonding constraint theory calculations have indicated that the defect formation can be closely related to bonding coordination at the interface of two materials.[14,15] The average bonding coordination per atom of SiO_2–Si interface (including the suboxide region) is 2.9. This is sufficiently low and promotes the formation of low defect density interfaces. On the other hand, the average bonding coordination per atom at the Si_3N_4–Si interface (including the transition region) was calculated to be ~3.5 which can promote a high defect density. In this work it is believed that this overcoordination of the Si_3N_4–Si compared to the Si–SiO_2 interface may result in interface trap formation that negatively impacts the electrical properties. To further verify this, a thin layer of oxide (<0.6 nm as measured by *in situ* Auger) was interposed between the nitride and the Si substrate. Transconductance values, shown in Fig. 8, now indicate normal characteristics for P and N devices suggesting that ~0.6 nm of interfacial oxide is sufficient to reduce the interface state density both near the conduction and valence band edges. It should also be noted that the presence of oxygen may assist in reducing the trap density by reducing the average bonding coordination and systematic experiments are being performed to verify this.

V. CONCLUSIONS

Data presented in this article demonstrate that degradation of PMOS and NMOS devices with nitride gate dielectrics derive from interfacial defects. It is believed that a markedly different distribution of interfacial traps near the valence band versus the conduction band edges results in the inability to form a PMOS channel. The observation that these traps are not present in PMOS devices in which an ultrathin oxide layer <0.6 nm is interposed between the Si and the nitride film demonstrates that these defects are a property of the Si–Si_3N_4 interface. Therefore, a pure silicon nitride film cannot be used directly on the silicon surface due to a high density of positively charged, donor-like defects in the lower half of the band-gap near the valence band edge of Si which prevents the formation of a conducting channel. The addition of oxygen as an ultrathin interfacial oxide layer and/or in the bulk of nitride film to form an oxynitride alloy, may be necessary if nitride interface layers are to be integrated into aggressively scaled CMOS devices.

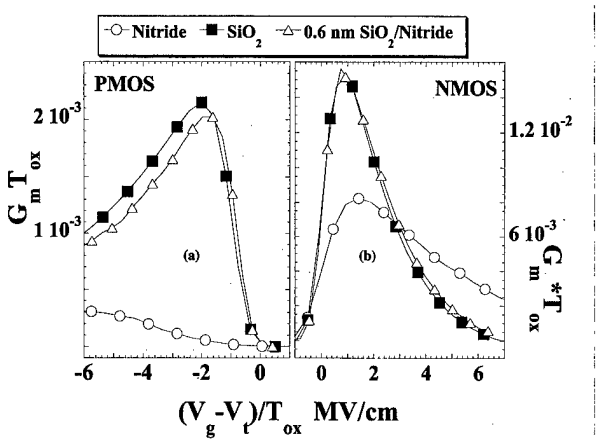


FIG. 8. $G_m{*}T_{ox}$ vs oxide field for silicon nitride, SiO_2 and stacked O/N dielectrics for 0.6 μm channel length (a) PMOSFETs and (b) NMOSFETs. The oxide field was calculated as $(V_g-V_t)/T_{ox}$.

ACKNOWLEDGMENTS

This research is supported in part by NSF, SRC, SEMATCH, and ONR.

[1]Semiconductor Industry Association, The National Technology Roadmap for Semiconductors.
[2]S. A. Campbell, D. C. Gilmer, X.-C. Wang, M.-T. Hsieh, H.-S. Kim, W. L. Gladfelter, and J. Yan, IEEE Trans. Electron Devices **44**, 104 (1997).
[3]Y. Ma, T. Yasuda, and G. Lucovsky, Appl. Phys. Lett. **64**, 2226 (1994).
[4]T. P. Ma, IEEE Trans. Electron Devices **45**, 680 (1998).
[5]B. Y. Kim, H. F. Luan, and D. L. Kwong, Tech. Dig. Int. Electron Devices Meet. **1997**, 463 (1997).
[6]Z. Lu, M. J. Williams, P. F. Santos-Filho, and G. Lucovsky, J. Vac. Sci. Technol. A **13**, 607 (1995).
[7]C. G. Parker, G. Lucovsky, and J. R. Hauser, IEEE Electron Device Lett. **19**, 106 (1998).
[8]Y. Wu and G. Lucovsky, IEEE Electron Device Lett. **19**, 367 (1998).
[9]J. R. Hauser and K. A. Ahmed (unpublished).
[10]H. Momose, S. Kitagawa, K. Yamabe, and H. Iwai, Tech. Dig. Int. Electron Devices Meet. **1989**, 267 (1989).
[11]T. Hori, Y. Naiot, H. Iwasaki, and H. Esaki, IEEE Electron Device Lett. **7**, 669 (1986).
[12]T. Hori, IEEE Trans. Electron Devices **37**, 2058 (1990).
[13]A. S. Ginovker, V. A. Gristsenko, and S. P. Sinitsa, Phys. Status Solidi A **26**, 489 (1974).
[14]J. Philips, J. Non-Cryst. Solids **34**, 153 (1979).
[15]G. Lucovsky, Y. Wu, H. Niimi, V. Misra, and J. Philips, Appl. Phys. Lett. **74**, 2005 (1999).

Remote plasma enhanced chemical vapor deposition SiO_2 in silicon based nanostructures

M. J. Rack,[a] L. L. Hilt, D. Vasileska, and D. K. Ferry
Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6206

(Received 19 January 1999; accepted 18 May 1999)

In the depletion gate approach to silicon-based nanostructures, a deposited oxide covering tiny metallic or silicide gate structures must function as a metal-oxide-semiconductor field effect transistor gate oxide. Depending upon the implementation, it may form the Si/SiO_2 interface or be placed upon a very thin thermal oxide. In the former case, bonding between the silicon and SiO_2 must be nearly perfect in order to achieve the high mobility required for observing quantum effects. In the latter case, the deposited oxide must not allow leakage current to obscure effects being observed, nor degrade the previously established thermal interface during deposition. Remote plasma enhanced chemical vapor deposited (RPECVD) silicon dioxide has been studied for use in silicon-based nanostructures. For thin oxides deposited at low temperature, oxide surface roughness has been shown to perturb the potential landscape seen by an electron traveling in a silicon inversion layer. [M. J. Rack, D. Vasileska, D. K. Ferry, and M. Siderov, J. Vac. Sci. Technol. B **16**, 2165 (1998).] We have further explored processing parameters that influence the deposited oxide surface roughness, presumably roughened primarily by gas phase nucleation, and examined the correlation between roughness and oxide reliability. We found that processing conditions that reduce the oxide roughness and unwanted oxidation of the depletion gate structures are not necessarily those that produce the best silicon/oxide interfaces, nor the most defect-free bulk oxides in our RPECVD system. Specific processes tailored for particular device strategies are presented. The low temperature process at 175 °C is extensively explored. © *1999 American Vacuum Society.* [S0734-211X(99)08304-3]

I. INTRODUCTION

Three-dimensional confinement of electrons in silicon based nanostructures is achievable through the use of nano-scaled gates that pattern the two-dimensional electron gas (2DEG) of a metal-oxide-semiconductor field effect transistor (MOSFET) inversion layer.[1–3] As shown in Fig. 1(a), these depletion gates may be positioned directly on the silicon substrate if they form a Schottky barrier with p-type silicon. Alternatively, they may be placed slightly above the silicon substrate as shown in Fig. 1(b), where the depletion gates are separated from the silicon substrate by a thin dielectric. The source and drain are not shown in this figure. When negatively biased, these tiny gates locally deplete carriers in the inversion layer below them, creating confinement barriers in the conduction band energy for electrons.

The shape of a well created by depletion gates influences the quantized energy spacing and ranges from effectively square to parabolic, depending upon a number of factors. These include the spacing and size of the depletion gates, the bias of the depletion gates, V_{depl}, the bias of the top inversion gate, V_G, the substrate doping, as well as the lower thermal oxide thickness which allows great flexibility in the types of structures that may be built. Simulations of a wire structure utilizing the remote gate implementation, described in Fig. 1(b), are provided in Fig. 2, which illustrate some of these effects.[3–5] The negatively biased depletion gates in this simulation have a 75 nm gap that allows a conducting channel in the silicon inversion layer. The thickness of the deposited oxide is 50 nm. The closer the depletion gate is to the silicon inversion layer, the more precisely the barrier is defined. This effect is shown in Fig. 2(a). The top inversion gate is positively biased in order to create an inversion layer density of $5\times10^{12}\,cm^{-2}$ in regions unaffected by the depletion gates. The depletion gates were negatively biased creating a 300 meV conduction band offset that confines carriers in the channel. As the lower oxide becomes thicker, electrostatic fringing effects soften the shape of the well. Although impractical due to electron direct tunneling, a 1 nm oxide is included in this simulation to demonstrate the advantage of very close proximity to the inversion layer. This softening is more pronounced at lower inversion layer densities as shown in Fig. 2(b). Here, a 5 nm thermal oxide has been used, and the top gate is varied to achieve inversion layer densities of 1 to $5\times10^{12}\,cm^{-2}$ in regions unaffected by the depletion gates. The channel may be further restricted by increasing the magnitude of the negative depletion gate bias as shown in Fig. 2(c). All of these factors interact so that careful three-dimensional modeling of any device structure is a necessity.

The oxide deposition process itself must not negatively impact the previously fabricated depletion gates. Deposited oxides benefit from high temperature post deposition annealing both to relax the bulk oxide[5] and to reduce interfacial suboxides.[6] The effects of this post deposition annealing (PDA) upon several candidate depletion gate materials have

[a]Electronic mail: jo.rack@asu.edu

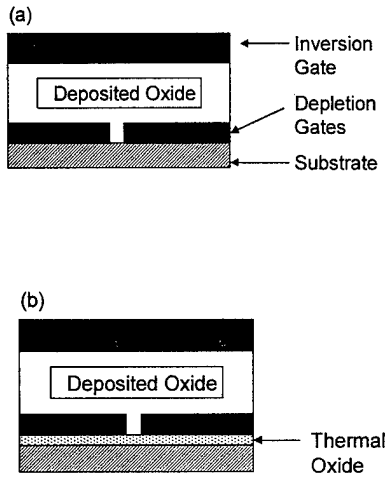


FIG. 1. Depletion gate approach to silicon based nanostructures. Quantum wire structure using (a) Schottky barrier depletion gates and (b) conducting depletion gates separated from inversion layer by thin thermal oxide.

previously been examined.[7] If the Schottky barrier approach of Fig. 1(a) is used, the deposited dielectric forms the Si/SiO_2 interface—a challenging job for a low temperature deposited dielectric. For the more remote depletion gate approach of Fig. 1(b), a thin thermally grown oxide may form the Si/SiO_2 interface. The deposited dielectric in this case must simply be electrically robust and not degrade the previously established Si/SiO_2 interface. We require the deposited oxide to withstand an applied field of about 2.5 MV/cm for the inversion layer density of $5\times10^{12}\,cm^{-2}$ at a deposited oxide thickness of 50 nm. For the implementation of Fig. 1(b), we assumed a 5 nm thermal oxide for these studies.

The use of remote plasma enhanced chemical vapor deposition (RPECVD) silicon dioxide from a microwave reactor for the implementation of both of the depletion gate strategies depicted in Fig. 1 has been investigated in this work. Processing conditions have been evaluated, and the suitability of the process and the resultant oxide is assessed relative to the requirements of the depletion gate approach employed.

We have particularly concentrated on two factors that we have found significantly influence measured oxide surface roughness and electrical reliability: deposition temperature and the ratio R_0 of oxidant (N_2O) to silane mass flow rates. In a recent study of a remote radio-frequency (rf) PECVD process, also using N_2O and silane, it was found that 50–100 nm sized particles were formed as soon as R_0 exceeded 4, or was high enough to form stoichiometric SiO_2.[8] This was attributed to gas phase nucleation. This study involved a fairly high flow rate of silane and, consequently, higher deposition rates than those used here, which might account for the greater severity of the problem reported in that study. The composition of plasma excited oxidant in a microwave reactor will differ from that of a rf reactor, and so it was important to examine this problem for the microwave type reactor. The effects of postdeposition annealing (PDA) using a rapid thermal anneal (RTA) upon oxide electrical properties are also presented. The low temperature regime of RPECVD deposition at or below 175 °C is extensively explored.

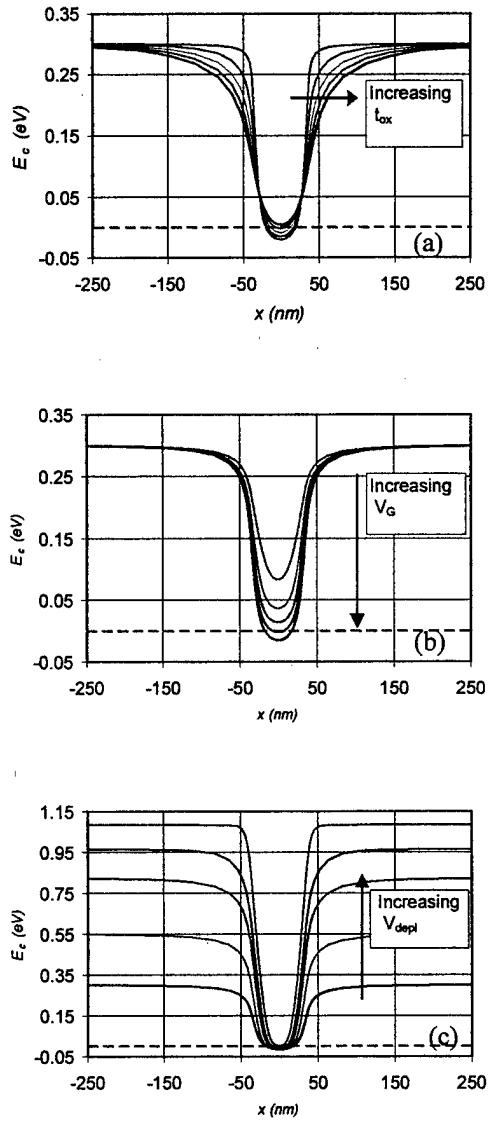


FIG. 2. Simulations of confinement energy in conduction band due to reversed biased depletion gates separated from the silicon substrate by a thin thermal oxide. (a) Effects of thermal oxide thickness variation from 1 to 20 nm. Depletion gates negatively biased at −0.04 to −0.24 V in order to achieve a barrier height of 300 mV. Top inversion gate positively biased (13.6 V) in order to achieve an inversion layer density of $N_s=5\times10^{12}\,cm^{-2}$ in regions unaffected by depletions gates. (b) Effects of inversion gate voltage variation. N_s ranging from 1 to $5\times10^{12}\,cm^{-2}$, $t_{ox}=5$ nm. (c) Pinch off capability of depletion gates, V_{depl} from −0.3 to −1.5 V. $V_G=13.6$ V.

The character of the bulk oxide is assessed with Auger electron spectroscopy (AES), Rutherford backscattering spectrometry (RBS), Fourier transform infrared spectroscopy (FTIR), secondary ion mass spectroscopy (SIMS), ellipsometry, and etch rates in buffered HF. Surface roughness measurements used atomic force measurements (AFM) and laser scattering. Electrical characterization includes capacitance–voltage measurements of MOS capacitors using the depos-

ited oxide and current–voltage measurements indicating leakage and breakdown characteristics of the deposited oxide.

II. EXPERIMENT

RPECVD SiO_2 was deposited in a remote microwave system. All processes were run at a pressure of 750 mTorr and a helium flow of 1000 sccms using N_2O and a flow of dilute silane (5% silane in helium) at 10 sccms. The ratios of N_2O to SiH_4 mass flow rates reported in this article are exact assuming that the actual flow of silane is 0.5 sccms. Unless otherwise noted, the microwave power was set at 250 W. The deposition temperature T_D reported is the heated chuck set point temperature. Calibrations indicate that this is always less than the wafer surface temperature, deviating more significantly at higher temperature. We have measured the temperature deviation in 100% He at 750 mTorr to be 15 °C at a T_D set point of 150 °C up to deviation of 60 °C at T_D of 350 °C.

Both *p*- and *n*-type 100 mm Si (100) wafers were used. Wafers were cleaned initially with H_2O_2:H_2SO_4 (3:7), carefully rinsed, etched in buffered HF, and rinsed again in DI water. A 120 nm sacrificial oxide was then grown at 1050 °C. Prior to deposition, these wafers were etched in 50:1 H_2O:HF beyond becoming hydrophobic. A 30 s exposure to plasma excited N_2O preceded the deposition. High temperature annealing for selected samples was done in a Tamarack model 180-*M* rapid thermal processor (RTA). Etch rate ratios (ERR) were measured by comparing the RPECVD oxide etch rate to that of dry thermal SiO_2 grown at 1050 °C in 20:1 buffered HF (BHF). The oxide thickness was measured using a Woolam *M*-44 Spectroscopic ellipsometer. The infrared Si–O stretching peak position and width were measured with a Nicolet 740 SX Fourier transform infrared spectroscopy (FTIR) system using 4 cm^{-1} resolution and a MCT detector.

AFM measurements were performed using a Digital Instruments scanning probe microscope. We used silicon nitride tips in contact mode. High attractive forces were encountered when these measurements were done in air. In the series of experiments involving R_0, a second set of data was collected in a nitrogen environment. This resulted in greater tip longevity, greater resolution and slightly higher roughness measurements. The trends and shape of the measured curves, however, were identical. Tip forces were measured at $15–50\times10^{-9}$ N. Data points are the mean of 5 to 25 scans with each scan containing 512×512 data points. Images are flattened on a line by line basis, root-mean-square (rms) height deviations are then calculated. Because roughness tends to scale with measured length, only 500 nm×500 nm areas were used. A Tencor SFS 4500 surface scanner was used to detect particles and haze at the oxide surface. This scanner cannot detect particles below 100 nm. Electrical measurements were done with aluminum gates deposited using a shadow mask as well as using optical lithography. When the deposited oxide is sufficiently thick, the roughness is fully established making it more easily quantifiable and repeatable. For this reason, all roughness measurements were done on 120-nm-thick oxides, while electrical measurements are done primarily on our target oxide thickness of 50 nm. Standard post metallization anneals were performed before electrical characterization at 430–450 °C both in nitrogen and forming gas.

AES measurements were made using a PHI 600 scanning multiprobe. The electron beam was set 5 kV, 50 nanoamps, and rastored over a 100 μm×100 μm square area. The samples were sputtered using argon at 3.0 kV, rastored over a 2 mm×2 mm area. The differential peak to peak height of the oxygen *KLL* Auger transition at 503 eV, and the silicon *KLL* Auger transition at 1619 eV were used. The higher energy silicon transition was used because it is less vulnerable to error due to charging effects and line shape changes that accompany varying silicon bonding arrangements. Thermally grown oxide was used as a reference. Oxygen to silicon ratios were averaged over the profiles.

Rutherford backscattering spectrometry was performed at the Ion Beam Analysis of Materials (IBeaM) Facility at Arizona State University. A He^{++} ion beam was accelerated to 3.05 MeV by using a Cockroft–Walton tandem electrostatic accelerator. Backscattered particles were detected at an angle of 173° with a beam current of approximately 60 nA. Spectra were collected at room temperature for an integrated charge current of 14 μC over an approximate spot size of 1 mm^2. Analysis of the data was performed using the RUMP[9] (REF) simulation program.

The SIMS profiles were done using a Cs^+ ion beam at 10 kV, 15 nA rastored over a 125×125 μm^2 area. Secondary ions were collected from a 16-μm-diam circular area. An electron beam was used to reduce charging effects and surfaces were coated with gold. Negative ions ^{30}Si, ^{18}O, and ^{1}H were detected.

III. OXIDE ROUGHNESS AND RELIABILITY VS DEPOSITION AND ANEALING TEMPERATURE

We have previously shown that the surface of RPECVD oxide deposited with an R_0 of 150 becomes rough as the deposition temperature T_D is reduced.[10] This roughness is a consequence of bumps whose size is in the tens of nanometers, making them difficult to detect optically.[11] As described in Fig. 3(a) this roughening behavior is accompanied by an increase in etch rates relative to the etch rate of thermally grown SiO_2. High temperature annealing does little to reduce the roughness, although the etch rates are continuously lowered with PDA temperature approaching that of thermally grown SiO_2 after the 900 °C anneal. This is shown in Fig. 3(b).

Depending upon R_0, we find that the electrical integrity of the 50 nm oxide deposited at temperatures below 250 °C is poor. The frequency of bumps on the 50 nm oxide is relatively low making them even more difficult to detect, but, possibly, more injurious. As seen in Fig. 4(a), the oxide deposited at 350 °C is as reliable as thermally grown oxide, while the oxide deposited at 225 °C using the same process conditions ($R_0 = 150$) is unusable. High temperature annealing offers only slight improvement for the breakdown behav-

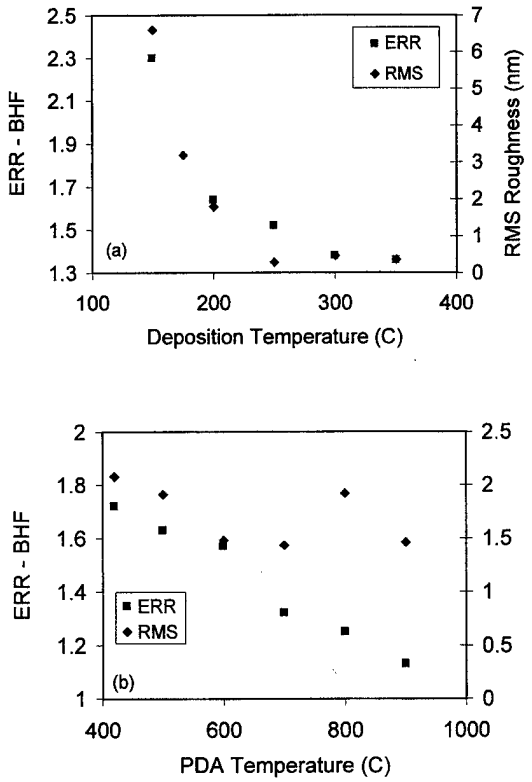


FIG. 3. rms oxide surface roughness measured with AFM and etch rate ratios of RPECVD to thermally grown oxide vs (a) T_D and (b) T_{PDA}. R_0 is 150, $t_{ox}=120$ nm.

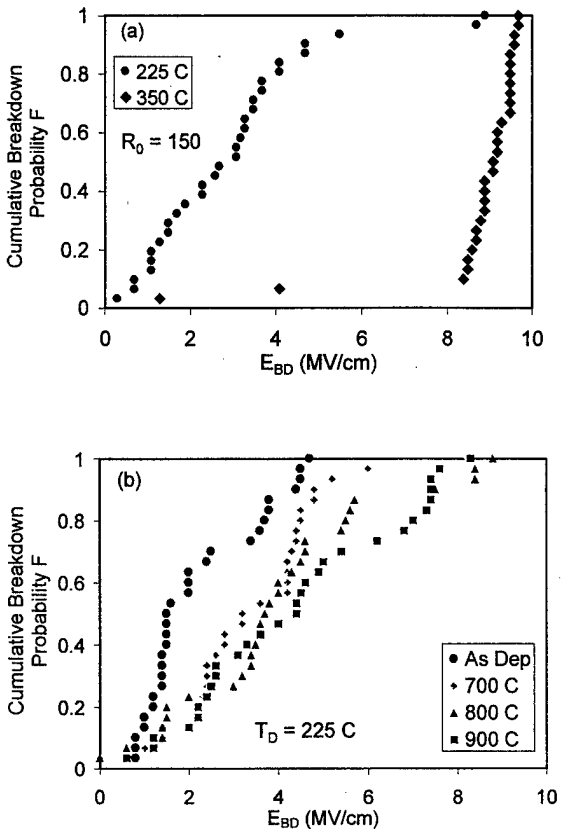


FIG. 4. Cumulative breakdown probability vs applied electric field for 50-nm-thick RPECVD oxide on *n*-type silicon having $R_0=150$. (a) Compares E_{BD} for oxide deposited at 225 and 350 °C under a positive applied voltage at a ramp rate of 1 V/s, (b) PDA effects upon E_{BD} for the 50 nm oxide deposited at 225 °C under a negative applied voltage using same ramp rate.

ior of oxides that are rough after deposition. This is evident in Fig. 4(b). We have seen that the addition of a small amount of oxygen during the anneal will provide some improvement in reliability, but this addition is not acceptable for our particular devices where unwanted oxidation of tiny depletion gates would undoubtedly occur.

While the oxide breakdown behavior is continuously, albeit only slightly, improved with the PDA temperature, interface properties, as expressed by flat band voltage shifts, may be initially degraded at temperatures between 700 and 800 °C, and improved by the 900 °C anneal. This may be seen in Fig. 5(a). In Fig. 5(b) the initial fixed charge is higher but we see continuous improvement in N_f as well as interface traps N_{it} with increasing PDA temperature. The interface properties of deposited oxide on silicon have been shown to be largely influenced by the *ex situ* surface preparation and plasma pretreatments used to clean the surface *in situ*.[12] Current recommended plasma excited N_2O pretreatments[13] are effective in our non-UHV system at high N_2O flow rates, and higher deposition temperatures. As shown in Fig. 5, the oxide charge is in the range of $10^{10}\,cm^{-2}$ for the oxide deposited at 350 °C, compared to $10^{11}\,cm^{-2}$ for the oxide deposited at 225 and 270 °C. These were done with an N_2O flow of 65–75 sccms. When the flow is reduced to 37.5 sccms for the pretreatment and deposition, fixed charge is measured at $3\times10^{11}\,cm^{-2}$ for the 350 °C deposition. While a change in N_2O flow between the pretreatment and deposition might allow more flexibility, the possibility of losing the plasma resonance and jeopardizing the depletion gates with additional oxygen exposure has made this tactic less attractive.

Unwanted oxidation of the metallic depletion gates during the oxide deposition process has proven problematic, particularly if subsequent 900 °C PDAs are required,[8] and so reduced deposition temperatures that might avoid the depletion gate oxidation are attractive if reliability may be achieved. This problem is illustrated in Fig. 6. These are AES sputter profiles of chromium deposited on oxide. In Fig. 6(a), the normal native oxide for as deposited chromium can be seen by the existence of oxygen at the surface (sputter time=0). The second plot (b) is one of chromium after being exposed to plasma excited N_2O for 30 s in the RPECVD deposition chamber. The native oxide has been thickened by approximately a factor of 3 actually consuming a significant percentage of the gate metal. There is the risk of oxidizing very thin depletion gates completely if care is not taken to prevent this.

For the remote interface implementation of Fig. 1(b), the deposited oxide must not disturb the previously established interface, and charges in the deposited oxide must be far enough removed from the interface so as not to scatter inversion layer electrons. Previous studies have indicated that "subcutaneous oxidation" or a thermal-like oxidation pro-

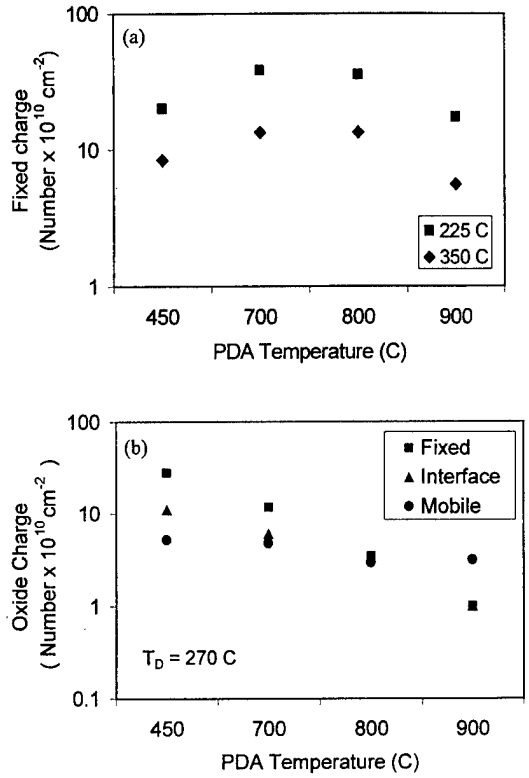


FIG. 5. Oxide charge as a function of one minute RTA temperature for 50 nm RPECVD oxide. (a) Deposited at 225 and 350 °C, 250 W, $R_0 = 150$. (b) deposited at 270 °C, 150 W, $R_0 = 130$.

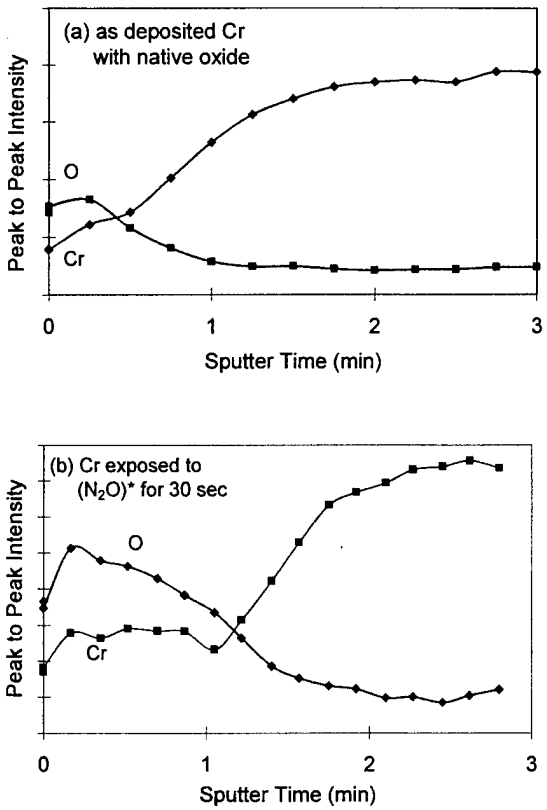


FIG. 6. AES sputter profiles of (a) as deposited chromium, and (b) exposed to plasma excited N_2O in the RPECVD deposition chamber at 270 °C for 30 s.

cess occurs during the RPECVD oxide deposition that will degrade a well constructed interface.[14] This was shown to be a problem when deposition temperatures exceeded 250 °C. The interface traps and fixed charge increased with temperature above 250 °C. It was not clear whether the 5 nm thermal oxide employed in our work was thick enough to prevent a similar degradation to the thermally established interface during the deposition of 50 nm of RPECVD oxide. In order to investigate this issue, RPECVD oxide was deposited at deposition temperatures ranging from 200 to 350 °C on 5 nm of thermal oxide that had been grown at 900 °C in dry nitrogen. The capacitance–voltage curves are shown in Fig. 7. The calculated fixed charge monotonically decreased with increasing deposition temperature from 1.16 to 1.39 $\times 10^{11}$ cm^{-2}. There is slightly more stretch out in the 200 °C and the 350 °C deposited samples.

IV. DEPOSITION AT 175 °C: EFFECTS OF R_0

In both direct and remote PECVD processes depositing SiO_2, the ratio of oxidant to silane flow R_0 is an important process parameter. Values ranging from 3 to 200 have been reported in the literature.[15–17] High R_0 ratios and low deposition rates used in the direct PECVD process allowed a breakthrough in oxide performance, where for the first time device quality oxide became possible.[18,19] For the remote process, there is a great deal more leeway in process parameters that can be used and still maintain high quality oxide with very little detectable hydrogen and nitrogen impurities.[16] More recently it has been reported that in a rf remote PECVD oxide deposition process, particulate formation was problematic for R_0 exceeding 4.[9]

In the remote microwave RPECVD process used for this study at 175 °C, for a fixed flow of silane, measured oxide surface roughness is sensitive to the ratio R_0 of N_2O to silane during deposition. As shown in the previous section, roughness is small and poses little problem for RPECVD oxide when it is deposited at temperatures above 250 °C. However,

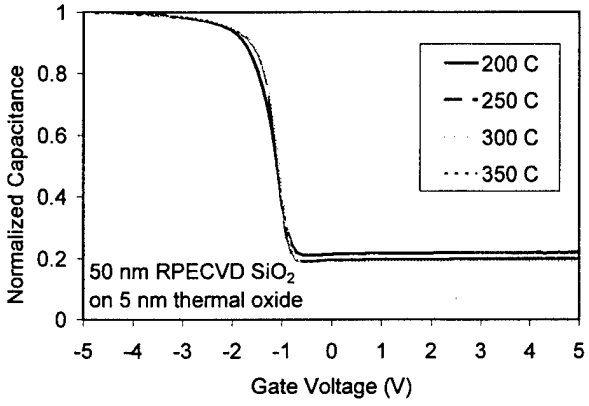


FIG. 7. Capacitance–voltage plots for MOS structures fabricated using 50-nm-thick RPECVD oxide deposited on 5-nm-thick thermally grown SiO_2. The thermal oxide was grown at 900 °C. The RPECVD oxide was deposited using $R_0 = 150$ at temperatures ranging from 200 to 350 °C.

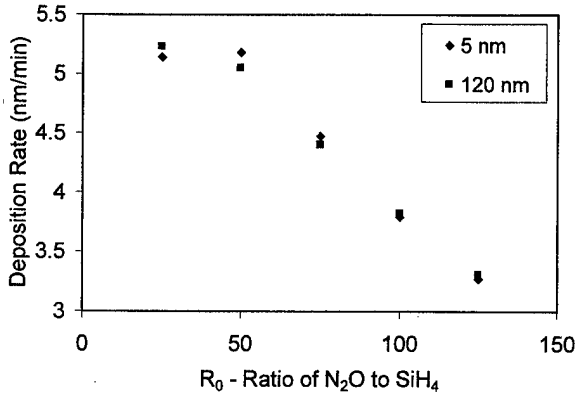


FIG. 8. RPECVD oxide deposition rate vs R_0 for a fixed silane flow rate of 10 sccms at 5% SiH_4 in He. T_D is 175 °C.

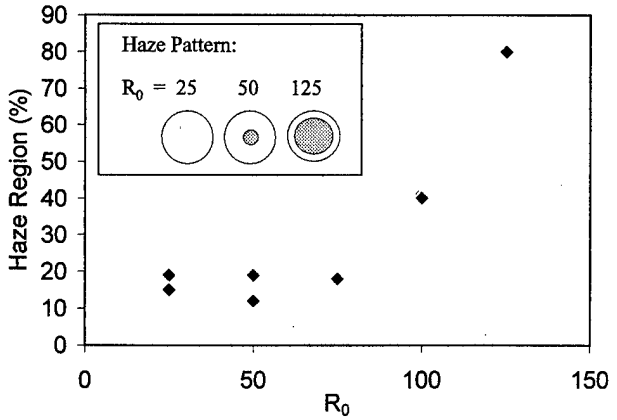


FIG. 9. Optical scanner measured haze coverage and insert of simplified pattern vs R_0 for 120-nm-thick RPECVD oxide deposited at 175 °C.

when lower deposition temperatures are required, the process window is reduced, and only a limited range of N_2O flows will result in oxide offering acceptable performance. In order to explore this effect, a series of 50-and 120-nm-thick RPECVD oxides with R_0 ranging from 25 to 125 were deposited at 175 °C. At the 120 nm oxide thickness, the deposited SiO_2 surface roughness is more fully established and more easily quantified. The 50 nm oxide thickness was the target thickness for our devices, and electrical reliability is more easily compromised at this thickness. The temperature of 175 °C was chosen so that roughening associated with a given R_0 would be more obvious. In addition, it is intrinsically important for novel applications using materials with relatively poor thermal stability.

Deposition rates provide insights into the reaction behavior. For R_0 exceeding 50, the deposition rates in this study drop with increasing N_2O flow as shown in Fig. 8. In general, we see deposition rates decrease with deposition temperature, or a negative activation energy.

A Tencore surface scanner was used to observe particles and haze on the whole wafers having 120 nm of oxide. The minimum particle dimension that this scanner can observe is 100 nm. The particle counts were not correlated with R_0. However, the percentage of haze region and observed haze patterns were useful. At an R_0 of 25, no pattern was observable on the wafer. At an R_0 of 50, a circular region in the center could be seen which became larger with increasing R_0, until at R_0 of 125, it covered most of the wafer. The circular pattern coincides with the shape of the microwave tube. High plug flow velocities and low chamber pressure make it difficult to alter the flow pattern of arriving plasma excited oxidant. This is shown in Fig. 9, with the inset showing a simplified pictorial of the haze pattern of several wafers. While gas phase nucleation has been attributed to water vapor in the deposition chamber,[16] the pattern of this indicates that the source of particles is not necessarily water vapor which would be more active at the outside of the wafer away from the stream of arriving oxidant.

While etch rates in HF containing solutions have often been considered a measure of PECVD oxide quality, we find that roughness is not minimized under conditions that yield the lowest etch rates. This is evident in Fig. 10. The rms roughness, as measured by AFM at the center of the wafer, continuously increases with R_0. This increase was primarily due to greater frequency of 5–12 nm high bumps. The rms roughness, measured with AFM is more sensitive to the surface than the Tencore scanner, and provides insights into the process not otherwise available. The ratio of RPECVD oxide etch rates to that of thermally grown SiO_2 behaves oppositely, indicating that continuously reducing R_0 introduces other defects.

The electrical quality of RPECVD SiO_2 deposited at 175 °C was analyzed using capacitance–voltage and current–voltage measurements. At this low temperature, fixed oxide charges are higher than at 350 °C, and are independent of R_0. For both the 50- and 120-nm-thick oxides, N_f averaged $5\times10^{11}\,cm^{-2}$, with no evident trend in R_0. For the device structure of Fig. 1(b), this oxide may still be useful since the Si/SiO_2 interface is established thermally prior to the RPECVD oxide deposition. All of the 120 nm oxide samples showed good reliability, having four incidences of breakdown at applied fields less than 7.5 MV/cm out of 75

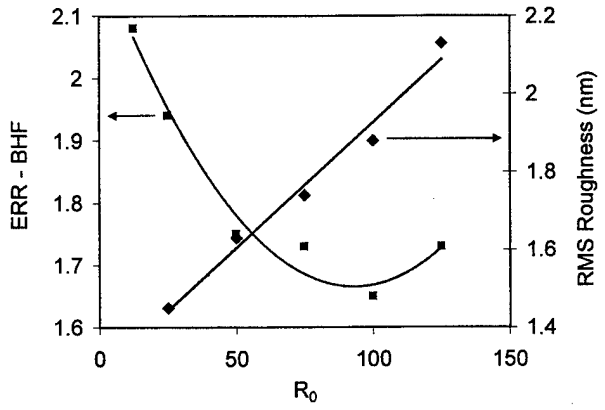


FIG. 10. rms oxide surface roughness measured with AFM and etch rate ratios of RPECVD to thermally grown oxide vs R_0. T_D is 175 °C No PDA was included.

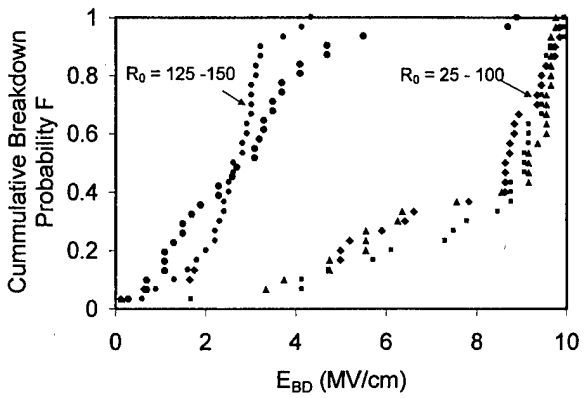


FIG. 11. Oxide electrical reliability vs R_0 for 50-nm-thick RPECVD oxide deposited at 175 °C (with the exception of the $R_0 = 150$ wafer which was deposited at 225 °C).

trials using 0.005 cm^2 aluminum contacts. The 50-nm-thick series showed greater sensitivity to R_0. This can be seen in Fig. 11. All of these wafers had similar particle counts according to the surface scanner. An obvious loss in reliability occurs, however, when R_0 is increased above 100. For $R_0 \leq 100$, the leakage current density was below the detection limit of 10^{-9} A/cm^2 at an applied field of 2.5 MV/cm. At an rms roughness of about 2 nm, reliability is seen to be compromised at both 225 and 175 °C regardless of the value of R_0 when MOS capacitors are formed with 50-nm-thick RPECVD oxide. Given the continuous roughening with R_0 that is shown in Fig. 10, it might be assumed that as the oxide thickness is further reduced, a narrower range of R_0 will be acceptable.

The RPECVD oxide stoichiometry for this series in R_0 was assessed using four methods: AES, RBS, SIMS, and FTIR. These results are shown in Tables I and II. In Table I, the refractive index and the FTIR results are presented. The frequency of the IR SiO stretching absorption peak, v, has been associated with stoichiometry[20] and the Si–O–Si bond angle.[21] The RPECVD oxide typically has v in the range of 1047–1056 cm^{-1} and there is some uncertainty as to why this deviates from the position of v in thermal oxide at ~1080 cm^{-1}. The closer v is to 1080 cm^{-1}, the more structurally and chemically relaxed the film is assumed to be.[22] For this deposition series in R_0, v lies between 1056 and 1058 cm^{-1}, the differences shown within the measurement error. Using the bond angle interpretation, the position of the PECVD peak below that of thermal oxide has been attributed to a smaller Si–O–Si bond angle and, subsequently, a denser amorphous network. The refractive index, although slightly higher than thermal SiO_2, is not high enough to predict the density that the Si–O–Si bond angle attributed to v might suggest.[23] This, and the tendency for PECVD oxides to shrink upon annealing, has suggested the presence of porosity or micropores that would create bond free volume.[21] The peak width Δv is considered a measure of bond homogeneity. There would seem to be some improvement in this as R_0 is increased from 25 to 125, and Δv decreases slightly. We investigated the stoichiometry of these RPECVD films using three very different techniques in order to determine if there was a consistent difference in stoichiometry between the deposited films and thermally grown SiO_2 as might be suggested by the IR SiO stretching frequency v, or any trend in stoichiometry with R_0. Using the analysis of Ref. 20, an IR absorption peak at 1056–1058 cm^{-1} might indicate that all of the SiO_x RPECVD films are silicon rich with $x = 1.82–1.86$, or x_{CVD}/x_{therm} of 0.91–0.93 assuming $x_{therm} = 2$. This analysis is limited in accuracy at the endpoint of 1065 cm^{-1}. Although there is scatter in the results, none of the techniques suggest the type of bias that silicon rich films would have. The SIMS data show greater variance in x for the RPECVD films than for the thermal oxide, which were quite reproducible. The standard deviation of six thermal oxide ratios of O/Si was 0.036 compared with six RPECVD samples having σ of 0.08. The results were independent of the sputter rates. The analyzed area of the SIMS samples was the smallest of the three techniques. It is possible that the ratio of O to Si varies across a sample. Given the lack of correlation between methods, it is likely that the films are all stoichiometric within the accuracy of the measurements techniques employed. In any case, there is no obvious trend in x with R_0 or consistent indication that any of the deposited oxides are silicon rich. The hydrogen content in the deposited films was measured at approximately 2–4 times that of thermal SiO_2 with no clear trend in R_0. The highest concentration did occur for the $R_0 = 25$ sample suggesting the source of defects responsible for increasing etch rates.

TABLE I. Optical data vs R_0.

R_0	Refractive index	IR SiO stretching Frequency v (cm^{-1})	IR SiO stretching Peak width Δv (cm^{-1})
25	1.464	1058	95
50	1.460	1058	93
75	1.460	1056	91
100	1.460	1056	92
125	1.461	1057	89
Thermal	1.454	1080	75

TABLE II. AES, RBS, and SIMS results vs R_0—stoichiometry and hydrogen content.

R_0	x_{CVD}/x_{therm} AES	x_{CVD}/x_{therm} RBS	x_{CVD}/x_{therm}[a] SIMS	H_{CVD}/H_{therm} SIMS
25	1.02	1.00	1.07	3.65
50	1.01	1.00	1.01	2.00
75	1.02	1.03		
100	0.98	.97	1.12	2.78
125	1.00	1.04	1.02	1.93
Avg.	1.01	1.01	1.06	2.60

[a]x defined as ratio of O counts/Si counts.
[b]Hydrogen normalized to oxide silicon count in each sample in order to compensate for any differences in sputter rates.

V. DISCUSSION

The fall off in deposition rates with increasing oxidant for a given flow of silane in thermal CVD has been attributed to reduced availability of surface sites for silane as the surface becomes saturated with oxygen species.[24] As the available reactants exceed the consumption at the surface from the chemical reactions taking place, it becomes possible for gas phase nucleation to occur. These reactions produce particulate that will degrade the deposited film. Therefore, as the surface temperature is increased, and more reactants are consumed, the production of particulate may be decreased. For a plasma process, the energy required to produce reactants is provided by the plasma rather than heating the gas. However, surface temperature would affect the surface reaction rate and surface mobility. So, similar behavior in gas phase nucleation may be seen. Excessive reactants would lead to gas phase nucleation. At higher temperature, more reactants would contribute to planar SiO_2 growth rather than gas phase nucleation. This is consistent with the observation that deposited oxides are better when the deposition rate is slow.[19] The slight reduction in deposition rate with temperature has been attributed to desorption of reactive species from the surface,[12] compression of the oxide as it densifies with temperature,[24] and the loss of atomic oxygen due to more active O_2 recombination at the surface.[25] In a plasma reactor, deposition continues even at room temperature although the film structure becomes more and more obviously granular. So the deposition rate permissible that provides high quality oxide might simply be limited by the surface reaction rate and surface mobility, and thus an inherent maximum deposition rate is established for a given temperature.

VI. CONCLUSIONS

RPECVD oxide from a microwave reactor has been evaluated for use in silicon based nanostructures. While all of the depositions conditions in this study provide stoichiometric oxide with low etch rates and impurities, there are significant differences in the bulk oxide quality, as indicated by surface roughness, that correlate strongly with oxide breakdown behavior. The strength of this relationship depends upon the oxide thickness. If the oxide surface exceeds about 2 nm rms roughness, electrical reliability is compromised for 50-nm-thick oxide MOS capaciors. It is presumed that the surface roughens as particles are incorporated into the amorphous oxide network through gas phase nucleation or some other type of clustering phenomenon. Deposition conditions that reduce surface roughness, that is a low oxidant to silane ratio, do not produce the best interface characteristics or the lowest etch rates.

If the deposited oxide is used to form the Si/SiO_2 interface, and the depletion gates cannot tolerate a 900 °C anneal, then the deposition temperature must be high enough to reduce the oxide and interface charges, and high ratios of N_2O to silane may be used which have provided improved interface quality compared with the lower values which reduce particulate formation below 250 °C. However, under these conditions, the gate material used to form the Schottky barrier must be resistant to oxidation.

If the deposited oxide is to be placed upon a thermally grown oxide, then a deposition temperature as low as 175 °C may be tolerated as long as R_0 is kept below 100. However, the issue of remote charges in the deposited oxide interfering with mobility is unresolved and requires further study.

[1]M. Koury, A. Gunther, M. J. Rack, D. Piven, Jr., and D. K. Ferry, Microelectron. Eng. (to be published).
[2]A. D. Gunther, S. M. Goodnick, M. D. Khoury, A. E. Krishnaswamy, M. J. Rack, and T. J. Thornton (unpublished).
[3]Simulations done with SHRED, a 1D Shrödinger/Poisson solver described in IEEE Trans. Electron Devices **44**, 584 (1997) and with the 3D Poisson solver described in Ref. 5.
[4]D. Vasileska, W. J. Gross, V. Kafedziski, and D. K. Ferry, VLSI Design (in press).
[5]J. T. Fitch, S. S. Kim, C. H. Bjorkman, and G. Lucovsky, J. Electron. Mater. **19**, 151 (1989).
[6]J. Schafer, A. P. Young, L. J. Brillson, H. Niimi, and G. Lucovsky, Appl. Phys. Lett. **73**, 791 (1998).
[7]M. J. Rack, A. D. Gunther, M. Khoury, D. Vasileska, M. Siderov, and D. K. Ferry, Semicond. Sci. Technol. **13**, A71 (1998).
[8]Y-B. Park, J-K. Kang, and S-W. Rhee, Thin Solid Films **280**, 43 (1996).
[9]L. R. Doolittle, Nucl. Instrum. Methods Phys. Res. B **9**, 344 (1985).
[10]M. J. Rack, D. Vasileska, D. K. Ferry, and M. Siderov, J. Vac. Sci. Technol. B **16**, 2165 (1998).
[11]N. P. Rao, S. Nijhawan, T. Kim, Z. Wu, S. Campbell, D. Kittelson, P. McMurry, C. C. Cheng, and E. Mastromatteo, J. Electrochem. Soc. **145**, 2051 (1998).
[12]G. Lucovsky, S. S. Kim, and J. T. Fitch, J. Vac. Sci. Technol. B **8**, 822 (1990).
[13]G. Lucovsky, H. Niimi, K. Koh, D. R. Lee, and Z. Jing, *The Physics and Chemistry of SiO_2 and the Si-SiO_2 Interface—3*, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (The Electrochemical Society, Pennington, NJ, 1996).
[14]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).
[15]R. L. Jackson, J. E. Spencer, J. L. McGuire, and A. M. Hoff, Semicond. Sci. Technol. **30**, 107 (1987).
[16]G. Lucovsky, D. V. Tsu, S. S. Kim, R. J. Markumas, and G. G. Fountain, Appl. Surf. Sci. **39**, 33 (1989).
[17]J. D. Chapple-Sokol, W. A. Pliskin, R. A. Conti, E. Tierney, and J. Batey, J. Electrochem. Soc. **138**, 3723 (1991).
[18]J. Batey and E. Tierney, J. Appl. Phys. **60**, 3136 (1986).
[19]J. Batey, E. Tierney, and T. N. Nguyen, IEEE Electron Device Lett. **8**, 148 (1987).
[20]D. V. Tsu, G. Lucovsky, and B. N. Davidson, Phys. Rev. B **40**, 1795 (1989).
[21]G. Lucovsky, J. T. Fitch, E. Kobeda, and E. A. Irene, *The Physics and Chemistry of SiO_2 and the Si-SiO_2 Interface* (Plenum, New York, 1975), p. 139.
[22]B. J. Hinds, F. Wang, D. M. Wolf, C. L. Hinkle, and G. Lucovsky, J. Vac. Sci. Technol. B **16**, 2171 (1998).
[23]R. A. B. Devine, J. Electron. Mater. **19**, 1299 (1990).
[24]C. Cobianu and C. P. Pavelescu, Thin Solid Films **102**, 361 (1983).
[25]G. Raupp, D. A. Levedakis, and T. Cale, J. Vac. Sci. Technol. A **13**, 676 (1995).

Self-organization in Si/$CoSi_2$(111) heteroepitaxy

T. Meyer,[a] M. Klemenc, T. Graf, and H. von Känel
Laboratorium für Festkörperphysik, Eidgenössische Technische Hochschule Zürich-Hönggerberg, CH-8093 Zürich, Switzerland

(Received 19 January 1999; accepted 5 May 1999)

We present a scanning tunneling microscopy study of the initial stages of epitaxial Si overgrowth of $CoSi_2$(111) films. Under suitable growth conditions the silicon was found to decorate the steps of the silicide. This opens up the possibility of creating arrays of parallel silicon wires on a $CoSi_2$ substrate. After oxidation of the silicon by exposing it to air the silicide in between the wires can be removed by wet etching, leaving an array of expitaxial $CoSi_2$ wires. The mechanism leading to the formation of the silicon wires was identified as strongly enhanced diffusion of adatoms on a silicide surface compared to a silicon surface and preferential nucleation on upward steps. © *1999 American Vacuum Society.* [S0734-211X(99)04804-0]

I. INTRODUCTION

Self-organization in epitaxial growth offers an easy and inexpensive way to fabricate ordered arrays of nanoscale structures, which are interesting both for their special physical properties as well as for the fabrication of devices. In heteroepitaxy it has been suggested to grow wires of one material running along the surface steps of a substrate made from another material in step-flow growth mode. The arriving adatoms would diffuse to the next upward step of the surface and nucleate there. By repeatedly depositing wire material and substrate material a formation of stacked wires embedded in the substrate material should occur (see, e.g., Ref. 1). It is crucial in this procedure that the diffusion of both materials is sufficiently fast, so that step-flow growth can be reached. In this article we suggest a different method to fabricate epitaxial $CoSi_2$ wires on a Si(111) substrate by a combination of self-organized growth of silicon and an *ex situ* wet-chemical etch, where Si is used as a mask. The Si mask formation is only possible because the diffusion of Si on the silicide surface is much faster than on Si itself. As a consequence, the step-flow growth mode can be reached at temperatures as low as 350 °C. This is crucial because the $CoSi_2$/Si(111) films used as substrates are not stable at temperatures much higher than 600 °C which are required to reach step-flow growth in Si/Si(111) homoepitaxy. In scanning tunneling microscopy (STM) experiments we have determined the temperature dependency of the diffusion length of Si/$CoSi_2$.

II. EXPERIMENT

It is well known that $CoSi_2$/Si(111) films can be epitaxially overgrown with thick silicon films ($d \geqslant 50$ nm) of good quality and devices have been fabricated out of $CoSi_2$/Si/$CoSi_2$ sandwich structures.[2] Contrary to this, little is known about the nucleation process and the early stages of Si growth on the silicide surface. In our experiments we have prepared epitaxial $CoSi_2$ films on Si(111) substrates by molecular beam epitaxy (MBE). The native oxide was removed from the wafer surface by thermal desorption. A 240 nm thick silicon buffer layer was grown at $T=700$ °C. The temperature of the wafer was then briefly raised to about 820 °C, where the 7×7 reconstruction vanishes according to reflection high-energy electron diffraction (RHEED). Subsequently the temperature was lowered to room temperature. This procedure yields a nicely 7×7 reconstructed surface with an arrangement of nearly parallel steps. A 3 nm thick $CoSi_2$ film was formed by stoichiometric codeposition of Si and Co and annealing it in several steps with a final anneal at 600 °C for 5 min. The films were grown in this manner in order to obtain a predominantly unreconstructed surface.[3] The 2×1 reconstruction present in some regions was found to hinder surface diffusion. The strain in the silicide was relaxed by the introduction of a network of interfacial dislocations. In our experiments we did not observe any influence of these dislocations on Si diffusion. The surface steps of the silicide were again nearly parallel. The step height is 0.31 nm, which corresponds to the height of 1 bilayer (BL) of silicon. Onto these surfaces we deposited Si at temperatures ranging from 200 to 430 °C and rates ranging from 0.017 to 0.105 nm/s.

III. RESULTS

In Fig. 1 a STM topography scan of a $CoSi_2$/Si(111) sample after deposition of 0.5 nm of Si at a rate of 0.105 nm/s and a substrate temperature of 350 °C is shown. The silicon has obviously accumulated along the surface steps of the silicide. The average terrace width was about 90 nm. No island formation is observed. Most of the Si lines are bounded by a sharp straight step on one side, and a rougher step on the other. The straight step is positioned where the underlying silicide surface step is located. The fuzzy line represents the growth front at the end of deposition. For most silicide steps, the Si is found to grow only at the lower step edge (see, e.g., terrace A), but some of them also decorate the upper terrace (see terrace B). It is evident that the wider the terrace, the more likely nucleation on the down step be-

[a]Electronic mail: meyer@solid.phys.ethz.ch

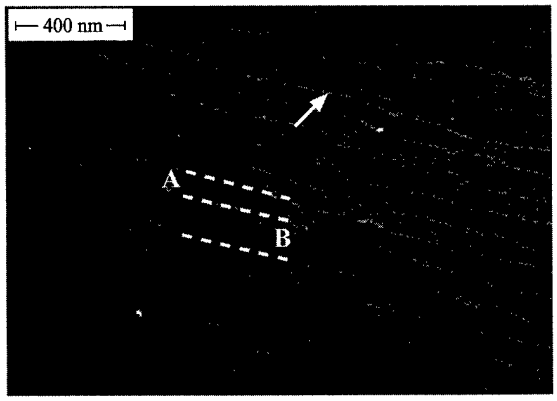


FIG. 1. STM topography image of the surface of a 3 nm thick $CoSi_2/Si(111)$ film after the deposition of 0.5 nm of Si. The dotted lines indicate the position of the silicide steps that bound terraces A and B. On the wider terraces both steps are decorated (B). For narrower terraces (A) only the upward step is decorated. The width of the silicon line is determined by the width of the terraces and can be quite small (shown by the arrow).

comes. The critical terrace width w_{crit} at which decoration of up and down steps starts to occur is a measure of the diffusion length of Si. Actually the diffusion length $d=2w_{crit}$, because adatoms are being reflected by a down step and have to cross the terrace twice before they attach to the up step. At lower temperatures the diffusion length is obtained from the average island spacing $n^{-1/2}$ for a given island density n. Another interesting feature visible in Fig. 1 is the fact that the fraction of the terrace that is covered by Si is nearly constant. An especially narrow terrace is indicated by an arrow. The Si line following the up step gets thinner as the terrace width decreases. Only material deposited on the same terrace contributes to the Si line. By choosing a substrate with greater misorientation the terrace width and consequently the width of the Si lines can be reduced at the same average coverage. The fraction of the terrace covered by silicon is controlled by the amount of material deposited. Earlier

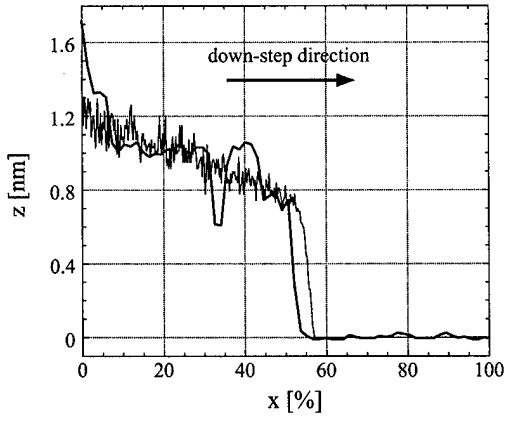


FIG. 2. Line section (thick line) taken across a terrace of the $CoSi_2$ surface bounded by an up step at $x=0\%$ and by a down step at $x=100\%$. Si has accumulated at the up step on the left. The thin line is a result of a Monte Carlo simulation of growth, assuming perfect diffusion of adatoms on the silicide and sticking on the silicon. The x axis is given in percent of the terrace width.

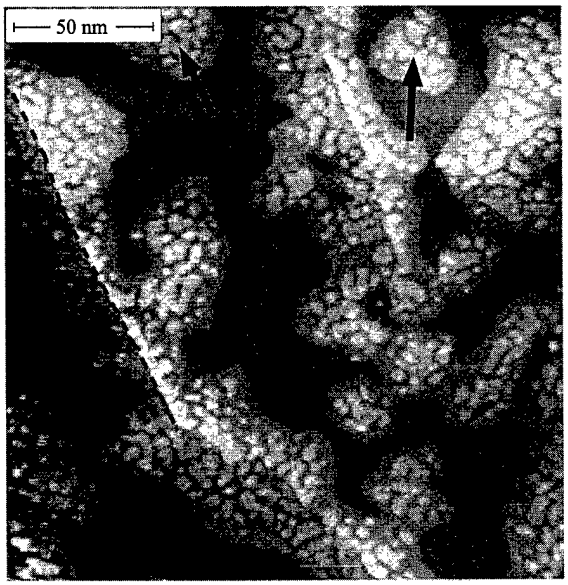


FIG. 3. STM topography image of a sample where 0.2 nm of silicon was deposited on a 3 nm thick $CoSi_2$ film at $T=235\,^{\circ}C$. The boundary of the silicon islands features a 1 BL step. The gray scale range is 2.2 nm.

it was mentioned that a total of 0.5 nm of silicon was evaporated onto this particular sample. This is more than a bilayer of silicon and still only about half of the surface is covered. The reason for this can be found from line sections like the one plotted as a thick line in Fig. 2, taken in the direction orthogonal to a surface step. The profile shows silicon attached to an up step on the left-hand side, whereas the down step on the right-hand side is not decorated. The Si line is found to always be at least 2 BL high, which explains why only half of the surface is covered. On top of this Si base layer there are silicon islands, which are 1 BL high. The thin line plotted in Fig. 2 along with the surface profile was obtained by a simple one-dimensional Monte Carlo simulation. The simulation assumes perfect diffusion on the silicide and no diffusion at all on top of the silicon. Adatoms diffusing to the down step at $x=100\%$ in Fig. 2 are reflected and atoms hitting the up step at $x=0\%$ stick to it. Further on it is assumed that, if an atom meets a site that is already occupied by silicon atoms, it hops on top of it until a thickness of 2 BL is reached. Twenty runs of this algorithm were averaged for the thin line plotted in Fig. 2. Although this is a very crude model, it nicely fits the surface profile, which suggests that the assumptions are basically correct: the diffusion on silicide is much faster than on silicon, the adatoms are likely to jump onto the first bilayer which forms, and almost all adatoms are reflected at a down step. Note that the model does not account for the small spike at $x=0\%$: there is additional material growing right next to the edge of the Si wire. This material is most likely due to adatoms diffusing on the upper terrace located at $x\leqslant 0\%$, that are not reflected. However, the amount of additional material is quite small and knowing that about half of the adatoms will be traveling towards the down step, we can conclude that the probability for reflection is indeed very close to 1. When a lower substrate tempera-

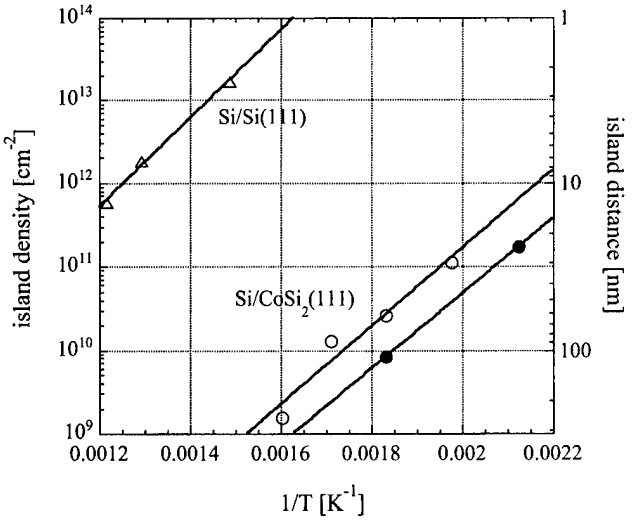


FIG. 4. Island density as a function of inverse growth temperature. For the measurements denoted by (●) a deposition rate of 0.016 nm/s was used. Silicon was deposited at a rate of 0.105 nm/s for the samples marked by (○). The data for Si/Si(111) (△) are taken from Ref. 4 and extrapolated to a rate of 0.105 nm/s.

ture is chosen during Si deposition, the diffusion length decreases and islands also nucleate in the middle of the silicide terraces. As an example, a STM image of a silicide surface onto which 0.2 nm of Si was evaporated at $T=235\,^{\circ}C$ is shown in Fig. 3. On this sample silicon has attached itself to both sides of the surface steps (one step is indicated by a dashed line). In addition islands have nucleated on the terraces (arrows). In this smaller scale image the difference in diffusion length between the two materials is clearly visible: on top of the Si terraces, which have formed on the silicide, there are small, closely packed silicon islands. Surprisingly, the step from the silicide to the first Si terrace is only 1 BL high. Contrary to the sample grown at higher temperature, the second bilayer has not been completed. After a brief annealing step at $T=570\,^{\circ}C$ the second bilayer was found to be present on the entire terrace. This indicates that the activation energy for the hopping of adatoms adsorbed at the border of a terrace onto the first bilayer is higher than the activation energy for diffusion on the silicide surface. In order to quantify the difference in diffusion length for $Si/CoSi_2(111)$ compared to Si/Si(111), the island density n was measured as a function of temperature. The results for two different deposition rates are plotted in Fig. 4. The average island spacing, $n^{-1/2}$, also called the diffusion length, can be seen from the scale on the right-hand side of the plot. The data for Si/Si(111) were taken from Ref. 4 and extrapolated from a growth rate of 0.2 monolayer (ML)/min to a rate of 0.105 nm/s to allow for a direct comparison with the measurements made for $Si/CoSi_2(111)$ (plotted by an open circle) deposited at the same rate. Obviously there is not much difference in the temperature dependence of the island density for the two types of substrates. At a given temperature, however, the island spacing is larger by more than two orders of magnitude for silicon deposited on $CoSi_2$. According to classical nucleation theory[5] for two-dimensional (2D)

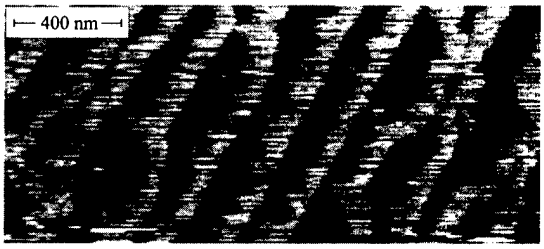


FIG. 5. AFM topography image of the sample shown in Fig. 1 after a brief HF etch. Only the silicide below the oxidized silicon lines remains on the sample surface.

islands and complete condensation the island density n depends on the growth rate R and temperature T,

$$n \sim \left(\frac{R}{\nu}\right)^{i/(i+2)} \exp\left(\frac{E_{\mathrm{eff}}}{kT}\right), \quad E_{\mathrm{eff}}=\frac{E_c+iE_d}{i+2}. \tag{1}$$

E_c is the energy gained by the formation of the stable critical cluster, which contains i adatoms. The activation energy for diffusion is denoted by E_d and ν is a typical atomic vibration frequency. The slope of the curves in Fig. 4 gives the effective activation energy E_{eff}. The size of the critical cluster is obtained by fitting the rate dependency of n to a power law. Our data suggest a value of $i=3$ atoms in the critical cluster. From this and the slope of the n vs $1/T$ curves, we obtain an estimate for the activation energy for diffusion of $E_d=0.95\pm0.1\,\mathrm{eV}$. This value depends heavily on the choice of the cluster binding energy E_c, which we have set equal to the energy $E_b=1.7\,\mathrm{eV}$ of a Si–Si bulk bond. It is very difficult to make a comparison with the E_d found for Si/Si(111) epitaxy, because both higher and lower values can be found in literature. Voigtländer *et al.* deduced a value of $E_d=0.75\pm0.2\,\mathrm{eV}$,[4] Nakahara and Ichikawa found $E_d=2\,\mathrm{eV}$,[6] and Latyshev *et al.* suggested $E_d=1.3\pm0.2\,\mathrm{eV}$.[7] The experimentally well defined quantity is the effective activation energy E_{eff}. In their experiments Latyshev *et al.* observed a dependence of E_{eff} on growth rate R, with higher values for small R (3 eV at 2.3 ML/min) and lower values for high R (1.3 eV at 1.2 ML/s). The data of Voigtländer *et al.* result in about $E_{\mathrm{eff}}=1.06\,\mathrm{eV}$ at a deposition rate of 0.2 ML/min, which contradicts the measurements of Latyshev *et al.*, but might be suitable as a lower limit to E_{eff} for Si/Si(111). The slope found in Fig. 4 results in $E_{\mathrm{eff}}=0.91\pm0.03\,\mathrm{eV}$ for silicon diffusion on the surface of $CoSi_2$, which is slightly lower than for Si/Si(111). Still most of the difference in the island densities stems from the pre-exponential factor and not from a different temperature dependence.

The large diffusion lengths of $Si/CoSi_2$ at temperatures far below the maximum annealing temperature of the silicide are essential for the self-ordered growth of silicon, since there is no danger of pinhole formation or intermixing effects. The structure shown in Fig. 1 was removed from the UHV environment. Afterwards the sample was etched in 5% HF for 15 s. In this etching step the silicon lines acted as a mask and only stripes of epitaxial $CoSi_2$ remained on the wafer. Figure 5 shows an atomic force microscopy (AFM) image of the sample after etching. The measured corrugation of the $CoSi_2$

lines is about 4.5 nm. It should be emphasized that for the results presented in this article nominally singular substrates were used with an unknown actual misorientation. It can be expected that by carefully adjusting the wafer miscut and growth parameters, lines of nanometric width can be grown using this procedure, allowing for measurements of one-dimensional transport in epitaxial metallic wires. Also, the fact that $CoSi_2$ films are superconducting at sufficiently low temperatures[8] could provide interesting measurements. Another issue we will address in the near future is Si overgrowth of structured $CoSi_2$ films.

IV. SUMMARY

In summary, we have investigated the early stages of Si/$CoSi_2$(111) heteroepitaxy by means of STM and found a very greatly enhanced diffusion length compared to that of Si/Si(111) homoepitaxy. At low temperatures silicon can be grown in quasi-stepflow mode, where the steps of the silicide surface are decorated by silicon wires at least 2 BL in height. The width of the wires can be adjusted by varying the coverage and the wafer miscut. Using the deposited silicon as a mask in an etching step, $CoSi_2$ wires are obtained.

ACKNOWLEDGMENTS

This work was financed by the Swiss National Science Foundation (NFP 36) and by the Swiss Priority Program in Micro and Nano System Technology (MINAST).

[1]J. M. Gaines, P. M. Petroff, H. Kroemer, R. J. Simes, R. S. Geels, and J. H. English, J. Vac. Sci. Technol. B **6**, 1378 (1988).
[2]C. Schwarz and H. von Känel, J. Appl. Phys. **79**, 8798 (1996).
[3]H. von Känel, Mater. Sci. Rep. **8**, 193 (1992).
[4]B. Voigtländer, A. Zinner, T. Weber, and H. P. Bonzel, Phys. Rev. B **51**, 7583 (1995).
[5]A. Venables, G. D. T. Spiller, and M. Hanbücken, Rep. Prog. Phys. **47**, 399 (1984).
[6]H. Nakahara and M. Ichikawa, Appl. Phys. Lett. **61**, 1531 (1992).
[7]A. V. Latyshev, A. B. Krasilnikov, and A. L. Aseev, Phys. Rev. B **54**, 2586 (1996).
[8]P. A. Badoz, A. Briggs, E. Rosencher, and F. Arnaud d'Avitaya, J. Phys. (France) Lett. **46**, 979 (1985).

Comparison of nanomachined III–V semiconductor substrates

L. Grazulis
University of Dayton Research Institute, Dayton, Ohio 45469-0167

D. L. Kelly
Engineering Research Center, University of Cincinnati, Cincinnati, Ohio 45221

D. E. Walker, Jr.
Wright State University, Dayton, Ohio 45435-0001

D. H. Tomich, K. G. Eyink,[a] and W. V. Lampert
Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433-7750

(Received 19 January 1999; accepted 27 May 1999)

Emerging device applications demand surface features on the order of hundreds of angstroms. Nanolithography by machining with a diamond tip is proposed as a means to reproducibly pattern semiconductor surfaces on this scale. This technique has already been shown to produce controlled features with depths down to 10 Å on GaAs [S. H. Goss *et al.*, J. Vac. Sci. Technol. B **16**, 1439 (1998)]. In this technique, a diamond tip is scanned along the sample surface with a constant force to produce the desired features. In this article we show the observed quality and reproducibility achieved by this technique in the patterning of several semiconductor substrates. The substrates patterned and examined include GaAs, GaSb, GaP, and InP. The samples were machined at a series of loads ranging from 9.8 to 196 μN. After machining they were cleaned with an appropriate solvent and supercritical CO_2 to remove debris caused by the machining. The resulting patterned surfaces were characterized with an atomic force microscope. Lateral resolutions as good as 100 Å were successfully achieved illustrating the ability of this technique to achieve the dimensions required to form quantum dots. Line profiles indicated cut depths ranging from 5 to 500 Å. A near linear trend was observed in the depth of cut versus applied force over the investigated range for most of the substrates. The exact slope and intercepts were material dependent. © *1999 American Vacuum Society.* [S0734-211X(99)09704-8]

I. INTRODUCTION

Recently a considerable amount of research has focused on the formation of quantum dots due to their unique optical and electrical properties.[1–4] To a large part these properties are due to the quantum dots having a very narrow density of states. This leads to extremely sharp emission and absorption characteristics. A process that has been receiving considerable attention is the formation of quantum dots through self-assembly. In these systems, the self-assembled dot (SAD) formation is carried out using conditions that cause the growth to proceed in a Stranski–Kranstanov manner. An inherent characteristic of dots formed in this manner is that they have a statistical distribution of sizes. In fact, it is typically seen that the photoluminescence (PL) spectra from a layer of these dots are nearly as broad as those from a normal quantum well. This fact severely limits the usefulness of quantum dots in a laser structure. Recently researchers have been growing on novel ($n11$) oriented surfaces to help order the quantum dot structures and to decrease the spread in the size distribution.[5] As yet the distributions are too wide to realize the full enhancements expected from the quantum dots. Typical lithography has not been able to fabricate dimensions at the level needed to get the quantization of the energy levels in the dot, and/or the variability in etch rates is such that a large distribution in dot sizes occurs.[6] In this study we are looking at the technique of nanoscribing as a method to pattern semiconductor surfaces to precisely produce the dimensions needed to induce quantum effects.

Under normal loading, III–V semiconductors are quite brittle; however when they are loaded with a large hydrostatic stress with a shear component they exhibit plasticity.[7] These conditions are found in nanoscribing and diamond turning. At high loads surface cracks can be formed. Bifano *et al.* developed a model which determined criteria to ensure that all microcracks formed were removed in the turning or scribing process.[8] This condition is given below where G is the Griffith crack propagation parameter, σ_y is the yield stress of the material, and K is a constant.

$$d_c = K\frac{G}{\sigma_y}.$$

Using approximate values for GaAs, this equation gives a depth of less than 1 μm. Therefore nanoscribing surface features to a depth of less than 1 μm will result in only residual plastic deformation. Previously nanoscribing was used to pattern GaAs(001) with features as fine as 50 Å linewidth and a 2–3 Å depth.[9] It was also shown that the damage produced by diamond turning a micron of material from a GaAs(001) substrate could be removed by annealing to

[a]Electronic mail: eyinkkg@ml.wpafb.af.mil

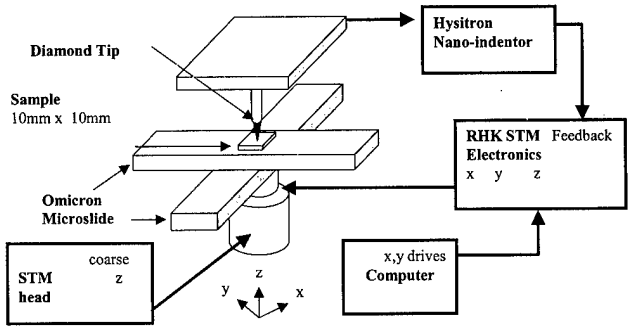


FIG. 1. Schematic diagram of the scribing apparatus used in this study.

600 °C.[9] The load used to diamond turn the material was several orders of magnitude larger than that used to nanoscribe the GaAs suggesting that the damage from nanoscribing should be annealed out. In this article we discuss the ability of nanoscribing to fabricate patterns in several III–V semiconductors and discuss the reproducibility of this process.

II. EXPERIMENT

Epiready wafers of GaSb, GaAs, InP, and GaP were cleaved into roughly 10 mm squares. The GaSb wafer was nonintentionally doped and the other wafers were semi-insulating. Patterns were scribed into the surface with a diamond tip using an instrument developed primarily from commercial components. These tips were obtained from Hysitron Inc. and had a maximum tip radius of 1000 Å. A schematic diagram of this system with the various interconnects between the components is shown in Fig. 1. The samples were affixed to Omicron microslides, which allowed coarse movements of the samples in the $x-y$ plane. The Omicron slides were mounted on a Micro Pulse Systems scanning tunneling microscope (STM) body which contains an inertial piezomotor for coarse z movement as well as a STM scan head with fine x, y, and z control. The STM scan directions were oriented 45° to that of the Omicron slides' movement. Scanning of the STM head was controlled with RHK electronics. The samples were oriented such that the x and y scribe directions were ~45° to the 110 cleavage planes so that "hard" and "soft" cleavage directions of the substrates could be neglected (averaged). The samples were brought into contact with the diamond tip through the coarse z movement until the Hysitron nanoindentor detected a force between the tip and the sample. The force output from the Hysitron was used as a feedback signal for the RHK electronics, which control the sample to tip separation thereby maintaining a constant applied force. The scanning was accomplished by computer control of the RHK electronics. This program was developed in-house and allowed arbitrary patterns to be formed at the 12 bit resolution of the scanning digital-to-analog converters (DACs), resulting in a step resolution of 100 Å. The scanning apparatus was surrounded by a double walled Plexiglas enclosure to minimize temperature variation, atmospheric disturbances, and acoustic noise. Once enclosed, the system was allowed to equilibrate for one half hour to stabilize. After equilibrating, the scan plane was electronically adjusted to be parallel to the sample surface to compensate for any mounting error. This was essential to allow sufficient dynamic range in the x and y deflections and to minimize nonlinearities near maximum deflection. In this study forces from 9.8 to 196 μN were applied to the sample and the substrates were patterned by movement of the sample in the x and y directions. Multiple patterns were placed on the same sample at the end of a registration mark by stepping the microslides enough to prevent the patterns from overlapping. A much heavier (deeper) registration mark was used to facilitate locating the patterns for atomic force microscope (AFM) analysis.

After patterning, the samples were cleaned to remove machining debris. An ammonia hydroxide rinse was used to remove oxides that may have been left after machining. A rinse with ethanol was used to remove the etchant. Supercritical carbon dioxide from an Applied Surface Technologies CO_2 snow jet cleaner was sprayed on for 30 s to remove large debris loosely bound to the substrate. Finally the surface was dried with dry nitrogen.

The AFM measurements were performed in ambient conditions, using a Park Scientific Instruments (PSI) AutoProbe CP research scanning probe microscope. The measurements were done in intermittent contact (IC), noncontact (NC), and contact modes to check for any artifacts. Scan artifacts were also minimized by using multiple scans, different scan directions, different size scans, and sample rotation relative to the scan axes. The probes used were etched silicon probes from PSI with a nominal tip curvature of 10 nm. To maximize the lateral resolution, all images were taken in 512×512 mode with an image size ($<5\times5$ μm) that allowed several scribe lines to be included. Zeroth order background removal followed by a sixth order fit was used to flatten the images. Line profiles were formed by averaging several lines which minimized noise and surface roughness in the analysis of the depth of cut.

III. RESULTS AND DISCUSSION

Figure 2 is a plot of the scribe depth in GaAs(001) for two different tips as a function of load. In each case, a linear dependence was observed between the load and the depth of cut. For these two samples, a positive x intercept was observed. This intercept is the point at which deformation remained in the material and hence should be related to the material's yield stress.[10] The slope or shape of the curve beyond yielding is complicated by dependencies on the tip geometry and the work hardening of the material. For most III–V a linear work hardening region has been found.[7] As can be seen from Figs. 1 and 2, the slope and the intercept for these two tips were different. In addition, analysis of the scribe morphology for positive and negative cut directions consistently showed a variation in the depth of cut. This difference ranged from values slightly above detection to ones easily observable. We believe that the difference in these curves as well as that observed for the scan direction is due to different tip geometries. This assumption agrees with

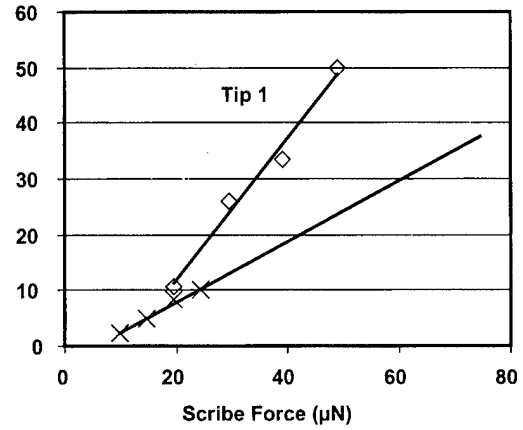


FIG. 2. Plot of the depth of cut in GaAs(001) for various loads using two different tips.

nanoindentation experiments in which it has been determined that the result for any given tip needs to be standardized to a known material for reproducible measurements. From analysis of the forces under a tip, it is known that the load force, *LC*, and the tangential forces, *TG*, are dependent on the horizontal area, A_h, and the vertical area, A_v. Using S_0 and P_0 for the critically resolved shear stress and the hardness of the material, respectively, we can write the load force and the tangential force as

$$LC = A_v P_0, \quad TG = A_v S_0 + A_h P_0.$$

These equations explicitly show that the tip geometry definitely influences the forces present and hence the depth of cut in these experiments. In our experiments, we load a sample with a force from the tip and scribe in a particular direction. At the point where scribing begins the area (A_h) will decrease to that portion which still is in contact with the sample. An assumption of a spherical tip would result in a decrease in contact area by a factor of 2 independent of scan direction, however a noncircular tip would result in A_h depending on direction. Therefore, we believe that the differences determined in the depth of cut with the direction of scribing and differences between tips are due to different areas.

Therefore to study the depth of cuts of scribing in different materials we used a single scan direction and the same tip with a particular orientation. Using this technique we scribed GaAs, GaSb, GaP, and InP substrates. The data from these experiments are plotted in Fig. 3. An interesting trend with the chemistry was seen. As the cohesive energy increased, the slope of the line decreased. This trend was true for all materials except in a comparison of InP and GaAs. This discrepancy could be due to the heavy doping with Fe that was used to make InP semi-insulating and thereby increased the InP hardness. A positive load intercept (or negative depth of cut intercept) is physically realistic since this is a measure of the elastic strain energy present prior to plastic deformation. For the GaAs and GaP data this was observed, however for GaSb and InP the data intersect the depth of cut axis at a positive value indicating material was removed without a load. We do not believe this discrepancy is due to a calibration error in the zero point of the force sensor and in fact calibration of the zero point was done prior to scribing each pattern. A possible explanation is that an attractive force exists between the diamond tip and the substrate. Since GaSb and InP are softer, this attractive force would cause a more pronounced effect on these curves. A calculation of the force due to the surface tension of absorbed water indicates that it

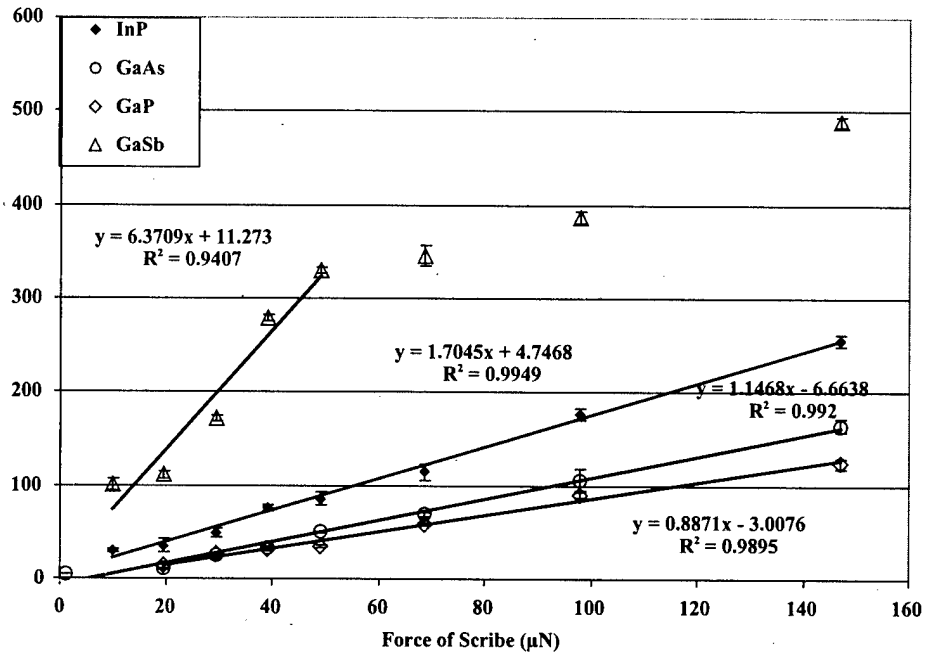


FIG. 3. Scribe depth as a function of load for various III–V semiconductors. The error bars represent the standard deviation in the measured depth in multiple patterns and lines.

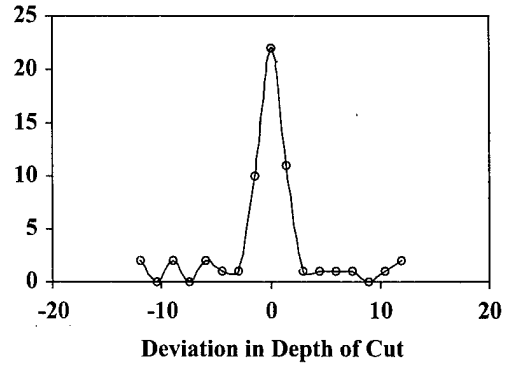


FIG. 4. Deviation from the average for all the depths scribed into a GaAs substrate.

is at least an order of magnitude too low. This intercept value was reproducible in the short term but was found to vary for long times between scribing. We believe that environmental changes may be the cause of this variation and we are currently investigating these effects.

Although the depth of cut was not consistent for different tips, the depth of cut was extremely consistent when a single tip was used and scanned with a particular orientation. Figure 4 shows a typical histogram of the spread in scribe depths for a series of loads. The standard deviation of this data suggests that the depths are within ±3 Å of the average in 95% of the scribes. The occurrences of a large deviation could generally be traced to a particle or contaminate on the surface. This analysis clearly demonstrates the extremely high degree of precision achievable by this technique. A similar analysis for the linewidth using a single tip indicated that less than a 10% variation in linewidth occurred by nanoscribing. A comparison of the reproducibility of a quantum dot formation between nanoscribing and self-assembly can now be made. Typically self-assembly produces a quantum dot distribution with sizes that vary by a factor of 2 or more. This variation in size gives rise to the composite PL transition linewidth of ~100 meV.[11] In our case assuming a particle in an infinite three-dimensional (3D) box, the 10% variation in size present would give rise to a spread of only 6 meV. Note that this is a worst case assumption since the reproducibility in depth is significantly better than the lateral 10%. This analysis suggest that nanoscribing is potentially a very accurate method to form quantum dots. We are currently fabricating structures to test this observation.

IV. CONCLUSION

The depth of cut was found to depend on the tip geometry causing variations between tips and with scan direction. Trends with the depth of cut and cohesive energy were found. Scribing was found to produce consistent depths and widths with any given tip. We were able to reproduce a pattern for a given semiconductor surface from a minimum depth of 5 Å with a width of 50 Å to 500 Å with a maximum width of 3500 Å. A near linear dependence was found for load versus depth for all the substrates except GaSb, which deviated at deeper scribes. Analysis of the depth and width of the scribes showed that dimensions could be fabricated to a couple of angstroms deep and 10 Å lateral control. This control is extremely attractive for the formation of quantum dots.

[1]D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. Denbaars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993).
[2]T. L. Reinecke, P. A. Knipp, and S. N. Walck, J. Vac. Sci. Technol. B **15**, 1040 (1997).
[3]A. Polimeni, A. Patane, M. Capizzi, F. Martelli, L. Nasi, and G. Salvaiti, Phys. Rev. B **53**, R4213 (1996).
[4]J. Phillips, K. Kamath, X. Zhou, N. Chervela, and P. Bhattacharya, J. Vac. Sci. Technol. B **16**, 1343 (1996).
[5]S. C. Fortina, S. Sanguinetti, E. Grilli, M. Henini, A. Polimeni, and L. Eaves, J. Cryst. Growth **187**, 126 (1998).
[6]Y. Wada, Microelectron. J. **29**, 601 (1998).
[7]V. Swaminatham and A. T. Macrander, *Materials Aspects of GaAs and InP* (Prentice–Hall, Englewood Cliffs, NJ, 1991), Chap. 6.7, p. 427.
[8]R. R. Kuntz, H. R. Clark, P. M. Nitishin, M. Rothschild, and B. S. Ahern, J. Mater. Res. **11**, 1228 (1996).
[9]S. H. Goss, L. Grazulis, D. H. Tomich, K. G. Eyink, S. D. Walck, T. W. Haas, D. R. Thomas, and W. V. Lampert, J. Vac. Sci. Technol. B **16**, 1439 (1998).
[10]M. V. Swain, Mater. Sci. Eng., A **253**, 160 (1998).
[11]Y. Nakata, Y. Sugiyama, T. Futatsugi, and N. Yokoyama, J. Cryst. Growth **175/176**, 713 (1997).

Properties of nanometer-sized metal–semiconductor interfaces of GaAs and InP formed by an *in situ* electrochemical process

Hideki Hasegawa,[a] Taketomo Sato, and Chinami Kaneshiro
Research Center for Interface Quantum Electronics and Graduate School of Electronics and Information Engineering, Hokkaido University, Sapporo 060-8628, Japan

(Received 19 January 1999; accepted 3 May 1999)

The properties of GaAs and InP Schottky diodes having nanometer-sized metal dots were investigated in order to clarify whether or not strong Fermi level pinning is an intrinsic property of the metal–semiconductor interface. Macroscopic Schottky diode samples having many nanometer-sized metal dots as well as single-dot Schottky diode samples were prepared by an *in situ* electrochemical process which consisted of pulsed anodic etching of the semiconductors followed by subsequent dc or pulsed cathodic deposition of the metal. Strong Fermi level pinning was not seen in the GaAs and InP macroscopic samples. The Schottky barrier height SBH values were strongly dependent on the metal work function and on the electrochemical processing conditions. Of particular interest, the difference in the dot size changed the SBH almost 340 meV in Pt/InP macroscopic Schottky diodes, indicating that Fermi level pinning disappears as the dot size is sufficiently reduced. X-ray photoelectron spectroscopy and Raman measurements indicated that these interfaces are oxide and stress free. Use of an atomic force microscope with a conductive probe allowed direct I–V measurements on single-dot samples. The metal work function and dot size dependencies of the SBHs in these samples were similar to those in macroscopic samples. Large ideality factors observed in the single-dot sample were explained in terms of environmental Fermi level pinning which produces a saddle point potential. The metal work function dependence of the SBHs measured as well as the relationship between the SBH and the ideality factor were both far from what was found in recent predictions based on the metal-induced gap state model. All the experimental results were consistently explained by the disorder-induced gap state model which asserts that strong Fermi level pinning is an extrinsic property of the metal–semiconductor interface.
© *1999 American Vacuum Society.* [S0734-211X(99)04704-6]

I. INTRODUCTION

Traditionally it has been widely accepted that Schottky barrier heights (SBHs) of electrons and holes at metal–semiconductor interfaces for Si and III–V materials are almost independent of the metal work function due to the Fermi level pinning phenomenon. The mechanism for this phenomenon, i.e., how the contact potential difference at the metal–semiconductor interface is almost completely screened so as to produce metal-independent SBH values, has been a long standing issue. Several models exist, including the unified defect model (UDM),[1] the metal-induced gap state (MIGS) model,[2,3] the disorder-induced gap state (DIGS) model,[4–6] and the effective work function (EWF) model.[7] But none of them has so far been established as the most suitable one.

The key question in the understanding of the metal–semiconductor interface is whether or not the strong Fermi level pinning is an intrinsic property of an intimate metal–semiconductor contact. One possible experimental approach to resolve this issue is to reduce the size of the metal contact into the nanometer range so that extrinsic defect formation in the macroscopic scale can be minimized, where possible complications from quantum size effects in metal and semiconductors are still not significant due to the short Fermi wavelengths of electrons and holes.

The purpose of this article is to investigate the properties of GaAs and InP Schottky diodes having nanometer-sized metal dots in order to clarify whether or not strong Fermi level pinning is an intrinsic property of an intimate metal–semiconductor contact.

Two types of samples were prepared. They were macroscopic Schottky diode samples having many nanometer-sized metal dots, and single-dot Schottky diode samples having a single nanometer-sized Schottky dot electrode. They were prepared by an *in situ* electrochemical process.[8–10] The process consisted of pulsed anodic etching of the semiconductors followed by subsequent dc or pulsed cathodic deposition of the metal *in situ* in the same electrolyte containing metal ions (Ag, Sn, Co, Ni, or Pt).

Strong Fermi level pinning was not seen in the GaAs and InP macroscopic samples. The SBH values were strongly dependent on the metal work function and on the electrochemical processing conditions. Of particular interest, the difference in the dot size changed the SBH almost 340 meV in Pt/InP macroscopic Schottky diodes, indicating that Fermi level pinning disappears as the dot size is sufficiently reduced. X-ray photoelectron spectroscopy (XPS) and Raman measurements indicated that these interfaces were oxide and stress free.

[a]Electronic mail: hasegawa@rciqe.hokudai.ac.jp

Use of an atomic force microscope (AFM) with a conductive probe allowed direct $I-V$ measurements on single-dot samples. The metal work function and dot size dependencies of the SBHs in these samples were similar to those in macroscopic samples. The large ideality factors observed in single-dot sample were explained in terms of environmental Fermi level pinning which produces a saddle point potential.

The metal work function dependence of the SBHs measured as well as the relationship between the SBH and the ideality factor were both far from what was found in recent predictions[11,12] based on the MIGS model. All the experimental results were consistently explained by the DIGS[4–6] model which asserts that strong Fermi level pinning is an extrinsic property of the metal–semiconductor interface.

II. EXPERIMENT

A. Electrochemical formation of Schottky barriers having nanometer-sized metal dots

The present work was motivated by our recent finding[9,10] that nanometer-sized metal dots are formed in a self-assembled fashion on the semiconductor surface at the initial stage of the electrochemical process. Longer plating times do not increase the dot size but increase the number of such grains, leading to uniform coverage of the semiconductor surfaces with nanometer-sized dots. It has further been found[13] that isolated dots can be formed by mask patterning the electrode surface with photoresist film (OFPR8600, Tokyo Ohka Kogyo Co., Ltd.) or electron beam (EB) resist film (OEBR1000, Tokyo Ohka Kogyo Co., Ltd.) using standard lithography techniques. Such organic films are not attacked by electrolytes, and electrochemical reactions occur only at the unmasked conductive area of the electrolyte/semiconductor interface.

In this study, GaAs and InP Schottky barrier samples having nanometer-sized metal dots were fabricated, and their electrical characteristics were investigated. As mentioned already, two types of samples were prepared. One type was a macroscopic Schottky diode sample. In this sample, a large circular metal electrode with a diameter of 400 μm was formed by the electrochemical process in the self-assembled mode to a thick uniform coverage. Selective deposition into macroscopic circular dots was achieved by using a photoresist mask. This sample contained many nanometer-sized metal dots. The other type was a single-dot Schottky diode sample where a single metal dot was formed either in the self-assembled mode or the masked selective mode. For the masked deposition at desired positions in single-dot samples, open windows were patterned on top of the semiconductor surface with periods of 200–300 nm by standard EB lithography using a JBX-5000LS (JEOL). An electron beam with an accelerating voltage of 50 kV and a beam diameter of 50 nm was used. In this study, altogether 38 samples including GaAs macroscopic samples, 29 InP macroscopic samples, and 5 InP single-dot samples were prepared and measured. From theses, details of the data are shown in Tables I and II for 12 major macroscopic samples, M1–M12, and 5 single-dot samples, S1–S5, respectively.

B. Electrochemical process

The electrochemical process used in this study for dot formation consisted of controlled anodic etching of the semiconductors followed by subsequent cathodic deposition of the metal *in situ* in the same electrolyte. The electrolyte bath contained three electrodes, i.e., a semiconductor electrode onto which metal is deposited, a Pt counterelectrode, and a reference saturated calomel electrode (SCE). The overpotential of the semiconductor electrode with respect to the SCE reference was controlled by a potentiostat.

In this study, Pt, Ni, Co, Ag, and Sn were deposited onto n-type (001) GaAs and n-type (001) InP, surfaces using the following electrolytes.

Pt:1 M HCl (200 ml)+H_2PtCl_6+NH_4OH [pH=1];
Co:1 M HCl (200 ml)+$CoSO_4$ (10 g)+NH_4OH [pH=1];
Ni:1 M HCl (200 ml)+$NiSO_4$ (24 g)+$NiCl_2$ (3 g)
+NH_4OH[pH=2];
Ag:1 M HCl (200 ml)+AgCl (1 g) [pH=0];
Sn:1 M H_2SO_4 (200 ml)+$SnSO_4$ (8 g) [pH=0.3].

The electrical supply wave forms used for the etching and plating modes are shown in Fig. 1. Etching was done in a pulsed mode where holes were produced near the surface of the semiconductor by avalanche pulses with height V_{he}, width t_{we}, and period t_{pe}. These holes oxidize the semiconductor surface where oxides are immediately dissolved into the electrolyte resulting in anodic dissolution. Typical values of V_{he}, t_{we}, and t_{pe} used in this study were 15 V, 10 μs, and 400 μs, respectively. A separate study[14] has shown that this pulsed anodic etching mode can achieve excellent etch depth uniformity and extremely fine etch depth control of a 3 $\times 10^{-5}$ nm/pulse. After etching of the surface, plating was initiated by changing the polarity of the potential supply. This was done either in dc mode with a constant potential V_{rd} or in the pulsed mode where pulses with height V_{hd}, width t_{wd}, and period t_{pd} were superposed onto an offset dc potential of V_{rd}.

C. Interface characterization

The electrical properties of the macroscopic Schottky diode samples containing many metal dots were investigated by standard $I-V$ and $C-V$ measurements. In order to measure the $I-V$ characteristics of a single-dot Schottky diode having a single metal dot, an AFM (Nanaoscope II) with a conductive cantilever was used. This conductive AFM system enables one to perform normal AFM observation by the optical lever method as well as $I-V$ measurements between the surface and the back surface of the sample using a HP4156A parameter analyzer. The conductive probe was fabricated from highly doped Si with a radius of curvature of about 10 nm. During the electrical measurements, the laser light was turned off in order to avoid the photovoltaic effect.

To determine the value of the SBH, ϕ_B, and the ideality factor, n, from the $I-V$ characteristics, the current density J was fitted into the following standard thermionic equation as a function of voltage V.

TABLE I. Preparation conditions, size, and electrical properties of the GaAs and InP macroscopic Schottky diode samples.

Sample No.	Semiconductor	Metal	Plating conditions	$\bar{d}$ (nm)	σ_d (nm)	ϕ_{Bn}^{I-V} (eV)	n value	ϕ_{Bn}^{C-V} (eV)
M1	GaAs	Sn	dc $V_{rd}=0.68$ V	···	···	0.69	1.06	0.70
M2	GaAs	Co	Pulse $V_{rd}=0.5$ V, $V_{hd}=0.7$ V $t_{pd}=400$ μs, $t_{wd}=50$ μs	···	···	0.78	1.03	0.80
M3	GaAs	Ni	dc $V_{rd}=0.9$ V	···	···	0.81	1.05	0.79
M4	GaAs	Pt	Pulse $V_{rd}=0.4$ V, $V_{hd}=1.0$ V $t_{pd}=400$ μs, $t_{wd}=10$ μs	···	···	0.92	1.08	0.91
M5	InP	Ag	Pulse $V_{rd}=0.3$ V, $V_{hd}=0.8$ V $t_{pd}=400$ μs, $t_{wd}=10$ μs	···	···	0.35	1.15	···
M6	InP	Co	Pulse $V_{rd}=0.4$ V, $V_{hd}=1.0$ V $t_{pd}=400$ μs, $t_{wd}=10$ μs	···	···	0.48	1.07	0.45
M7	InP	Ni	Pulse $V_{rd}=0.4$ V, $V_{hd}=1.0$ V $t_{pd}=400$ μs, $t_{wd}=10$ μs	···	···	0.53	1.12	0.50
M8	InP	Pt	Pulse $V_{rd}=0.4$ V, $V_{hd}=1.0$ V $t_{pd}=400$ μs, $t_{wd}=10$ μs	41	14	0.86	1.08	0.84
M9	InP	Pt	Pulse $V_{rd}=0.4$ V, $V_{hd}=1.0$ V $t_{pd}=200$ μs, $t_{wd}=10$ μs	33	9	0.73	1.13	0.72
M10	InP	Pt	dc $V_{rd}=0.7$ V	29	9	0.63	1.03	0.64
M11	InP	Pt	Pulse $V_{rd}=0.4$ V, $V_{hd}=1.0$ V $t_{pd}=400$ μs, $t_{wd}=100$ μs	26	6	0.60	1.10	0.62
M12	InP	Pt	dc $V_{rd}=1.4$ V	26	3	0.52	1.10	···

$$J=A^{**}T^2\exp(-q\phi_B/kT)[\exp(qV/nkT)-1], \quad (1)$$

where A^{**} is the effective Richardson constant. As values of A^{**}, 8.16 and 9.6 $A\,cm^{-2}K^{-2}$ were used for n-type GaAs and n-type InP, respectively.

The size and size distribution of the metal dots were investigated by the AFM and a scanning electron microscope (SEM) using the Nanoscope II and the Hitachi S-4100, respectively.

The chemical composition and profile of the interface were studied by XPS with Ar-ion sputtering using a Perkin-Elmer 5100(C) spectrometer. The excitation source was Al $K\alpha(h\nu=1486.6\,eV)$. Raman spectroscopy measurements of the plated surfaces were made using Jobin-Yvon TA 64000 Raman spectrometer. XPS and Raman spectrometers are connected either directly through an ultrahigh vacuum (UHV) transfer chamber or indirectly via a portable UHV chamber (a vacuum suitcase) to the main UHV-based growth, processing, and characterization system installed at the Research Center for Interface Quantum Electronics, Hokkaido University, so that reliable reference spectra from clean GaAs and InP surfaces freshly grown by molecular beam epitaxy (MBE) can be used for spectral analysis.

TABLE II. Preparation conditions, size, and electrical properties of InP single-dot Schottky diode samples.

Sample No.	Semiconductor	Metal	Preparation	Dot size (nm)	ϕ_{Bn}^{I-V} (eV)	n value
S1	InP	Sn	Masked	230	0.41	2.1
S2	InP	Pt	Self-assembled	190	0.47	1.5
S3	InP	Ag	Self-assembled	140	0.45	1.5
S4	InP	Pt	Masked	100	0.57	1.6
S5	InP	Pt	Self-assembled	70	0.69	1.27

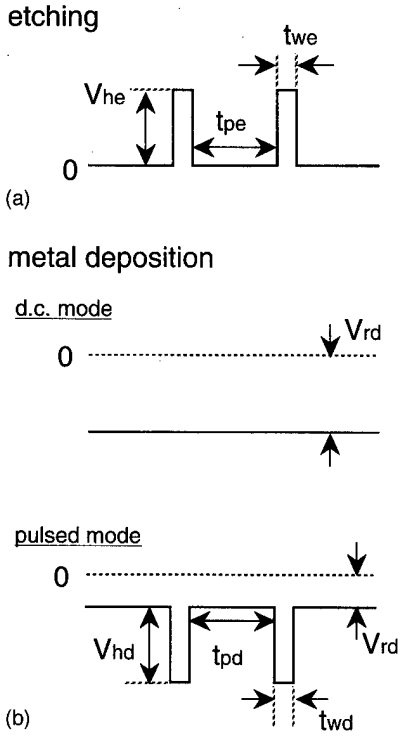


FIG. 1. Electrical supply wave forms for etching and plating modes of the electrochemical process.

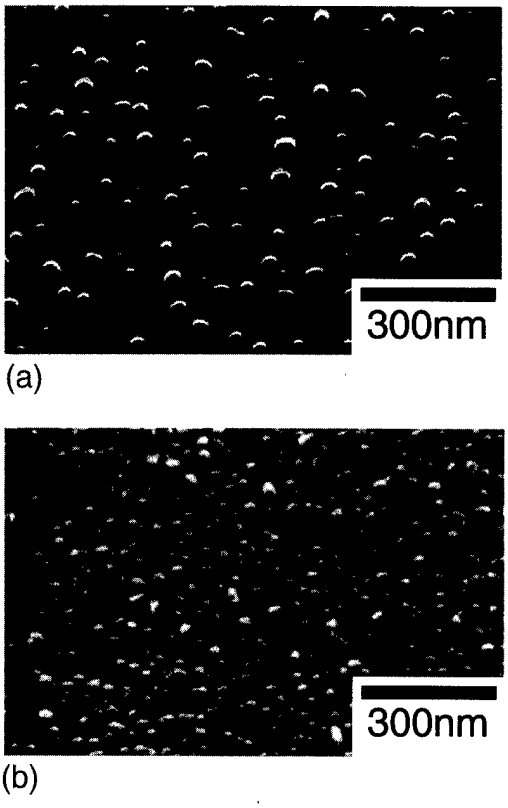


FIG. 2. SEM images of self-assembled nanometer Pt dots on (a) GaAs and (b) InP formed by the *in situ* electrochemical process.

III. EXPERIMENTAL RESULTS

A. SEM/AFM study of dot geometry in macroscopic samples

As mentioned already, nanometer-sized metal dots were formed in a self-assembled fashion on the surface of the semiconductor substrate at the initial stage of the electrochemical process. Examples of SEM images of Pt dots formed on GaAs and InP are shown in Fig. 2. In both cases, the formation of nanometer-sized Pt dots can be seen. Further deposition did not increase the size of the individual dots, but increased the number of the dots. Eventually, remarkably smooth Pt films were formed after uniform coverage by size-saturated Pt dots in both dc and pulsed modes. For example, the Pt-covered surfaces consisting of a dense assembly of small grains obtained by electrodeposition had a root mean square (rms) roughness of 2.4 nm whereas the Pt surface obtained by standard electron beam deposition showed a rms roughness of 7.9 nm and the presence of much larger and irregular grains.

The detailed shape, size, and size distribution of the dots were found to depend strongly on the metals deposited and the wave forms applied. For example, SEM images of four GaAs macroscopic Schottky samples, M1–M4, having Sn, Co, Ni, and Pt dots are shown in Figs. 3(a)–3(d), respectively. Co dots were rather similar to the Pt dots, and the shape of the Ni dots was more irregular. For Sn, only large elliptic or circular dots were obtained. A similar tendency existed in the InP macroscopic samples, M5–M8.

Through the initial experiments, Pt was found to consistently produce the smallest dot sizes. Since the main motivation for the present study was to clarify the properties of Schottky diodes having small dots, a more detailed study was carried out on the size distribution of Pt dots as a function of electrochemical conditions. The distributions were characterized in terms of the average in-plane diameter, the average vertical height, and their standard deviations. The average in-plane diameter, $\bar{d}$, of the dot and its uniformity strongly depended on the electrochemical conditions. Compared with the dc mode, the pulsed mode produced smaller, more uniform Pt particles on GaAs and InP substrates. In the pulsed mode, the diameter distributions were strongly dependent on the pulse conditions. Both the dot diameter, $\bar{d}$, and the its standard deviation, σ_d, decreased with smaller t_{wd} and longer t_{pd}, indicating the importance of the supply of metal-ion species at the semiconductor/electrolyte interface. A uniform distribution of small dots could be obtained by quickly depositing the metal in a short pulse-on time and waiting for recovery of the reduced metal-ion density to a sufficient level during a long pulse-off time. On the other hand, the vertical height of the dot was generally smaller than the in-plane diameter, roughly keeping the ratio of the height to the diameter within the range of 0.5–0.8.

The dot size distributions measured on actual Pt/InP macroscopic samples, M8–M12, are shown in Fig. 4. As already mentioned, the electrochemical preparation conditions and the measured average dot diameter, $\bar{d}$, and its standard deviation, σ_d, are summarized in Table I.

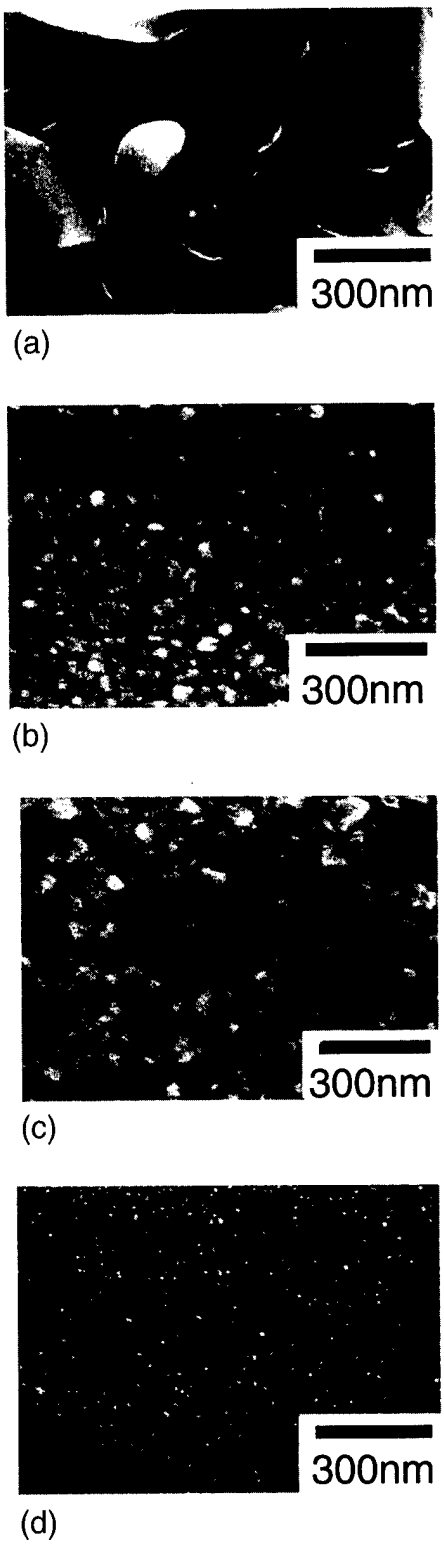


FIG. 3. SEM images of (a) Sn, (b) Co, and (c) Ni and Pt dots on GaAs surfaces prepared by the *in situ* electrochemical process.

B. Electrical, XPS, and Raman studies of macroscopic samples

The measured forward and reverse $I-V$ characteristics of n-GaAs and n-InP macroscopic Schottky samples (M1–M8) are shown in Fig. 5. As seen in Fig. 5, all the samples showed nearly ideal thermionic emission characteristics. The values of the SBH and the ideality factor n using Eq. (1) are shown in Table I. The result showed much metal-dependent behavior. The highest SBH values were obtained by Pt deposition in the cases of both GaAs and InP.

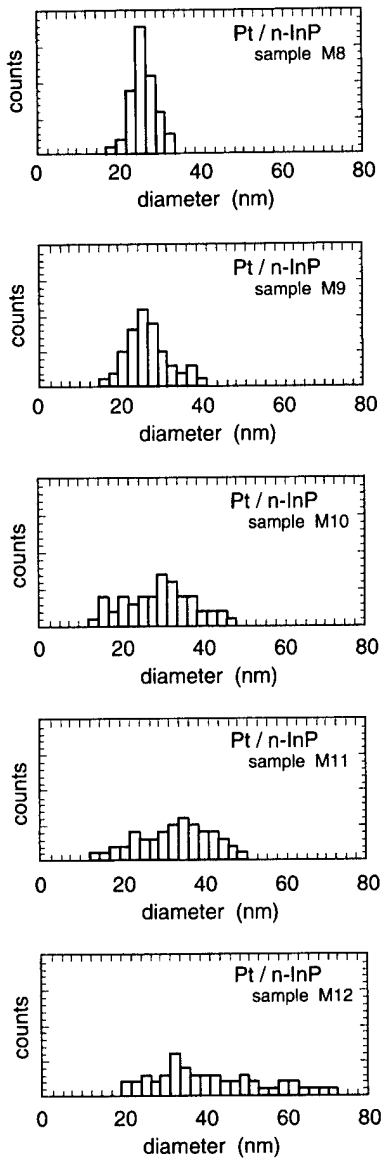


FIG. 4. Pt dot size distributions measured by AFM on five InP macroscopic Schottky samples, M8–M12, prepared under different electrochemical conditions.

The measured forward and reverse $I-V$ characteristics of InP macroscopic Schottky diode samples (M8–M12), having the different dot size distributions shown in Fig. 4, are summarized in Fig. 6. The values of the SBH and ideality factor n of these samples determined from the $I-V$ characteristics are again summarized in Table I for the various Schottky diodes studied here. As seen in Fig. 6 and Table I, there exists an extremely strong correlation between the dot size distribution and the SBH value. For the same combination of InP and electroplated Pt, variation of the SBH of 340 meV took place, depending on the size distribution. That is, the

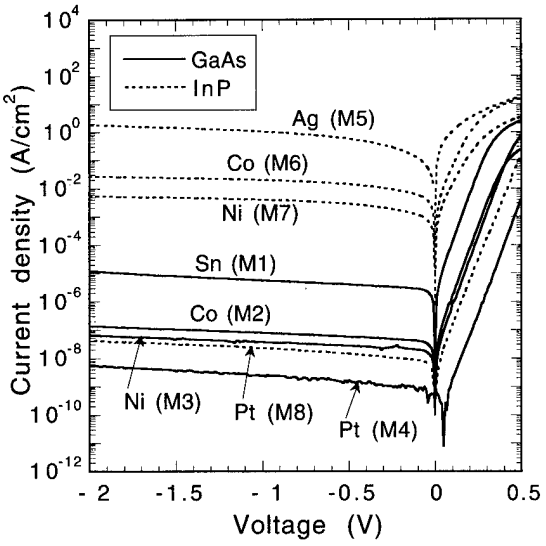


FIG. 5. Measured forward and reverse $I-V$ characteristics of GaAs and InP macroscopic Schottky diode samples.

SBH value became larger as the size distribution approached a more uniform distribution of smaller dots. The highest SBH value of 0.86 eV obtained by the optimized pulse condition corresponds to the smallest Pt dots with a minimal standard deviation.

$C-V$ measurements were also made on the macroscopic GaAs and InP Schottky samples in order to investigate the possible existence of an interfacial layer or near-surface band modification which may give the apparent high SBH values. Very straight $1/C^2-V$ plots were obtained in all the samples, and their slopes were consistent with the electron concentration determined by Hall measurements. The values of SBHs determined by $C-V$ measurements are shown in Table I. Excellent agreement of the SBH values was obtained between the $I-V$ and $C-V$ methods. This result strongly indicates that the *in situ* electrochemical process used in this study realizes intimate Schottky contacts without an interfacial layer and near-surface modification of the band profile.

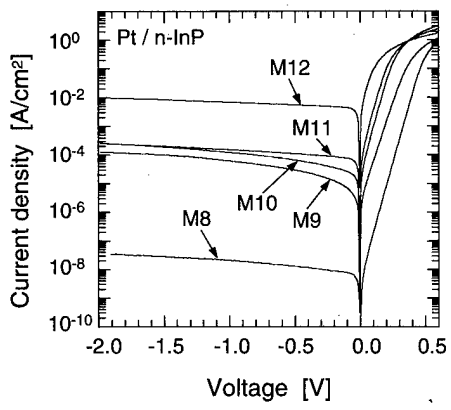


FIG. 6. Forward and reverse $I-V$ characteristics of Pt/n-InP macroscopic Schottky diode samples, M8–M12, having various dot size distributions that are shown in Fig. 4.

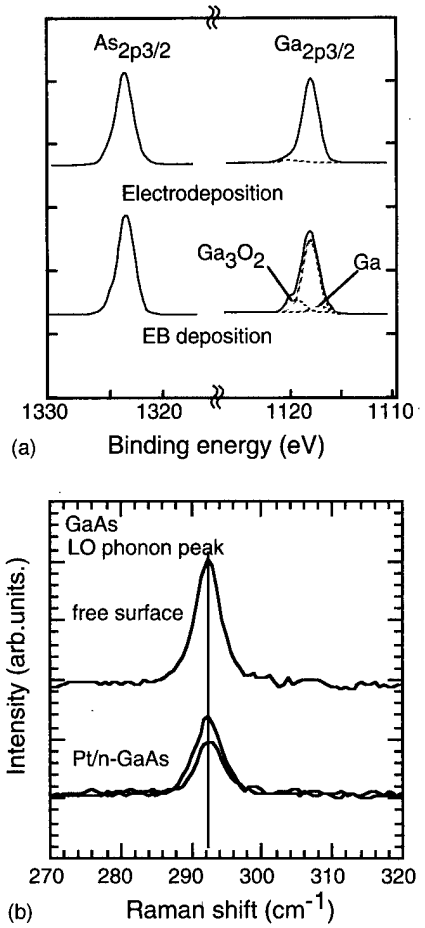


FIG. 7. (a) XPS and (b) Raman spectra from a Pt/GaAs macroscopic Schottky sample. XPS spectra from an EB deposited sample and Raman spectra from an InP free surface are shown for comparison.

An XPS in-depth profile analysis using Ar-ion sputtering was made on the GaAs and InP macroscopic Schottky samples in order to further check the presence or absence of an interfacial layer. The $As_{2p3/2}$ and $Ga_{2p3/2}$ spectra observed at the interface are compared in Fig. 7(a) for electrodeposited and EB-deposited Pt/GaAs interfaces. These spectra were deconvoluted using the reference spectra obtained from a fresh MBE grown (001) GaAs surface subjected to Ar-ion sputtering. As seen in Fig. 8(a), the spectra from the Pt-electroplated interface were almost completely the same as those from the Ar-sputtered reference free surface. In contrast to this, EB-deposited Pt/GaAs interfaces possessed spectra that are significantly different from the reference, showing broadening, peak shifts, and inclusion of oxides. Thus, the Pt-electroplated interface is an oxide free intimate contact without inclusion of an appreciable interfacial oxide layer. Exactly the same conclusion was obtained on electroplated InP samples as already reported.[10]

The Raman spectra obtained from the surface of a Pt/GaAs sample that was prepared by electrodeposition in pulsed mode under the same conditions as sample M4 are shown for two Pt different thicknesses in Fig. 7(b). It is seen in Fig. 7(b) that no appreciable Raman peak shift of the GaAs longitudinal optical (LO) phonon peak and no broad-

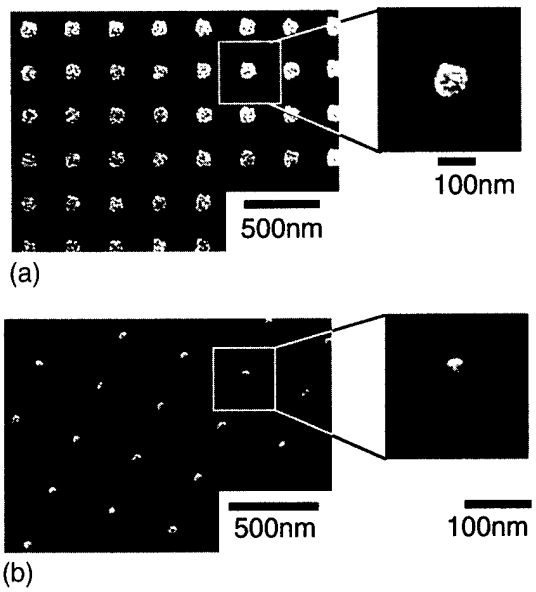


FIG. 8. SEM images of nanometer-sized metal dot arrays prepared by EB lithography and electrochamical deposition. (a) GaAs and (b) InP.

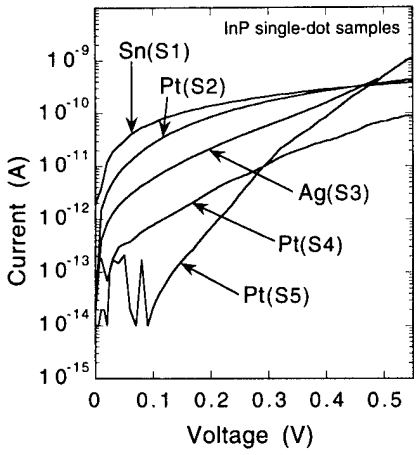


FIG. 9. Forward $I-V$ characteristics of InP single-dot Schottky diode samples, S1–S5, measured by and AFM with a conductive tip.

ening take place during electrodeposition of Pt with respect to the reference spectra taken from a MBE grown free (001) GaAs surface. Since Raman peak shifts and broadening are related to the presence of stress, the result here shows that the Pt-electrodeposited GaAs surface is remarkably free from stress. Exactly the same conclusion was obtained on electroplated InP samples reported earlier.[10]

C. SEM/AFM and electrical studies of single-dot samples

In single-dot samples, either a single dot chosen from the assembly of isolated dots appearing at the initial stage of the self-assembled mode, or a single dot chosen from the dot array formed by selective deposition using a patterned EB resist mask was used as the Schottky electrode. Examples of self-assembled Pt dots on GaAs and InP were already given in Fig. 2. Figures 8(a) and 8(b) show examples of SEM images of the Pt dot arrays formed on n-GaAs and n-InP, respectively. It is seen that Pt dots were selectively formed only within the open windows. The size of the Pt dot in the arrays can be changed not only by changing the window size, but also by changing the number of pulses applied. In fact, an extremely small dot with a diameter of 20 nm could be formed at the center of 100 nm window on GaAs just by supplying a single pulse. Details of this technique are described elsewhere.[13]

In this study, the $I-V$ characteristics of five single-dot InP Schottky samples (S1–S5) were measured using an AFM system with a conductive probe. The dot size and preparation conditions of these samples are shown in Table II. Due to a rather large radius of curvature of the probe tip of 10 nm, the smallest dot measured in this study had a diameter of 60 nm. This is because it became increasingly difficult to accurately bring the probe tip right above the dot top and make and maintain a secure electrical contact during the measurement. It should also be mentioned that the dot diameter values shown in Table II were obtained by taking the full width at half maximum (FWHM) of the AFM line profile of each dot. Due to the limited spatial resolution of the AFM tip, these values are not accurate. A comparison with the SEM images has indicated that these values seem to overestimate the dot diameters by about 15%.

All the single-dot samples showed clearly rectifying behavior. The measured forward $I-V$ characteristics of five single-dot InP Schottky samples (S1–S5) are shown in Fig. 9. The oscillatory behavior of the current in sample S5 in Fig. 9 was due to residual current noise of several tens of fA. The values of the SBH and the ideality factor, n, obtained by fitting them to Eq. (1) are shown in Table II. Since the measured $I-V$ curves in the $\log I$ vs V plots in Fig. 9 were highly nonlinear, the fitting was done with the slope at low forward voltages near zero bias.

It should be noted that qualitatively the dependencies of the $I-V$ characteristics of the single-dot samples on the contact metal and on the dot size are similar to those observed in the macroscopic samples. That is, the SBH value increased in the order of Ag, Sn, and Pt, and for the same Pt, the SBH value increased with a reduction of the dot size.

However, the measured values of the ideality factor are extremely large, being in the range of 1.3–2.1, and one is not quite sure to what extent the SBH values obtained by forcing Eq. (1) to nonlinear $I-V$ plots in Fig. 9 are reliable for quantitative discussion. This point will be discussed later.

IV. DISCUSSION

The observed SBH values of n-type GaAs and InP macroscopic Schottky diodes produced by the electrochemical process are plotted in Figs. 10(a) and 10(b) versus the metal work function using the work function values of Michaelson.[15] The result for InP basically reproduces our previous result.[10] However, there is an important new finding that the SBH value for Pt can change 340 meV depending on the size distribution of the Pt dots. In Fig. 11, the measured SBH values for Pt/InP macroscopic samples are plotted ver-

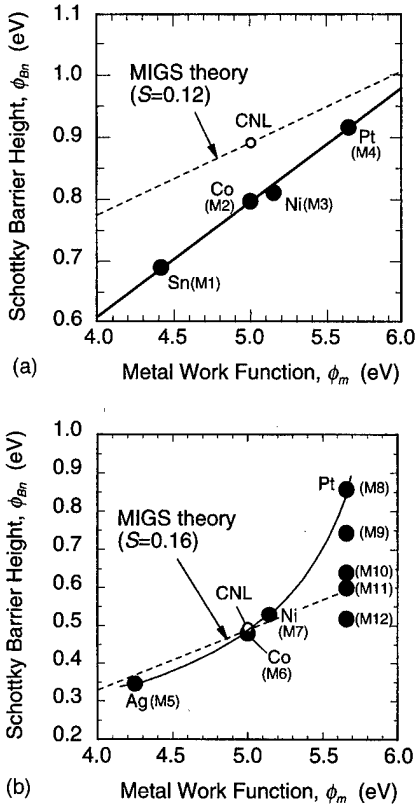


FIG. 10. Metal work function dependence of SBHs of (a) GaAs and (b) InP macroscopic Schottky diode samples. Dotted lines show prediction by the formula given in Ref. 11 based on the MIGS model.

sus the average dot diameter, $\bar{d}$. As the dot size decreased, the SBH rapidly increased toward the value of the Mott–Schottky limit. It should, however, be mentioned here that the result in Fig. 11 does not imply that the average dot diameter, $\bar{d}$, is the only controlling parameter that determines

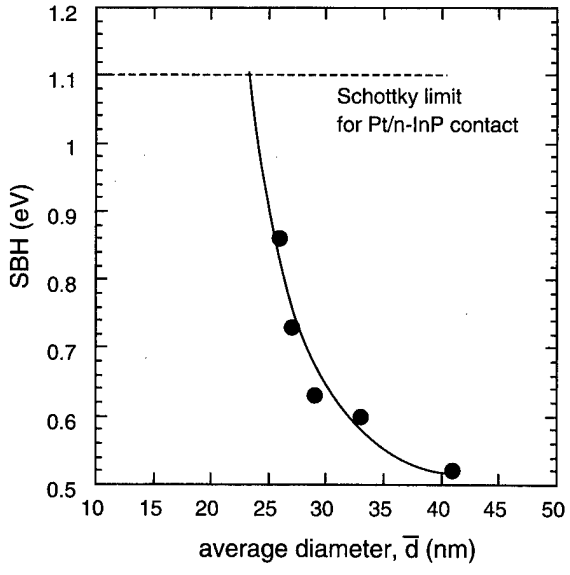


FIG. 11. Schottky barrier heights plotted vs average dot diameter, $\bar{d}$, for Pt/InP macroscopic Schottky samples.

the SBH, since the average diameter, $\bar{d}$, and the standard deviation, σ_d, could not be changed independently in the present study.

In both cases of GaAs and InP, the SBHs are strongly dependent on the work function, and there is no indication that the Fermi level is "pinned." Furthermore, the result on Pt/InP shows that, even if one considers the electrodeposition process, the SBH is strongly dependent on details of the interface formation process which determines the size distribution. Although not systematically investigated here, it is highly likely that a similar situation exists for other combinations of barrier metals and semiconductors. In fact, our first report[8] on electroplated InP Schottky barriers included a brief result on GaAs Schottky barriers for comparison where the Fermi level was pinned close to 900 meV over the work function range of 4.4–5.6 eV. By having more experience on electrodeposition, the SBH values for GaAs became much more work function dependent for low work function metals.

Traditionally the major models for Schottky barriers such as the UDM,[1] MIGS model,[2,3] DIGS model,[4–6] and EWF model[7] have been tried to explain why the SBH values are so strongly "independent" of the metal work function and the method and conditions of the metal deposition. Ironically, a more pertinent question related to the present result now is why they are so strongly "dependent" on them. Since the nonapplicability of the UDM and the EWF model to the electrochemically produced Schottky diodes was discussed in our previous paper,[10] model discussion here concentrates on the applicability of the MIGS and DIGS models to the present data.

According to the MIGS model, penetration of the metal wave function into the semiconductor produces a metal-induced gap state continuum with a characteristic charge neutrality level (CNL), E_{CNL}. If the state density of the MIGS continuum is sufficiently high, the Fermi level is pinned at E_{CNL}. Actually, this is not the case, and this makes the screening capability of the MIGS continuum against the contact potential difference a limited one. To take this effect into account, Mönch[11] recently derived a semiemprical formula for the SBH.

$$\phi_{\mathrm{Bn}}=(E_C-E_{\mathrm{CNL}})+S(\phi_m-\phi_{\mathrm{CNL}}), \tag{2}$$

with

$$S=\frac{1}{1+\dfrac{q^2}{\epsilon_i\epsilon_0}D_{\mathrm{MIGS}}(E_{\mathrm{CNL}})\,\delta_i} \tag{3}$$

$$=\frac{1}{1+0.29\dfrac{(\epsilon_\infty-1)^2}{\epsilon_i}}, \tag{4}$$

where ϕ_m is the metal work function, ϕ_{CNL} is the energy location of E_{CNL} from the vacuum level, E_{VAC}, S is the slope factor, ϵ_0 is the permittivity of vacuum, ϵ_i is the relative permittivity of the interface, q is the electronic charge, $D_{\mathrm{MIGS}}(E_{\mathrm{CNL}})$ is the density of the MIGS at E_{CNL}, and δ_i is the thickness of the interface dipole layer (typically 0.2 nm)

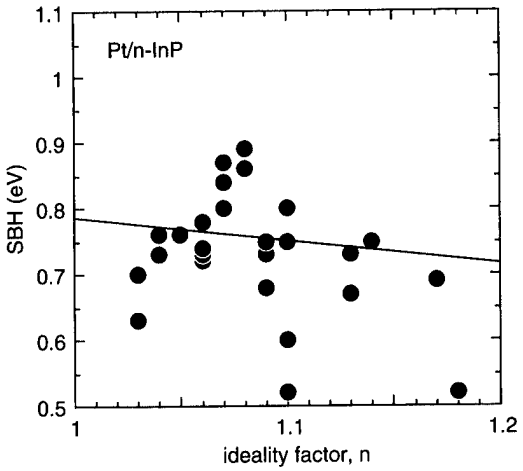


FIG. 12. Schottky barrier heights plotted vs ideality factor, n, for 26 Pt/InP macroscopic Schottky samples prepared by the electrochemical process.

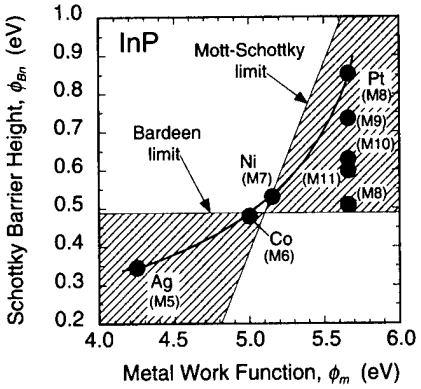


FIG. 13. SBH diagram for n-InP based on the DIGS model.

determined by the penetration depth of the MIGS. Here, the notations are changed from those of Ref. 11 so as to use the work function rather than the electronegativity.

Using, $E_{\mathrm{CNL}}=E_V+0.55\,\mathrm{eV}$,[16] $\phi_{\mathrm{CNL}}=5.0\,\mathrm{eV}$,[11] and $\epsilon_\infty=10.9$ (Ref. 17) for GaAs and $E_{\mathrm{CNL}}=E_V+0.87\,\mathrm{eV}$,[16] $\phi_{\mathrm{CNL}}=5.0\,\mathrm{eV}$,[11] $\epsilon_\infty=9.5$ (Ref. 18) for InP, and $\epsilon_i=4$,[11] one obtains $S=0.12$ for GaAs and $S=0.16$ for InP. The relationship given by Eqs. (2) and (4) are plotted by dotted lines in Figs. 10(a) and 10(b). As seen in Figs. 10(a) and 10(b), these lines are far from the experimental data obtained here. Thus, the SBH values obtained here cannot be explained by the MIGS model.

However, the above formula assumes an intimate and homogeneous interface, and one has to consider possible deviation due to inhomogeneities. As we have seen, our diodes are not homogeneous and consist of nanometer-sized metal dots with statistical distributions of the dot sizes. Based on a general theory of inhomogeneous patchy Schottky barriers developed by Tung,[19] Kampen *et al.*[12] recently utilized the following equation to deduce the "true" SBH, $\phi_{\mathrm{Bn}}^{\mathrm{hom}}$, of a homogeneous diode from measurements of the SBH, $\phi_{\mathrm{Bn}}^{\mathrm{eff}}$, and the ideality factor, n_{eff}, of inhomogeneous diodes of the same type.

$$\phi_{\mathrm{Bn}}^{\mathrm{eff}}=\phi_{\mathrm{Bn}}^{\mathrm{hom}}-(n_{\mathrm{eff}}-1.01)qV_i^{\mathrm{hom}}/\xi, \tag{5}$$

where qV_i^{hom} is the band bending at the uniform interface, and ξ is a characteristic parameter related to the shape of the "patch."

In order to whether see such a relationship exists in our macroscopic Pt/InP Schottky diodes prepared by electrodeposition, the SBH and n were measured for 26 diodes including samples M5–M12. Except for some details of the electrochemical current supply conditions, the diodes were prepared under exactly the same conditions. The result is plotted in Fig. 12. As seen in Fig. 12, there is no clear indication of linear correlation between the SBH and n. Additionally, if one forces a linear relation, using the method of least squares, the line indicated in Fig. 12 is obtained which gives a value of the SBH of the homogeneous interface of 780 meV. This value is much larger than the predicted value of 584 meV using Eqs. (2) and (4) with $\phi_m=5.6\,\mathrm{eV}$ for Pt.[15] Thus, consideration of inhomogeneity does not further improve the agreement.

On the other hand, according to the DIGS model for Fermi level pinning[4–6] proposed by our group, deposition of a metal or an insulator on the semiconductor surface disturbs the crystalline perfection of the semiconductor surface and forms a thin disordered semiconductor layer with a thickness, δ, whose electronic properties are featured by the DIGS continuum. The continuum has continuous energy and spatial distribution of gap states of acceptor type and donor type with a characteristic charge neutrality level E_{HO}. E_{HO} agrees with E_{CNL} of the MIGS model.

The DIGS model gives the same formula of Eq. (2) for the SBH, but with the following expression for the slope factor.[5,6]

$$S=\mathrm{sech}(\delta/\lambda) \tag{6}$$

$$\fallingdotseq\frac{1}{1+\dfrac{q^2}{2\epsilon_i\epsilon_0}N_{ss}^{\mathrm{DIGS}}(E_{\mathrm{HO}})\delta}, \tag{7}$$

where

$$\lambda=[\epsilon_0\epsilon_\infty/q^2N_{\mathrm{DIGS}}(E_{\mathrm{HO}})]^{1/2}. \tag{8}$$

Here, λ is the DIGS screening length, $N_{\mathrm{DIGS}}(E_{\mathrm{HO}})$ is the volume density of the DIGS at the charge neutrality level, E_{HO}, $N_{ss}^{\mathrm{DIGS}}(E_{\mathrm{HO}})=N_{\mathrm{DIGS}}(E_{\mathrm{HO}})\cdot\delta$. δ is the surface density of the DIGS at the charge neutrality level E_{HO}. According to this model, Fermi level pinning is extrinsic and the value of S is process dependent. When the degree of disorder is high, S becomes zero and the Fermi level is firmly pinned at E_{HO}. The situation is different from the recent MIGS model and complete pinning can take place here. However, when an intimate metal–semiconductor (M–S) interface is formed without causing disorder, S becomes unity and the ideal Mott–Schottky limit is realized. Thus, the SBH values of real Schottky diodes should lie between these two limits. A SBH diagram based on these two limiting lines was made and the result is compared in Fig. 13 with the measured data for InP macroscopic Schottky diodes. Here, we used

ϕ_{CNL}=5.1 eV for a better fit, although our initial guess based on photoemission data was 5.0±0.1 eV.[4] All the data obtained in this study lie in the shaded region, and thus they are consistent with the DIGS model.

The observed remarkable change in the SBH with the Pt dot size can be explained in the following way. When the dot size is small, the metal dot can maintain a coherent bond relationship with the semiconductor underneath. Thus, no severe processing-induced disorder is caused and the Fermi level is not pinned. This is particularly expected for the present *in situ* electrochemical process which possesses an extremely low processing energy with the application of small voltages of 500–700 mV at room temperature. However, when a large dot is formed, it tends to generate stress at the interface and to introduce defects and disorder into the semiconductor surface region. Thus, the Fermi level starts to be pinned.

With regard to the relationship between the SBH and n, it should be noted that the data in Fig. 12, without a well-defined linear correlation, do not contradict the general theory itself reported by Tung. In deriving Eq. (5), it is assumed that all the diodes have the same average SBH. The data in Fig. 12 simply indicate that such an assumption is not true in our samples, and that the average SBH may be different from one diode to another. Our diodes are obviously inhomogeneous, and, from the viewpoint of the DIGS model, the average SBH will be different from one diode to another, depending on the degree of process-induced disorder they received.

Finally, measurements on the single-dot samples also confirmed more directly that the SBH for Pt increases as the dot size is reduced. There was, however, a problem in that the ideality factor was too large to be acceptable from a conventional point of view. One possible mechanism for such high n is hole injection from the metal, which was indeed the operational principle of the classical "point" contact transistor. However, from a careful survey of old literature on minority carrier injection in Schottky diodes, we came to the conclusion that this is not the case here, because the barrier against hole injection is too high in InP Schottky barriers to make this process dominant. We believe that the reason for the observed highly nonlinear $\log I$ vs V plots with large values of n can be explained in terms of environmental Fermi level pinning, schematically shown in Fig. 14. Namely, the Fermi level on the surfaces of GaAs and InP is usually pinned, irrespective of whether the surface is free or covered with metal. When a small metal Schottky electrode whose size is smaller than the depletion width is placed on such a surface with an "environmental" Fermi level pinning, the potential distribution around the Schottky dot is very much influenced and tends to produce a "saddle" point potential whose height is dependent on the applied voltage, V. In Fig. 14, a 20 nm dot with SBH=0.6 eV was placed on a InP surface with firm Fermi level pinning at $Ec-0.48$ eV, and the potential distribution under a forward bias of 400 mV was calculated using a successive overrelaxation (SOR) method. We clearly see the appearance of a saddle point potential. This is much like the situation that happens in inhomogeneous Schottky diodes discussed by Tung.[19] The appearance and the persistence of this saddle point potential result in highly nonlinear $\log I$ vs V plots with large values of n as in the cases of highly inhomogeneous Schottky diodes.

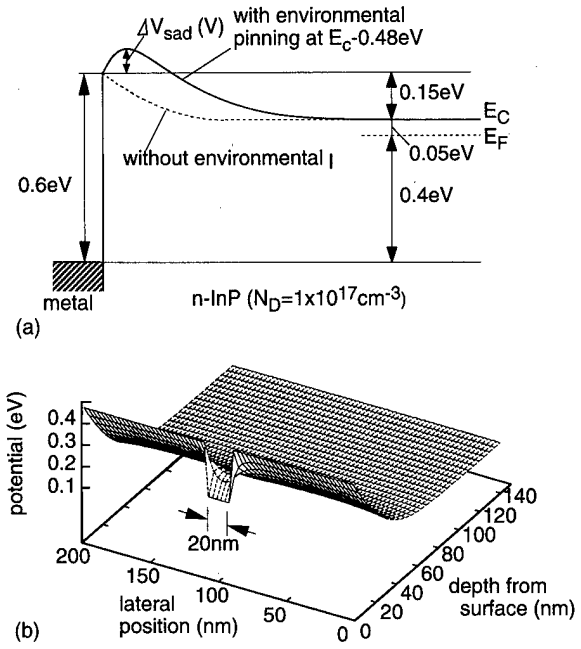


FIG. 14. Potential distribution in a InP single-dot Schottky diode with environmental Fermi level pinning: (a) potential profile underneath the gate and (b) potential distribution around the dot.

V. CONCLUSION

The properties of GaAs and InP Schottky diodes having nanometer-sized metal dots were investigated in an attempt to clarify whether or not strong Fermi level pinning is an intrinsic property of intimate metal–semiconductor contact. Macroscopic Schottky diode samples having many nanometer-sized metal dots as well as single-dot Schottky diode samples having a single nanometer-sized Schottky dot electrode were prepared by a novel *in situ* electrochemical process. The main conclusions are the following.

(1) Strong Fermi level pinning was not seen in the GaAs and InP macroscopic Schottky diode samples. The SBH values were strongly dependent on the metal work function and on the electrochemical processing conditions.

(2) Of particular interest, the difference in the dot size distribution changed the SBH by 340 meV in Pt/InP macroscopic Schottky diode samples, indicating that Fermi level pinning disappears as the dot size is sufficiently reduced.

(3) XPS and Raman measurements indicated that these interfaces are oxide and stress free.

(4) Use of an atomic force microscope with a conductive probe allowed direct $I-V$ measurements on single-dot samples. The metal work function and dot size dependencies of the SBH in these samples were similar to those in macro-

scopic samples. Large ideality factors obtained in single-dot samples were explained in terms of environmental pinning which produces a saddle point potential.

(5) The measured metal work function dependence of the SBH as well as the relationship between the SBH and the ideality factor were both far from the recent predictions based on the metal-induced gap state model.

(6) All the experimental results were explained by the disorder-induced gap state model which asserts that strong Fermi level pinning is an extrinsic property of the metal–semiconductor interface.

ACKNOWLEDGMENTS

The authors would like to thank Dr. H. Okada of the Research Center for Quantum Electronics for performing potential calculations for them. This work was supported in part by a Grant-in Aid for Scientific Research in Priority Areas A (Grant No. 08247101) and by a Grant-in-Aid for Scientific Research B (Grant No. 09450118) from the Ministry of Education, Science, Culture and Sports.

[1]W. E. Spicer, I. Lindau, P. Skeath, and C. Y. Su, J. Vac. Sci. Technol. **16**, 1422 (1979).
[2]V. Heine, Phys. Rev. A **138**, 1689 (1965).
[3]J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[4]H. Hasegawa and H. Ohno, J. Vac. Sci. Technol. B **4**, 1130 (1986).
[5]H. Hasegawa, *Proceedings of the 18th International Conference on the Physics of Semiconductors*, edited by O. Engström (World Scientific, Singapore, 1987), Vol. 1, p. 291.
[6]H. Hasegawa, *Metal–Semiconductor Interfaces*, edited by A. Hiraki (IOS, Tokyo, 1995), p. 280.
[7]J. M. Woodall and J. L. Freeouf, J. Vac. Sci. Technol. **19**, 794 (1981).
[8]N. J. Wu, T. Hashizume, H. Hasegawa, and Y. Amemiya, Jpn. J. Appl. Phys., Part 1 **34**, 936 (1994).
[9]T. Sato, S. Uno, T. Hashizume, and H. Hasegawa, Jpn. J. Appl. Phys., Part 1 **36**, 1811 (1997).
[10]H. Hasegawa, T. Sato, and T. Hashizume, J. Vac. Sci. Technol. B **15**, 1227 (1997).
[11]W. Mönch, Appl. Surf. Sci. **92**, 367 (1996).
[12]T. U. Kampen, R. F. Schmitsdorf, and W. Mönch, Appl. Phys. A: Mater. Sci. Process. **60A**, 391 (1995).
[13]T. Sato, C. Kaneshiro, H. Okada, and H. Hasegawa, Jpn. J. Appl. Phys., Part 1 **38**, 2448 (1999).
[14]C. Kaneshiro, T. Sato, and H. Hasegawa, Jpn. J. Appl. Phys., Part 1 **38**, 1147 (1999).
[15]H. B. Michaelson, J. Appl. Phys. **48**, 4729 (1997).
[16]M. Cardona and N. E. Christensen, Phys. Rev. B **35**, 6182 (1987).
[17]*Properties of Gallium Arsenide*, EMIS data reviews Ser. 2 (INSPEC, London, 1990).
[18]*Properties of Indium Phosphide*, EMIS data reviews Ser. 6 (INSPEC, London, 1991).
[19]R. T. Tung, Phys. Rev. B **45**, 13509 (1992).

Barrier heights of real Schottky contacts explained by metal-induced gap states and lateral inhomogeneities

Winfried Mönch[a)]
Laboratorium für Festkörperphysik, Gerhard-Mercator-Universität Duisburg, D-47049 Duisburg, Germany

(Received 19 January 1999; accepted 3 May 1999)

Most metal–semiconductor contacts are rectifying. For moderately doped semiconductors, the current transport across such Schottky contacts occurs by thermionic emission over the Schottky barrier. The current–voltage characteristics of real Schottky contacts are described by two fitting parameters that are the effective barrier heights Φ_B^{eff} and the ideality factors n. Due to lateral inhomogeneities of the barrier height, both parameters differ from one diode to another. However, their variations are correlated in that Φ_B^{eff} becomes smaller with increasing n. Extrapolations of such Φ_B^{eff}-versus-n plots to the corresponding image-force-controlled ideality factors n_{if} give the barrier heights of laterally homogeneous contacts. They are then compared with the theoretical predictions for ideal Schottky contacts. Data of Si, GaN, GaAs, and CdTe Schottky contacts reveal that the continuum of metal-induced gap states is the fundamental mechanism that determines the barrier heights. However, there are additional but then secondary mechanisms. As an example, contacts with $(7\times7)^i$-reconstructed interfaces have smaller barrier heights than diodes with $(1\times1)^i$-unreconstructed interfaces. This lowering of the Schottky barrier is caused by the electric dipole associated with the stacking fault in one of the triangular halves of the (7×7) unit mesh.
© *1999 American Vacuum Society.* [S0734-211X(99)04604-1]

I. INTRODUCTION

In 1874, i.e., 125 years ago Braun[1] first reported on rectifying metal–semiconductor contacts. Schottky[2] explained this behavior by space-charge layers on the semiconductor side of the interface that are depleted of mobile majority carriers. The most characteristic parameter of such a Schottky contact is its barrier height, i.e., the energy separation between the Fermi level and the edge of the majority-carrier band right at the interface.

A most common method to determine barrier heights of Schottky contacts is the analysis of their current–voltage characteristics. The current transport across metal–semiconductor contacts occurs by thermionic emission over the Schottky barrier for doping levels of the semiconductor up to approximately 10^{18} donors or acceptors per cm^3, while above this limit tunneling through the then narrower depletion layers dominates. Fits of current–voltage (I/V) curves measured with real Schottky contacts to the relation derived for thermionic emission require a second fitting parameter, the ideality factor n, as it is called, in addition to the barrier height.

The effective barrier heights and ideality factors differ from one diode to another even if they were identically prepared. However, these variations are correlated in that the effective barrier heights Φ_B^{eff} become lower with increasing ideality factors n.[3–6] A recent investigation demonstrated that the experimentally observed dependence of the effective barrier heights and the ideality factors of real metal–semiconductor contacts can be explained by lateral inhomogeneities of the barrier height.[7] The dimensions of these patches are in the order of the depletion-layer width.[8–11] The barrier heights of laterally homogeneous contacts may then be obtained by extrapolation of experimental $\Phi_B^{\text{eff}}-n$ relationships to the corresponding image-force-controlled ideality factor n_{if}.[6,7] Depending on the bulk doping level and the interface band-bending n_{if} ranges between 1.01 and 1.03.

However, laterally homogeneous contacts of the same kind but with different interface structures differ in their barrier heights. $NiSi_2$/Si(111) (Ref. 12) and Ag/Si(111) (Ref. 6) contacts are two of the most well-known examples. The two epitaxial $NiSi_2$/Si(111) structures differ in that the lattices are identically aligned or rotated by 180° on both sides of the interface. Their barrier heights deviate by 140 meV. Ag/Si(111) contacts may be prepared with $(1\times1)^i$-unreconstructed and $(7\times7)^i$-reconstructed interfaces. The difference of the barrier heights amounts to 80 meV. These observations support the early conclusion[13] that more than one mechanism has an effect on the barrier heights of Schottky contacts.

On the basis of simple physical arguments, Heine[14] first pointed out that the wave functions of the metal electrons tail into the semiconductor in the energy range where the metal conduction band overlaps the semiconductor band gap. He concluded that these metal-induced gap states (MIGS), as they were called later, represent the fundamental mechanism that determines the barrier heights of Schottky contacts. However, there exist other and then secondary mechanisms that may easily conceal the influence of the MIGS. Meanwhile, it seems to be agreed that most of the real Schottky contacts are laterally inhomogeneous. However, it is not generally conceived that the extrapolation of experimentally observed $\Phi_B^{\text{eff}}-n$ relationships to n_{if} gives the barrier height of the laterally homogeneous contact and that only these barrier heights should be compared with the MIGS model.

a)Electronic mail: w.moench@uni-duisburg.de

The present study will provide no new experimental barrier heights of any Schottky contact but it will rather combine individual sets of effective barrier heights and ideality factors published for many metal–semiconductor contacts. The barrier heights $\Phi_B^{\rm hom}$ of the laterally homogeneous contacts will be determined by extrapolation of the corresponding $\Phi_B^{\rm eff}-n$ relationships to $n_{\rm if}$ and by adding the image-force lowering $\delta\Phi_{\rm if}^0$ to this extrapolated value $\Phi_B^{\rm nif}$. Only these values are then compared with the predictions of the MIGS model. Many of these well-defined data will corroborate the concept that the MIGS are the fundamental mechanisms that determine the barrier heights of Schottky contacts. In addition to this, electric dipole layers induced by specific interface structures will be presented as one of the secondary mechanisms.

II. BARRIER HEIGHTS OF LATERALLY HOMOGENEOUS CONTACTS

A. Effective barrier heights and ideality factors of real Schottky contacts

The thermionic-emission current across intimate, abrupt, and laterally homogeneous Schottky contacts is derived as (see, for example Ref. 15)

$$I_{\rm te}=AA_R^{**}T^2\exp\left(-\frac{\Phi_B^{\rm hom}-\delta\Phi_{\rm if}^0}{k_BT}\right)\times\exp\left(\frac{e_0V_c}{n_{\rm if}k_BT}\right)\left[1-\exp\left(\frac{e_0V_c}{k_BT}\right)\right], \quad (1)$$

where A is the diode area, A_R^{**} is the effective Richardson constant, T is the temperature, and k_B and e_0 are Boltzmann's constant and the electronic charge, respectively. The externally applied bias V_a divides up into a voltage drop V_c across the space-charge region and an IR drop at the series resistance R_s of the diode so that one obtains $V_c=V_a-I_{\rm te}R_s$. The ideality factor $n_{\rm if}$ considers the image-force effect[16] at Schottky barriers and is given by

$$n_{\rm if}=(1-\delta\Phi_{\rm if}^0/4e_0V_i)^{-1}, \quad (2)$$

where e_0V_i is the interface band-bending. The image force lowers the barriers heights of Schottky contacts by

$$\delta\Phi_{\rm if}=e_0\left(\frac{2e_0^3N_{d,a}}{(4\pi)^2\epsilon_\infty^2\epsilon_b\epsilon_0^3}(V_i-V_c)\right)^{1/4}, \quad (3)$$

where $N_{d,a}$ is the bulk donor or acceptor density, ϵ_∞ and ϵ_b are the optical and the static dielectric constant, respectively, of the semiconductor, and ϵ_0 is the permittivity of vacuum. The zero-bias barrier height of the Schottky contact thus is the difference of the barrier height $\Phi_B^{\rm hom}$ and the zero-bias image-force lowering $\delta\Phi_{\rm if}^0$.

Fits of the I/V relation (1) to current–voltage curves measured at identically prepared Schottky contacts yield barrier heights and ideality factors that vary from one diode to another. Remarkably, the ideality factors are generally larger than the image-force-controlled values $n_{\rm if}$ that range between 1.01 and 1.03 depending on the bulk doping level and the interface band-bending. Therefore, the "standard" I/V relation for thermionic emission across intimate and abrupt Schottky contacts is written as

$$I_{\rm te}^{\rm stand}=AA_R^{**}T^2\exp\left(-\frac{\Phi_B^{\rm eff}}{k_BT}\right)\times\exp\left(\frac{e_0V_c}{nk_BT}\right)\left[1-\exp\left(\frac{e_0V_c}{k_BT}\right)\right]. \quad (4)$$

Furthermore, there exists a pronounced correlation between the two fitting parameters in that the effective barrier heights $\Phi_B^{\rm eff}$ get smaller with increasing ideality factors n. When the temperature of a diode is lowered, the effective barrier height of $\Phi_B^{\rm eff}(T)$ of the diode generally decreases and, correlated with this, its ideality factor $n(T)$ becomes larger.

Experimental data that exhibit such $\Phi_B^{\rm eff}-n$ relationships have been published for Al,[17–19] Ag,[6] Ir,[3,20] Au,[21,22] Pb,[23,24] $CoSi_2$,[5] $NiSi_2$,[25] Pd_2Si,[26] and PtSi (Refs. 4, 27, and 28) on Si substrates, for Al,[29] Au,[30] Ti,[31,32] Ni,[33,34] Pd,[35] and Pt (Refs. 36 and 37) on GaAs substrates, for Ag and Pb on 2H–GaN substrates,[38] for Fe,[39] Pd,[40] and Au (Ref. 41) on CdTe substrates, as well as for Au (Ref. 42) on $3C-$, Ni,[43,44] Pd, and Pt (Ref. 45) on $4H-$ and Pd and Pt on $6H-$SiC (Ref. 44) substrates. Figures 1–4 show some of the experimental data. First of all, the $\Phi_B^{\rm eff}-n$ correlations are linear up to $n\approx1.4$. The extrapolation of the experimental data to $n_{\rm if}$ then gives the zero-bias barrier heights $\Phi_B^{\rm nif}$, and by considering Eq. (1), the barrier heights $\Phi_B^{\rm hom}=\Phi_B^{\rm nif}+\delta\Phi_{\rm if}^0$ of the laterally homogeneous contacts. This conclusion will be further justified in the next section.

The two data sets shown in Fig. 1 were obtained with Au contacts on (111) (Ref. 21) and (001)-oriented n-Si (Ref. 22) substrates. The Si wafers used in both investigations were

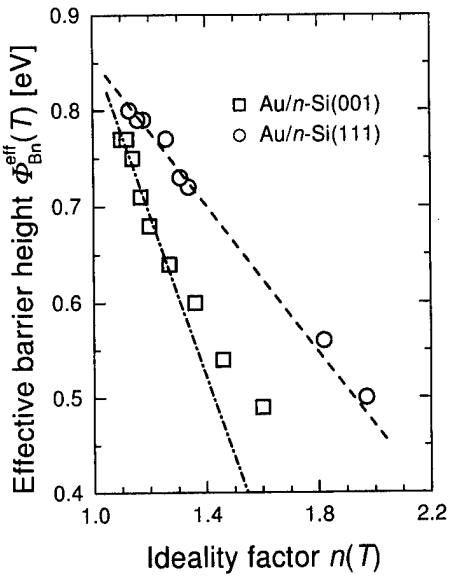


FIG. 1. Effective barrier heights and ideality factors recorded as a function of temperature with one Au/n-Si(001) and Au/n-Si(111) diode each. The dashed and the dash-dotted lines are linear least-squares fits to the data with ideality factors smaller than 1.4. Data are from Refs. 21 and 22, respectively.

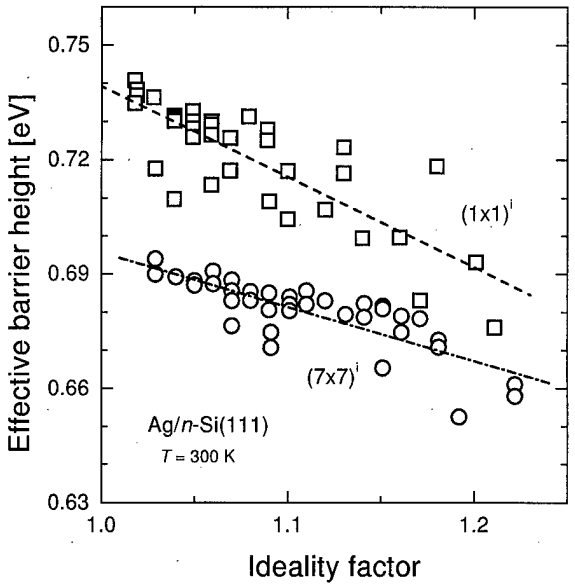


FIG. 2. Effective barrier heights and ideality factors of Ag/n-Si(111) diodes with $(1\times1)^i$-unreconstructed and $(7\times7)^i$-reconstructed interfaces. The dashed and the dash-dotted lines are linear least-squares fits to the data. Data are from Ref. 6.

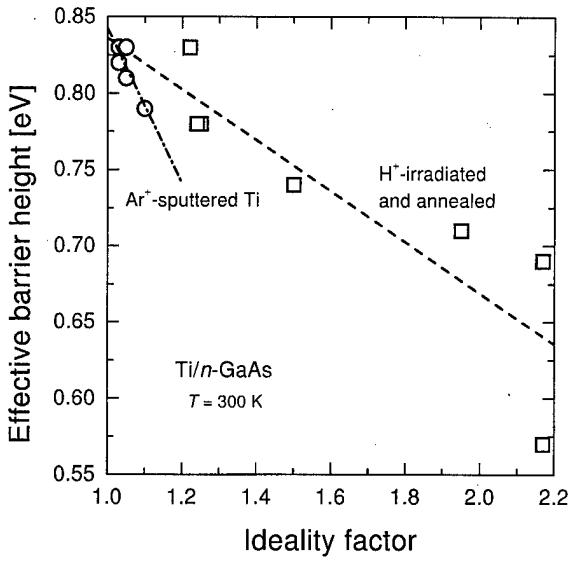


FIG. 4. Effective barrier heights and ideality factors of Ti/n-GaAs diodes irradiated with 150 keV protons and prepared by Ti sputtering with increasing sputtering power, respectively. The dashed and the dash-dotted lines are linear least-squares fits to the data. Data are from Refs. 31 and 32, respectively.

finally dipped into diluted HF and, thereby, became hydrogen-terminated. During the following Au evaporation the hydrogen most probably desorbed so that the interfaces were $(1\times1)^i$-unreconstructed. Otherwise, much lower barrier heights or even Ohmic I/V characteristics should have been observed.[23,46–48] The dashed lines in Fig. 1 are linear least-squares fits to the data points. Their slopes are different but the extrapolations to n_{if} give virtually identical barrier heights Φ_B^{nif} of 0.85 ± 0.07 and 0.86 ± 0.06 eV, respectively. Adding of the corresponding image-force lowering $\delta\Phi_{if}^0$ to these extrapolated values gives the same barrier heights $\Phi_B^{hom}=0.87\pm0.07$ eV of laterally homogeneous Au/Si contacts irrespective of the substrate orientation. An investigation of Ir/n-Si contacts[3] reveals a difference of 0.027 ± 0.014 eV between the barrier heights on (001)- and (111)-oriented substrates that are obtained from the corresponding Φ_{Bn}^{eff}-versus-n plots. These observation indicate that the surface orientation *per se* has, even if at all, only a minor influence on the barrier height.

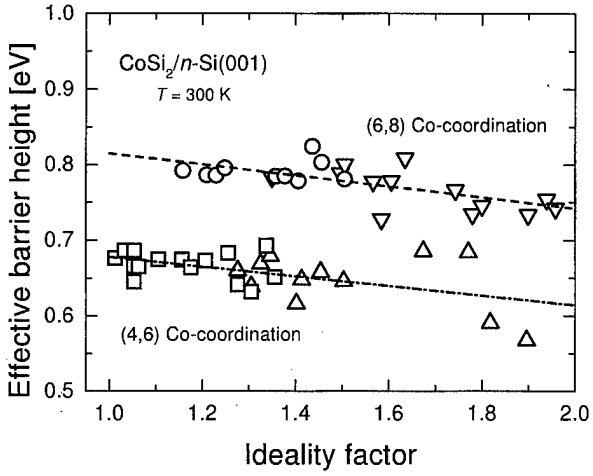


FIG. 3. Effective barrier heights and ideality factors of $CoSi_2/n$-Si(001) diodes with six- and eightfold (type-B) and four- and sixfold coordinated (type-A) Co interface atoms. The dashed and the dash-dotted lines are linear least-squares fits to the data. Data are from Ref. 5.

The interface structure, on the other hand, plays a crucial role as the data displayed in Figs. 2 and 3 demonstrate. Figure 2 shows Φ_{Bn}^{eff}-versus-n plots obtained with two sets of differently prepared Ag/n-Si(111) diodes.[6] They differ in that their interfaces are either $(1\times1)^i$-unreconstructed or have a $(7\times7)^i$ interface structure. The dashed lines are again linear least-squares fits and their extrapolations to n_{if} give barrier heights of the laterally homogeneous contacts that differ by 47 ± 20 meV. The interfaces of the $CoSi_2/n$-Si(001) contacts[5] considered in Fig. 3 also have different interface structures. The larger extrapolated barrier height of 0.82 eV is characteristic of six- and eight-fold coordinated or type-B Co interface atoms while that by 0.14 eV smaller value of 0.68 eV correlates with four- and sixfold or type-A coordination. Consideration of the image-force correction yields barrier heights of 0.83 and 0.69 eV for laterally homogeneous type-A and type-B $CoSi_2/n$-Si(001) contacts, respectively.

Variations of the metal deposition may modify the barrier heights and ideality factors of the Schottky diodes fabricated. Figure 4 displays data of Ti/n-GaAs diodes[31] that were prepared by sputter deposition of Ti. The increase of the sputtering voltage was found to reduce the barrier heights of the diodes and to enlarge their ideality factors simultaneously. In another study,[32] Ti/n-GaAs contacts were irradiated with 150 keV protons and subsequently annealed. The irradiation-induced defects decreased the barrier heights of the contacts and increased their ideality factors. Annealing treatments

caused a partial recovery. In both investigations the Ti/n-GaAs diodes show a linear reduction of their effective barrier heights with increasing ideality factors. The slopes of the linear least-squares fits are different but the extrapolations of the straight lines to n_{if} give virtually the same laterally homogeneous barrier height of 0.87 ± 0.04 eV. The extrapolated barrier heights of the Ti/n-GaAs contacts are obviously the same irrespective of the preparational details. Consideration of the image-force correction yields a barrier height of 0.87 eV for laterally homogeneous Ti/n-GaAs contacts.

B. Laterally inhomogeneous Schottky contacts

In the forward direction, the exponential in the square brackets of Eq. (2) dominates and one obtains

$$I_{te}^{stand} \approx AA_R^{**}T^2 \exp\left(-\frac{\Phi_B^{eff}+(1-1/n)e_0V_c}{k_BT}\right) \times \exp\left(\frac{e_0V_c}{k_BT}\right) \quad \text{for } V_c > 3k_BT/e_0. \tag{5}$$

The ideality factor n or, to be more precise $(1-1/n)$, thus describes the dependence of the barrier height on the applied voltage. Freeouf *et al.*[8] and Bastys *et al.*[9] proposed that ideality factors larger than 1 may be caused by lateral inhomogeneities of the barrier height that are on the length scale of the depletion layer width. At the border between regions of different barrier heights the equipotential lines have to align. Laterally, the transition occurs within approximately one depletion-layer width. Thus, the transition regions of opposite sides will overlap in front of patches with lower barrier heights and with lateral dimensions comparable to or even smaller than the depletion-layer width. Consequently, the potential distribution in front of such patches exhibits a saddle point that depends on the externally applied voltage.

The saddle-point potentials in front of such small-size patches are lower than the barrier height of the surrounding regions. The properties of the patches will thus strongly influence the current transport across laterally inhomogeneous Schottky contacts. Tung[10] derived an analytical expression for the I/V characteristics of laterally inhomogeneous Schottky contacts. Schmitsdorf and co-workers[7,24] applied Tung's as well as the "standard" forward I/V relationship (5) to experimental data of Ag–, Pb–, and Sn/Si(111) contacts. They found that the linear extrapolation of the experimentally observed Φ_B^{eff}-versus-n curves to n_{if} indeed gives the image-force-lowered barrier heights $\Phi_B^{nif}=\Phi_B^{hom}-\delta\Phi_{if}^0$ of the laterally homogeneous contacts.

III. MIGS-AND-ELECTRONEGATIVITY MODEL

A. Chemical trend of barrier heights

As all heteropolar bonds, the interface bonds at metal–semiconductor contacts will be partly ionic. Pauling[49] described the ionicity of covalent single bonds in diatomic molecules by the difference of the electronegativities of the atoms involved. In generalizing this concept, the charge transfer at metal–semiconductor contacts will be characterized by the electronegativity difference X_m-X_s of the metal and the semiconductor in contact.

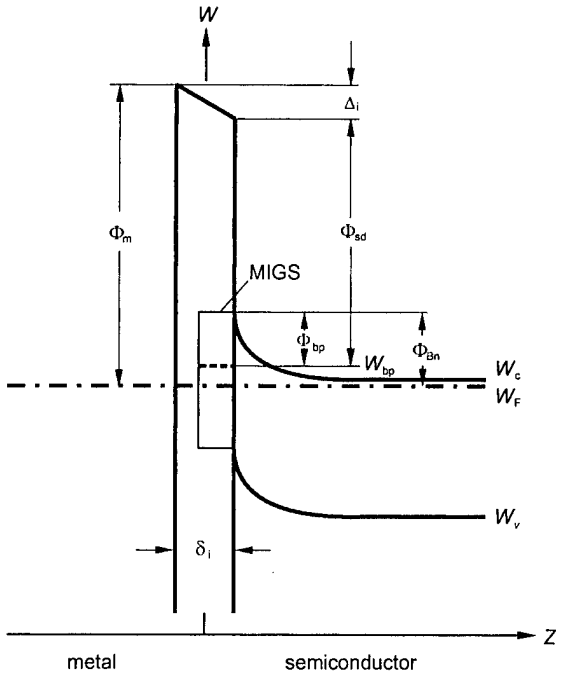


FIG. 5. Energy diagram of a metal–semiconductor contact with a continuum of interface states (schematically).

The charge transfer at metal–semiconductor interfaces may be also described by electronic interface states. The schematic energy diagram of Fig. 5 shows that the conduction-band states of the metal overlap not only the valence band, but in parts, also the band gap of the semiconductor. Below the valence-band top the wave functions are bulk-like on both sides of the contact. In the energy range from the valence-band maximum up to the Fermi level, however, the wave functions of the metal electrons tunnel into the semiconductor, as Heine[14] first pointed out. These wave-function tails derive from the continuum of virtual gap states (ViGS) of the semiconductor. They represent the continuum of intrinsic interface states at metal–semiconductor contacts[50] and they are named metal-induced gap states.[51] The MIGS derive from the nearby bands and they are predominantly donor-like closer to the valence-band maximum W_v but become mostly acceptor-like nearer to the conduction-band minimum W_c. The energy level at which their dominant character changes is called the branch point of the gap states. The net charge in these intrinsic interface states is positive, vanishes, and becomes negative if the branch point W_{bp} is above, coincides with, or drops to below the Fermi level W_F, respectively. In combining the physical concept of metal-induced gap states and the chemical idea of partly ionic covalent interface bonds, Schottky barrier heights $\Phi_{Bn}=W_{ci}-W_F$ are predicted to be smaller or larger than the zero-charge-transfer barrier height $\Phi_{bp}=W_{ci}-W_{bp}$ if the electronegativity difference X_m-X_s is negative and

positive, respectively. Subscripts i indicate properties right at the interface.

Metal–semiconductor interfaces ar electrically neutral in thermal equilibrium. The charge density Q_m on the metal side of the contact is compensated by a charge density Q_s of equal size but opposite sign on the semiconductor side, i.e.,

$$Q_m+Q_s=0. \tag{6}$$

On the semiconductor side, both the intrinsic metal-induced gap states and the space charge of the depletion layer must be considered so that the condition of charge neutrality (6) may be rewritten as[52]

$$Q_m+Q_s=Q_m+Q_{\rm gs}^{\rm mi}+Q_{\rm sc}=0. \tag{7}$$

The space-charge density $Q_{\rm sc}$ may be safely neglected in relation (7) as the following numerical example shows. For silicon with a bulk donor density of 10^{16} cm^{-3} and a barrier height of 0.8 eV, one obtains $Q_{\rm sc}=2.8\times10^{11}\ e_0/{\rm cm}^2$. This value is by a factor of approximately 100 smaller than a charge transfer of only 0.01 e_0 per interface bond at metal–Si(001) interfaces.

The charge densities Q_m and $Q_{\rm gs}^{\rm mi}$ may be described as an electric double layer of thickness δ_i. If the density of states $D_{\rm gs}^{\rm mi}$ of the MIGS around their branch point is assumed to be constant then the net charge density in these intrinsic interface states results as

$$Q_{\rm gs}^{\rm mi}=e_0D_{\rm gs}^{\rm mi}(W_{\rm bp}-W_F)=e_0D_{\rm gs}^{\rm mi}(\Phi_{Bn}-\Phi_{\rm bp}). \tag{8}$$

The voltage drop across this intrinsic interfacial dipole layer may then be written as

$$\Delta_i=(e_0^2/\epsilon_i\epsilon_0)D_{\rm gs}^{\rm mi}\delta_i(\Phi_{Bn}-\Phi_{\rm bp}), \tag{9}$$

where ϵ_i is an appropriate interface dielectric constant. The schematic energy diagram of Fig. 5 gives

$$\Phi_m=\Delta_i+\Phi_{sd}-\Phi_{\rm bp}+\Phi_{Bn}. \tag{10}$$

The work functions $\Phi_m=W_{\rm vac}-W_F$ of metals[53] as well as the dielectric work functions $\Phi_{sd}=W_{\rm vac}-W_{\rm bp}$ of semiconductors[54] vary linearly as a function of the corresponding electronegativities, i.e.,

$$\Phi_{m,sd}=A_XX_{m,s}+B_X. \tag{11}$$

The coefficient A_X amounts to 0.86 eV/Miedema-unit and 1.79 eV/Pauling-unit, respectively. Combining relations (9)–(11) finally yields[54,55]

$$\Phi_{Bn}=\Phi_{\rm bp}+S_X(X_m-X_s) \tag{12}$$

and

$$A_X/S_X-1=(e_0^2/\epsilon_i\epsilon_0)D_{\rm gs}^{\rm mi}\delta_i. \tag{13}$$

The Schottky barrier heights turn out to vary linearly as a function of the electronegativity difference X_m-X_s. The two parameters describing the chemical trend are the zero-charge-transfer barrier height $\Phi_{\rm bp}=W_{ci}-W_{\rm bp}$ and the slope parameter S_X.

Figure 6 displays barrier heights of laterally homogeneous silicon Schottky contacts as a function of the electronegativity difference $X_m-X_{\rm Si}$. Except for the Na data, the $\Phi_{Bn}^{\rm hom}$ values

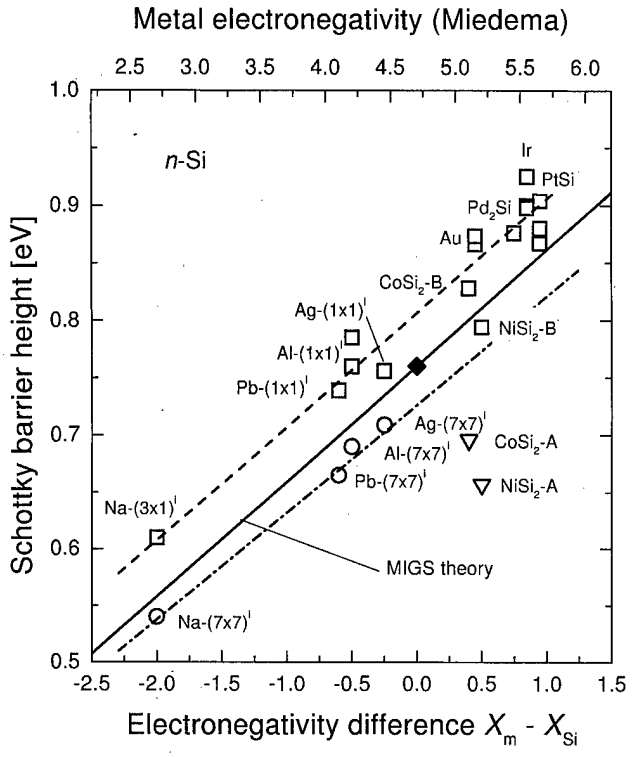


FIG. 6. Barrier heights of laterally homogeneous silicon Schottky contacts as a function of the difference of the metal and the silicon electronegativity. The data points were obtained by extrapolation of $\Phi_{Bn}^{\rm eff}$-vs-n plots to $n_{\rm if}$ and adding of the image-force lowering $\delta\Phi_{\rm if}^0$, the experimental $\Phi_{Bn}^{\rm eff}$ and n values are from Refs. 3–6 and 17–28. The Na data area from Ref. 56. The dashed and the dash-dotted lines are linear least-squares fits to the ○ and □ data, respectively. The MIGS line is drawn with $\Phi_{\rm bp}=0.76$ eV (Ref. 64) and $S_X=0.101$ eV/Miedema-unit (see the text).

shown were obtained by extrapolating $\Phi_{Bn}^{\rm eff}$-versus-n plots to $n_{\rm if}$ and adding the image-force lowering $\delta\Phi_{\rm if}^0$ to the extrapolated values $\Phi_{Bn}^{\rm nif}$. The Na data were determined using x-ray photoemission spectroscopy.[56] Miedema's electronegativities[57] are taken since they were derived from properties of solids. There are three groups of data that are marked as squares, circles, and triangles. The barrier heights represented as squares increase linearly as a function of the electronegativity difference in agreement with relation (12). The dashed line is the linear least-squares fit

$$\Phi_{Bn}^{(1\times1)^i}=0.806+0.099(X_m-X_{\rm Si})\ ({\rm eV}) \tag{14}$$

to these data. This set of data includes Ag and Pb contacts, the interface structure of which is definitely $(1\times1)^i$-unreconstructed as revealed by surface x-ray diffraction.[58–60] The barrier heights of $(7\times7)^i$-reconstructed Pb–, Al–, and Ag/Si(111) interfaces as well as of the epitaxial type-A $NiSi_2$/Si(111) and $CoSi_2$/Si(001) contacts are set off towards lower barrier heights. The dash-dotted line is the linear least-squares fit

$$\Phi_{Bn}^{(7\times7)i}=0.726+0.094(X_m-X_{\rm Si})\ ({\rm eV}) \tag{15}$$

to the data of the four $(7\times7)^i$ interfaces. The zero-charge-transfer barrier height of the $(7\times7)^i$-reconstructed contacts is, thus, by 80 meV lower than the respective value of the

$(1\times1)^i$-unreconstructed interfaces. However, the slope parameters are equal within the limits of experimental error, ±0.012 eV/Miedema-unit.

B. Zero-charge-transfer barrier heights

In the one-dimensional case, the branch point of the ViGS is slightly below the middle of the band gap. Both the density of states and the decay length of the one-dimensional ViGS at their branch point very proportional to the width of the band gap. These results cannot be directly applied to the three-dimensional case. However, we derived simple relations for the branch-point energy[61] and the slope parameter[53] by applying Baldereschi's concept[62] of mean-value points in the Brillouin zone and Penn's isotropic model of semiconductors.[63]

First, we observed that the band gap of a zinc-blende-structure semiconductor at the mean-value point indeed equals Penn's dielectric band gap W_{dg}. It is defined by

$$\epsilon_\infty - 1 = (\hbar\omega_p / W_{dg})^2, \tag{16}$$

where ϵ_∞ and $\hbar\omega_p$ are the optical dielectric constant and the plasmon energy of the bulk valence electrons, respectively, of the semiconductor. Second, we obtained the branch-point energies of binary compound semiconductors as[61]

$$W_{bp} = W_v(\mathbf{k}_{mv}) + 0.449 W_{dg}, \tag{17}$$

where $W_v(\mathbf{k}_{mv})$ is the valence-band energy at the mean-value point $\mathbf{k}_{mv}$ in the Brillouin zone. As in the one-dimensional case, the branch points are slightly below the middle of the average gap. For calibration, we used branch-point energies calculated by Tersoff.[64] The zero-charge-transfer barrier heights then result as

$$\Phi_{bp} = W_g - \{0.449 W_{dg} - [W_v(\Gamma) - W_v(\mathbf{k}_{mv})]_{ETB}\}, \tag{18}$$

where W_g is the fundamental energy gap of the semiconductor and the energy difference $[W_v(\Gamma) - W_v(\mathbf{k}_{mv})]_{ETB}$ is calculated using an empirical tight-binding approach.

C. Slope parameters

Equation (13) relates the slope parameter S_X to the density of states D_{gs}^{mi} of the MIGS and the effective width δ_i/ϵ_i of the corresponding dipole layer. The charge densities Q_m and Q_{gs}^{mi} that constitute the intrinsic MIGS interface dipole reside on the metal and the semiconductor side of the interface, respectively. Therefore, the effective dipole length may be written as[65]

$$\delta_i/\epsilon_i = \delta_m/\epsilon_m + \delta_i/\epsilon_b. \tag{19}$$

The screening lengths δ_m/ϵ_m in the metal and δ_i/ϵ_b in the semiconductor are approximated by the Thomas–Fermi screening length and the decay length $1/2q_{gs}^{mi}$ of the MIGS screened by the bulk dielectric constant ϵ_b of the semiconductor, respectively. A typical value of the Thomas–Fermi screening length is $\delta_m/\epsilon_m = 0.05$ nm, and theoretical calculations yielded decay lengths $1/2q_{gs}^{mi}$ of 0.137 nm for diamond[66] ($\epsilon_b = 5.7$) and 0.3 (Ref. 65) for GaAs ($\epsilon_b = 12.8$) and silicon ($\epsilon_b = 11.9$). With these values relation (19) yields an interface dielectric constant ϵ_i of 2 for diamond, 4 for GaAs, and 4.6 for silicon. It has to be mentioned that Eq. (19) is a very crude approximation in that it combines microscopic lengths and macroscopic dielectric constants.

In the one-dimensional case, both the density of states and the decay length of the ViGS at their branch point are determined by the width of the band gap. Penn's relation (16), on the other hand, relates the dielectric band gap to the optical dielectric constant. Therefore, we searched for a correlation between $D_{gs}^{mi}/2q_{gs}^{mi}$ and $\epsilon_\infty - 1$. For that purpose we used theoretical values of D_{gs}^{mi} and $1/2q_{gs}^{mi}$ that are available for some semiconductors. As a result, we found[54]

$$(e_0^2/\epsilon_0) D_{gs}^{mi}/2q_{gs}^{mi} = 0.29(\epsilon_\infty - 1)^{1.94}. \tag{20}$$

Considering this result and assuming $\epsilon_i \approx 3$ as a reasonable approximation, one obtains

$$A_X/S_X - 1 \approx 0.1(\epsilon_\infty - 1)^2, \tag{21}$$

for binary compound semiconductors.

D. Comparison with experimental data

Experimental barrier heights of laterally homogeneous Schottky were determined by extrapolation of Φ_B^{eff}-versus-n plots for Si, GaN, GaAs, CdTe, and the three SiC polytypes 3C, 4H, and 6H. The experimental Si and GaAs data were extrapolated to the corresponding n_{if} values and the respective image-force lowering $\delta\Phi_{if}^0$ of the barrier heights were calculated using Eq. (3). This procedure could not be applied for GaN, CdTe, and SiC since for these semiconductors the effective densities of states of their conduction bands are not well known. Therefore, the Φ_B^{eff}-versus-n plots of these semiconductors were extrapolated to $n = 1$. This extrapolation to $n = 1$ rather than to the correct n_{if} value underestimates the laterally homogeneous barrier heights by approximately 10 meV. The Si, GaAs, GaN, and CdTe data are compared with the predictions of MIGS-and-electronegativity model in Figs. 6–9, respectively. The zero-charge-transfer barrier heights Φ_{bp} and the slope parameters S_X of the binary compounds were calculated using Eqs. (17) and (21), respectively. For silicon, relation (18) gives a branch-point energy $W_{bp} - W_v(\Gamma) = 0.03$ eV. Tersoff,[64] on the other hand, calculated a value that is by 0.33 eV larger. With $\epsilon_i = 4.6$, the value of the interface dielectric constant derived for silicon in the preceding section, one finds $S_X = 0.101$ eV/Miedema-unit.

For the silicon data represented by squares in Fig. 6, the experimental zero-charge-transfer barrier height is by 46 meV larger than Tersoff's MIGS value, $\Phi_{bp} = 0.76$ eV,[64] but the slope parameter of the MIGS model agrees with the least-squares fit to the experimental data within the limits of experimental error. The experimental results of contacts with $(7\times7)^i$-reconstructed interfaces and of the type-A silicides are not explained by the MIGS model. However, the dash-dotted fitting line also runs parallel to the MIGS line. This finding indicates the existence of a secondary mechanism that is caused by the interfacial $(7\times7)^i$ reconstruction and

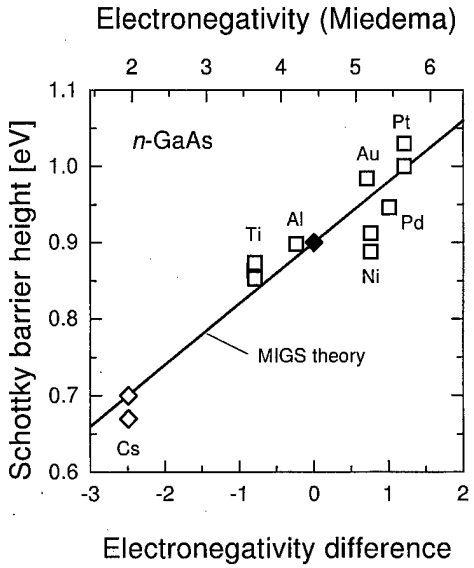


FIG. 7. Barrier heights of laterally homogeneous GaAs Schottky contacts as a function of the difference of the metal and the GaAs electronegativity. The □ data points were obtained by extrapolation of Φ_{Bn}^{eff}-vs-n plots to n_{if} and adding of the image-force lowering $\delta\Phi_{if}^{0}$; the experimental Φ_{Bn}^{eff} and n values are from Refs. 29–37. The data points for Cs were determined by using ultraviolet light photoemission spectroscopy and a Kelvin probe, Refs. 67 and 68. The MIGS line is drawn with $\Phi_{bp}=0.9$ eV (Ref. 61) and $S_X=0.08$ eV/Miedema-unit.

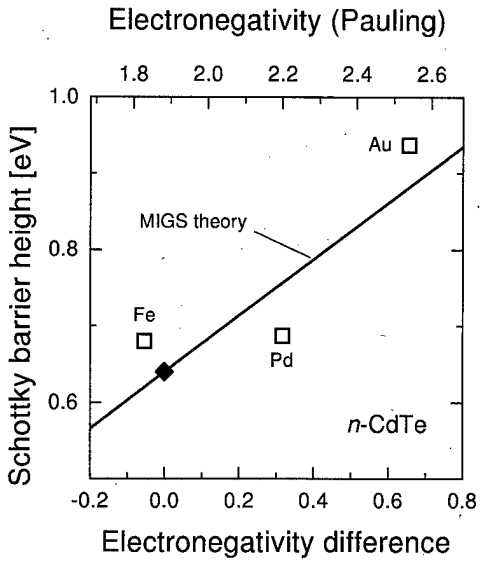


FIG. 9. Barrier heights of laterally homogeneous CdTe Schottky contacts as a function of the difference of the metal and the CdTe electronegativity. The data points were obtained by extrapolation of Φ_{Bn}^{eff}-vs-n plots to $n=1$; the experimental Φ_{Bn}^{eff} and n values are from Refs. 39–41. The MIGS line is drawn with $\Phi_{bp}=0.64$ eV (Ref. 61) and $S_X=0.37$ eV/Pauling-unit.

that accounts for the reduction of the zero-charge-transfer barrier heights by 80 meV. This observation will be explained in the next section.

The barrier heights $\Phi_{Bn}^{hom}=\Phi_{Bn}^{nif}+\delta\Phi_{if}^{0}$ of laterally homogeneous Ti–, Au–, and Pt/GaAs contacts, that are shown in Fig. 7, are excellently reproduced by the MIGS model. The Cs data points, on the other hand, were determined using ultraviolet light photoemission spectroscopy (UPS) (Ref. 67) and an additional Kelvin probe (CPD).[68] They are also on the MIGS line. While Pt shows no evidence of reaction with GaAs below approximately 300 °C, the other two transition metals, Ni and Pd, are known to decompose GaAs even at room temperature.[69] This most probably leads to nonabrupt Ni– and Pd/GaAs interfaces so that the extrapolated barrier heights of Pt– but not of Ni– and Pd/GaAs contacts can be compared with the MIGS model of abrupt interfaces.

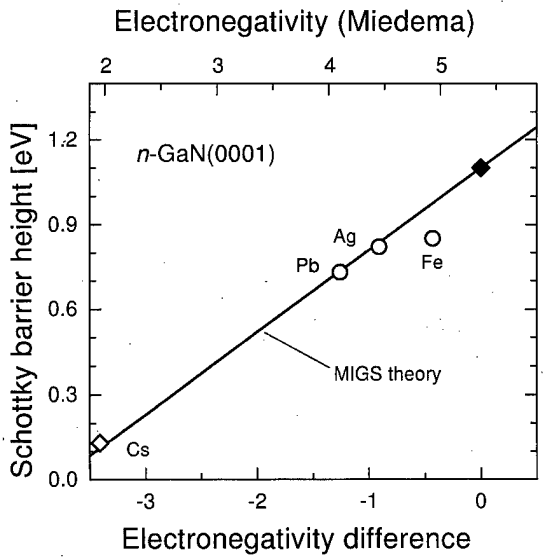


FIG. 8. Barrier heights of laterally homogeneous GaN Schottky contacts as a function of the difference of the metal and the GaAs electronegativity. The data points were obtained by extrapolation of Φ_{Bn}^{eff}-vs-n plots to $n=1$, from Ref. 38. The data point for Cs was determined by using ultraviolet light photoemission spectroscopy and a Kelvin probe, Ref. 69. The MIGS line is drawn with $\Phi_{bp}=1.1$ eV (Ref. 61) and $S_X=0.29$ eV/Miedema-unit.

Figure 8 shows that the MIGs model excellently reproduces the extrapolated barrier heights of Pb– and Ag/2H–GaN contacts. The Cs data point was again determined by UPS-CPD measurements.[70] The Fe data point, on the other hand, was also obtained by extrapolation of a Φ_{Bn}^{eff}-versus-n plot to $n=1$, but the contacts were prepared on oxygen-contaminated GaN surfaces. For this reason, one should not compare this value with the MIGS model since it requires clean interfaces.

As a fourth example, Fig. 9 compares extrapolated barrier heights of CdTe Schottky contacts with the predictions of the MIGS model. Here, Pauling's electronegativities are used since Miedema's scale contains no value for Te. The experimental data deviate by ±25 meV only, i.e., within the limits of experimental error from the predictions of the MIGS model.

IV. DISCUSSIONS AND CONCLUSIONS

The MIGS model assumes "ideal" interfaces. First of all, "ideal" means that the contacts should be intimate, abrupt, and laterally homogeneous. All diodes discussed here are intimate in that they are free of any interlayers such as, for

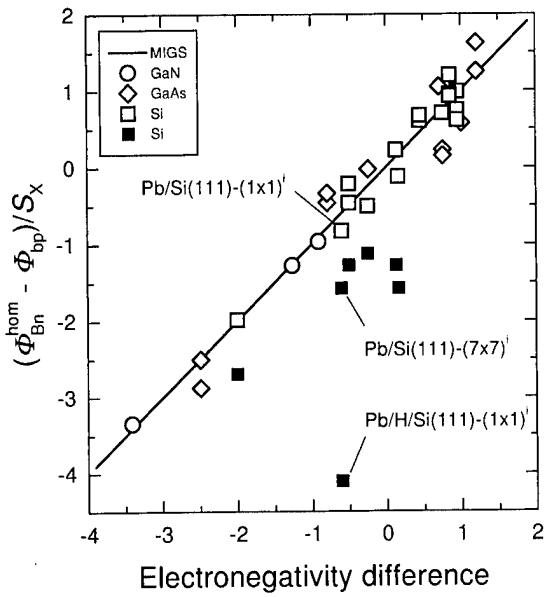


FIG. 10. Differences of laterally homogeneous barrier heights Φ_{Bn}^{hom} of Si, GaAs, and GaN Schottky contacts and the zero-charge-transfer barrier heights Φ_{bp} of these semiconductors normalized by the corresponding slope parameters S_X as a function of the difference of the metal and semiconductor electronegativities. The data are the same as in Figs. 6–9. The barrier height of laterally homogeneous Pb/H/n-Si(111)-$(111)^i$ contacts was taken from Ref. 23.

example, native oxides. For this reason, the Fe/2H–GaN data point will not be considered further. Furthermore, only abrupt interfaces are considered so that reacted interface such as Ni– and Pd/GaAs are excluded. Lateral uniformity of the barrier heights was "obtained" by extrapolation of Φ_B^{eff}-versus-n plots to n_{if}. Figures 6–9 display barrier heights of such Si, GaN, GaAs, and CdTe Schottky contacts separately. However, a rearrangement of relation (12) yields

$$(\Phi_{Bn}^{hom} - \Phi_{bp})/S_X = X_m - X_s . \tag{22}$$

Figure 10 now sums up the experimental barrier heights Φ_{Bn}^{hom} of the intimate, abrupt, and laterally homogeneous Si, GaN, and GaAs Schottky contacts that are shown in Figs. 6–9. The straight line is the prediction of the MIGS-and-electronegativity model. It excellenty describes the experimental data represented by open symbols. The MIGS model thus explains the Schottky barrier heights of these experimentally well-characterized metal–semiconductor interfaces and no other mechanisms have to be invoked. This finding confirms the earlier conclusion that the MIGS are the fundamental mechanism that determines the barrier heights of ideal Schottky contacts while additional and then secondary mechanisms may be effective at many of the real metal–semiconductor interfaces.[13]

The dependence of the effective barrier heights Φ_B^{eff} on the ideality factor n clearly demonstrates the existence of secondary mechanisms in addition to the MIGS. One set of data shown in Fig. 4 was obtained with Ti/n-GaAs contacts prepared by sputter deposition of Ti. The diodes deteriorated with increasing sputtering power. This observation points at fabrication-induced defects. These might be excess anions[71] or antisite defects[72] at interfaces of compound semiconductors and vacancies or self-interstitials at silicon contacts.[73]

The data points depicted by closed squares in Figs. 5 and 10 represent metal–silicon interfaces with interface structures that contain extrinsic interface dipoles in addition to the intrinsic MIGS dipole. A most simple case of extrinsic interface dipoles are a layer of hydrogen atoms at metal–semiconductor interfaces.[23,74] Figure 10 displays the barrier height of laterally homogeneous Pb/H/n-Si(111)-$(1\times)^i$ contacts.[23] It was again obtained by extrapolation of a Φ_{Bn}^{eff}-versus-n plot to n_{if}. The Pb deposition does not remove the hydrogen from the Si(111):H-(1×1) surfaces[48] that were prepared by a HF dip. Clean but otherwise identical Pb/n-Si(111)-$(1\times1)^i$ interfaces have barrier heights that are larger by 0.32 eV. The hydrogen-induced reduction of the barrier height is explained by $H^{-\Delta q}$–$Si^{+\Delta q}$ interface dipoles.[23] The sign of the change may be easily explained. Extrinsic interface dipoles add a charge density Q_{id} to the charge density on the semiconductor side, i.e.,

$$Q_s = Q_{gs}^{mi} + Q_{id} + Q_{sc} . \tag{23}$$

The charge density Q_s will be the same at clean and hydrogen-doped interfaces so that one obtains from Eqs. (8) and (23)

$$e_0 D_{gs}^{mi}(\Phi_{Bn}^{clean} - \Phi_{Bn}^{H}) = Q_{id} . \tag{24}$$

Since hydrogen is more electronegative than silicon, the charge density Q_{id} induced by the chemical H–Si dipoles will have a positive sign. This means $\Phi_{Bn}^{clean} > \Phi_{Bn}^{H}$ in agreement with the experimental observations.

The barrier heights of $(7\times7)^i$-reconstructed Na–, Al–, Ag–, and Pb/Si(111) contacts are by 80 meV smaller that the values of the corresponding $(1\times1)^i$-unreconstructed interfaces [see Figs. 2 and 6 and relations (14) and (15)]. This reduction is attributed to structure-induced interface dipoles. The (7×7) unit mesh contains a stacking fault in one of its triangular halves. The distribution of the valence charge at the stacking faults differs from the undisturbed bulk. The bonds of perfectly ordered silicon are purely covalent and, therefore, the charge distribution at the stacking faults is dipole like in silicon.[75] The stacking fault of the 7×7 reconstruction has its positive and negative charge shifted towards the bulk and the surface or interface, respectively. Analogous to what happens due to the chemical $H^{-\Delta q}$–$Si^{+\Delta q}$ dipoles just discussed, the reconstruction-induced $Si^{-\Delta\Delta q}$–$Si^{+\Delta q}$ dipoles at $(7\times7)^i$-reconstructed metal–semiconductor interfaces will also lower the barrier compared to $(1\times1)^i$-unreconstructed contacts.

The differences of the charge distributions at the type-A and type-B interfaces of the $NiSi_2$– and $CoSi_2$/Si(111) contacts cannot be easily distinguished. Density-functional calculations within the local-density approximation and using the linear-muffin-tin-orbital method in the atom-spheres approximation have been performed for the two types of $NiSi_2$/Si(111) contacts.[76–78] In agreement with Pauling's electronegativity concept the calculations revealed electric interface dipoles with the positive charge on the silicon and

the negative charge on the $NiSi_2$ side. The strength of the interface dipoles turned out to be such that the type-B interfaces have larger barrier heights than the type-A contacts. This finding again reproduces the experimental results even if the values of the barrier heights are sensitive to details of the atomic arrangements.

The preceding discussions clearly demonstrate that the continuum of the metal-induced gap states indeed explains the barrier heights experimentally observed with intimate, abrupt, defect-free, and laterally homogeneous metal–semiconductor contacts. The barrier heights of laterally homogeneous contacts may be obtained by extrapolation of Φ_{Bn}^{eff}-versus-n plots to n_{if} and adding the image-force lowering $\delta\Phi_{\mathrm{if}}^{0}$. Defect-free does not only mean the absence of foreign atoms but also of "displaced" substrate atoms even if they constitute a definite interface reconstruction. Most instructive examples of this kind are silicon Schottky contacts with $(1\times1)^i$-unreconstructed and $(7\times7)^i$-reconstructed interfaces. Reconstructions are accompanied by distributions of the valence charge that differ from the undisturbed bulk. This then gives rise to extra interface dipoles that will alter the barrier heights of reconstructed interfaces in comparison to $(1\times1)^i$-unreconstructed interfaces, the barrier heights of which are solely determined by the continuum of the metal-induced gap states.

[1]F. Braun, Pogg. Ann. Phys. **153**, 556 (1874).
[2]W. Schottky, Naturwissenschaften **26**, 843 (1938).
[3]I. Ohdomari, T. S. Kuan, and K. N. Tu, J. Appl. Phys. **50**, 7020 (1979).
[4]V. W. L. Chin, J. W. V. Storey, and M. A. Green, Solid-State Electron. **32**, 475 (1989).
[5]P. Werner, W. Jäger, and A. Schüppen, J. Appl. Phys. **74**, 3846 (1993).
[6]R. F. Schmitsdorf, T. U. Kampen, and W. Mönch, Surf. Sci. **324**, 249 (1995).
[7]R. F. Schmitsdorf, T. U. Kampen, and W. Mönch, J. Vac. Sci. Technol. B **15**, 1221 (1997).
[8]J. L. Freeouf, T. N. Jackson, S. E. Laux, and J. M. Woodall, J. Vac. Sci. Technol. **21**, 570 (1982); Appl. Phys. Lett. **40**, 634 (1982).
[9]A. I. Bastys, V. B. Bikbaev, J. J. Vaitkus, and S. C. Karpinskas, Litov. Fiz. Sb. **28**, 191 (1998).
[10]R. T. Tung, Phys. Rev. B **45**, 13509 (1992).
[11]U. Rau, H. H. Güttler, and J. H. Werner, Mater. Res. Soc. Symp. Proc. **260**, 245 (1992).
[12]R. Tung, Phys. Rev. Lett. **52**, 461 (1984).
[13]W. Mönch, Phys. Rev. Lett. **58**, 1260 (1987).
[14]V. Heine, Phys. Rev. **138**, A1689 (1965).
[15]E. H. Rhoderick and R. H. Williams, *Metal–Semiconductor Contacts*, 2nd ed. (Clarendon, Oxford, 1988).
[16]W. Schottky, Phys. Z. **15**, 872 (1914).
[17]I. Landmann and D. Jäger, Thesis, Universität Münster (1975) (unpublished).
[18]D. Jäger and R. Kassing, J. Appl. Phys. **48**, 4413 (1977).
[19]Y. Miura, F. Fujieda, and K. Hirose, Phys. Rev. B **50**, 4893 (1994).
[20]O. Nur, M. R. Sardela, Jr., M. Willander, and R. Turan, Semicond. Sci. Technol. **10**, 551 (1991).
[21]T. P. Chen, T. C. Lee, C. C. Ling, C. D. Beling, and S. Fung, Solid-State Electron. **36**, 949 (1993).
[22]K. Maeda and E. Kitahara, Appl. Surf. Sci. **130**, 925 (1998).
[23]T. U. Kampen and W. Mönch, Surf. Sci. **331**, 490 (1995).
[24]R. F. Schmitsdorf and W. Mönch, Eur. Phys. J. B **7**, 457 (1999).
[25]R. T. Tung, J. Vac. Sci. Technol. B **2**, 465 (1984).
[26]S. Chand and J. Kumar, Appl. Phys. A: Mater. Sci. Process. **63A**, 171 (1996).
[27]P. G. McCafferty, A. Sellai, P. Dawson, and E. Elabd, Solid-State Electron. **39**, 583 (1996).
[28]J. H. Werner and H. H. Güttler, J. Appl. Phys. **69**, 1522 (1991).
[29]A. S. Bhuiyan, A. Martinez, and D. Esteve, Thin Solid Films **161**, 93 (1988).
[30]S. Hardikar, M. K. Hudait, P. Modak, S. B. Krupanidhi, and N. Padha (unpublished).
[31]M. Di Dio, A. Cola, M. G. Lupo, and L. Vasanelli, Solid-State Electron. **38**, 1923 (1995).
[32]S. Arulkumaran, J. Arokiaraj, N. Dharmarasu, and J. Kumar, Nucl. Instrum. Methods Phys. Res. B **119**, 519 (1996).
[33]R. Hackham and P. Hardrop, IEEE Trans. Electron Devices **ED-19**, 1231 (1972).
[34]M. Nathan, Z. Soshani, G. Ashkinazi, B. Meyler, and O. Zolotarevski, Solid-State Electron. **39**, 1457 (1996).
[35]N. Dharmarsu, S. Arulkumaran, R. R. Sumathi, P. Jayavel, J. Kumar, P. Magudapathy, and K. G. M. Nair, Nucl. Instrum. Methods Phys. Res. B **140**, 119 (1998).
[36]W. O. Barnard, G. Myburg, F. D. Auret, S. A. Goodman, and W. E. Meyer, J. Electron. Mater. **25**, 1695 (1996).
[37]H.-W. Hübers and H. P. Röser, J. Appl. Phys. **84**, 5326 (1998).
[38]T. U. Kampen and W. Mönch, Appl. Surf. Sci. **117**, 388 (1997).
[39]A. B. McLean, I. M. Dharmadasa, and R. H. Williams, Semicond. Sci. Technol. **1**, 137 (1986).
[40]I. M. Dharmadasa, A. B. McLean, M. H. Patterson, and R. H. Williams, Semicond. Sci. Technol. **2**, 404 (1987).
[41]D. Sands and C. G. Scott, J. Appl. Phys. **77**, 3295 (1995).
[42]G. Constantinidis, J. Kuzmic, K. Michelakis, and K. Tsagaraki, Solid-State Electron. **42**, 253 (1998).
[43]K. J. Schoen, J. M. Woodall, J. A. Cooper, Jr., and M. R. Melloch, IEEE Trans. Electron Devices **ED-45**, 1595 (1998).
[44]V. Khemka, T. P. Chow, and R. J. Gutmann, J. Electron. Mater. **27**, 1128 (1998).
[45]H.-J. Im, B. Kaczer, J. P. Peiz, and W. J. Choyke, Appl. Phys. Lett. **72**, 839 (1989); (private communication).
[46]M. Tsunoda, K. Fukutani, Y. Murata, H. Yamashita, K. Komaki, and K. Kobayashi, in *Formation of Semiconductor Interfaces*, edited by B. Lengler, H. Lüth, W. Mönch, and J. Pollmann (World Scientific, Singapore, 1994), p. 162.
[47]T. U. Kampen, R. F. Schmitsdorf, and W. Mönch, Appl. Phys. A: Mater. Sci. Process. **60A**, 391 (1995).
[48]K. Fukutani, H. Iwai, H. Yamashita, Y. Murata, S. Hatori, and K. Kobayashi, Surf. Sci. **377**, 1010 (1997).
[49]L. N. Pauling, *The Nature of the Chemical Bond*, 2nd ed. (Cornell University Press, Ithaca, NY, 1960).
[50]E. Louis, F. Yndurain, and F. Flores, Phys. Rev. B **13**, 4408 (1976).
[51]S. G. Louie and M. L. Cohen, Phys. Rev. B **13**, 2461 (1976).
[52]J. Bardeen, Phys. Rev. **71**, 717 (1947).
[53]W. Gordy and W. J. O. Thomas, Phys. Rev. **24**, 439 (1956).
[54]W. Mönch, Appl. Surf. Sci. **92**, 367 (1996).
[55]A. M. Cowley and S. M. Sze, J. Appl. Phys. **36**, 3212 (1965).
[56]J. J. Paggel, G. Neuhold, H. Haak, and K. Horn, Surf. Sci. **414**, 221 (1998).
[57]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. Pettifor (North-Holland, Amsterdam, 1988), p. 699.
[58]R. D. Aburano, H. Hong, J. M. Roesler, K. Chung, D.-S. Lin, P. Zschack, H. Chen, and T.-C. Chiang, Phys. Rev. B **52**, 1839 (1995), and articles cited therein.
[59]C. A. Lucas and D. Loretto, Surf. Sci. **344**, L1219 (1995).
[60]P. B. Howes, K. A. Edwards, D. J. Hughes, J. E. Macdonald, T. Hibma, T. Bootsma, and M. A. James, Phys. Rev. B **51**, 17740 (1995).
[61]W. Mönch, J. Appl. Phys. **80**, 5076 (1996).
[62]A. Baldereschi, Phys. Rev. B **7**, 5212 (1973).
[63]D. R. Penn, Phys. Rev. **128**, 2093 (1962).
[64]J. Tersoff, J. Vac. Sci. Technol. B **4**, 1066 (1986), and references cited therein.
[65]S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B **15**, 2154 (1977).
[66]J. Ihm, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **40**, 1208 (1978).
[67]W. E. Spicer, P. E. Gregory, P. W. Chye, J. A. Babaola, and T. Sukegawa, Appl. Phys. Lett. **27**, 617 (1975).
[68]F. Grunwald, Diploma thesis, Universität Duisburg (1987) (unpublished).
[69]T. Sands, V. G. Keramidas, A. J. Yu, K.-M. Yu, R. Gronsky, and J. Washburn, J. Mater. Res. **2**, 262 (1987).

[70]M. Eyckeler, W. Mönch, T. U. Kampen, R. Dimitrov, O. Ambacher, and M. Stutzmann, J. Vac. Sci. Technol. B **16**, 2224 (1998).
[71]J. L. Freeouf and J. M. Woodall, Appl. Phys. Lett. **39**, 727 (1981).
[72]E. R. Weber, H. Ennen, U. Kaufmann, J. Windscheid, and J. Schneider, J. Appl. Phys. **53**, 6140 (1982).
[73]B. G. Svensson, M. O. Aboelfotoh, and J. L. Lindström, Phys. Rev. Lett. **66**, 3028 (1991).
[74]W. Mönch, Europhys. Lett. **27**, 479 (1994).
[75]M. Y. Chou, M. L. Cohen, and S. G. Louie, Phys. Rev. B **32**, 7979 (1985).
[76]G. P. Das, P. Blöchl, O. K. Anderson, N. E. Christensen, and O. Gunnarsson, Phys. Rev. Lett. **63**, 1168 (1989).
[77]H. Fujitani and S. Asano, Phys. Rev. B **42**, 1696 (1990).
[78]S. Ossicini, O. Bisi, and C. M. Bertoni, Phys. Rev. B **42**, 5735 (1990).

In situ electrical determination of reaction kinetics and interface properties at molecular beam epitaxy grown metal/semiconductor interfaces

L. C. Chen[a] and C. J. Palmstrøm
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

(Received 22 March 1999; accepted 7 May 1999)

In situ patterning and electrical characterization are used to study the electrical properties, reaction kinetics, and interface properties during reactions at Ni/*n*-GaAs interfaces. Ni contacts were formed *in situ* by deposition through a removable molybdenum shadow mask onto a GaAs(100) $c(4\times4)$ As-rich surface. Annealing at 300 °C resulted in Ni_3GaAs formation. Subsequent exposure of the Ni_3GaAs to an As_4 flux at 350 °C resulted in the formation of NiAs at the surface and the epitaxial regrowth of GaAs at the Ni_xGaAs/GaAs interface. The GaAs regrowth thickness and the regrown GaAs electrical properties were determined electrically by *in situ* capacitance–voltage and current–voltage measurements. An interlayer model was applied to explain the Schottky barrier height discrepancy between capacitance–voltage and current–voltage measurements for metal/GaAs contacts with regrown GaAs. The model predicts that the regrown GaAs interlayer is *p* type with a hole concentration of $\sim3\times10^{15}$ holes/cm^3. © *1999 American Vacuum Society.*
[S0734-211X(99)04504-7]

I. INTRODUCTION

The combination of molecular beam epitaxy (MBE) growth, *in situ* metal/semiconductor reactions, and *in situ* electrical measurements enable the reaction kinetics and electrical properties of metal/semiconductor to be determined and the interlayer effects to be studied. Most electrical studies of metal/semiconductor reactions have involved metal deposition on chemically etched GaAs surfaces followed by *ex situ* reactions and electrical measurements.[1–5] These have concentrated on changes in the Schottky barrier height as a result of annealing and reactions without controlling the phase sequence. Only a limited number of studies has involved *in situ* ultrahigh vacuum (UHV) deposition on atomically clean surfaces and *ex situ* electrical characterization.[6–10] In the case of Ni and Pd, the reaction kinetics has been determined electrically for deposition on *ex situ* chemically etched GaAs.[11] Caldwell *et al.*[12,13] have demonstrated that the Ni/GaAs reaction sequence and phase formation can be controlled *in situ* by the exposure to Ga or As_4 fluxes during annealing. They demonstrate that the exposure to As_4 or Ga fluxes results in GaAs regrowth at the reacted metal/semiconductor interface. The electrical properties of the contacts will depend upon the properties of the regrown GaAs. Interlayers have been used between a metal contact and a semiconductor to alter the transport properties.[14,15] The interlayer could be a semiconductor with different electrical properties,[16–18] such as band gap,[17,18] or a different type of doping.[14] Recently, limited *in situ* electrical characterization and reaction kinetics studies have been reported by Chen *et al.*[19,20] for the As_4-flux induced solid phase GaAs regrowth reaction. In this article, a similar reaction sequence is used and a detailed study of the electrical properties and reaction kinetics is presented together with an interlayer model, which is used to explain the Schottky barrier height discrepancy between capacitance–voltage ($C-V$) and current–voltage ($I-V$) measurements for samples with regrown GaAs.

In order to perform *in situ* $I-V$ and $C-V$ measurements during the different reaction steps, individual contacts were formed *in situ* on MBE grown GaAs using a shadow mask deposition process. By incorporating an n^+ marker layer in the MBE grown *n*-type GaAs buffer layer, a doping density spike will be generated in the $C-V$ doping depth profile. Hence, the distance between the metal/*n*-GaAs interface and the marker layer can be determined electrically.[19,20] By tracing this marker layer position with respect to the metal/*n*-GaAs interface, the penetration depth of metal/*n*-GaAs interface can be monitored *in situ*.

For Ni, the metal/GaAs reaction starts at an annealing temperature of less than 200 °C[21] with the formation of Ni_xGaAs [reaction (1)]

$$x\mathrm{Ni}+\mathrm{GaAs}\rightarrow \mathrm{Ni}_x\mathrm{GaAs}. \tag{1}$$

The composition of Ni_xGaAs was reported to range from $x=2$ to $x=4$.[21–24] Similar annealing conditions corresponding to those used in this study were found to yield $x\approx3$.[12,13] Subsequent *in situ* exposure of the Ni_xGaAs/GaAs structure to arsenic results in the formation of NiAs and GaAs [reaction (2)]

$$x\mathrm{As}+\mathrm{Ni}_x\mathrm{GaAS}\rightarrow \mathrm{GaAs}+x\mathrm{NiAs}. \tag{2}$$

The GaAs formed through reaction (2) is reported to regrow epitaxially on the unreacted-GaAs substrate[12,13] in a manner similar to the solid phase regrowth observed in Si/Ni/GaAs,[25] Al/Ni/GaAs,[26] and Si/Pd/GaAs[27] systems.

[a]Author to whom correspondence should be addressed; electronic mail: chenx102@tc.umn.edu

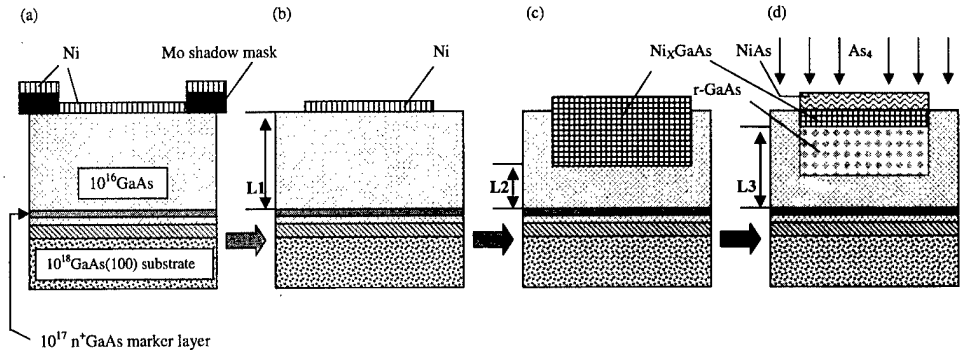


FIG. 1. Sequential reaction procedure: (a) depositing Ni through a Mo shadow mask; (b) after mask removal; (c) after reaction (1) induced by 300 °C annealing; (d) after regrowth reaction (2) induced by As_4 exposure at 350 °C.

II. EXPERIMENTAL PROCEDURE

Some n^+-doped GaAs (100) substrates were indium bonded to molybdenum sample holders and loaded into a modified VG Semicon V80H MBE system with a base pressure of $\leq 5\times10^{-11}$ mbar equipped with conventional effusion cells and an arsenic valved cracker. The indium served as the bonding agent as well as the back ohmic contact to the GaAs wafer. For each substrate, after the native oxide was removed by heating to 640 °C under As_4 flux, a Si doped 0.2-μm-thick $n^{++}(10^{18}/cm^3)$ GaAs buffer layer was grown at 585 °C, followed by sequential layers of $n(2\times10^{16}/cm^3, 0.10\,\mu m)$, $n^+(10^{17}/cm^3, 0.10\,\mu m$, marker layer), and $n(2\times10^{16}/cm^3, 0.6\,\mu m)$ doping. After GaAs growth, the sample was cooled to 350 °C before the As_4 flux was shut off. This resulted in a $c(4\times4)$ As-rich surface reconstruction. In order to minimize unwanted interface reactions during Ni deposition, the samples were held in the preparation chamber under UHV $\leq 5\times10^{-11}$ mbar for more than 4 h in order to cool to room temperature. The samples were transported in UHV to an attached e-beam deposition chamber that is equipped with two electrostatically focused e-gun sources. Ni (~1000 Å) was deposited through a removable molybdenum shadow mask that was attached in front of the GaAs substrate. This produced an array of Ni contacts with ten different sizes ranging from ~500 to ~1400 μm in diameter. After mask removal, the samples were transported to an attached UHV electrical characterization chamber. With the help of the optical microscope, the metal dots on the sample were aligned to, and probed by, an electrochemically etched tungsten wire probe to allow *in situ* electrical characterization using $I-V$ and $C-V$ measurements. After electrical characterization, the samples were transported back to the MBE growth chamber for subsequent reactions.

The patterned Ni/GaAs samples were annealed at 300 °C to induce reaction (1) [Fig. 1(c)], transported to the electrical characterization chamber for electrical measurements and then transported back into the growth chamber for further reactions. The samples were left in the preparation chamber for 4 h after each reaction in order to cool to room temperature prior to electrical characterization. Annealing at 350 °C under an As_4 flux for the desired exposure time was used to drive the regrowth reaction [reaction (2), Fig. 1(d)]. An effective As_4 flux (~7.2×10^{14} atoms/cm^2 s) was determined from the As_4-induced reflection high-energy electron diffraction (RHEED) oscillations for GaAs growth which corresponded to 1.14 bilayer/s GaAs growth.

III. RESULTS AND DISCUSSION

A. Schottky barrier height and ideality factor

In situ forward biased $I-V$ measurements of the same contact after different reaction stages are shown in Fig. 2. Similar results were found for all contacts. The averages of the Schottky barrier heights and ideality factors are depicted in Fig. 3. The $C-V$ Schottky barrier heights ($\phi_{bn,C-V}$) were determined from plots of $1/C^2$ vs V with image force lowering correction and the $I-V$ Schottky barrier heights ($\phi_{bn,I-V}$) were determined from the saturation current density.[15] The high ideality factor of the as-deposited Ni/GaAs contacts may result from either nonuniform interfacial

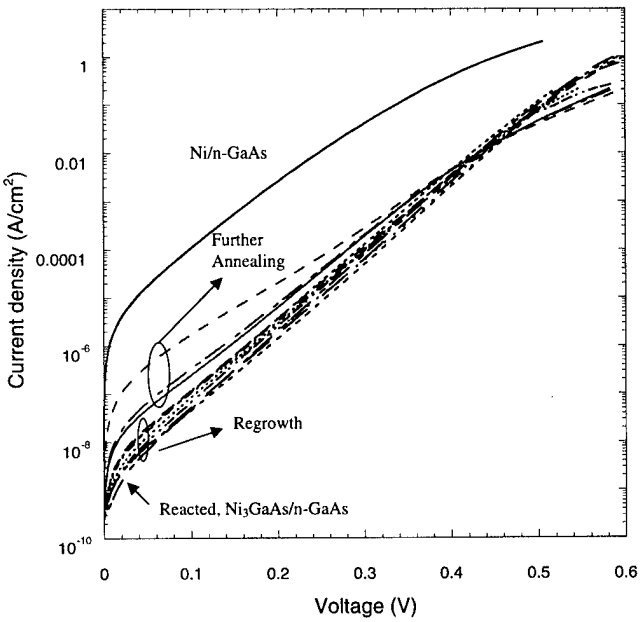


FIG. 2. *In situ* current–voltage characterization of the same Ni/GaAs contact for as-deposited, annealed at 300 °C, exposure to As_4 at 350 °C for different times.

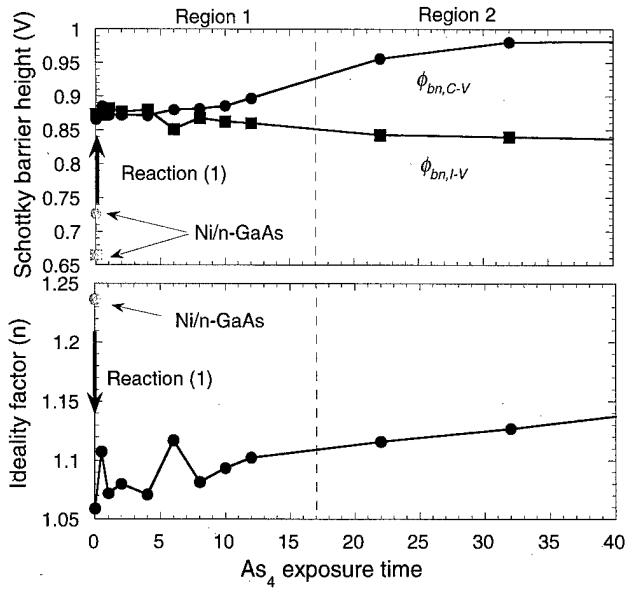


FIG. 3. Summary of averaged Schottky barrier height and ideality factor obtained from *in situ* $C-V$ and $I-V$ measurements.

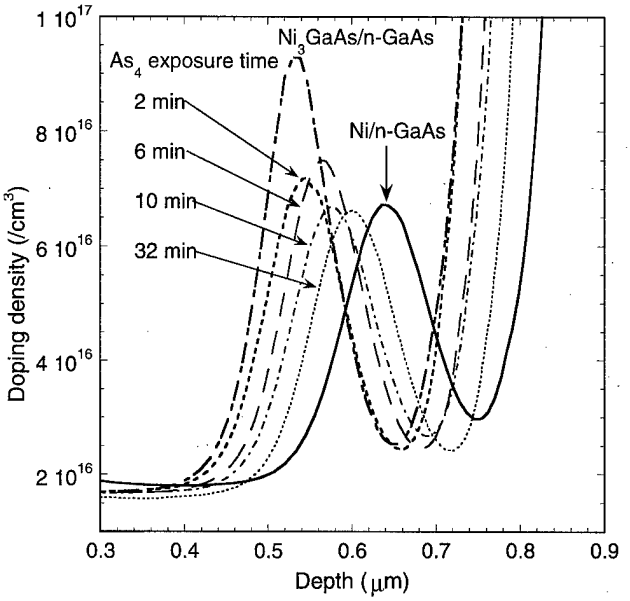


FIG. 4. *In situ* $C-V$ doping depth profiles from Ni/GaAs contacts of as-deposited, reacted at 300 °C, and subsequent As_4 exposure at 350 °C for different times. The depletion depth was obtained from Eq. (3).

reaction during contact formation[28,29] or possibly defects in the near surface region of the GaAs induced during Ni deposition. For the NiAs regrown GaAs, the low ideality factor of $n \leqslant 1.11$ indicates that conduction is primarily dominated by thermionic emission[30] and that the electrical quality of the regrown GaAs is high.

The Schottky barrier height of Ni on MBE grown $c(4\times4)$ GaAs(100) was determined to be $\phi_{bn,I-V}=0.67$ V and $\phi_{bn,C-V}=0.73$ V, from $I-V$ and $C-V$ measurements, respectively. When Ni was deposited on substrates which had not been cooled completely to room temperature (estimated substrate temperature ~100–150 °C) a significantly higher barrier height was observed ($\phi_{bn,I-V}=0.77$ V and $\phi_{bn,C-V}=0.82$ V). As the barrier height increases with Ni_xGaAs formation, the increased barrier height observed is consistent with the onset of interfacial reactions during the deposition. This may explain the discrepancy in barrier height between the as-deposited Ni/n-GaAs(100) ($\phi_{bn,I-V}=0.67$ V and $\phi_{bn,C-V}=0.73$ V) reported here and the Ni deposited on the thermally cleaned GaAs (100) ($\phi_{bn,I-V}=0.77$ V and $\phi_{bn,C-V}=0.91$ V).[6,7]

When the Ni/GaAs structure is annealed at 300 °C, Ni_xGaAs forms and the Schottky barrier height increases ($\phi_{bn,I-V}=0.87$ V and $\phi_{bn,C-V}=0.87$ V). This value is similar to that reported for *ex situ* annealing ($\phi_{bn,I-V}=0.83$ V[31] and $\phi_{bn,C-V}=0.82$ V[11]). After the As_4-induced regrowth reaction at 350 °C, the barrier height from $I-V$ measurements remains approximately constant for different As_4 exposure times ($\phi_{bn,I-V}\approx0.85$ V), whereas the barrier height from $C-V$ measurements increased. As shown in Fig. 3, the Schottky barrier height discrepancy between $I-V$ and $C-V$ measurements increased with As_4 exposure time. A possible explanation is an increase in the inhomogeneity of the Schottky barrier height in the contacts.[28,29] However, the low ideality factor, $n \leqslant 1.11$, suggests a relatively homogeneous contact. An alternative explanation could result from the properties and increased thickness of the regrown GaAs, which is discussed below.

B. Doping depth profiles

Figure 4 shows the doping depth profiles obtained from $C-V$ measurements using the conventional assumptions for depletion width, w [Eq. (3)], and doping concentration[14,30] for the same sample after Ni deposition, reaction, and different As_4 exposure times. With these assumptions, the depletion width is given by

$$w=\frac{\epsilon_s}{C}, \tag{3}$$

where ϵ_s is the semiconductor dielectric constant and C the measured capacitance per unit area. The depth of the doping spike in the doping depth profile represents the distance between the marker layer and metal/GaAs interface. The profile for the as-deposited Ni/GaAs structure shows the n^+ marker layer at depth of 0.64 μm beneath the metal/GaAs interface [L1 in Fig. 1(b)]. Upon annealing at 300 °C without As_4 exposure, the marker layer shifts to 0.534 μm [L2 in Fig. 1(c)]. This corresponds to 1070 Å of GaAs consumption (the difference between L1 and L2 in Fig. 1). When subsequent annealing is performed with the As_4 flux, the marker layer moves deeper and closer to its original position, indicating that the distance between the marker layer and metal/GaAs interface increases with As_4 exposure time. Figure 5 shows the GaAs regrowth thickness obtained from the measurements of the n^+ marker layer position for different As_4 exposure times and from previous Rutherford backscattering spectrometry (RBS) results on other samples.[12,13] The solid data points came from doping depth profiles which used Eq.

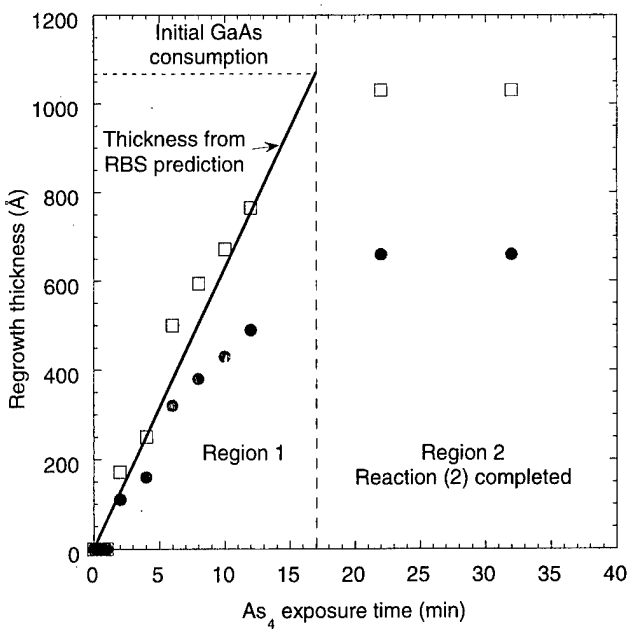


FIG. 5. GaAs regrowth thickness obtained from *in situ* $C-V$ doping depth profiles (data points) and *ex situ* RBS results (solid line) (see Refs. 12 and 13). The data points, ● and □, represent the regrown GaAs thickness calculated from Eqs. (3) and (4), respectively.

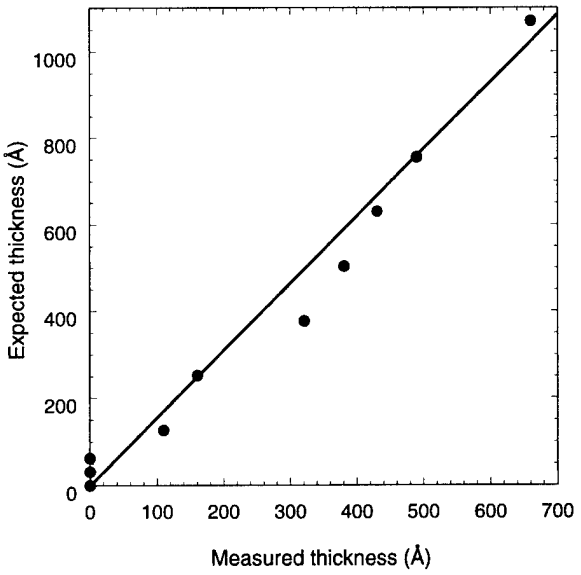


FIG. 6. Linear dependence of t on t_m shows the interlayer dielectric constant of $\epsilon_{\text{regrown}} \sim 20.5\epsilon_0$.

(3) for determining the depletion width. There are two distinct regions in the GaAs regrowth. It appears that no GaAs regrowth occurs for annealing with an As_4 flux for time $\leqslant 60$ s. This apparent lack of regrowth may result from the regrown layer being thinner than depth resolution of the $C-V$ depth profiling technique (Debye length ~300 Å) or from no actual regrowth. The latter possibility could occur if Ni outdiffused from the Ni_3GaAs to form NiAs on the surface, leaving behind a Ni_xGaAs ($x<3$) phase without inducing GaAs regrowth. Evidence for a change in Ni concentration during growth comes from the observed change in Ni_xGaAs lattice parameters with As_4 exposure.[12,13]

Longer As_4 exposure times result in an increase in the regrown layer thickness. For exposure times $\geqslant 20$ min (region 2), no more GaAs regrowth on the GaAs substrate occurs, indicating that the regrowth reaction has come to completion. This suggests that the GaAs regrowth reaction [reaction (2)] stops after 15–20 min of As_4 exposure time. The data obtained using Eq. (3) (solid dots in Fig. 5) indicate that 410 Å GaAs does not regrow on GaAs substrate. This may result from the metallurgical interface not being the same as the electrical interface, GaAs clusters, or Ga dissolution in the NiAs, or from the method used to obtain the doping depth profile from the $C-V$ measurements. As Ni is believed to be the dominant diffusing species at this temperature,[12,13,21,22] it is difficult to envisage a reaction mechanism that results in GaAs formation in the NiAs. The RBS data and Auger depth profiles from this sample show no Ga incorporation in the NiAs. In addition, our previous RBS and transmission electron microscopy (TEM) results[12,13] suggest that complete regrowth takes place. It is interesting to note that if the GaAs regrowth thickness were extrapolated from the RBS results, the complete GaAs regrowth would occur after ~17 min of As_4 exposure (Fig. 5), which is comparable to that determined from the $C-V$ depth profiles. Therefore, the 410 Å thickness discrepancy may arise from the method used to obtain depth profiles from $C-V$ measurements. Hence, the electrical properties of the regrown GaAs and their effects on the electrical measurements must be determined. If the regrown GaAs has a different dielectric constant from the unreacted GaAs, Eq. (3) is no longer valid. As the regrown GaAs and unreacted GaAs will act as series capacitors, the capacitance per unit area will be given by[17]

$$\frac{1}{C}=\frac{t}{\epsilon_{\text{regrown}}}+\frac{w-t}{\epsilon_s}, \tag{4}$$

where t is the regrown-GaAs layer thickness, $\epsilon_{\text{regrown}}$ the dielectric constant of regrown GaAs, and w the depletion width. The relationship between the measured depletion width ($w_m=\epsilon_s/C$) and measured regrown layer thickness (t_m) obtained using Eq. (3) and the true depletion width (w) and regrown layer thickness (t) obtained from Eq. (4) is given by

$$w_m-w=t_m-t=\left(\frac{\epsilon_s}{\epsilon_{\text{regrown}}}-1\right)t \tag{5}$$

and

$$t=\frac{\epsilon_{\text{regrown}}}{\epsilon_s}t_m. \tag{6}$$

By assuming that all the GaAs regrows on the GaAs substrate, the regrowth depends on the As_4 arrival, and the regrowth reaction completes in ~17 min, the true regrown-GaAs thickness can be predicted. The dielectric constant of the regrown GaAs can then be determined from the slope of t vs t_m plot (Fig. 6). A slope of 1.56 suggests a 56% increase in the dielectric constant of the regrown GaAs as compared

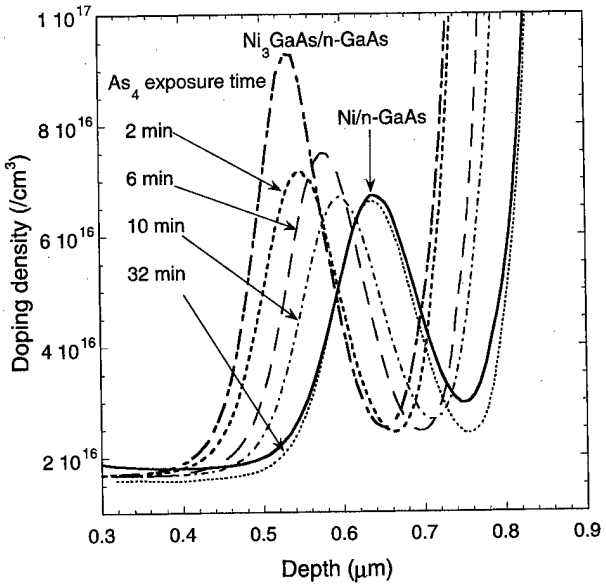


FIG. 7. Doping depth profiles obtained from $C-V$ measurements. The depletion depth for samples with regrown GaAs was calculated using Eq. (4), the dielectric constant modified depletion width.

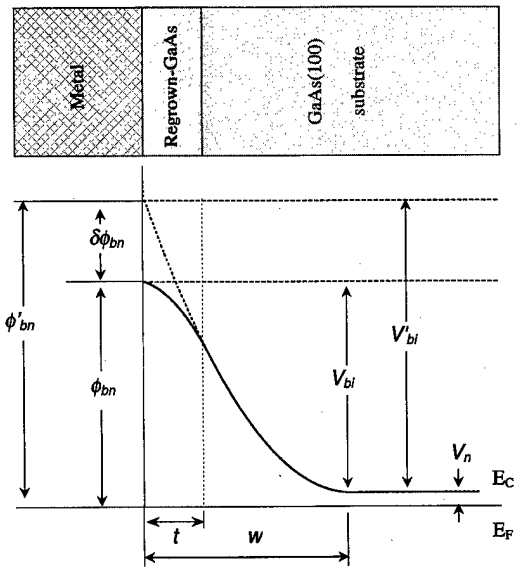


FIG. 8. Simulation of band bending with (solid line) and without (dashed line) interlayer ($N_D=2\times10^{16}$/ cm^3, $N_t=-3\times10^{16}$/cm^3 (p type), and $t=1000$ Å).

to GaAs, i.e., $\epsilon_{\text{regrown}}=20.5\epsilon_0$, which would result in the underestimation of the regrown GaAs thickness using Eq. (3). The recalculated doping depth profiles (Fig. 7) show that the n^+ marker layer has shifted back to its original position after 32 min of As_4 exposure, indicating that the GaAs is fully regrown on the substrate. The open square data points in Fig. 5 show the dielectric constant modified regrown-GaAs thickness with As_4 exposure.

In this solid phase regrowth mechanism, one may expect significant Ni incorporation, which is a deep acceptor in GaAs, and point and structural defects in the regrown GaAs. These would be expected to increase the leakage current and ideality factor due to trap-assisted tunneling through the barrier. The low ideality factor, $n\sim1.11$, for the regrown contacts suggests that this is not the case. This is particularly surprising as our previous TEM results indicated stacking faults and precipitates in the regrown GaAs, which were assumed to be NiAs.[12,13] Although the precipitates do not appear to influence the transport properties significantly, such precipitates may be the cause of the proposed high dielectric constant of the regrown GaAs. Unfortunately, the low doping concentration and thin layer thickness of the regrown GaAs made direct doping density determination from $C-V$ measurements impossible as it was always completely depleted.

Figure 8 shows schematic sketches of a metal/regrown-GaAs (interlayer)/n-GaAs structure and its band diagram. The doping and the dielectric constant of the regrown GaAs are different from the n-GaAs. In order to verify this model, the expected barrier height with interlayer thickness t must be determined by solving Poisson's equation with the appropriate boundary conditions. In addition to the usual boundary conditions,[14,30] the doping density and dielectric constants are assumed to be constant, N_t, $\epsilon_{\text{regrown}}$ in the interlayer and N_D, ϵ_s in the remaining semiconductor. Solving Poisson's equation gives the electric field at the metal/semiconductor interface, E_m:

$$-E_m=\frac{qN_D}{\epsilon_s}(w-t)+\frac{qN_t}{\epsilon_{\text{regrown}}}t \tag{7}$$

the built-in potential, V_{bi}:

$$V_{\text{bi}}=\frac{qN_D}{2\epsilon_s}w^2+\left(\frac{qN_t}{2\epsilon_{\text{regrown}}}-\frac{qN_D}{2\epsilon_s}\right)t^2 \tag{8}$$

for $0\leq x\leq t$, $V(x)$:

$$V(x)=-E_m x-\frac{qN_t}{2\epsilon_{\text{regrown}}}x^2 \tag{9}$$

and for $t\leq x\leq w$, $V(x)$:

$$V(x)=\left(-E_m+\frac{qN_D}{\epsilon_s}t-\frac{qN_t}{\epsilon_{\text{regrown}}}t\right)x-\frac{qN_D}{2\epsilon_s}x^2 + \frac{qN_t}{2\epsilon_{\text{regrown}}}t^2-\frac{qN_D}{2\epsilon_s}t^2, \tag{10}$$

where $V(x)$, q, t, and w are electric potential, electron charge, interlayer thickness, and depletion width, respectively.

If there is no interlayer, i.e., $t=0$, the equations reduce to the usual equations for a uniform Schottky barrier without an interlayer.[14,30] However, if the interlayer is heavily p doped, it can be shown that the energy band profile would reach a maximum at $x=\xi$, where

$$\xi=\frac{-E_m\epsilon_{\text{regrown}}}{qN_t}=\frac{N_D\epsilon_{\text{regrown}}}{N_t\epsilon_s}(w-t)+t \tag{11}$$

and $0\leq\xi<t$. This results in a Schottky barrier height given by

$$\phi_{bn}=V_{bi}+V_n+\frac{kT}{q}+E_m\xi+\frac{qN_t}{2\epsilon_{regrown}}\xi^2 \quad (12)$$

and the extrapolated Schottky barrier height obtained from extrapolating the band bending at w is given by

$$\phi_{bn}=V'_{bi}+V_n+\frac{kT}{q}, \quad (13)$$

where

$$V'_{bi}=\frac{qN_D}{2\epsilon_s}w^2. \quad (14)$$

This extrapolated barrier height is identical to that for a uniform Schottky contact to N_D doped GaAs with a depletion width, w, neglecting image force lowering. Therefore, the Schottky barrier height discrepancy between $I-V$ and $C-V$ measurements can be described as $\delta\phi_{bn}=\phi'_{bn}-\phi_{bn}$ where ϕ'_{bn} and ϕ_{bn} do not include an image force lowering term. Figure 8 shows the simulated band bending from the equations above. It is clear that the thicker the interlayer, the larger Schottky barrier height discrepancy between $C-V$ and $I-V$ measurements. Since Ni is a deep level acceptor in GaAs,[14] it is reasonable to assume that the interlayer will not be heavily p doped, so that the potential maximum will be close to the metal/semiconductor interface (i.e., $\xi\sim 0$) and the Schottky barrier height discrepancy can be simplified to

$$\delta\phi_{bn}=\frac{qN_D}{2\epsilon_s}\left(1-\frac{N_t\epsilon_s}{N_D\epsilon_{regrown}}\right)t^2. \quad (15)$$

This equation predicts that the Schottky barrier height discrepancy would be linearly dependent on the square of the regrown GaAs thickness. As inhomogeneous barrier heights will also result in barrier height discrepancy between $I-V$ and $C-V$ measurements, the intrinsic Schottky barrier heights (ϕ_{bn}), which take into account inhomogeneous barrier heights and image force lowering, were determined from the $I-V$ measurements using the procedure in Refs. 32 and 33. Figure 9 shows a plot of $\delta\phi_{bn}$ vs t^2. Since there is no inhomogeneous correction for the $C-V$ measurements, the data in Fig. 9 were fitted to

$$\delta\phi_{bn}=\frac{qN_D}{2\epsilon_s}\left(1-\frac{N_t\epsilon_s}{N_D\epsilon_{regrown}}\right)t^2+\delta\phi'_{bn}, \quad (16)$$

where $\delta\phi'_{bn}$ is to account for any barrier height inhomogeneity in the $C-V$ measurement. The fit (solid line) gives the regrown GaAs as p type with doping density of $\sim 3\times 10^{15}/\text{cm}^3$, which is consistent with Ni being a deep acceptor in GaAs and $\delta\phi'_{bn}\sim 0.005$ V. Hence, the barrier height discrepancy between $I-V$ and $C-V$ measurements can be explained by the regrown GaAs being p type. This procedure enabled the electrical properties of the regrown GaAs to be inferred without direct measurement.

IV. CONCLUSION

Combining MBE, *in situ* patterning and electrical measurements provide a powerful approach for studying metal/semiconductor interfaces and determining interlayer properties. $C-V$ and $I-V$ measurements showed that the Ni/GaAs Schottky barrier height increased after Ni_xGaAs formation [reaction (1)], then decreased during GaAs regrowth and NiAs formation [reaction (2)]. Doping depth profiles showed that the metal/semiconductor interface moved in after reaction (1) then retreated after reaction (2), confirming the regrowth of GaAs. The GaAs regrowth rate obtained from the doping depth profiles is in excellent agreement with previous RBS results indicating As_4 arrival rate limited mechanism during the initial regrowth stage. The discrepancy of Schottky barrier height measured from $C-V$ and $I-V$ measurements with regrowth layer thickness is consistent with the regrown GaAs being p type with a doping density of $\sim 3\times 10^{15}/\text{cm}^3$.

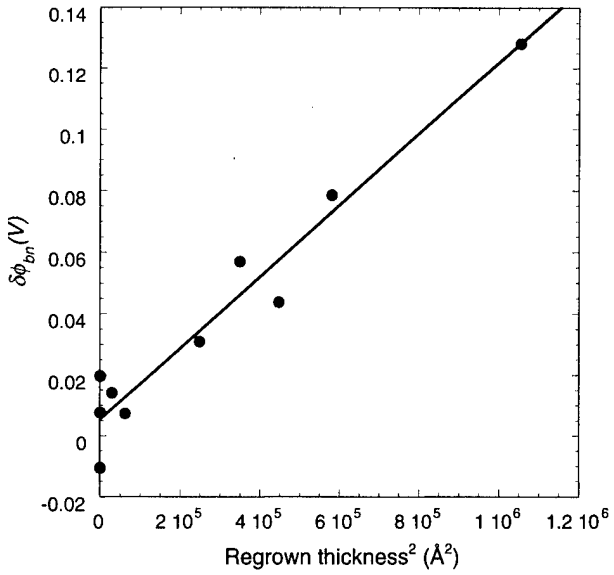


FIG. 9. Dependence of Schottky barrier height discrepancy ($\delta\phi_{bn}$) on the square of the interlayer thickness (t^2). The dots are data points obtained from *in situ* electrical measurements and the line is linear curve fit to Eq. (16), indicating a p-type interlayer with $\sim 3\times 10^{15}/\text{cm}^3$ doping and $\delta\phi_{bn}=0.005$ V.

ACKNOWLEDGMENTS

Special thanks to Professor J. R. Chelikowsky and M. I. Nathan for valuable discussions. This research was supported in part by the U.S. Air Force Office of Scientific Research, Contract Nos. AF/F49620-95-1-0360 and AF/F49620-97-1-0473.

[1] A. Piotrowska, A. Guivarc'h, and G. Pelous, Solid-State Electron. **26**, 179 (1983).
[2] C. J. Palmstrøm and D. V. Morgan, in *Gallium Arsenide Materials, Devices, and Circuits*, edited by M. J. Howes and D. V. Morgan (Wiley, Chichester, 1985), p. 195.
[3] T. Sands, V. G. Keramidas, A. J. Yu, R. Gronsky, and J. Washburn, J. Mater. Res. **2**, 262 (1987).
[4] T. Sands, Mater. Sci. Eng., B **1**, 289 (1989).
[5] A. Guivarc'h, J. Caulet, M. Minier, M. C. Le Clanche, S. Deputier, and R. Guérin, J. Appl. Phys. **75**, 5061 (1994).
[6] J. R. Waldrop, Appl. Phys. Lett. **44**, 1002 (1984).
[7] J. R. Waldrop, J. Vac. Sci. Technol. B **2**, 445 (1984).

[8]N. Newman, W. E. Spicer, T. Kendelewicz, and I. Lindau, J. Vac. Sci. Technol. B **4**, 931 (1986).
[9]N. Newman, M. Van Schilfgaarde, T. Kendelwicz, M. D. Williams, and W. E. Spicer, Phys. Rev. B **33**, 1146 (1986).
[10]A. B. Mclean and R. H. Williams, J. Phys. C **21**, 783 (1988).
[11]H. F. Chuang, C. P. Lee, and D. C. Liu, J. Electron. Mater. **24**, 767 (1995).
[12]D. A. Caldwell, L. C. Chen, A. H. Bensaoula, J. K. Farrer, C. B. Carter, and C. J. Palmstrøm, J. Vac. Sci. Technol. B **16**, 2280 (1998).
[13]D. A. Caldwell, L. C. Chen, A. H. Bensaoula, J. K. Farrer, C. B. Carter, and C. J. Palmstrøm, Mater. Res. Soc. Symp. Proc. **514**, 455 (1998).
[14]D. Redfield, Appl. Phys. Lett. **35**, 182 (1979).
[15]S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1985).
[16]M. O. Manasreh, D. C. Look, K. R. Evans, and C. E. Stuts, Phys. Rev. B **41**, 10272 (1990).
[17]T. J. Miller and M. I. Nathan, Appl. Phys. Lett. **61**, 2332 (1992).
[18]T. H. Lim, T. J. Miller, F. Williamson, and M. I. Nathan, Appl. Phys. Lett. **69**, 1599 (1996).
[19]L. C. Chen, D. A. Caldwell, T. A. Müller, T. G. Finstad, W. Schildgen, and C. J. Palmstrøm, J. Cryst. Growth **201/202**, 146 (1999).
[20]L. C. Chen, D. A. Caldwell, T. G. Finstad, and C. J. Palmstrøm, J. Vac. Sci. Technol. A **17**, 1307 (1999).
[21]S. H. Chen, C. B. Carter, and C. J. Palmstrøm, J. Mater. Res. **3**, 1385 (1988).
[22]S. H. Chen, C. B. Carter, C. J. Palmstrøm, and T. Ohashi, Mater. Res. Soc. Symp. Proc. **54**, 361 (1986).
[23]T. Sands, V. G. Keramidas, J. Washburn, and R. Gronsky, Appl. Phys. Lett. **48**, 402 (1986).
[24]A. Lahav, M. Eizenberg, and Y. Komem, J. Appl. Phys. **60**, 991 (1986).
[25]T. Sands, E. D. Marshall, and L. C. Wang, J. Mater. Res. **3**, 914 (1988).
[26]T. Sands, J. P. Harbison, W. K. Chan, S. A. Schwarz, C. C. Chang, C. J. Palmstrøm, and V. G. Keramidas, Appl. Phys. Lett. **52**, 1216 (1988).
[27]L. C. Wang, B. Zhang, F. Fang, E. D. Marshall, S. S. Lau, T. Sands, and T. Kuech, J. Mater. Res. **3**, 922 (1988).
[28]R. T. Tung, A. F. J. Levi, J. P. Sullivan, and F. Schrey, Phys. Rev. Lett. **66**, 72 (1991).
[29]R. T. Tung, Appl. Phys. Lett. **58**, 2821 (1991).
[30]E. H. Rhoderick and R. H. Williams, in *Metal-Semiconductor Contacts* (Clarendon, Oxford, 1988), p. 90.
[31]T. G. Finstad and J. S. Johannessen, Proceedings of the 10th Nordic Semiconductor Meeting, Elsinore, Denmark, 9–11 June, 1982, p. I4:1.
[32]R. F. Schmitsdrof, T. U. Kampen, and W. Mönch, Surf. Sci. **324**, 249 (1995).
[33]R. F. Schmitsdrof, T. U. Kampen, and W. Mönch, J. Vac. Sci. Technol. B **15**, 1221 (1997).

Photoemission spectromicroscopy studies on epitaxial lateral overgrowth GaN surfaces

Y. Yang
Physics Department, Montana State University, Bozeman, Montana 59717

S. Mishra and F. Cerrina
Department of Electrical and Computer Engineering, University of Wisconsin–Madison, Madison, Wisconsin 53706

S. H. Xu, H. Cruguel, and G. J. Lapeyre[a)]
Physics Department, Montana State University, Bozeman, Montana 59717

J. F. Schetzina
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695

(Received 28 April 1999; accepted 10 May 1999)

Photoemission spectromicroscopy is employed to investigate the inhomogeneities of surface electronic structures of epitaxial lateral overgrowth GaN material. The image, acquired on a clean surface, shows the surface morphology and agrees with the atomic force microscopy image. The dominant contrast mechanism is attributed to the angular dependence of the quantum yield for regions at different angles. Energy distribution curves localized to a submicron region for the Ga *3d* core level demonstrate that growth-front areas have different Fermi level pinning behavior compared with window areas and overgrowth regions. The sample exposed to atomic hydrogen shows the same Fermi level position for all areas of the surface. Photoemission spectromicroscopy reveals island formation when about 10 monolayers of Mg is deposited on the surface. © *1999 American Vacuum Society.* [S0734-211X(99)04404-2]

I. INTRODUCTION

Wide-band-gap III-V nitrides have attracted much attention because of their scientific significance and for their potential for many practical applications.[1] Among them, gallium nitride (GaN) is the most interesting material because of its suitable direct band gap of 3.4 eV at room temperature, notable chemical inertness, and great physical hardness. These attractive properties make it ideal for fabricating electronic and electro-optic devices. Considerable effort is underway to develop high power electronics. The optical applications include devices operated near the short wavelength end of the visible range [i.e., blue and ultraviolet light emitting diodes (LEDs), detectors, and laser devices].[2,3] However, the development of III-V nitride materials and devices has suffered from the lack of availability of low-dislocation-density, lattice-matched native nitride substrates for device synthesis using metalorganic vapor phase epitaxy (MOVPE) and molecular beam epitaxy (MBE) methods. As a consequence, growth of GaN on mismatched substrates such as sapphire or SiC produces a columnar-like material consisting of many small hexagonal-like grains.[4] When the latter materials are prepared by MOVPE they show high dislocation densities of 10^9-10^{10} per cm^2. Recently, there have been demonstrations of defect reduction in the growth of GaN layers on sapphire and SiC using an epitaxial lateral overgrowth (ELO) technique.[5] The technique uses GaN/sapphire (or SiC) layers patterned with SiO_2 as a substrate, which produces stripes ($\sim$10 μm wide) of GaN and a remarkable reduction in the dislocation density to about 10^4 per cm^2 or less. Laser diodes with very long continuous wave (cw) lifetimes have been fabricated from these ELO GaN materials.[6]

In this work, the synchrotron-radiation-based photoemission spectromicroscopy technique is employed to study the inhomogeneities of the surface electronic structure of MOVPE-grown ELO GaN materials. Photoemission spectromicroscopy is a combination of classic photoemission techniques and microscopy. The key point in the method is to reduce the size of the light spot on the sample to submicron size while keeping enough photon flux to obtain a good signal-to-noise ratio. In the past several years this technique has been highly developed at high-brightness third-generation synchrotron radiation sources. These experiments used the MAXIMUM microscope in the Advanced Light Source (ALS) in Berkeley.[7] The technique is capable of acquiring chemical and electronic information from a tiny local area so that all investigation of inhomogeneities of sample surfaces can be performed. The measured size of the beam spot is 0.1 μm. Due to a patterned substrate, ELO-grown GaN materials have three different areas, i.e., the window areas between masks (homoepitaxial growth), the areas over the SiO_2 stripes (lateral overgrowth) and the boundary regions where the growth fronts meet and coalesce. The GaN does not wet the SiO_2 film. To characterize these materials, photoemission spectromicroscopy with μm or sub-μm spatial resolution is a suitable technique for mapping these surfaces to obtain their electronic structures. In this article, the experiments have been done on clean surfaces, surfaces exposed to atomic hydrogen, and surfaces covered by about 10 monolayers (ML) of Mg. This article is organized as follows: In Sec. II are details of the experiment. In Sec. III we describe the results and their interpretations. The conclusions are given in Sec. IV.

[a)]Electronic mail: uphgl@gemini.oscs.montana.edu

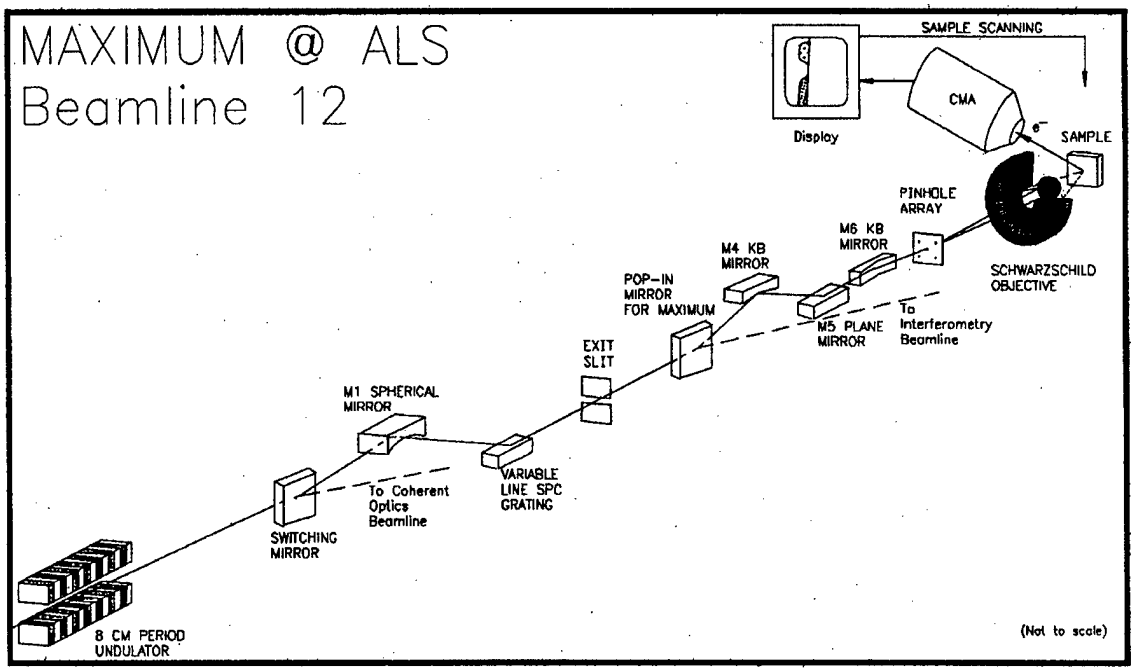


FIG. 1. Schematic layout of the MAXIMUM photoemission microscope and Beamline 12 at the ALS.

II. EXPERIMENTAL DETAILS

The ELO GaN films used in the experiments were grown on the basal plane of sapphire by MOVPE at North Carolina State University. The first step was to grow a conventional low-temperature (~500 °C) buffer layer of GaN on (0001) sapphire followed by the deposition of 1–2 μm of GaN at high temperature (~1050 °C). The GaN film contained the typical $\sim 10^9-10^{10}$ dislocations per cm^2. The next step was to make the SiO_2 pattern by covering the entire wafer with ~100 nm of SiO_2. Then the window stripes in SiO_2 were obtained by the conventional photolithography method where etching of the SiO_2 exposes parallel stripes to clean GaN separated by stripes of SiO_2. These stripes were oriented along a $\{1\bar{1}00\}$ GaN crystal direction. Because of the 30° rotation of the GaN epitaxy on basal-plane sapphire, the GaN stripe length corresponds to the direction of a line drawn from the center of the sapphire wafer that is perpendicular to the $\{11\bar{2}0\}$ sapphire flat. The window and SiO_2 stripe widths are 3 and 15 μm, respectively. The resultant substrate was used for final growth of a film of about 7 μm. Homoepitaxy occurs in the window areas with vertical growth. When the slots are filled the film also grows laterally.

The surface morphology of ELO GaN films was characterized by the atomic force microscopy (AFM) technique. The AFM image was taken in contact mode with a TopoMetrix Explore spectroscopic phase modulated scanning probe microscopy (SPM) system.

The photoemission spectromicroscopy investigations were conducted with MAXIMUM on the undulator Beamline 12 at ALS, Berkeley.[8] Figure 1 is a schematic layout of the MAXIMUM microscope. The monochromatic beam is focused onto a 2 μm pinhole, which is demagnified with a Schwarzchild lens. The 2 μm pinhole gives a sample surface spot size of 0.1 μm. To obtain high reflectivity the surfaces are coated with a multilayer film designed for $h\nu = 130$ eV. The band pass is quite narrow, so the photon energy is not tunable. Scanning of the sample surface to obtain an image is realized by rastering the sample in a plane *(X-Y)* perpendicular to the beam. Photoelectrons are detected by a cylindrical mirror analyzer (CMA). The symmetry axis of the CMA is in the plane defined by the sample surface. The polar angle, θ, of the emitted electrons varies from 60° to 90°. The sample is oriented so that the CMA "looks" across the stripes resulting in images of the groove with some asymmetry. Data are collected in two modes: (a) the scanning mode where two-dimensional (2D) image is acquired by measuring the photoelectrons at a given kinetic energy, and (b) the microprobe mode where an energy distribution curve (EDC) is acquired at the position of interest. Independent tests show that the beam spot is 0.1 μm.

The samples were cleaned by heating them in an ultrahigh vacuum (UHV) chamber with a recipe used in other studies.[9] The samples were exposed to atomic hydrogen with a typical hot W filament near the sample surface. Magnesium was evaporated from a Ta boat.[9]

III. RESULTS

A. Clean surface

The AFM image of the ELO GaN surface (sample M287) is depicted in the 100×100 μm^2 image of Fig. 2(a). The lateral growth fronts meet and leave grooves as clearly seen

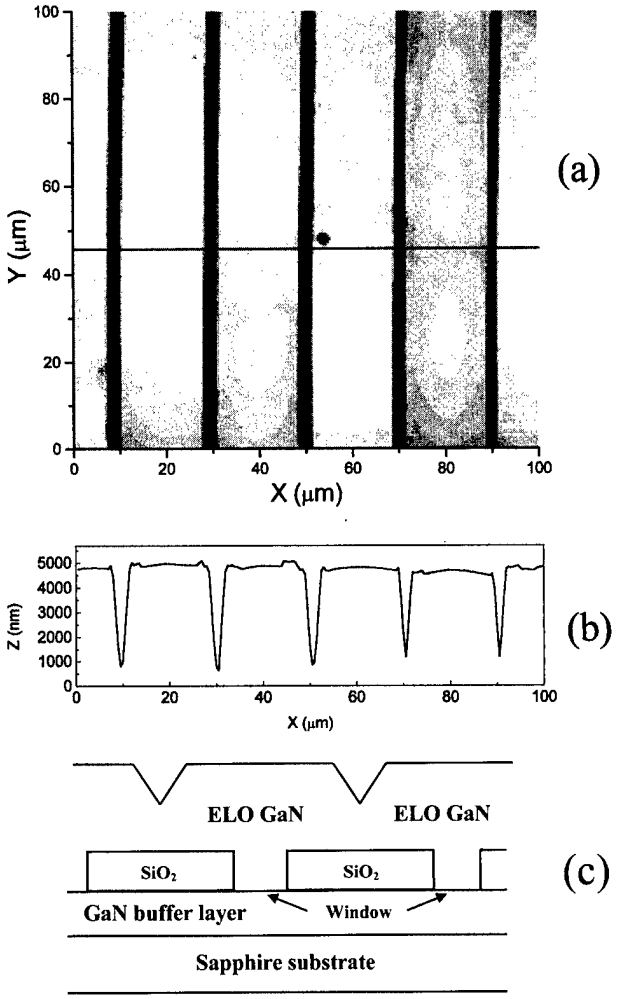


FIG. 2. (a) AFM image of a 100×100 μm^2 area of the ELO GaN surface (sample M287). The bright areas represent the flat growth stripes and the dark areas are the fronts where the growth stripes meet. (b) Line profile obtained from the above image at the position marked by the horizontal line across the image. (c) Schematic drawing of the structure of the ELO GaN material.

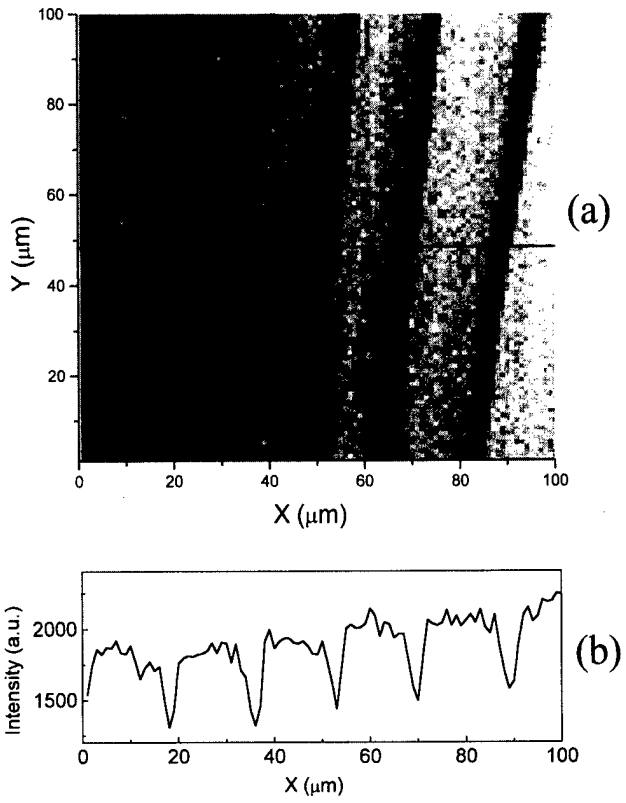


FIG. 3. (a) Photoemission spectromicroscopy image of a 100×100 μm^2 area of the ELO GaN surface (sample M287) formed by the signal at the peak of the Ga *3d* core-level emission at a kinetic energy of 104.6 eV. In the 2D map, the bright areas represent growth stripes and the dark areas represent the stripes where the growth fronts meet. The CMA detects the electron emission directed to the left. (b) Line profile obtained from the above image at the position of the straight line across the image.

in the image (dark stripes). The growth fronts have a partial coalescence. Figure 2(b) shows a line profile made across the sample at the location indicated by the horizontal line. Figure 2(c) shows a sketch of the cross section of the ELO GaN. The (0001) surfaces of the GaN stripes are very flat but differ in height from each other. The front-meeting grooves are 4 μm wide and about 4.3 μm deep. The angle between the walls of the fronts that meet and the sample's surface [i.e., the (0001) face] in the line profile is about 65°, which suggests that it is the $\{1\bar{1}01\}$ plane. The AFM tip used in this measurement is 20°. The growth side wall is $\{11\bar{2}0\}$ and depends on the growth conditions (the III–V ratio). The inclined walls at the growth-front areas is either in the $\{11\bar{2}2\}$ or the $\{1\bar{1}01\}$ facet, which is 50° or 62°, respectively, to the sample's surface. The measured face is attributed to the latter.

The scanning-mode measurement of the photoemission signal with the spectromicroscope is used to obtain an overview of the above sample. Figure 3(a) is the image and Fig. 3(b) the line profile. The 100×100 μm^2 image is acquired with a step size of 1 μm by recording the signal from Ga *3d* core-level emission peak at a kinetic energy of 104.6 eV. In the image, bright areas correspond to stripes of GaN and dark areas are stripes due to the front-meeting grooves. The photoemission signal (PESS) line profile shows the same general features as the AFM results shown in Fig. 2. The PESS has a slightly wider "V" groove than that in the AFM micrograph; they are about 4.2 and 3.6 μm, respectively.

The contrast mechanism producing the spectromicroscopy image is largely due to the angular dependence of the yield and the different orientations of the plateau areas and the valley areas. So the CMA analyzer "looks" at plateau and valley areas from different angles. The photoemission intensity falls off as the horizon is approached and is nominally proportional to $\cos^2\theta$ where θ is the photoelectrons' takeoff angle. Thus, the CMA will measure different emission intensity from surfaces at different angles. Other contrast mechanisms can be the inhomogeneity of the chemical environment, the electronic structures, etc. In addition to these properties that modulate the emission signal strength, they may also cause kinetic energy shifts that can be monitored.

For better understanding of inhomogeneities of the sample's surface electronic structure, high-spatial-resolution im-

ages are acquired. Then, based on this image, localized EDCs of the Ga *3d* core-level emission are obtained by the microprobe mode from different locations on the surface. Figure 4(a) shows a 30×30 μm^2 image with a step size of 0.3 μm made with the Ga *3d* peak emission at an electron kinetic energy of 104.6 eV. Figure 4(b) is a line profile across the sample surface at the position marked by the arrow where the CMA electron detector is on the left side. The PESS is shown by a gray scale and one can note that the average intensity decreases as the ALS storage ring's stored current decreases, i.e., the scan is from right to left. The curvature in the image is due to drift in the microscope. Localized EDCs for the Ga *3d* core level are obtained from the sample at positions labeled (a), (b), and (c) in the image with about a 0.1 μm spot size located at the front-meeting area, lateral overgrowth area, and vertical-growth window area, respectively. The spectra are shown in Fig. 4(c) where the three EDCs are normalized to the same peak intensity. The difference curve "(c)-(b)" is a flat featureless line, which demonstrates the uniformity in surface electronic structures for the vertical-growth window areas and the lateral overgrowth areas. However, curve (a), which is from a front-meeting area shows a shift of 0.15 eV. However, the difference curve "(b)-(a)" is also flat after curves (a) and (b) are lined up. These observations suggest no line shape changes and the shift is attributed to a change in the surface Fermi level position. The surface Fermi level pinning position for the GaN stripes (including the window and overgrowth areas) is about 0.15 eV lower than that for the region where the growth fronts meet.

Photoemission data from another sample (M273) are shown in Fig. 5, which has narrower growth-front grooves. The former sample is used for atomic H adsorption and the latter for Mg adsorption experiments. The Ga *3d* EDC set shows the same Fermi level pinning behavior for the front-meeting regions versus the flat stripes. The only difference is that the front-meeting grooves are about 2 μm wide. Note that the flat region near the groove, particularly the right side, has a slightly larger PESS (the white stripe). There is no present explanation for this unless the film "pushes" up a bit next to the coalescence line.

B. Atomic hydrogen adsorption

The clean ELO GaN surface is exposed to atomic H by the use of a 2000 °C W filament very close (5 cm) to the sample surface. The molecular H_2 dose was about 1000 langmuir. Figure 6(a) illustrates the photoemission spectromicroscopy formed from the peak emission of the Ga *3d* level, where the image is 30×30 μm^2 with a 0.5 μm step size. The same Ga *3d* peak position of kinetic energy, 104.6 eV, is used. Figure 6(b) depicts the line profile taken across the sample at the position marked by the arrow. The sample had to be moved to the preparation chamber for dosing, so the exact sample position is not reproducible. Furthermore, the sample stage in the microscope permits some variation of the angle between the light beam and the sample normal, which is the reason why the signal in Fig. 6(b) has a slope. Com-

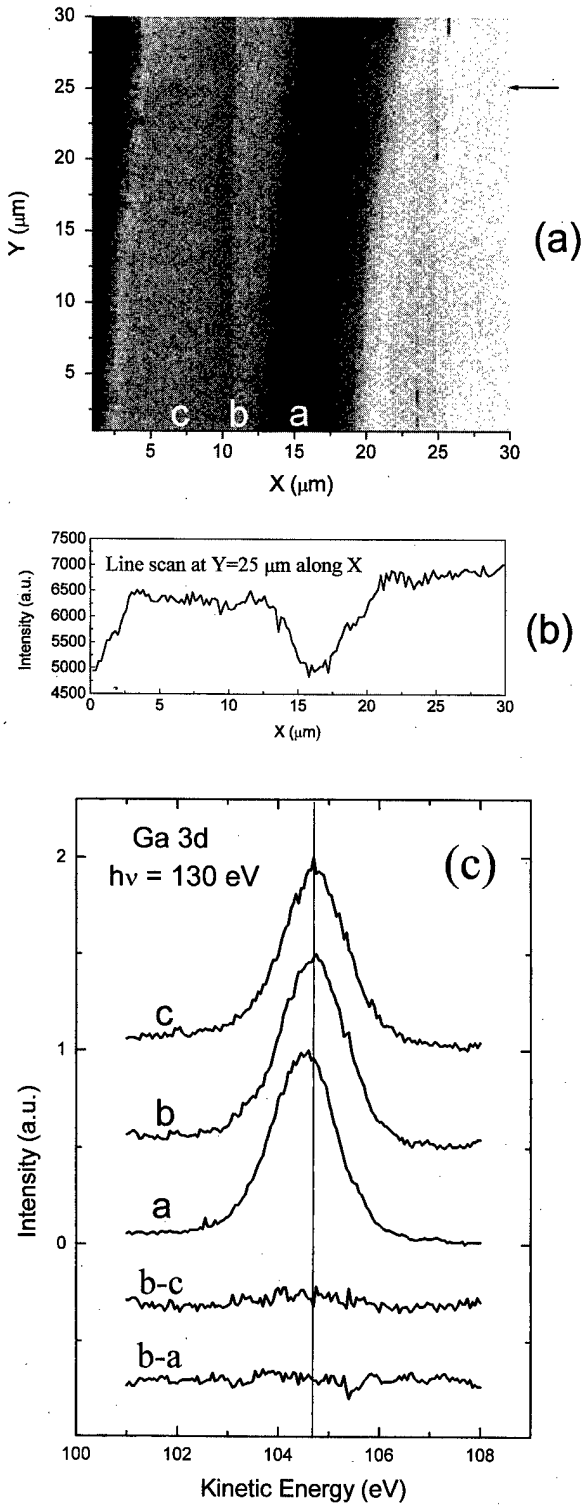


FIG. 4. (a) Photoemission spectromicroscopy image of a 30×30 μm^2 area of the ELO GaN surface (sample M287) formed by the signal at the peak position of the Ga *3d* core-level emission at a kinetic energy of 104.6 eV. (b) Line profile obtained from the above image at the position marked by the arrow. (c) Ga *3d* core-level EDCs collected by the microprobe mode with a 0.1 μm beam spot at positions (a), (b), and (c) shown in the image. The curve "(b)-(c)" is the difference between curves (b) and (c), where one is from vertical growth in the window and the other is from the lateral overgrowth area, respectively. The curve "(b)-(a)" is the difference between curves (b) and (a), which is from a front-meeting region after they are shifted to line up.

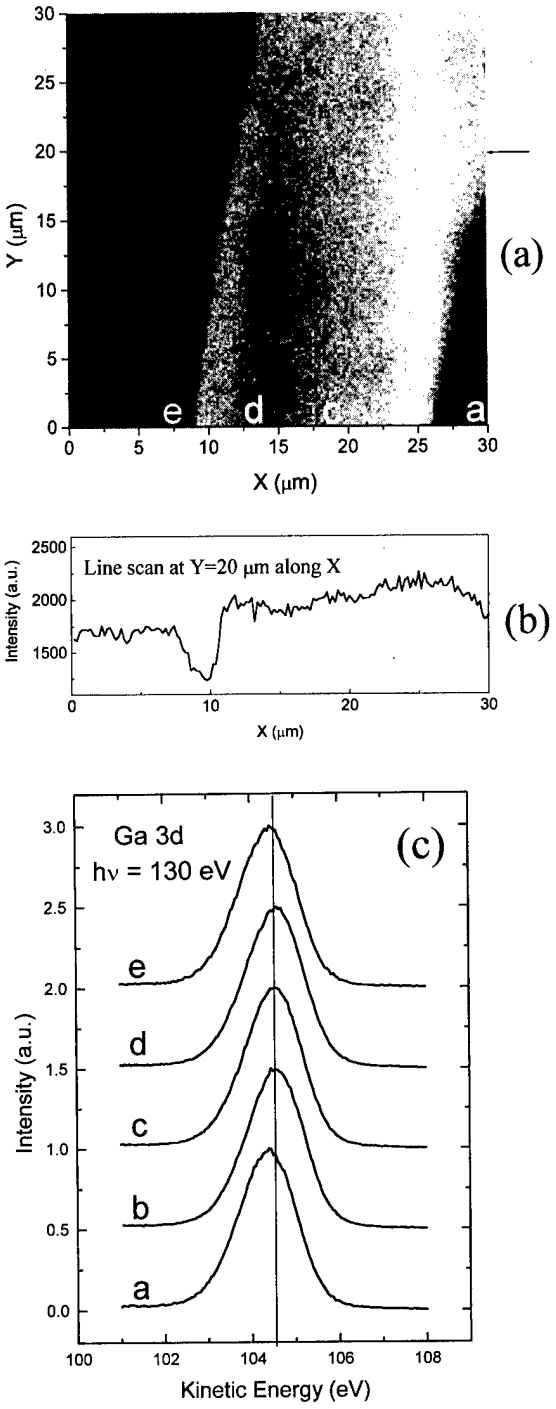


FIG. 5. Scanning mode image (a) with line profile (b) and microprobe Ga *3d* EDCs (c) obtained from sample M273. The results are basically the same as those in Figs. 3 and 4 (sample M287) except for the width of the front-meeting stripes which are 2 μm according to the AFM data.

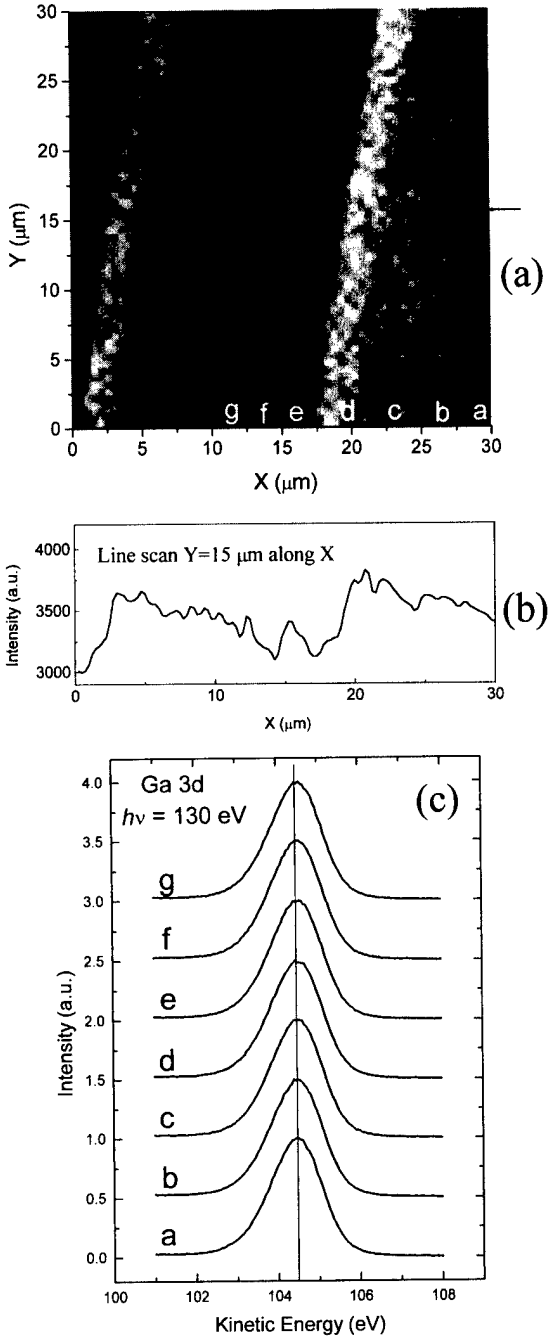


FIG. 6. (a) Photoemission spectromicroscopy image of a 30×30 μm^2 area of the ELO GaN surface (sample M287) exposed to atomic hydrogen. It is acquired by collecting the photoelectron signal from the peak of the Ga *3d* core-level emission at a kinetic energy of 104.6 eV. (b) Line profile obtained from the above image at the position marked by the arrow. The weakest emission (darkest) stripe is attributed to the bottom of the "V" in the front-meeting region. (c) Microprobe Ga *3d* EDCs collected at the position noted by (a)–(g) in the image. All the EDCs are essentially equivalent.

pared with the image and line profile from the clean surface in Fig. 4, the atomic hydrogen exposure causes dramatic changes for the ELO GaN surface. On clean surfaces, the window and overgrowth areas have uniform intensity, and the front-meeting signal looks like a V-shaped groove, which is analogous to the AFM "true" morphology. The ratio of the photoemission signal for the uniform area to the center of the front-meeting minimum is about 3/2. On the H-exposed surfaces, the window and overgrowth areas are still rather uniform, but the front-meeting region has a "W" shape. The darkest lines in the image are attributed to the darkest line in the image of the clean surface. The CMA is collecting photoelectrons going to the "left." The ratio of the signal between the "groove" and the uniform region appears to be

close to 1. The photoemission image of the growth-front region has a width of about 6.5 μm. The H appears to increase the emission signal at the groove centers, which face the CMA detector. An interpretation of this is not straightforward, but we suggest that H at the corner portions of the growth-front area increases the emission normal to the emitting surface which is off normal towards the groove with respect to the overall surface normal. This would increase the emission at the right corner and decrease it at the left corner, as observed in the data. If the corner has a weaker slope angle than the groove side wall, one could say that the H peaks the emission in the forward direction and the detector's orientation is such that we see a bigger signal on one side and a smaller one on the other. Another mechanism could be different H adsorption properties for the different crystalline surface and it would not be expected that one side wall would be different from the other side wall.

A set of localized Ga *3d* EDC data was taken on a line with a uniform distribution of positions, as shown by the letters in the image shown in Fig. 6(c). Some are from the flat region and some are from the groove. The spectra are essentially all the same shape and at the same energy position, as shown in Fig. 6(c). The peaks are at a kinetic energy of 104.45 eV, which is the same position as the Ga *3d* peak from the front-meeting regions for the clean surfaces. The H adsorption shifted peaks from the vertical and lateral growth areas from 104.6 to 104.45 eV (i.e., 0.15 eV). The uniform Fermi level is a bit of a surprise.

C. Magnesium films

Magnesium was vapor deposited onto a clean ELO GaN surface. About 10 ML on sample M273 gives an image that is quite similar to the one observed for the clean surface; see Figs. 7(a) and 5(a), respectively. As is seen in the line scan of Fig. 7(b) the signal strength ratio for the flat region versus the groove is about 3/2, the value for the clean surface. The result is quite different from the behavior for atomic H exposures. Figure 7(c) shows localized EDCs for the Ga *3d* level from the Mg-covered ELO GaN surface taken by the microprobe mode. The curves from points located at positions labeled (a)–(f) show different line shapes and kinetic energy positions. The differences show no obvious relationship with the sample's structure. The differences are attributed to the inhomogeneities in the Mg overlayer thickness on the ELO GaN material. Based on the previous data we have measured for Mg/GaN contacts[9] (done on conventional GaN samples). We found that Mg grows on the GaN surface by Stranski-Krastanov (SK) mode with one or two layers for wetting layer followed by island growth. The EDCs for positions (a) and (e) should be viewed differently than those for other positions since the absorption on crystalline faces in the groove can be different. Notice that the groove images are not uniform, which is again attributed to Mg island formation. If it is also the case for this ELO GaN material, then the photoemission spectromicroscope can "see" either island areas or the between-island areas. These two areas have different Mg coverages and this causes the differences in Ga

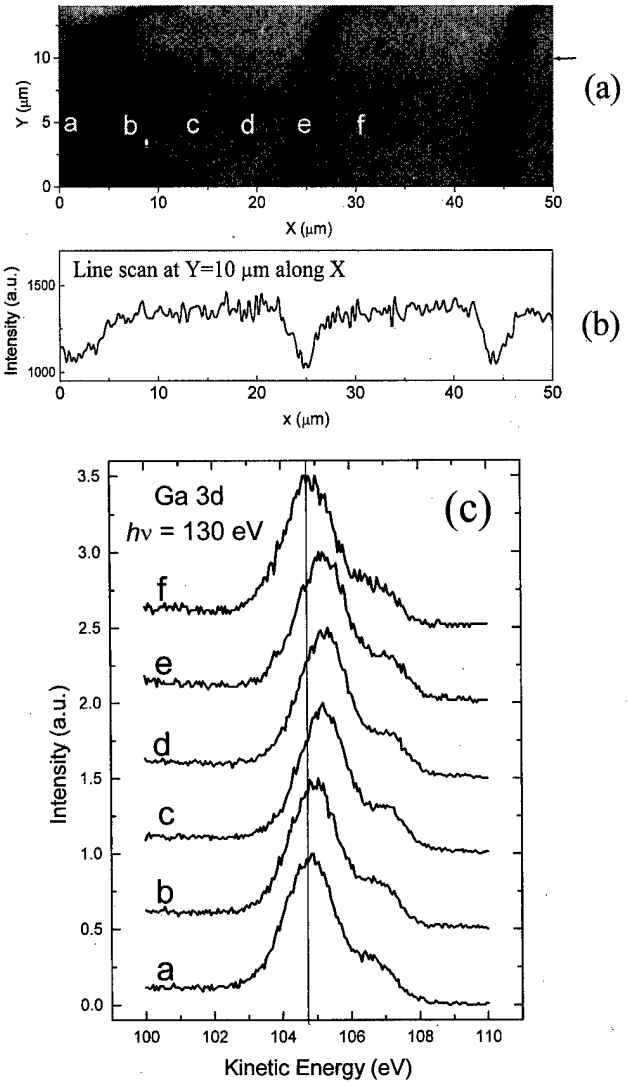


FIG. 7. (a) Photoemission spectromicroscopy image of a 50×15 μm^2 area of the ELO GaN surface (sample M287) with a 10 ML equivalent Mg deposition. (b) Line profile obtained from the above image at the position marked by the arrow. (c) Microprobe Ga *3d* EDCs collected at the positions from (a)–(f) noted in the image. All the curves have different energy positions that have no apparent relationship to the growth stripes on the film.

3d core-level emission lines. The island areas have a thicker Mg film so Mg-induced band bending can be larger and the Mg/GaN reaction can be more prevalent. As a result, the Ga *3d* core level peak shifts to the higher kinetic energy side. For the same reason, if the probe beam spot is on the areas between islands, which have less Mg coverage compared with the islands, the Ga *3d* core level will show a smaller energy shift. The small shoulder on the right-hand side of the main peak is attributed to metallic Ga atoms. It forms due to the reaction between GaN and Mg, where Mg atoms replace Ga atoms. The intensity of this shoulder also shows some weak variations in the EDCs from different positions.

IV. CONCLUSIONS

Photoemission spectromicroscopy has been employed to investigate the inhomogeneities of the surface electronic

structure of ELO GaN films. The PESS image acquired by the scanning mode on the clean surface shows a surface morphology which is similar to that of the AMF image. The contrast mechanism is attributed to the angular dependence of the photoemission yield. While Mg deposition has little effect on the mechanism, atomic H adsorption modifies it. Contrast changes due to the presence of adatoms appear to be quite weak. The relative orientation of the grooves and the CMA shows some differences in the groove areas. Nominally the emission is greater when the groove face is oriented towards the CMA. Localized EDCs for Ga *3d* core-level emission taken with the microprobe mode show morphological differences. On the clean surface, the Fermi level pinning position for the front-meeting regions shifts 150 meV compared with the window areas and overgrowth areas. However, the position dependent EDCs showed no difference for vertical growth from the windows and the lateral growth areas.

Hydrogen adsorption induced changes in the image and the EDCs energy position. The front-meeting area showed an increased signal strength for the corner facing the CMA and a decrease of emission from the corner on the other side. This is attributed to H changing the emission direction more toward the local-face normal. Effects due to crystalline face orientation or defect densities, if they exist, were not strong enough to be observed. All the EDCs for the hydrogenated surface have the same energy position, which is the one measured from the grooves on the clean surface.

The Mg-covered sample has an image quite similar to the clean case, indicating that the crystal orientation effects are weak. The energy dependence of the Ga or Mg EDC peak positions varies a good bit, but with no correlation to the GaN film's striped character. The variation is attributed to island formation in the SK growth mode.

ACKNOWLEDGMENTS

The help from ALS staff is gratefully acknowledged. This work is supported by ONR DESPCoR. ALS is supported by the DOE.

[1]S. Srtite and H. Morkoç, J. Vac. Sci. Technol. B **10**, 1237 (1992).
[2]M. A. Khan, J. N. Kuznia, A. R. Bhattarai, and D. T. Olson, Appl. Phys. Lett. **62**, 1786 (1993).
[3]S. Nakamura, T. Mukai, and M. Senoh, Appl. Phys. Lett. **64**, 1687 (1994).
[4]F. A. Ponce, MRS Bull. **22**, 51 (1997).
[5]Z. Yu, M. A. L. Johnson, T. McNulty, J. D. Brown, J. W. Cook, Jr., and J. F. Schetzina, Internet J. Nitride Semicond. Res. **3** (1998).
[6]S. Nakamura *et al.*, Appl. Phys. Lett. **72**, 211 (1998).
[7]H. H. Solak, G. G. Lourssо, S. Singh-Gasson, and F. Cerrina, Appl. Phys. Lett. **74**, 22 (1999).
[8]A. K. Ray-Chaudhuri *et al.*, J. Vac. Sci. Technol. A **11**, 2324 (1993).
[9]Y. Yang, S. H. Xu, H. Cruguel, G. J. Lapeyre, J. M. Van Hove, A. M. Wowchak, and P. P. Chow, Phys. Rev. B (submitted).

CUMULATIVE AUTHOR INDEX

All authors published to date in Volume 17 are listed alphabetically with the issue (in parentheses) and page numbers following the dash. An (E) after the page number indicates Erratum.

Abe, Kazuhide–(4) 1464
Abernathy, C. R.–(2) 366
Abraham-Shrauner, Barbara–(1) 158; (3) 961
Abrams, B.–(2) 758
Abrokwah, J.–(1) 49
Adachi, Hiroshi–(2) 620
Adamcyk, M.–(4) 1728
Adler, R. J.–(2) 883
Agius, B.–(4) 1430
Aguilar, L.–(3) 1158
Ahmari, D.–(3) 1185
Akahane, K.–(3) 1105
Aksenov, Igor–(1) 1; (4) 1525
Alberti, A.–(4) 1448
Alimova, A. N.–(2) 666, 670, 715
Allegrini, Maria–(2) 468
Allen, S. D.–(4) 1435
Allerman, A. A.–(4) 1781
Alport, M.–(2) 808
Amaratunga, G. A. J.–(2) 557
Amaratunga, Gehan A. J.–(2) 315
Ambacher, O.–(3) 1252
Anda, Yoshiharu–(2) 320
Andres, R. P.–(4) 1773
Andresen, A.–(4) 1757
Aoki, Koyu–(4) 1585
Aqariden, F.–(3) 1205
Arkles, Barry–(1) 182; (3) 1101
Arnal, Y.–(1) 118; (2) 879
Arslan, D.–(2) 784
Asbeck, P. M.–(4) 1742
Assanto, G.–(2) 465
Astolfi, D.–(2) 345
Auciello, O.–(2) 705
Ayón, A. A.–(4) 1589

Babic, D. I.–(3) 1272
Bacchetta, M.–(1) 44
Bacher, K.–(3) 1200
Baharav, Y.–(1) 253
Baik, Hong Koo–(1) 162, 174; (2) 627, 728, 731; (4) 1470
Baklanov, M. R.–(2) 372
Baklenov, O.–(3) 1124
Balogh, M. P.–(2) 828
Bandić, Z. Z.–(3) 1241; (4) 1750
Banerjee, S.–(2) 460
Baoping, Wang–(4) 1575
Barber, David–(2) 750
Bardeau, J.-F.–(2) 822
Bartlett, B.–(1) 90
Barvosa-Carter, W.–(4) 1786
Bass, R.–(3) 1022
Batistuta, M.–(4) 1773
Batke, E.–(4) 1738
Beach, R. A.–(4) 1753
Beaudoin, M.–(3) 1233, 1237
Beaudoin, Mario–(3) 1223
Bechstedt, F.–(4) 1691
Belobrov, P. I.–(2) 666, 715
Belova, N. G.–(1) 68
Bender, H.–(2) 372
Bennett, B. R.–(3) 1022; (4) 1786
Bensaoula, A.–(3) 1209
Berge, Torunn–(1) 217
Bergeron, D. J.–(4) 1309
Bergman, L.–(2) 700, 734
Berishev, I.–(3) 1209
Berkowitz, A. E.–(4) 1702
Bernholc, J.–(4) 1691
Besombes, C.–(4) 1510
Betsui, K.–(2) 604
Bhan, M.–(2) 477
Bhasin, G.–(3) 1163
Bhattacharya, P.–(3) 1116, 1276
Biefeld, R. M.–(4) 1781
Bilbro, G. L.–(2) 552
Bimberg, D.–(4) 1632, 1639
Bindley, U.–(3) 1214
Bird, J.–(4) 1757
Birkhahn, R.–(3) 1195
Blick, R. H.–(4) 1594
Block, T. R.–(3) 1136, 1163
Boit, C.–(4) 1570
Bolchouchine, V. A.–(2) 765
Bolduc, M.–(2) 859
Bonanni, A.–(4) 1697, 1722
Boonzaayer, M. D.–(3) 1233, 1237
Boos, J. B.–(3) 1022
Borenstein, J. T.–(4) 1336
Borghesi, A.–(1) 44
Borovsky, Brian–(1) 7
Bottomley, D. J.–(1) 12; (2) 259
Boucaud, P.–(3) 1301
Bowers, B.–(1) 49
Bracker, A. S.–(4) 1786
Braff, R. A.–(4) 1589
Braun, W.–(2) 474
Brett, M. J.–(1) 186
Bricard, L.–(4) 1510
Bridger, P. M.–(3) 1241; (4) 1750
Brown, A. D.–(2) 323
Brückl, H.–(4) 1350
Brueck, S. R. J.–(3) 908, 921
Brueckner, J.–(2) 867
Buckley, L. J.–(1) 90
Bulger, J. M.–(1) 194; (2) 410
Bullock, D. W.–(4) 1778
Burm, J.–(3) 905
Busnaina, A. A.–(2) 443

Cai, W. Z.–(3) 1190
Cai, W.-Z.–(3) 1180
Campbell, A. B.–(2) 828
Campbell, Albert B.–(2) 813
Campbell, J. C.–(3) 1124, 1175
Cao, X. A.–(4) 1540
Capellini, G.–(2) 465
Carelli, P.–(2) 500
Cargill, G. S., III–(3) 1255
Casey, Jeffrey A.–(2) 888
Castellano, M. G.–(2) 500
Celotta, R. J.–(3) 1087
Cerrina, F.–(4) 1884
Chabal, Y. J.–(4) 1795
Chadda, G. K.–(4) 1549
Chalamala, Babu R.–(1) 233; (2) 303
Chang, Huicheng–(1) 224
Chang, K. L.–(3) 1120
Chang, R. P. H.–(2) 705
Chang, Y. A.–(2) 432
Chang, Ying-Lan–(3) 1272
Chao, L. C.–(3) 1056
Chayahara, Akiyoshi–(2) 840
Chayka, V. E.–(2) 655
Chen, C. K.–(2) 345
Chen, C.-P.–(2) 432
Chen, Chih-Liang–(4) 1545
Chen, E. H.–(4) 1773
Chen, E.-H.–(4) 1767
Chen, H.–(2) 688; (3) 1139
Chen, J.–(2) 592, 710; (3) 1059
Chen, Jiaxing–(1) 224
Chen, Junghui–(1) 145
Chen, Kunji–(3) 989
Chen, L.–(2) 638
Chen, L. C.–(4) 1877
Chen, L. J.–(2) 385; (3) 939
Chen, M. C.–(2) 385
Chen, P. J.–(4) 1702
Chen, Xiaolan–(3) 908, 921
Chen, Xiaomeng–(1) 182
Chen, Yong–(3) 1080
Cheng, K. Y.–(3) 1120
Cheng, Shiou-Ying–(4) 1477
Cheung, D.–(2) 477
Chi, Eung Joon–(2) 728, 731
Chin, T. P.–(3) 1136, 1163
Chirovsky, L. M.–(3) 905
Cho, H.–(4) 1540
Cho, Kyeongjae–(4) 1612
Cho, S. J.–(2) 380
Choi, H. S.–(3) 1017
Choi, Junhee–(2) 547
Choi, S. S.–(2) 583
Choi, W. B.–(2) 632, 700, 734
Choi, Y. S.–(2) 744
Choi, Yearn-Ik–(2) 635
Choy, K. H.–(3) 1281
Chu, A. K.–(2) 455
Chu, K.–(3) 1252
Chu, P. K.–(2) 851, 875
Chua, S. J.–(4) 1507
Chuang, M. S.–(2) 385
Chubun, N. N.–(2) 715
Chun, M.–(2) 863
Chun, M. C.–(2) 385
Chung, S. J.–(3) 1151
Chung, Suk Jae–(1) 241
Chyr, I.–(2) 362
Cirelli, R. A.–(1) 101
Cirelli, Raymond A.–(3) 930
Clark, E. J.–(2) 883
Clark, J.–(1) 194
Clark, Joseph–(2) 410
Clauss, W.–(4) 1309
Clement, M.–(2) 397
Clevenger, Marvin B.–(3) 1289
Colace, L.–(2) 465
Comrie, C.–(2) 808
Conrad, J. R.–(2) 863
Corrigan, T. D.–(2) 705
Cowie, B. C. C.–(4) 1666
Cowles, J.–(3) 1136
Craighead, H. G.–(1) 29; (3) 1087; (4) 1366
Crespo, J.–(3) 1085
Cruguel, H.–(4) 1884
Culbertson, J. C.–(4) 1786
Cuomo, J. J.–(2) 632, 700, 734
Currie, J. F.–(2) 423

Dähne-Prietsch, M.–(4) 1632, 1639, 1682
Dai, N.–(3) 1297
Dalton, T. J.–(4) 1435
Dang, G. T.–(4) 1540
Dang, X. Z.–(4) 1742
Dashiell, M. W.–(3) 1301
Davidson, J. L.–(2) 740
Davies, Dominic–(2) 750
Davies, M. C.–(2) 273
Davis, P. R.–(2) 613
Däweritz, L.–(4) 1716
Deacon, T.–(2) 477
Dehé, A.–(2) 784
De La Rue, Richard M.–(1) 113
Delmotte, F.–(4) 1430
DenBaars, S. P.–(1) 29
Deng, S. Z.–(2) 592, 710
Denton, John P.–(3) 994
Detkov, P. Y.–(2) 666
Dew, S. K.–(1) 186
Di Carlo, Aldo–(4) 1674
Dicke, J.–(4) 1773
Di Gaspare, L.–(2) 465
Ding, M. Q.–(2) 705
Ding, Weiping–(3) 1050
Dip, A.–(4) 1424
Döbeli, M.–(3) 945
Dorsey, D. L.–(3) 1218, 1227
Droopad, R.–(1) 49
Duchenois, A. M.–(4) 1510
Ducroquet, F.–(4) 1553
Dumont, A.–(1) 194
Dumont, Andrew–(2) 410
Dunning, F. B.–(3) 1087
Dunstan, P. R.–(4) 1666

Ea, C. S.–(2) 323
Eastman, L. F.–(3) 1252
Eberhardt, W.–(4) 1728
Ebersberger, B.–(4) 1570
Egami, N.–(3) 1127
Egelhoff, W. F., Jr.–(4) 1702
Eisebitt, S.–(4) 1728
Eisele, H.–(4) 1639
Eisinger, S.–(1) 90
El-Gomati, M. M.–(2) 638
Elmoursi, A. A.–(2) 828
Elmoursi, Alaa A.–(2) 813
Emmerling, M.–(1) 86
Engemann, J.–(2) 696
Ensinger, W.–(2) 799
Esser, N.–(4) 1691
Etrillard, J.–(4) 1510
Eustis, T. J.–(3) 1252
Evangelisti, F.–(2) 465
Evoy, S.–(1) 29
Evtukh, A. A.–(2) 655, 679
Eyink, K. G.–(3) 1218; (4) 1852

Fabian, J.–(4) 1708
Fair, R. B.–(4) 1813
Fairbairn, K.–(2) 477
Fan, S.–(3) 1171
Fan, Z.–(2) 311
Fang, X. M.–(3) 1263, 1297
Fedirko, V. A.–(1) 68
Fedotova, J. A.–(2) 836
Feng, M.–(3) 1139, 1185
Feng, M. S.–(4) 1456
Feng, Z. C.–(4) 1507
Ferry, D. K.–(4) 1757, 1840
Fetherston, P.–(2) 863

Filip, V.–(2) 520
Filip, Valeriu–(2) 542
Fimland, Bjørn-Ove–(1) 217
Fitzgerald, Eugene A.–(4) 1485
Flanagin, Lewis W.–(4) 1371
Flebbe, O.–(4) 1639
Fonstad, C. G., Jr.–(3) 1281
Fontenot, C.–(1) 90
Foord, John S.–(2) 719
Forbes, Richard G.–(2) 526, 534
Forchel, A.–(1) 86
Forrest, R. D.–(2) 596
Forster, John–(2) 416
Fortin, V.–(2) 423
Fretwell, J. L.–(2) 460
Freundlich, A.–(3) 1158
Friedrich, L. J.–(1) 186
Friedrich, M.–(4) 1733
Frisch, A. M.–(4) 1691
Frisch, Harry L.–(1) 182
Fritze, M.–(2) 345
Fuhrmann, H.–(3) 945
Fujino, T.–(3) 983
Fujita, Jun-ichi–(3) 933
Fujita, K.–(3) 1127
Fujita, Kazushi–(3) 957
Furdyna, J. K.–(3) 1214
Furutake, N.–(2) 567
Fuse, T.–(3) 983
Fuso, Francesco–(2) 468
Futai, M.–(2) 288
Futatsugi, Toshiro–(4) 1317

Gagnon, G.–(2) 423
Galdetskiy, A.–(2) 778, 781
Galluzzi, F.–(2) 465
Ganser, P.–(4) 1738
Ganz, Eric–(1) 7
Gao, C. X.–(2) 684
Garcia-Rocha, M.–(4) 1697
Gardella, Joseph A., Jr.–(1) 224
Garner, David M.–(2) 315
Gauch, Michel–(4) 1390
Gaudreau, P. E., Marcel P. J.–(2) 888
Gauvin, R.–(2) 423
Ge, F.–(4) 1757
Gessner, T.–(4) 1733
Giapis, Konstantinos P.–(3) 999
Gibbons, Carol–(2) 750
Gibson, P. N.–(4) 1329
Gill, W. N.–(1) 205
Gnade, Bruce E.–(2) 303
Godbey, D. J.–(1) 90
Göhl, A.–(2) 670, 696
Goldammer, K. J.–(3) 1151
Goldberg, Cindy–(3) 1101
Goldburt, E. T.–(2) 765
Goncharuk, N. M.–(2) 655
Goodman, R. B.–(3) 1045
Goodman, T.–(1) 60
Goorsky, M. S.–(3) 1163, 1289
Gopal, V.–(4) 1767
Goshima, Hidekazu–(3) 1028
Goto, Toshio–(3) 949, 957, 1098
Gotoh, Y.–(2) 604, 623
Graf, T.–(4) 1848
Grant, J. M.–(4) 1424
Grassi, E.–(3) 1237
Grassi, Elena–(3) 1223
Gratteau, N.–(3) 1180
Gratteau, Navid–(3) 1285
Graves, C. A.–(2) 773
Grazulis, L.–(4) 1852
Gregg, C. W.–(2) 883
Grein, C. H.–(3) 1205
Grepstad, Jostein K.–(1) 217
Gröning, O.–(2) 666; (3) 1064
Gröning, P.–(3) 1064
Gruen, D. M.–(2) 705
Gu, C.–(2) 688
Gu, C. Z.–(2) 684
Guenzel, R.–(2) 836
Gujrathi, S. C.–(2) 423
Günzel, R.–(2) 832, 895
Guo, S. P.–(3) 1255
Guo, Y. L.–(2) 448
Gupta, V. K.–(3) 1246, 1249
Gutierrez-Aitken, A. L.–(3) 1136

Habermann, T.–(2) 670, 696
Hahn, Y. B.–(2) 366
Haich, C. A.–(3) 1087
Halder, N. C.–(1) 60
Hamai, Chiho–(4) 1313
Hamdi, A. H.–(2) 828
Hamdi, Aboud H.–(2) 813
Hamm, R. A.–(3) 905
Hammadi, Zoubida–(4) 1390
Han, A. C.–(3) 1136
Han, Eun Jung–(1) 241
Han, I. K.–(3) 1017
Han, J.–(3) 1087
Han, L.–(2) 688
Han, Min-Koo–(2) 635
Han, S.–(2) 863
Hanchun, Yin–(4) 1575
Hannappel, Th.–(4) 1691
Hänni, W.–(4) 1570
Hansen, Marie–(2) 297
Hansing, C.–(3) 1175
Harada, Hiroshi–(2) 575
Harada, T.–(3) 983
Harada, Yusuke–(4) 1464
Harris, C. I.–(3) 1040
Harris, C. T.–(2) 773
Harrison, M. T.–(2) 580
Hartnagel, H. L.–(2) 784
Hasegawa, Hideki–(4) 1856
Hatfield, C. W.–(2) 552
Hattendorf, M.–(3) 1185
Hauser, J. R.–(4) 1836
Hawkey, Timothy J.–(2) 888
Haworth, L.–(4) 1659
Hayafuji, Norio–(3) 1011
Hays, D. C.–(2) 366
He, Xiao-Ming–(2) 822, 845
Hebert, K. J.–(1) 53
Heinlein, Christian–(1) 217
Heinrichsdorff, F.–(4) 1632, 1639
Heiss, W.–(4) 1722
Heremans, J.–(3) 1267
Hernández-Calderón, I.–(3) 1259
Heroux, J. B.–(3) 1144
Hetzler, K.–(3) 1131
Hey, R.–(4) 1791
Hickman, R.–(4) 1540
Hill, P.–(4) 1659
Hill, S. B.–(3) 1087
Hiller, K.–(4) 1733
Hilt, L. L.–(4) 1840
Hingerl, K.–(4) 1697, 1722
Hinrichs, C. H.–(2) 297
Hiramatsu, Mineo–(3) 949
Hiraoka, T.–(3) 1167
Hiraoka, Tetsuya–(4) 1482
Hirashita, Norio–(2) 355
Hirose, Masataka–(2) 306
Hirose, Nobumitsu–(2) 320
Hirvonen, J.-P.–(4) 1329
Hiyamizu, S.–(3) 1127, 1167
Hiyamizu, Satoshi–(3) 1155; (4) 1482
Hoffmann, H.–(4) 1570
Hogg, S. M.–(4) 1502
Hoke, W. E.–(3) 1131
Holl, S.–(4) 1722
Holland, E. R.–(2) 580
Hollis, M. A.–(2) 773
Holloway, P. H.–(2) 758
Holmes, A. L., Jr.–(3) 1124, 1175
Hong, J. P.–(2) 489
Hong, M.–(3) 1294
Hopwood, J.–(2) 405
Hori, Masaru–(3) 949, 957, 1098
Horibata, Shinji–(2) 575
Höricke, M.–(4) 1791
Horino, Yuji–(2) 840
Horn, M. W.–(3) 1045
Hosono, Akihiko–(2) 575
Houlihan, F. M.–(1) 101
Howell, M.–(2) 740
Hren, J. J.–(2) 632, 700, 734
Hsia, H.–(3) 1139
Hsieh, K. C.–(3) 1120
Hu, Evelyn L.–(4) 1622
Hu, M. C.–(2) 392
Hu, Xiaoming–(2) 265
Hua, X. D.–(1) 138
Huang, Feng–(2) 269
Huang, M.–(2) 580
Huang, Xinfan–(3) 989
Huang, Y. H.–(4) 1507
Huang, Y. S.–(2) 455
Huang, Zhongping–(2) 570
Hudait, Mantu Kumar–(3) 1003
Hudgins, R.–(3) 1195
Hugon, M. C.–(4) 1430
Huq, S. E.–(2) 592
Hwang, C. J.–(3) 1017
Hwang, Gyeong S.–(3) 999
Hwang, Ki-Hyun–(1) 213

Ichikawa, Masakazu–(3) 978
Ikegami, Naokatu–(2) 355
Ikoma, Hideaki–(4) 1516
Inayoshi, Muneto–(3) 949
Inoue, Kaoru–(2) 320
Ippen, E. P.–(3) 1171
Ireland, Terry–(2) 750
Irene, E. A.–(1) 33, 53; (4) 1430
Ishihara, K.–(4) 1589
Ishii, Kenichi–(1) 77
Ishikawa, J.–(2) 604, 623
Ishitani, Tohru–(4) 1400
Isono, Ryota–(2) 871
Itani, Toshiro–(1) 101
Ito, Masafumi–(3) 949, 957, 1098
Itoh, J.–(2) 520, 608, 788
Itoh, Junji–(2) 542, 588
Itsumi, Manabu–(1) 201
Ivanov, I. G.–(3) 1040
Iwai, Hidekazu–(4) 1525
Iwami, M.–(1) 12
Izumi, Shigekazu–(3) 1011

Jackman, Richard B.–(2) 719
Jain, A.–(1) 205
Jan, C.-H.–(2) 432
Jandaska, A.–(3) 1205
Janes, D. B.–(4) 1773
Jang, H. G.–(3) 1017
Jang, J. E.–(2) 489
Jang, Jin–(1) 241; (2) 486
Jang, Shi-Shang–(1) 145
Jayatissa, A. H.–(1) 237
Jazwiecki, M.–(3) 1289
Jensen, John A.–(3) 1080
Jensen, K. L.–(2) 515
Jeon, D. Y.–(2) 489
Jeon, H.–(2) 583
Jessing, J. R.–(2) 601
Jeyakumar, R.–(4) 1549
Ji, H.–(2) 684, 688
Jiang, H.–(2) 688; (3) 1276
Jiménez-Leube, F. J.–(2) 397
Jin, C.–(2) 688
Jin, W. C.–(2) 684
Jin, Y.–(2) 688
Jin, Y. W.–(2) 489, 744
Jin, Z. S.–(2) 684
Jing, Xiping–(2) 750
Joannopoulos, J. D.–(3) 1171; (4) 1612
Johanson, Robert E.–(1) 73
Johnson, A. T.–(4) 1309
Johnson, M. B.–(3) 1151
Johnson, S. R.–(2) 474; (3) 1233, 1237
Johnson, Shane R.–(3) 1223
Johs, B.–(3) 1233
Jones, S. L.–(2) 758
Jonhson, A.–(2) 477
Joubert, O.–(4) 1406
Ju, Byeong Kwon–(1) 241; (2) 486
Juan, W. H.–(4) 1336
Juang, M. H.–(2) 392
Jung, J. E.–(2) 489, 744
Jung, Jae Hoon–(1) 241, 246; (2) 486
Jung, M. Y.–(2) 583
Jurkovic, M. J.–(3) 1144

Kalka, T.–(4) 1639, 1682
Kaloyeros, Alain E.–(1) 182; (3) 1101
Kamimoto, Hitoshi–(3) 1155
Kamo, Mutsukazu–(2) 723
Kamp, M.–(1) 86
Kanamori, Jun–(2) 355
Kanazawa, Keisuke–(4) 1516
Kaneko, Y.–(3) 1272
Kanemaru, Seigo–(1) 77; (2) 588
Kaneshiro, Chinami–(4) 1856
Kaneshiro, E. N.–(3) 1136
Kang, D. H.–(2) 632
Kang, Ho Kwan–(1) 246
Kang, Sang-Woo–(1) 154
Kang, W. P.–(2) 740
Kao, C. R.–(2) 432
Karl, A.–(4) 1728
Kasama, Kunihiko–(1) 101
Kasap, S. O.–(1) 73
Kasibhatla, B.–(4) 1773
Kästner, Martin–(2) 294
Katayama, M.–(3) 983
Kavanagh, K. L.–(4) 1649
Kawabe, M.–(3) 1105
Kawabuchi, Shinji–(2) 575
Kawai, Akira–(3) 1090
Kawai, T.–(2) 288, 494
Kawai, Tomoji–(4) 1313
Kawanami, Yoshimi–(4) 1400
Kawata, Satoru–(2) 620
Keister, J. W.–(4) 1831
Keller, S.–(1) 29
Kelly, D. L.–(4) 1852
Kelsey, Jean E.–(3) 1101
Kempkes, Michael A.–(2) 888
Kerns, D. V.–(2) 740
Khairurrijal, –(2) 306
Khodasevich, V. V.–(2) 836
Kim, Andrew Y.–(4) 1485
Kim, B.–(2) 863
Kim, D. W.–(2) 583
Kim, Dong Joon–(4) 1598
Kim, E.–(3) 1209
Kim, Edwin–(2) 416
Kim, Gi Bum–(1) 162; (2) 627
Kim, Hong-Ryong–(2) 601

Kim, Hyo-Uk–(1) 154
Kim, J. M.–(2) 489, 744
Kim, Jaewan–(2) 547
Kim, Jong Kyu–(2) 497
Kim, Jongmin–(2) 547
Kim, K. W.–(2) 380
Kim, N. Y.–(2) 380
Kim, S. Y.–(2) 380
Kim, Taeil–(2) 497
Kim, To Hoon–(1) 246
Kim, Yi-Tae–(3) 1034
Kim, Yong Tae–(4) 1598
Kimura, Chiharu–(2) 723
Kinomura, Atsushi–(2) 840
Kinoshita, Haruhisa–(1) 109
Kinsey, G.–(3) 1175
Kitada, T.–(3) 1167
Kitada, Takahiro–(3) 1155; (4) 1482
Kitano, Yoshiaki–(3) 1155
Kitching, S.–(2) 273
Klemenc, M.–(4) 1848
Klotzkin, D.–(3) 1276
Klyui, N. I.–(2) 679
Knobel, R.–(3) 1147
Koch, M. W.–(3) 1246, 1249
Koga, Keisuke–(2) 588
Köhler, K.–(4) 1738
Koizumi, Satoshi–(2) 723
Kolagunta, V. R.–(4) 1773
Kolodzey, J.–(3) 1301
Kolodziej, J. J.–(4) 1831
Kolodziejski, L. A.–(3) 1171
Komoda, Takuya–(3) 1076
Kondoleon, C.–(2) 758
Koops, H. W.–(4) 1350
Kopf, R. F.–(3) 905
Koshida, Nobuyoshi–(3) 1076
Kostial, H.–(4) 1791
Kouji, Yoshiharu–(3) 1011
Kozicki, Michael N.–(4) 1561
Krauss, A. R.–(2) 705
Krauss, Thomas F.–(1) 113
Krenn, H.–(4) 1722
Kretz, J.–(4) 1350
Kriele, A.–(4) 1594
Krishna, S.–(3) 1116
Kropfeld, P.–(4) 1553
Krost, A.–(4) 1632, 1639
Krueger, Michael–(1) 7
Kruger, Joachim–(3) 1272
Krupanidhi, S. B.–(3) 1003
Krupenin, V. A.–(1) 230
Kruppa, W.–(3) 1022
Kruse, O.–(2) 855
Kubiak, C. P.–(4) 1773
Kubo, O.–(3) 983
Kudzinovski, S. Yu.–(2) 679
Kuhn, M.–(1) 194
Kuhn, Markus–(2) 410
Kuhn, S.–(1) 86
Kuleshov, A, K.–(2) 836
Kummel, A. C.–(1) 49
Kung, H.–(4) 1329
Kuo, H. C.–(3) 1139, 1185
Kuriyama, Kenji–(2) 723
Kuriyama, R.–(3) 1167
Kurooka, Shunji–(2) 840
Küttel, O. M.–(2) 666; (3) 1064
Kuzmin, L. S.–(4) 1413
Kvam, E. P.–(4) 1767
Kwak, Joon Seop–(1) 162
Kwo, J.–(3) 1294

LaBella, V. P.–(4) 1778
Lagarde, T.–(1) 118
Lai, L.–(1) 33, 53
Lakshmikumar, S. T.–(4) 1549
Lambers, E. S.–(2) 366
Lampert, W. V.–(4) 1852
Lan, S.–(3) 1105
Langan, J.–(2) 477
Langouche, G.–(4) 1502
Lantier, Roberta–(4) 1674
Lapeyre, G. J.–(4) 1884
Lardizabal, S. M.–(3) 1131
Lau, S. S.–(4) 1742
Lauterbach, J.–(4) 1773
La Via, F.–(4) 1448
Lazar, H.–(4) 1836
Leavitt, R. P.–(3) 965
Le Coeur, F.–(2) 879
Lee, D.–(3) 1195
Lee, D. H.–(2) 818, 822
Lee, Eunhwa–(1) 224
Lee, Gi Young–(3) 1034
Lee, H. W.–(2) 744
Lee, Hangwoo–(2) 547
Lee, Heon Young–(2) 731
Lee, J. W.–(2) 366
Lee, Jae Won–(2) 497
Lee, Jeong-Sik–(4) 1341
Lee, Jong-Lam–(2) 497; (3) 1034
Lee, N. S.–(2) 489, 744
Lee, S.–(3) 1214
Lee, Sang-In–(4) 1470
Lee, Se Jong–(2) 731
Lee, Sung Man–(2) 731
Lee, Sung-Man–(1) 162, 174; (4) 1470
Lee, T. D.–(2) 690
Lee, T. S.–(3) 1205
Lee, Takhee–(4) 1773
Lee, Y. S.–(1) 138
Lee, Yao-Sheng–(3) 961
Lee, Yun Hi–(2) 486
Lemonias, P. J.–(3) 1131
Lenox, C.–(3) 1175
Leoni, R.–(2) 500
Lesaint, O.–(2) 879
Lester, L. F.–(3) 965
Leung, C. H.–(2) 828
Leung, Chi H.–(2) 813
Levonovitch, B. N.–(2) 765
Li, G.–(4) 1507
Li, Guohui–(4) 1575
Li, Heng-De–(4) 1567
Li, J.–(4) 1502
Li, Q.–(2) 311, 740
Li, W.–(2) 311
Li, Wei–(3) 989
Li, Wen-Zhi–(4) 1567
Liao, Jin-Yuan–(4) 1545
Lim, Kuo-Yi–(3) 1171
Lim, Moo-Sup–(2) 635
Lim, S. H.–(2) 583
Lin, C.-F.–(2) 432; (4) 1456
Lin, C. H.–(3) 1139
Lin, Po-Hung–(4) 1477
Lin, W.–(3) 1255
Lin, Wen-Kuan–(4) 1545
Lin, Y. C.–(2) 385
Linder, K. K.–(3) 1116
Lindsay, J. H.–(2) 828
Lindstrom, S. C.–(3) 1151
Litovchenko, V. G.–(2) 655, 679
Litvin, Yu. M.–(2) 655
Liu, A. G.–(2) 875
Liu, B. B.–(2) 684
Liu, Guo Lin–(2) 355
Liu, H.–(2) 345
Liu, Jia–(4) 1773
Liu, Jianhong–(3) 989
Liu, R.–(4) 1424
Liu, W. K.–(3) 1200
Liu, Wen-Chau–(4) 1477
Liu, X. F.–(3) 1116
Liu, X. H.–(2) 311
Livengood, Richard H.–(1) 40
López-López, M.–(3) 1259
Lorbacher, J.–(4) 1632
Lorenz, H.–(4) 1594
Lourtioz, J.-M.–(3) 1301
Lu, Guangquan–(4) 1417
Lu, Hai–(3) 989
Lu, J.–(4) 1659
Lu, X. Y.–(2) 684
Lu, Y.-C.–(4) 1329
Lubyshev, D.–(3) 1180, 1285
Lubyshev, D. I.–(3) 1190
Lucovsky, G.–(4) 1806, 1813, 1831, 1836
Ludeke, R.–(4) 1823
Lugli, Paolo–(4) 1674
Lullo, G.–(3) 1281
Luo, Hong–(1) 224
Luo, Y. Y.–(3) 1255
Lüth, Hans–(4) 1674
Lutz, R. C.–(3) 1200
Luysberg, M.–(3) 1200
Lyman, P. S.–(3) 1131

Ma, Xianyun–(2) 769; (4) 1580
Macdonald, J. E.–(4) 1659
Macfarlane, P. J.–(4) 1627
MacGibbon, B. S.–(2) 443
Mackie, W. A.–(2) 613
Madey, T. E.–(4) 1831
Maeda, K.–(2) 608
Maeda, Masahiko–(1) 201
Maeda, Tatsuro–(1) 77
Maeda, Y.–(2) 288, 494
Maex, K.–(2) 372
Magno, R.–(4) 1786
Mahalingam, K.–(3) 1218
Maillard-Schaller, E.–(2) 666
Majewski, J. A.–(4) 1617
Malaczynski, G. W.–(2) 828
Malaczynski, Gerard W.–(2) 813
Malik, S. M.–(2) 863
Manako, Shoko–(3) 933
Mändl, S.–(2) 832
Manke, I.–(4) 1632
Mannaerts, J. P.–(3) 1294
Mao, D. S.–(2) 311
Marchenko, R. I.–(2) 679
Marqués, M. I.–(2) 788
Mars, D. E.–(3) 1272
Marsh, P. F.–(3) 1131
Marsh, Paul–(2) 750
Martien, D.–(4) 1702
Maruyama, Toshiro–(2) 840
Masaitis, R. L.–(3) 1294
Masamune, Sadao–(2) 871
Masini, G.–(2) 465
Masuda, T.–(1) 237
Masumoto, Yasuaki–(4) 1341
Matsui, Toshiaki–(2) 320
Matsukawa, Takashi–(2) 588
Matsumoto, T.–(2) 494
Matsuno, Toshinobu–(2) 320
Matsuo, P. J.–(4) 1435
Matyi, R. J.–(2) 863
Maulat, O.–(2) 879
Mayer, A.–(2) 506
McAlister, D.–(3) 1263
McCann, P. J.–(3) 1263, 1297
McClelland, J. J.–(3) 1087
McCullough, Wendy S.–(4) 1485
McGill, T. C.–(3) 1241; (4) 1750, 1753
McGregor, D. R.–(2) 700
McKeown, Patrick J.–(1) 224
McMichael, R. D.–(4) 1702
McTaggart, R. A.–(3) 1131
Melloch, M. R.–(4) 1773
Menagarishvili, Vladimir M.–(1) 95
Méndez-García, V. H.–(3) 1259
Mescheryakova, A. L.–(2) 670
Meyer, K.–(2) 808
Meyer, R.–(3) 1185
Meyer, T.–(4) 1848
Micovic, M.–(3) 1180
Milekhin, A.–(4) 1733, 1738
Militello, M. C.–(2) 828
Millecchia, M. R.–(4) 1654
Miller, D. L.–(3) 1180, 1190, 1285
Mishra, S.–(4) 1884
Mishra, U. K.–(1) 29
Misra, V.–(4) 1806, 1836
Mitchell, William J.–(4) 1622
Mitsushima, Yasuichi–(4) 1585
Mittereder, J.–(3) 1022
Miyano, S.–(2) 567
Miyata, Noriyuki–(3) 978
Miyazaki, Seiichi–(2) 306
Mizuno, Seigi–(2) 620
Moeller, W.–(2) 855
Moll, N.–(4) 1612
Möller, H.–(2) 474
Möller, W.–(2) 832
Molnar, R. J.–(3) 1195
Mönch, Winfried–(4) 1867
Mondol, M.–(3) 1171
Monget, C.–(4) 1406
Monguchi, Toshihide–(2) 871
Monier, C.–(3) 1158
Monteiro, Othon R.–(3) 1094
Monti, Fulvia–(4) 1674
Moon, H. S.–(3) 1017
Moon, Jong Hyun–(1) 241
Moon, Sung–(1) 246
Moore, C. D.–(3) 1163
Morin, Roger–(4) 1390
Mosca, J. J.–(3) 1131
Moser, B. G.–(3) 1139, 1185
Mu, J.–(3) 1185
Mühle, R.–(3) 945
Müller, G.–(2) 670, 696
Mulvaney, James M.–(2) 888
Murad, Saad K.–(1) 113
Murashita, Tooru–(1) 22
Murata, Kazuya–(3) 1098
Murphy, D. W.–(3) 1294
Murphy, M. J.–(3) 1252
Murphy, R. A.–(2) 773
Muzykov, P. G.–(2) 769
Myers, T. H.–(4) 1654

Na, Myunghee–(1) 224
Nagao, M.–(2) 604, 623
Nakada, Yoshinobu–(1) 1; (4) 1525
Nakagiri, Nobuyuki–(4) 1605
Nakajima, Anri–(4) 1317
Nakamura, T.–(2) 288, 494
Nakane, Hideaki–(2) 620
Nakanishi, Y.–(1) 237
Nakao, Hiroshi–(4) 1317
Nakasuji, Mamoru–(4) 1395
Nakatani, T.–(2) 604
Nalamasu, O.–(1) 101
Nastasi, M.–(2) 818, 822; (4) 1329
Natarajan, K.–(3) 1227
Nau, D.–(2) 670, 696
Nejoh, H.–(4) 1323
Nelson, M. W.–(4) 1354
Nemani, A.–(3) 1205
Nemanich, R. J.–(2) 700, 734

Neudeck, Gerold W.–(3) 994
Newman, F.–(3) 1158
Newport, Anne–(2) 750
Ngai, T.–(2) 460
Nicolaescu, D.–(2) 520, 788
Nicolaescu, Dan–(2) 542
Nicoll, C.–(4) 1728
Nie, H.–(3) 1124, 1175
Niedermann, Ph.–(4) 1570
Niemeyer, J.–(1) 230
Nihei, Masayasu–(3) 1028
Nii, K.–(3) 1167
Nii, Keisuke–(4) 1482
Niimi, H.–(4) 1806, 1813, 1831, 1836
Nishii, Katsunori–(2) 320
Nishikawa, Kazuyasu–(1) 127
Nishikawa, O.–(2) 608
Nitta, S. V.–(1) 205
Nomura, Eiichi–(3) 933
Nomura, Setsuo–(1) 82
Nordell, N.–(3) 1040
Nordheden, K. J.–(1) 138
Nordheden, Karen J.–(3) 961
Nosho, B. Z.–(4) 1786
Nuesca, Guillermo–(3) 1101

Obraztsov, A. N.–(2) 674
Ochiai, Yukinori–(3) 933
Oehrlein, G. S.–(4) 1435
Ogino, T.–(4) 1346
Oh, Myung Hwan–(1) 241, 246; (2) 486
Ohfuji, Shin-ichi–(1) 201
Ohta, K.–(3) 1127
Ohtani, Y.–(2) 608
Oka, Masayoshi–(2) 620
Oka, T.–(2) 288
Okada, Y.–(3) 1105
Okamoto, A.–(2) 567
Okamoto, Nariaki–(4) 1516
Okano, K.–(2) 557
Oki, A. K.–(3) 1136
Okuda, Soichiro–(2) 575
Okumura, Hajime–(1) 1; (4) 1525
Okuyama, F.–(2) 520; (4) 1602
Okuyama, Fumio–(2) 542
Olbrich,[a)], A.–(4) 1570
Olkhovets, A.–(4) 1366
Ono, Kouichi–(1) 127
Onoda, Hiroshi–(4) 1464
Onuki, Jin–(3) 1028
Oomori, Tatsuo–(1) 127
Opila, R. L.–(3) 1294
Osipowicz, T.–(4) 1507
Ouellet, L.–(2) 423
Oura, K.–(3) 983
Overgaard, C.–(1) 49

Pae, Sangwoo–(3) 994
Palange, E.–(2) 465
Palmstrøm, C. J.–(4) 1877
Pan, M.–(4) 1666
Pang, S. W.–(4) 1336
Parameswaran, L.–(2) 773
Parascandola, S.–(2) 855
Park, Cheol-Min–(2) 635
Park, D.–(3) 1022
Park, Jin-Won–(1) 213
Park, Jong-Wan–(4) 1598
Park, M.–(2) 700, 734
Park, N. S.–(2) 744
Park, Yong Jo–(2) 497
Park, Youngjun–(2) 547
Parker, D. L.–(2) 601
Parkinson, B. A.–(4) 1354
Parks, D.–(4) 1702
Partin, D. L.–(3) 1267
Pashkin, Yu. A.–(4) 1413
Paskova, T.–(3) 1040
Passlack, M.–(1) 49
Pavlovsky, I. Yu.–(2) 674
Pavolotsky, A. B.–(1) 230
Paynter, R. W.–(2) 859
Pearce, R.–(2) 477
Pearton, S. J.–(2) 366; (4) 1540
Peckham, E. L.–(4) 1773
Peiris, F. C.–(3) 1214
Pekola, J. P.–(4) 1413
Pelletier, J.–(1) 118; (2) 879
Penn, Cecil W.–(1) 233
Pennelli, Giovanni–(2) 468
Perlin, Piotr–(3) 1272
Peterson, Gregory–(3) 1101
Petrich, G. S.–(3) 1171
Pham, F.–(2) 477
Phillips, A. M.–(4) 1424
Phillips, J.–(3) 1116, 1276
Phillips, J. C.–(4) 1803, 1806
Pinnington, T.–(4) 1728
Piquette, E. C.–(3) 1241; (4) 1750
Pisupatti, V.–(1) 205
Pivac, B.–(1) 44
Pivin, D.–(4) 1757
Plavitu, C. N.–(2) 520
Plawsky, J. L.–(1) 205
Pletcher, T.–(4) 1773
Ploog, K. H.–(4) 1716
Ponnekanti, H.–(2) 477
Popovici, D.–(2) 859
Postigo, P. A.–(3) 1281
Po-Tao Chu, Paul–(1) 145
Powell, C. J.–(4) 1702
Prabhakaran, K.–(4) 1346
Preble, D.–(2) 345
Preble, E. A.–(2) 632
Prechtl, G.–(4) 1722
Preinesberger, C.–(4) 1682
Preobrazhenskii, V.–(4) 1738
Pristovsek, M.–(4) 1691
Probert, Simon J.–(2) 719
Proffitt, Simon S.–(2) 719
Prozesky, V. M.–(2) 808
Ptak, A. J.–(4) 1654

Qasaimeh, O.–(3) 1116
Qin, L.–(3) 1289
Queeney, K. T.–(4) 1795
Quintana, J. A.–(3) 1085

Rack, M. J.–(4) 1840
Raghavachari, K.–(4) 1795
Raiko, V.–(2) 696
Rakvin, B.–(1) 44
Rao, Valluri R.–(1) 40
Raoux, S.–(2) 477
Rasmussen, D. H.–(2) 443
Rastogi, A. C.–(4) 1549
Ray, O.–(3) 1180
Reifenberger, R.–(4) 1773
Reiss, G.–(4) 1350
Ren, C. X.–(2) 311
Ren, F.–(4) 1540
Ren, Hong-Wen–(4) 1341
Reynolds, Geoffrey W.–(2) 334
Rhee, Shi-Woo–(1) 154
Richter, E.–(2) 832, 836, 855
Richter, W.–(4) 1691
Richter-Sand, R. J.–(2) 883
Rider, Keith B.–(3) 1080
Rimini, F.–(4) 1448
Ringel, S. A.–(3) 1289
Rinne, C. Lizzul–(2) 734
Ripin, D. J.–(3) 1171
Rizzi, Angela–(4) 1674
Roberts, C. J.–(2) 273
Robertson, J.–(2) 659
Roche, M.–(2) 879
Rodionov, A. N.–(2) 647
Roe, K. J.–(3) 1301
Roedel, Ronald J.–(4) 1561
Rogers, D.–(4) 1323
Rösch, M.–(4) 1738
Ross, G. G.–(2) 859
Rothschild, M.–(3) 1045
Rowe, J. E.–(4) 1831
Rubloff, Gary W.–(4) 1417
Rujirawat, S.–(3) 1205
Rusalsky, D. P.–(2) 836
Ryan, R. W.–(3) 905
Ryu, J. T.–(3) 983

Sacks, R. N.–(3) 1289
Saito, N.–(1) 237
Saka, M.–(4) 1380
Sakakibara, Y.–(4) 1361
Sakashita, T.–(2) 604
Sala, Fabio Della–(4) 1674
Salamo, G. J.–(1) 253
Salas, R., Jr.–(3) 1263
Saleh, A.–(3) 1195
Salmeron, Miquel–(3) 1080
Samarth, N.–(3) 1147
Samavedam, S. B.–(4) 1424
Sandhu, R. S.–(3) 1163
Sano, Masanori–(2) 840
Santos, M. B.–(3) 1151
Sanz-Maudes, J.–(2) 397
Saranin, A. A.–(3) 983
Sari, Huseyin–(4) 1761
Sarkissian, A. H.–(2) 859
Sarma, S. Das–(4) 1708
Sassella, A.–(1) 44
Sato, F.–(1) 237
Sato, M.–(4) 1602
Sato, Taketomo–(4) 1856
Savage, S.–(3) 1040
Sawada, K.–(1) 237
Sawai, Mikio–(1) 109
Sawin, H. H.–(4) 1589
Scansen, D.–(1) 73
Schaff, W. J.–(3) 1252
Schenk, Andreas–(4) 1823
Scherer, H.–(1) 230
Scherer, R.–(4) 1728
Schetzina, J. F.–(4) 1884
Schinagl, F.–(4) 1722
Schippan, F.–(4) 1716
Schlaf, R.–(4) 1354
Schlapbach, L.–(2) 666; (3) 1064
Schlesser, R.–(2) 734
Schmid, M.–(4) 1697
Schmidt, M. A.–(4) 1589
Schmidt, W. G.–(4) 1691
Schroeder, M. J.–(1) 90
Schroeder, P. G.–(4) 1354
Schuelke, T.–(2) 867
Schultz, J. A.–(3) 1209
Scott, D.–(3) 1185
Sealy, B. J.–(2) 596
Seamons, M.–(2) 477
Sebastian, J. S.–(2) 758
Sekine, T.–(2) 608
Semyagin, B.–(4) 1738
Serdiukova, I.–(3) 1158
Serena, P. A.–(2) 788
Sergent, A. M.–(3) 1294
Serpa, G.–(4) 1702
Shalish, I.–(1) 166
Shanabrook, B. V.–(4) 1786
Shan Hill Wong, David–(1) 145
Shannon, J. M.–(2) 596
Shapira, Yoram–(1) 166
Sharma, R.–(2) 460
She, J. C.–(2) 592
Shealy, J. R.–(3) 1252
Shen, Y.-L.–(2) 448
Sheng, Xia–(3) 1076
Shi, X. H.–(3) 1205
Shi, Xuelong–(2) 350
Shi, Z.–(3) 1297
Shim, Jae Yeob–(2) 728, 731
Shimizu, Hiroyasu–(4) 1395
Shimomura, S.–(3) 1127, 1167
Shimomura, Satoshi–(3) 1155
Shimomura, Satoshi–(4) 1482
Shin, I. H.–(2) 690
Shpuniarsky, D.–(2) 828
Shu, L.–(2) 845
Shul, R. J.–(2) 366; (3) 965; (4) 1540
Silva, S. R. P.–(2) 557, 596
Silver, Jack–(2) 750
Silvetti, D.–(2) 477
Simko, S. J.–(2) 828
Simmel, F.–(4) 1594
Sin, Johnny K. O.–(2) 570
Singh, J.–(3) 1276
Singh, Vivek K.–(4) 1371
Sitar, Z.–(2) 734
Sitter, H.–(4) 1697, 1722
Sivananthan, S.–(3) 1205
Smart, J.–(3) 1252
Smith, Bruce G.–(1) 233
Smith, Christopher J. M.–(1) 113
Smith, Henry I.–(3) 1171
Smith, J. A.–(4) 1424
Smorchkova, I. P.–(3) 1147
Smy, T.–(1) 186
Sochtine, N. P.–(2) 765
Solntsev, V. A.–(2) 647
Somorjai, Gabor A.–(3) 1080
Song, H. Z.–(3) 1105
Sowers, A. T.–(2) 734
Specht, P.–(3) 1200
Spithoven, J. L.–(4) 1632
SpringThorpe, A. J.–(3) 1180, 1190
Standaert, T. E. F. M.–(4) 1435
Stansfield, B. L.–(2) 859
Stecki, A. J.–(2) 362; (3) 1056, 1195
Stefanov, B. B.–(4) 1795
Stewart, T. R.–(3) 1200
Stifter, D.–(4) 1697, 1722
Stillman, G. E.–(3) 1139, 1185
Stinespring, C. D.–(4) 1654
Streater, R. W.–(3) 1180, 1190
Streetman, B. G.–(3) 1124, 1175
Streit, D. C.–(1) 138; (3) 1136, 1163
Su, Tai-chi–(3) 994
Su, Yan-Kuin–(4) 1545
Subramanya, Sudhir–(3) 1272
Sudarshan, T. S.–(2) 769; (4) 1580
Sugie, H.–(4) 1602
Sugimura, Hiroyuki–(4) 1605
Sugino, Takashi–(2) 642, 723
Sugou, Shigeo–(4) 1341
Sullivan, G. J.–(4) 1742
Sullivan, John–(1) 182
Sumstine, M.–(3) 1205
Suntharalingam, V.–(2) 345
Suter, M.–(3) 945
Suwa, Masateru–(3) 1028
Suzuki, Eiichi–(1) 77
Suzuki, Yasuo–(2) 840
Swart, Hendrik–(2) 758
Swenson, D.–(2) 432

Tabata, H.–(2) 288, 494
Taferner, W. T.–(3) 1218
Tagawa, Shigeru–(2) 642
Takai, Mikio–(2) 575
Takai, Osamu–(4) 1605
Takami, Taichi–(2) 620
Takeda, T.–(4) 1380
Takemura, H.–(2) 567
Tamargo, M. C.–(3) 1255
Tanabe, Hiroyoshi–(1) 101
Tanabe, Mitsuru–(2) 320
Tanaka, Hiroyuki–(4) 1313
Tanaka, K.–(2) 608
Tanaka, T.–(2) 477
Tanaka, Yoichiro–(2) 416
Tanenbaum, D. M.–(3) 1087
Tang, B. Y.–(2) 851, 875
Tang, S. H.–(2) 455
Tang, Yongming–(2) 570
Tang, Z.–(3) 1139
Tani, H.–(3) 983
Tani, T.–(4) 1361
Tatarenko, N. I.–(2) 647
Tate, A.–(3) 905
Tatsuoka, Y.–(3) 1127
Tatsuoka, Yasuaki–(3) 1155
Taylor, James W.–(2) 334
Taylor, W. J.–(4) 1424
Tedder, Laura L.–(4) 1417
Tendler, S. J. B.–(2) 273
Teng, J. H.–(4) 1507
Terreault, B.–(2) 859
Theirich, D.–(2) 696
Thibado, P. M.–(1) 253; (4) 1778
Thomes, W. J.–(2) 758
Thornton, T. J.–(4) 1757
Thrush, C. M.–(3) 1267
Tian, X. B.–(2) 851, 875
Tian, Y.–(2) 688
Tiedje, T.–(4) 1728
Tilke, A.–(4) 1594
Timko, A. G.–(1) 101
Tobin, P. J.–(4) 1424
Tomich, D. H.–(4) 1852
Tomizawa, Kazutaka–(1) 77
Tong, Linsu–(2) 570
Torabi, A.–(3) 1131
Torri, P.–(4) 1329
Torrioli, G.–(2) 500
Towner, F. J.–(3) 1200
Trampert, A.–(4) 1716
Tremblay, Y.–(2) 423
Trottier, T. A.–(2) 758
Trusso, Sebastiano–(2) 468
Tsai, Jung-Hui–(4) 1477
Tsakalis, K. S.–(3) 1237
Tsakalis, Kostas S.–(3) 1223
Tseng, W. T.–(4) 1456
Tsuji, H.–(2) 604, 623
Tsutsumi, Toshiyuki–(1) 77
Tu, C. W.–(4) 1649
Turner, Anthony–(4) 1561
Tuszewski, M.–(2) 822
Twigg, M. E.–(3) 1022

Uchida, Hidetsugu–(2) 355
Uchimura, K.–(2) 494
Uda, Tomoya–(2) 320
Uehara, Y.–(1) 12
Ueng, H. J.–(4) 1773
Ueno, Hiroaki–(4) 1317
Uglov, V. V.–(2) 836
Ura, T.–(2) 623
U'Ren, G. D.–(3) 1163
Ushioda, S.–(1) 12
Utsumi, K.–(2) 604

Valcheva, E.–(3) 1040
Vallés-Abarca, J. A.–(3) 1085
Vancea, J.–(4) 1570
Vandré, S.–(4) 1682
Vanhaelemeersch, S.–(2) 372
Van Hove, J. M.–(4) 1540
Vanoverschelde, A.–(4) 1553
Vantomme, A.–(4) 1502
Vasi, Cirino–(2) 468
Vasileska, D.–(4) 1840
Vawter, G. A.–(2) 366
Vecht, Aron–(2) 750
Velicu, S.–(3) 1205
Venkat, R.–(3) 1227
Vigneron, J.-P.–(2) 506
Vilela, M. F.–(3) 1158
Villalvilla, J. M.–(3) 1085
Villeneuve, P. R.–(3) 1171
Vinogradov, Georgy K.–(1) 95
Visbeck, S.–(4) 1691
Vogel, M.–(4) 1594
Vogl, P.–(4) 1617
Vogt, P.–(4) 1691
Voigtländer, Bert–(2) 294
Volkov, A. P.–(2) 674
von Blanckenhagen, P.–(2) 265
von Känel, H.–(4) 1848

Wada, Satoshi–(4) 1516
Wahab, Q.–(3) 1040
Walker, D. E., Jr.–(4) 1852
Wallace, Robert M.–(2) 303; (3) 970
Walsh, B. L.–(4) 1773
Walter, K. C.–(2) 818, 822
Walters, G. K.–(3) 1087
Wamsley, C. C.–(3) 1246, 1249
Wang, Baoping–(2) 570
Wang, Chen–(2) 570
Wang, Ching-Wu–(4) 1545
Wang, E. G.–(3) 1059
Wang, J. Y.–(2) 684
Wang, Jing–(4) 1567
Wang, M. T.–(2) 385
Wang, S. Y.–(2) 851, 875
Wang, Tairan–(4) 1612
Wang, W.–(2) 688; (4) 1507
Wang, W. B.–(2) 684
Wang, W. I.–(3) 1144
Wang, X.–(2) 311, 688
Wang, X. F.–(2) 851, 875
Wang, Y. L.–(4) 1456
Wang, Z.–(4) 1836
Wassermeier, M.–(4) 1791
Watanabe, Heiji–(3) 978
Watanabe, M.–(2) 608
Wayner, P. C., Jr.–(1) 205
Weber, E. R.–(3) 1200
Weber, Eicke R.–(3) 1272
Wei, Lei–(4) 1575
Wei, Yi–(1) 233; (3) 970
Weigold, J. W.–(4) 1336
Weimann, Th.–(1) 230
Weinberg, W. H.–(4) 1786
Weldon, M. K.–(4) 1795
Westwood, D. I.–(4) 1659
Whang, C. N.–(2) 380
Wharam, D. A.–(4) 1594
Whelan, C. S.–(1) 194; (2) 410
Whitfield, Michael D.–(2) 719
Whitman, L. J.–(4) 1786
Wicks, G. W.–(3) 1246, 1249
Wiebicke, E.–(4) 1791
Wieder, H. H.–(4) 1761
Wiemer, M.–(4) 1733
Wilkinson, Christopher D. W.–(1) 113
Wilks, S. P.–(4) 1666
Williams, P. M.–(2) 273
Williams, R. H.–(4) 1666
Willison, C. G.–(3) 965
Willson, C. Grant–(4) 1371
Wilshaw, P. R.–(2) 580
Wilt, David–(3) 1289
Windt, David L.–(3) 930; (4) 1385
Winer, Paul–(1) 40
Wisitsora-at, A.–(2) 740
Witke, T.–(2) 867
Wohlert, D. E.–(3) 1120
Wojak, G. J.–(2) 632
Wojtowicz, M.–(3) 1163
Woo, S. H.–(3) 1017
Wood, Troy D.–(1) 224
Woodall, J. M.–(4) 1767, 1773
Wortman, J. J.–(4) 1836
Wu, H.–(3) 1252
Wu, H. Z.–(3) 1263, 1297
Wu, Kehui–(3) 1059
Wu, Lan–(2) 269
Wu, M. F.–(4) 1502
Wu, S. S.–(2) 562
Wu, W. F.–(3) 939
Wu, Y.–(4) 1806, 1813, 1836
Wyatt, P. W.–(2) 345

Xia, L.-Q.–(2) 477
Xie, Tianbao–(2) 613
Xie, Z. W.–(2) 845
Xin, H. P.–(4) 1649
Xing, H.–(3) 1255
Xu, F.–(2) 740
Xu, J. F.–(2) 311
Xu, N. S.–(2) 562, 592, 710; (3) 1059
Xu, S. H.–(4) 1884
Xu, Zheng–(2) 416

Ya Detkov, P.–(2) 715
Yakimova, R.–(3) 1040
Yamamoto, Ei-ichi–(1) 201
Yamauchi, Atsushi–(1) 109
Yanagihara, Manabu–(2) 320
Yang, B.–(3) 1205
Yang, C. J.–(2) 392
Yang, H.–(4) 1778, 1813
Yang, H. S.–(3) 1017
Yang, L. W.–(1) 138
Yang, Su-Hua–(3) 1072
Yang, X.–(3) 1144
Yang, Y.–(4) 1884
Yang, Zhonghai–(1) 250
Yao, Shude–(4) 1502
Yaradou, O.–(4) 1553
Yen, H. C.–(1) 138
Yeo, W.–(3) 1252
Yew, J. Y.–(3) 939
Yi, W. K.–(2) 489
Yokota, Yuuko–(2) 642
Yokoyama, Meiso–(3) 1072; (4) 1545
Yokoyama, Naoki–(4) 1317
Yoo, Hyung Mo–(3) 1034
Yoon, Dong-Soo–(1) 174; (4) 1470
Yoon, Euijoon–(1) 213
Yoon, H.-S.–(2) 380
Yoon, Young Joon–(2) 627
Yoshida, Tomoyuki–(4) 1585
Yoshiki, M.–(2) 567
Yoshino, Hiroshi–(1) 101
Yoshioka, Kenji–(2) 871
Young, Robert M.–(2) 280
Yu, E. T.–(4) 1742, 1781
Yu, Z.–(1) 49
Yu, Z. X.–(2) 562
Yuan, G.–(2) 684, 688
Yukimura, Ken–(2) 840, 871
Yun, Minhee–(4) 1561

Zahn, D. R. T.–(4) 1733, 1738
Zandler, G.–(4) 1617
Zanotti, L.–(1) 44
Zavada, J. M.–(3) 1195
Zeng, Baoqing–(1) 250
Zeng, L.–(3) 1255
Zeng, Z. M.–(2) 851, 875
Zhang, A. P.–(4) 1540
Zhang, G. Y.–(4) 1502
Zhang, Haijun–(2) 269
Zhang, J.–(4) 1309
Zhang, L.–(3) 965; (4) 1540
Zhang, Ning–(3) 1050
Zhang, Y. H.–(3) 1237
Zhang, Y.-H.–(2) 474; (3) 1233
Zhao, H.–(2) 688
Zhao, J.–(2) 311
Zhao, Jin–(1) 224
Zhao, R.–(3) 1200
Zhen, X. G.–(2) 710
Zhirnov, V. V.–(2) 632, 666, 670, 700, 715, 734
Zhong, G.–(2) 405
Zhong, Wei–(3) 1050
Zhou, J. Y.–(2) 311
Zhou, T.–(2) 688
Zhu, Y. K.–(2) 311
Zhu, Z. Q.–(4) 1649
Ziemer, K. S.–(4) 1654
Zorin, A. B.–(1) 230
Zorn, M.–(4) 1691
Zotov, A. V.–(3) 983
Zuo, S. L.–(4) 1781
Zvanut, M. E.–(4) 1627

JOURNAL OF VACUUM SCIENCE AND TECHNOLOGY
INFORMATION FOR CONTRIBUTORS

1. Articles. This journal publishes two types of articles: original research reports comprising a complete, detailed, self-contained description of a research effort and Review papers that provide an updating of an area of science or technology pertinent to the scope of this journal. All articles must contain an abstract (see item **5**).

2. Brief Reports and Comments. These are short submissions that either report a brief item of research or technology that does not require extensive descriptive writing or that comment on an item previously published in the journal. Brief Reports and Comments do not require the format structure of regular journal articles and do not require an abstract.

3. Rapid Communications. These are short submissions which contain important new results and are intended for accelerated publication. Rapid Communications are given priority in handling and do not require an abstract.

4. Shop Notes. 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. An abstract is not required for a Shop Note.

5. Abstracts. An *abstract must accompany each article.* The abstract should be typed (double-spaced) *on a separate page.* It should be adequate as an index and as a summary. It should give all subjects, major and minor, concerning which new information is presented. It should give the conclusions of the article and all numerical results of general interest. An abstract is usually reprinted verbatim in abstract journals. Therefore, great care should be used in writing it.

6. Manuscripts. Manuscripts are accepted with the understanding that they have not been copyrighted, published, or accepted for publication elsewhere. Manuscripts must be in English, typewritten, double-spaced on one side of the page, on $8\frac{1}{2}\times 11$ in. durable, opaque, white paper. *An original and two* (2) *copies must be submitted.* Wide margins should be provided to permit editorial instructions to the printer. *Authors should adopt a particular way of writing their names in the "by-line" and use the same name for all publications.* This practice makes indexes more useful and less confusing. For submission of electronic files for publication see item **15** below.

7. Style. Authors are expected to follow the conventional writing, notation, and illustration style prescribed in the American Institute of Physics *Style Manual* which is available from the AIP. For ordering information, visit http://www.aip.org/epub/styman.html. Authors should also study the form and style of printed material in this journal. The SI units should be used.

8. Illustrations: Illustrations published in the journal are either scanned by AIP using a digital scanner or received electronically from the author, and integrated with the text of the article, creating completely electronic pages. To receive optimal quality, we strongly encourage you to send electronic graphics files to AIP, rather than laser output. (Note: If you are submitting electronic graphics files, you are still required to send hardcopies of the figures to the Editorial Office. Adherence to electronic submission instructions is crucial. If your electronic files are received after AIP has already processed the hardcopy illustrations, the electronic files will not be used.) Please adhere to the following guidelines when preparing your illustrations for submission:

Sizing Illustrations (Electronic Graphics Files and Hardcopy)

- **Prepare illustrations in the final published size, not oversized or undersized.** Size your illustrations according to your journal's specifications. Submit each illustration at the final size in which it will appear in the journal. The standard is 8.5-cm maximum width (3-3/8″ or 20.5 picas) for one column. This is especially important for screened or shaded illustrations; reduction of screened/shaded originals during the digitizing process introduces an unacceptable moiré pattern.
- **Ensure a minimum of 8-point type size (2.8 mm high; 1/8″ high) and 1-point line width within illustrations.** Ensure that line weights will be **0.5 points** or greater in the final published size. **Line weights below 0.5 points will reproduce poorly.** Avoid inconsistencies in lettering within individual figures, and from one figure to the next. Lettering and symbols cannot be handwritten. Avoid small open symbols that tend to fill in if any reduction is necessary.

Preparation of Hardcopy Illustrations for Scanning

- **Number figures in the order in which they appear in text.** Label illustrations with their number, the name of the first author, and the journal, on the **front** of the figure well outside the image area.
- **Place only one figure per page (including all parts).** Place all parts of the same figure on one sheet of white bond paper, spaced 1/4 in. apart, using a glue stick or wax on the back of the illustration and leaving a 2-in. bottom margin. Label all figure parts with (a), (b), etc. Make sure each figure is straight on the page. *Photocopies of artwork are not acceptable.*
- **Do not use correction fluid or tape on illustrations.** The scanner is extremely sensitive and reproduces all flaws (e.g., correction fluid, tape, smudges, dust). *Do not write on the back of the figure* because it will be picked up by the scanner.
- **Authors' laser-generated graphics are acceptable only if the lettering and lines are dark enough, and thick enough, to reproduce clearly,** especially if reduction is required. Maximum black–white contrast is necessary. Choose a laser printer with the highest dot-per-inch (dpi) available (i.e., the highest resolution possible). Remember that fine lines in laser-generated graphics tend to disappear upon reduction, even if the oversized original looks acceptable.
- **Submit continuous-tone photographs in final published size on white glossy or matte paper.** Avoid glossy paper stock that is off-white, ivory, or colored because contrast within the illustration will be lost in reproduction. Print the photograph with more contrast than is desired in the final printed journal page. Avoid dull, textured paper stock, which will cause illustrations to lose contrast and detail when reproduced.

Preparation of Electronic Graphics Files

- We recommend that all halftone art (screened art), shaded figures, and combinations (line art + halftone) be submitted electronically. Computer-generated illustrations output to desktop printers produce a screen. These figures are most problematic in the scanning process, because scanning screened output produces an unacceptable moiré pattern.
- **Acceptable formats:** Graphics must be submitted as PostScript, EPS (using either Arial or Times Roman fonts), or TIFF (lzw compressed). Do not send application files, e.g., Corel Draw, etc.
- **Settings:** Set the graphic for **600 dpi** resolution for line art, **264 dpi** for halftones (noncompressed), and **600 dpi** for combinations (line art + halftone). Save the files to grayscale (B/W), not color.
- Make sure there is only **ONE** figure per file. Each figure file should include all parts of the figure. For example, if Figure 1 contains three parts (a, b, c), then all of the parts should be combined in a single file for Figure 1.
- You are still required to send hardcopies of all figures to the Editorial Office, along with a hardcopy of the manuscript.
- Do not FTP the graphics files to the Editorial Office or AIP unless otherwise instructed to do so. Full instructions will be sent to you twice: Once on a hardcopy form after initial submission of your article, and once via e-mail after your article has been accepted for publication.

Detailed instructions for submitting electronic graphics to AIP and a glossary of terms may be found on the AIP Physics Information Netsite at www.aip.org/epub/submitgraph.html

9. Mathematics. Mathematical expressions should be typewritten as completely as possible, with unavailable symbols being carefully inserted in ink. Special care should be given to make equations and formulas clear to the typesetter, and all handwritten letters and symbols that cannot be typed should be identified *in the margin and only the first time that such occur.* Clearly distinguish capital and lowercase letters where there could be confusion. Use fractional exponents to *avoid root signs.* Extra symbols should be introduced to avoid complicated exponents or where it is necessary to repeat a complicated expression a number of times. Use the solidus (/) wherever possible for fractions. Do not repeat mathematical derivations that are easily found elsewhere in the literature; merely cite the references.

10. References. *References should be collected and typed (double-spaced) on a separate page.* A list of the abbreviations for the names of journals appears in the *Style Manual.* They should be numbered consecutively and arranged thus:

[1]A. B. Smith, Phys. Rev. **41**, 852 (1932).
[2]H. Lamb, *Hydrodynamics,* 6th ed. (Cambridge University Press, Cambridge, England, 1940), pp. 573, 645.

11. Tables. All but the simplest tabular material should be organized into separate tables. Tables should be numbered with Roman numerals and typed on sheets at the end of the running text. Each table must have a caption typed at the top of the table, which makes the data in the table intelligible without reference to the text. Avoid complicated column headings. If necessary, use symbols which are explained in the caption. Type a double horizontal line below the caption, a single line below the headings, and another double line at the end of the table. Long tables should, if possible, be submitted in a form ready for direct photoreproduction. Detailed instructions for preparing camera-ready tables are available from the Editorial Office.

12. Publication Charges comprise the major source of income for this journal. Consequently, author's institutions or companies are asked *and urged* to pay a publication charge—which entitles them to 100 reprints without covers—of $95 per printed page.

13. Correspondence and Alterations. *Send manuscripts to the Editorial Office, Journal of Vacuum Science and Technology, 10 Park Plaza, Suite 4A, Caller Box 13994, Research Triangle Park, North Carolina 27709-3994;* telephone: (919) 361-2787; FAX: 919-361-1378; E-mail: jvst@jvst.org. Manuscripts to be published in special issues (papers from conferences) should be sent to the Guest Editor appointed for the particular Conference. Please do not send the Editor any correspondence about proofs, reprints, or publication charges.

If your paper is accepted for publication, you will be notified by the Editor. You will also be notified when your paper is sent to the American Institute of Physics. *All* subsequent correspondence about your paper should be addressed to the Editorial Supervisor, JVSTA *or* JVSTB, American Institute of Physics, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502. Reference must be made to the author, journal, and scheduled date of issue.

14. Proofs of Articles will be sent to you and should *be returned promptly to the Editorial Supervisor, JVSTA or JVSTB, AIP.* A few alterations in proof are unavoidable, but the cost of making extensive alterations or of correcting mistakes caused by careless preparation of the manuscript will be charged to the author. *Unless corrected proofs are returned punctually, your paper will be published in uncorrected form and will be so labeled.*

15. Compuscripts. AIP is accepting author-prepared computer files for use in production. If you have used REVTeX, LaTeX, Microsoft Word, or WordPerfect to compose your manuscript, AIP may be able to use your file to produce author proofs. If you are interested in participating in the program, please indicate so in the cover letter that accompanies your original submission. Please also include an electronic mail address if available. Do not send a disk with your original submission. Wait for further instructions from the editorial office on when to send the disk (or, if the editorial office accepts electronic submissions, when to e-mail the file). Do not send the file directly to AIP Production unless otherwise requested. AIP uses translation software to convert REVTeX, LaTeX, MS Word, or WordPerfect files into Xyvision composition files for production. For this project to be of benefit to both the author and the production staff, if is imperative that the guidelines as documented in either the REVTeX toolbox or the MS Word/WordPerfect toolbox be followed precisely. The REVTeX toolbox is available via anonymous FTP on the Internet from ftp.aip.org in the directory /pub/revtex. The MS Word/WordPerfect toolkit is also available via anonymous FTP from ftp.aip.org. Move to the directory /ems, then follow the instructions given on the screen. If electronic retrieval is not possible, you may receive the toolbox or toolkit on disk by contacting toolkits@aip.org. The REVTeX toolbox and the Word/WP toolkit, as well as general information regarding the compuscript program, are accessible via the AIP Physics Information Netsite (www.aip.org/epub/compuscripts.info.html). The Word/WP author toolkit FAQ ("Frequently Asked Questions") may be found at www.aip.org/pubservs/toolkit.html."

16. Electronic Mail. Authors may include their internet e-mail addresses along with all other footnotes in the following format: Electronic mail: smith@foo.bar.com

INDEX TO ADVERTISERS

CERAC Inc. A15
Hiden Analytical, Inc. A9
Innovac Corp. A1
International Radiation Detectors A4
Kurt J. Lesker Co. . . . . Cover 3
Leybold Inficon, Inc. . . . . Cover 2
LK Technologies A11
Omicron A9
Pfeiffer Vacuum A3
Quad Group, Inc. . . . . Cover 4
RBD Enterprises A17
Skion Corp. A11
Stanford Research Systems A7
Surface Science Spectra A16
VLSI Standards, Inc. A13

Advertising Sales Office

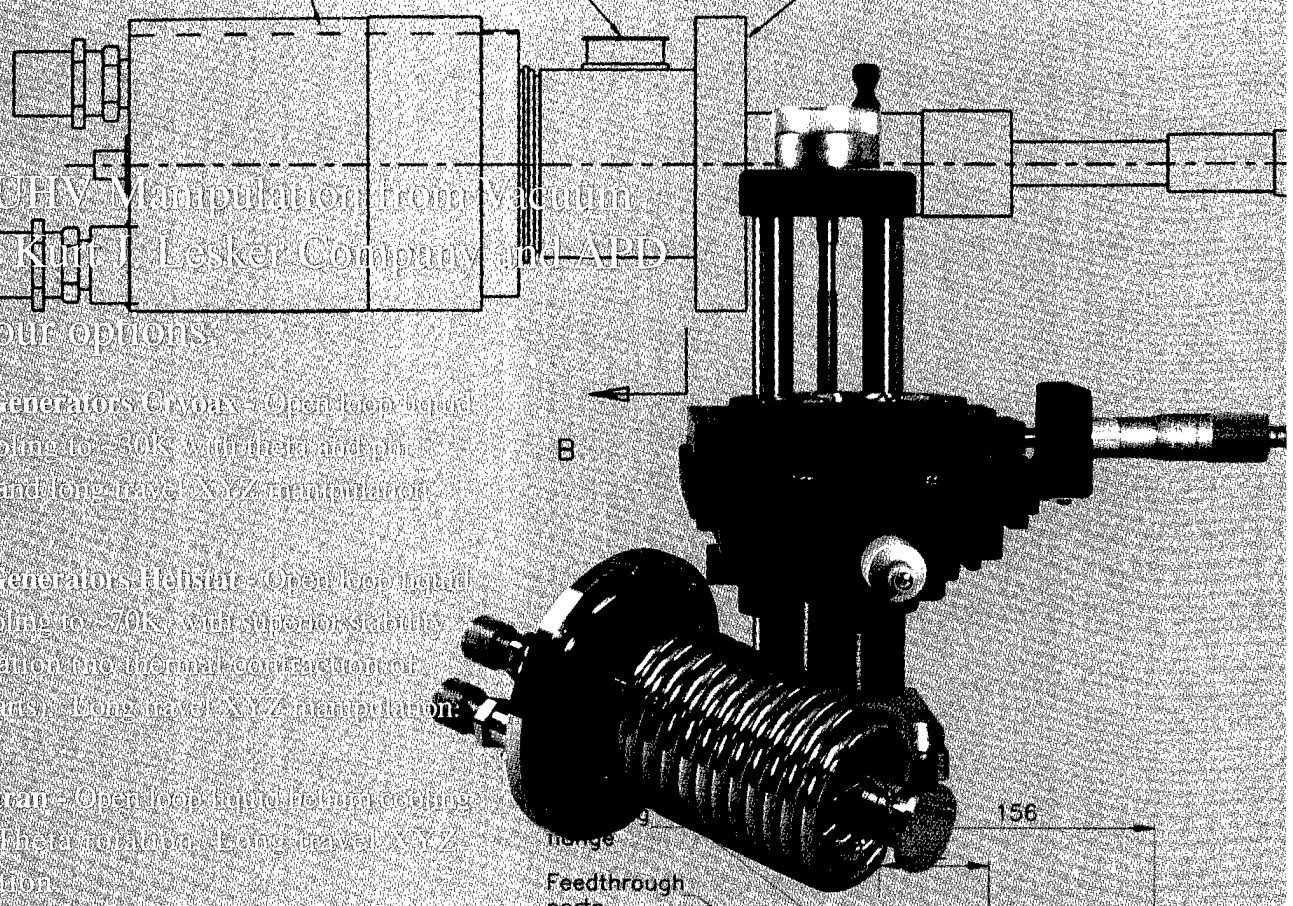
American Institute of Physics
2 Huntington Quadrangle
Suite 1N01
Melville, NY 11747--4502
Telephone (516) 576-2440
Fax (516) 576-2481

Advertising Manager: Richard T. Kobel

Advertising Sales Representative: Robert G. Finnegan

Production Manager: Betty Aroesty

Hey, Chill Out!



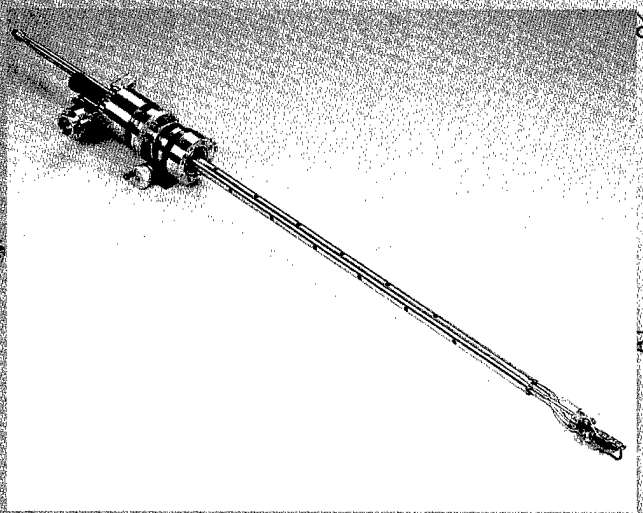
Cryogenic UHV Manipulation from Vacuum Generators, Kurt J. Lesker Company and APD.

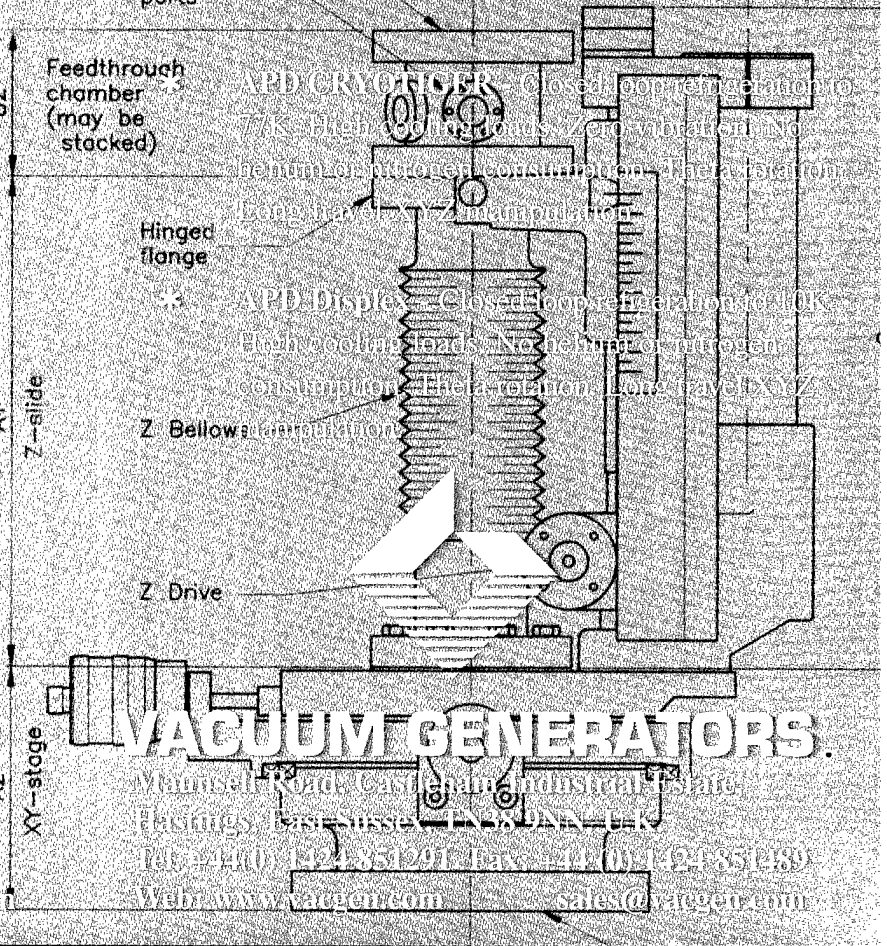
Examine your options:

* **Vacuum Generators Cryovac** - Open loop liquid helium cooling to ~30K with theta and phi rotations, and long travel XYZ manipulation.

* **Vacuum Generators Helisim** - Open loop liquid helium cooling to ~70K with superior stability on phi rotation due to thermal contraction of working parts. Long travel XYZ manipulation.

* **APD Helitran** - Open loop liquid helium cooling to 4.5K. Theta rotation. Long travel XYZ manipulation.





* **APD CRYOTRIDER** - Closed loop manipulation to 7.7K. Bi-directional bonds. Zero vibration. No helium consumption. Theta rotation. Long travel XYZ manipulation.

* **APD Displex** - Closed loop refrigeration to 10K. Bi-directional bonds. No helium consumption. Theta rotation. Long travel XYZ manipulation.

Kurt J. Lesker
Company

1515 Worthington Avenue, Clairton, PA 15025
Sales: (800) 245-1656 (412) 233-4200
Fax: (412) 233-4275 sales@lesker.com
international@lesker.com Web: www.lesker.com

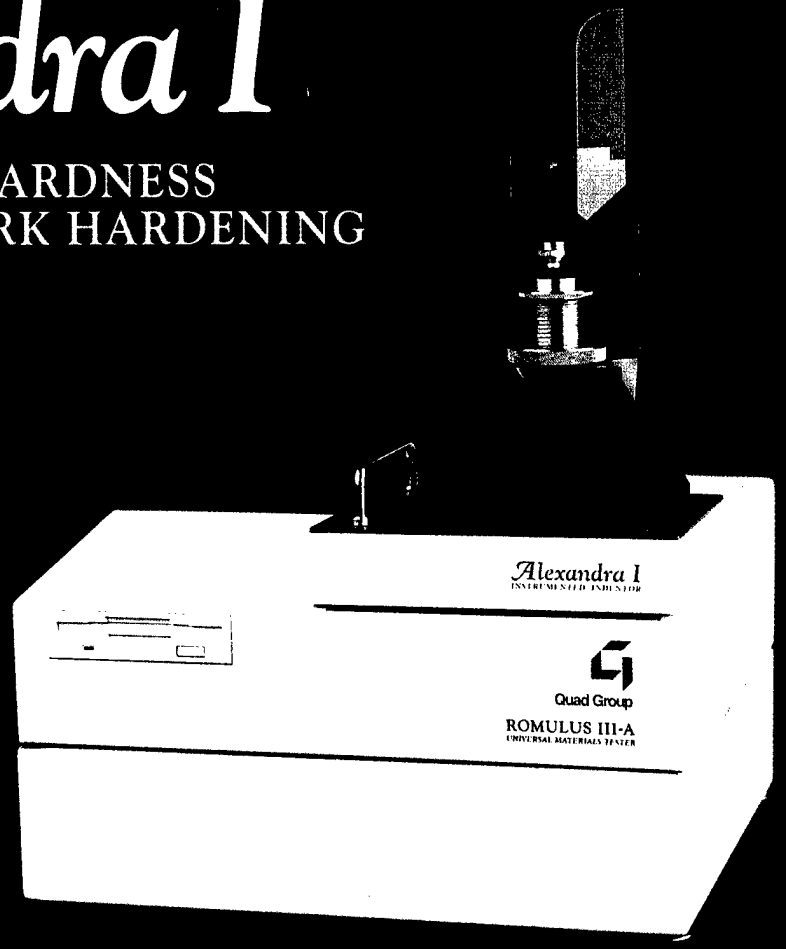
VACUUM GENERATORS

Manuel Road, Castleham Industrial Estate
Hastings, East Sussex TN38 9NA UK
Tel: +44(0) 1424 851291 Fax: +44(0) 1424 851489
Web: www.vacgen.com sales@vacgen.com

Instrumented Indentation

Alexandra I

YOUNG'S MODULUS HARDNESS
YIELD STRENGTH WORK HARDENING



Quad Group's *Alexandra I* [illegible] the hardness, yield strength, work hardening [illegible] of other parameters [illegible] of continuums [illegible] This is accomplished by [illegible] controlled indentation or a diamond into the sample's surface.

The ***Alexandra I*** is intended to characterize all materials of thickness above 5 micrometers and is especially valuable in determining the properties at various depths, such as with Ion-Beam Modified Materials, IBMM.

- Simply and quickly determine the mechanical properties associated with the TOUGHNESS of solid materials, or the capacity to absorb energy.
- Readily appraise the degradation of toughness (durability) as a function of exposure to various environments.
- Greatly reduce the amount of time involved in sample preparation and testing. ***Alexandra I*** requires only a flat sample specimen with a smooth surface.
- A single analysis can be accomplished in approximately 100 seconds. Statistically significant numbers of tests can be done in a few hours.
- Customize the loading and deloading schedule and continuously record load-displacement data.

"The *Alexandra I* is intended to characterize all materials of thickness above 5 micrometers"

MECHANICAL DESCRIPTION:

[illegible] III-A Universal Strength Tester. The [illegible] the indentor unit and the indention depth sensor. The indentor unit can be equipped with virtually any standard indenter shape, including a Vickers or Rockwell diamond. Stainless spheres measure plastics or durometer values.

Quad Group

Quad Group, Inc.
1815 South Lewis Street
Spokane, WA 99224
509-458-4558 / 800-342-2430
Fax: 509-458-4555
Email: quadgroup@spokane.net
www.spokane.net/quadgroup