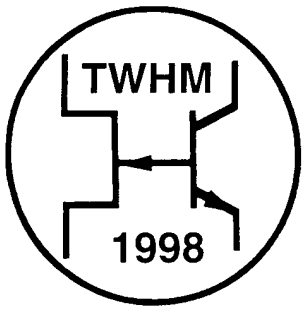
PROCEEDINGS OF THE TOPICAL WORKSHOP ON HETEROSTRUCTURE OF MICROELECTRONICS (TWHM'98)

30th August – 2nd September, 1998
Shonan Village Center, Hayama – Machi, Kanagawa, Japan

Proceedings of the Topical Workshop on Heterostructure Microelectronics



Editors:

Y. Arakawa
C.E. Hunt
F.M. Klaassen
S.J. Pearton

Guest Editors:

D. Pavlidis
T. Ishibashi
T. Mizutani
B. Bayraktaroglu





IEEE



DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The 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 the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
15/11/99	Proceedings	30 Aug - 2 Sep 98

4. TITLE AND SUBTITLE
"Proceedings of the Topical Workshop on Heterostructure of Microelectronics (TWHM '98)", 30 Aug - 2 Sep 98, Shonan Village Center, Hayama - Machi, Kanagawa, Japan.

5a. CONTRACT NUMBER
F6256298M9082 (Conference) - CSP-98-03)
F6256298M9196 (Publication)

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Editors: Y. Arakawa, C.E. Hunt, F.M. Klaassen, and S.J. Pearton

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Workshop Committee

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AOARD
UNIT 45002
APO AP 96337-5002

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
CSP-98-03

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

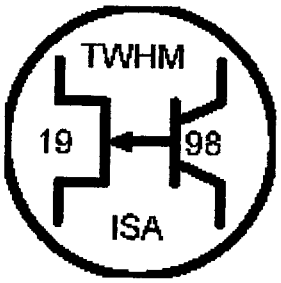
13. SUPPLEMENTARY NOTES

14. ABSTRACT
Topics Included in the Proceedings Include the Following::
"High-Speed and High-Frequency Devices"
"Novel Devices"
"Microwave/Millimeter Wave Devices"
"Wide Bandgap Devices"
"Circuit Applications, Device Fabrication and Characterization"
"Systems Applications"
"High-Speed ICS for Optical Communication Systems"
"Manufacturing"

15. SUBJECT TERMS

Semiconductor heterostructure growth, solid-state electronics, microwave power, opto electronic integrated circuits

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Ms. Joanne Maurice
UNCL	UNCL	UNCL	UL	1,663	19b. TELEPHONE NUMBER *(Include area code)* 81-3-5410-4409 (Tokyo, Japan)



SPECIAL ISSUE

PROCEEDINGS OF THE TOPICAL WORKSHOP ON HETEROSTRUCTURE MICROELECTRONICS

Shonan Village Center, Hayama-Machi, Kanagawa, Japan
30th August–2nd September, 1998











19991123 145

PERGAMON

DTIC QUALITY INSPECTED 4

AIMS AND SCOPE

It is the function of this Journal to bring together in one publication outstanding papers reporting original work in the following areas: (1) applications of solid-state physics and technology to electronics, including theory and design, measurement techniques, preparation of semiconductor devices, and also materials growth, measurement and evaluation; (2) the physics and modelling of submicron and nanoscale microelectronic devices, including methods of processing, measurement and evaluation; and (3) applications of numerical methods to the modelling and simulation of solid-state devices and processes.

Also of interest are the design and performance of power semiconductor devices, the physics and applications of compound semiconductors, and thermoelectric, ferroelectric and galvanomagnetic devices. Of equal interest are solid-state optical devices, including photoconductors and photodetectors, solar cells, electroluminescent devices, light emitters and lasers, and optoelectronic devices for the storage, transfer and processing of information.

Papers covering novel topics extending the frontiers of solid-state electronics are of course encouraged. Review papers covering important topics in the above areas will be presented at regular intervals. Papers will be published in the format of regular Research papers, Review papers and Letters.

It is expected that, because of its international scope and because it emphasizes the association of theory and practice, the Journal will advance further understanding of the relationship between solid-state physics and semiconductor device engineering.

© 1999 Elsevier Science Ltd. All rights reserved.

This journal and the individual contributions contained in it are protected under copyright by Elsevier Science Ltd, and the following terms and conditions apply to their use:

Photocopying

Single photocopies of single articles may be made for personal use as allowed by national copyright laws. Permission of the publisher and payment of a fee is required for all other photocopying, including multiple or systematic copying, copying for advertising or promotional purposes, resale, and all forms of document delivery. Special rates are available for educational institutions that wish to make photocopies for non-profit educational classroom use.

Permissions may be sought directly from Elsevier Science Rights & Permissions Department, PO Box 800, Oxford OX5 1DX, UK; phone: (+44) 1865 843830, fax: (+44) 1865 853333, e-mail: permissions@elsevier.co.uk. You may also contact Rights & Permissions directly through Elsevier's home page (http://www.elsevier.nl), selecting first 'Customer Support', then 'General Information', then 'Permissions Query Form'.

In the USA, users may clear permissions and make payments through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA; phone: (978) 7508400, fax: (978) 7504744, and in the UK through the Copyright Licensing Agency Rapid Clearance Service (CLARCS), 90 Tottenham Court Road, London W1P 0LP, UK; phone: (+44) 171 631 5555; fax: (+44) 171 631 5500. Other countries may have a local reprographic rights agency for payments.

Derivative Works

Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the publisher is required for resale or distribution outside the institution.

Permission of the publisher is required for all other derivative works, including compilations and translations.

Electronic Storage or Usage

Permission of the publisher is required to store or use electronically any material contained in this journal, including any article or part of an article.

Except as outlined above, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the publisher.

Address permissions requests to: Elsevier Science Rights & Permissions Department, at the mail, fax and e-mail addresses noted above.

Notice

No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

Although all advertising material is expected to conform to ethical (medical) standards, inclusion in this publication does not constitute a guarantee or endorsement of the quality or value of such product or of the claims made of it by its manufacturer.

Back Issues

Back issues of all previously published volumes are available direct from Elsevier Science Offices (Oxford and New York). Complete volumes and single issues can be purchased for 1990–1998. Earlier issues are available in high quality photo-duplicated copies as complete volumes only.

®™The text paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992.

PERGAMON

Solid-State Electronics 43 (1999) iii–vi

SOLID-STATE ELECTRONICS

SPECIAL ISSUE

PROCEEDINGS OF THE TOPICAL WORKSHOP ON HETEROSTRUCTURE OF MICROELECTRONICS

CONTENTS

Preface to special issue 1321

Committee 1323

HIGH-SPEED & HIGH-FREQUENCY DEVICES

J. C. M. HWANG: Relationship between gate lag, power drift, and power slump of pseudomorphic high electron mobility transistors 1325

I. ADESIDA, A. MAHAJAN, G. CUEVA and P. FAY: Novel HEMT processing technologies and their circuit applications 1333

Y. OHNO, S. OHKUBO, K. KASAHARA, K. KUNIHIRO and Y. TAKAHASHI: Suppression of G_{DS} frequency dispersion in heterojunction FETs with a partially depleted p-type buffer layer 1339

T. NOZU, T. SUGIYAMA, S. HONGO and K. MORIZUKA: A new method for evaluation of surface recombination in heterojunction bipolar transistors by magnetotransport 1347

NOVEL DEVICES

A. SEABAUGH, B. BRAR, T. BROEKAERT, F. MORRIS, P. VAN DER WAGT and G. FRAZIER: Resonant-tunneling mixed-signal circuit technology 1355

J. N. SCHULMAN: Analysis of Sb-based resonant interband tunnel diodes for circuit modeling 1367

A. SCHÜPPEN: SiGe-HBTs for mobile communication 1373

U. KÖNIG, M. ZEUNER, G. HÖCK, T. HACKBARTH, M. GLÜCK, T. OSTERMANN and M. SAXARRA: n- and p-Type SiGe HFETs and circuits 1383

INDEXED IN Res. Alert, Cam. Sci. Abstr., Chem. Abstr. Serv., Curr. Cont./Phys. Chem. & Earth Sci., Curr. Cont./Eng. Tech. & Appl. Sci., Curr. Tech. Indx, Eng. Indx, INSPEC Data., PASCAL-CNRS Data., Curr. Cont. Sci. Cit. Indx, Curr. Cont. SCISEARCH Data., SSSA/CISA/ECA/ISMEC, Mater. Sci. Cit. Indx, Appl. Sci. & Tech. Indx, Wilson Appl. Sci. & Tech. Abstr.

PERGAMON

ISSN 0038-1101

F. SATO, T. HASHIMOTO, T. TATSUMI and T. TASHIRO: Process design for SiGe-HBTs prepared using cold-wall UHV/CVD 1389

Y. MIYAMOTO, H. TOBITA, K. OSHIMA and K. FURUYA: Barrier thickness dependence of peak current density in GaInAs/AlAs/InP resonant tunneling diodes by MOVPE 1395

MICROWAVE/MILLIMETER WAVE DEVICES

N-L. L. WANG, B. LIN, F. H-F. CHAU, G. JACKSON, Z. CHEN and C. P. LEE: Advanced heterostructure transistor technologies for wireless communications 1399

T. TANAKA, H. FURUKAWA, H. NAGATA and D. UEDA: A novel GaAs flip-chip power MODFET with high gain and efficiency 1405

K. FUJIMOTO, H. MASATO, K. KAWASHIMA, M. NISHITSUJI, M. TAKATA, Y. KUDO and O. ISHIKAWA: A low power dissipation $0.4 \sim 7$ GHz transimpedance amplifier IC for SCM optical communication system 1413

T. B. NISHIMURA, N. IWATA, Y. BITO and G. HAU: 44% efficiency operation of power heterojunction FET at near pinch-off for 3.5 V wide-band CDMA cellular phones 1419

K. MOCHIZUKI, T. OKA, K. OUCHI and T. TANOUE: Reliability investigation of heavily C-doped InGaP/GaAs HBTs operated under a very high current-density condition 1425

S. S. H. HSU, B. BAYRAKTAROGLU and D. PAVLIDIS: Comparison of conventional and thermally-stable cascode (TSC) AlGaAs/GaAs HBTs for microwave power applications 1429

T. LOW, T. SHIRLEY, C. HUTCHINSON, G. ESSILFIE, W. WHITELEY, B. YEATS and D. D'AVANZO: InGaP HBT technology for RF and microwave instrumentation..... 1437

R. DRIAD, S. R. LAFRAMBOISE, Z. H. LU, S. P. MCALISTER and W. R. MCKINNON: Passivation of InP-based HBTs 1445

WIDE BANDGAP DEVICES

M. S. SHUR, R. GASKA and A. BYKHOVSKI: GaN-based electronic devices..... 1451

A. A. BURK Jr., M. J. O'LOUGHLIN, R. R. SIERGIEJ, A. K. AGARWAL, S. SRIRAM, R. C. CLARKE, M. F. MACMILLAN, V. BALAKRISHNA and C. D. BRANDT: SiC and GaN wide bandgap semiconductor materials and devices 1459

A. HOKAZONO, K. TSUGAWA, H. UMEZANA, K. KITATANI and H. KAWARADA: Surface p-channel metal-oxide-semiconductor field effect transistors fabricated on hydrogen terminated (001) surfaces of diamond 1465

D. E. GRIDER, N. X. NGUYEN and C. NGUYEN: GaN MODFET microwave power technology for future generation radar and communications systems 1473

J. C. ZOLPER: Progress towards ultra-wideband AlGaN/GaN MMICs..... 1479

Y. KOYAMA, T. HASHIZUME and H. HASEGAWA: Formation processes and properties of Schottky and ohmic contacts on *n*-type GaN for field effect transistor applications..... 1483

CIRCUIT APPLICATIONS, DEVICE FABRICATION AND CHARACTERIZATION

M. RODWELL, Q. LEE, D. MENSA, J. GUTHRIE, S. C. MARTIN, R. P. SMITH, R. PULLELA, B. AGARWAL, S. JAGANATHAN, T. MATHEW and S. LONG: Transferred-substrate HBT integrated circuits 1489

E. H. C. PARKER and T. E. WHALL: SiGe heterostructure CMOS circuits and applications 1497

D. SAWDAI and D. PAVLIDIS: InP-based complementary HBT amplifiers for use in communication systems 1507

N. HARADA, Y. AWANO, K. HIKOSAKA and N. YOKOYAMA: A 77 K analog monolithic HEMT amplifier for high-speed Josephson-semiconductor interface circuit 1513

T. OHSHIMA, R. SHIGEMASA, M. SATO, M. TSUNOTANI and T. KIMURA: Improvement of 0.1 μm-gate InGaAs/AlGaAs HEMT performance by suppression of electro-chemical etching in deionized water 1519

D. XU, T. SUEMITSU, H. YOKOYAMA, Y. UMEDA, Y. YAMANE, T. ENOKI and Y. ISHII: Short gate-length InAlAs/InGaAs MODFETs with asymmetry gate-recess grooves: electrochemical fabrication and performance 1527

M. KUNZE, L. H. LEE, H. Y. CHUNG and E. KOHN: First demonstration of low temperature grown InP-channel HFET transferred onto GaAs substrate 1535

H. SAI, H. FUJIKURA, A. HIRAMA and H. HASEGAWA: Growth of device quality InGaP/GaAs heterostructures by gas source molecular beam epitaxy using tertiarybutylphosphine 1541

K. MATSUNAMI, T. TAKEYAMA, T. USUNAMI, S. KISHIMOTO, K. MAEZAWA, T. MIZUTANI, M. TOMIZAWA, P. SCHMID, K. M. LIPKA and E. KOHN: Potential profile measurement of GaAs MESFETs passivated with low-temperature grown GaAs layer by Kelvin probe force microscopy 1547

F. SCHUERMEYER, P. J. ZAMPARDI and P. M. ASBECK: Photoemission studies on heterostructure bipolar transistors 1555

H. TAKAHASHI, T. YOSHIDA, M. MUTOH, T. SAKAI and H. HASEGAWA: In-situ characterization technique of compound semiconductor heterostructure growth and device processing steps based on UHV contactless capacitance-voltage measurement 1561

T. SUGIMURA, T. TSUZUKU, T. KATSUI, Y. KASAI, T. INOKUMA, S. HASHIMOTO, K. IIYAMA and S. TAKAMIYA: A preliminary study of MIS diodes with nm-thin GaAs-oxide layers 1571

SYSTEMS APPLICATIONS

M. C. COMPARINI, M. FEUDALE and A. SURIANI: Applications of HEMT devices in space communication systems and equipment: a European perspective 1577

M. MURAGUCHI: RF device trends for mobile communications 1591

P. QUENTIN, C. DOURLENS, M. CAMIADE and D. PONS: MM-wave integrated circuits and their applications to communication and automotive systems 1599

J. DICKMANN, M. BERG, V. ZIEGLER, C. WÖLK and R. DEUFEL: MM-wave HEMT based circuits and their system applications 1607

HIGH-SPEED ICS FOR OPTICAL COMMUNICATION SYSTEMS

K. EMURA: Technologies for making full use of high-speed IC performance in the development of 40-Gb/s optical receivers 1613

K. WASHIO: SiGe HBTs and ICs for optical-fiber communication systems 1619

L. LUNARDI: Heterostructure circuit applications in optical communications 1627

P. J. ZAMPARDI, K. RUNGE, R. L. PIERSON, J. A. HIGGINS, R. YU, B. T. MCDERMOTT and N. PAN: Heterostructure-based high-speed/high-frequency electronic circuit applications 1633

MANUFACTURING

H. ROHDIN, A. WAKITA, A. NAGY, V. ROBBINS, N. MOLL and CHUNG-YI SU: A 0.1-μm MHEMT millimeter-wave IC technology designed for manufacturability 1645

H. WADA, H. SASAKI and T. KAMIJOH: Wafer bonding technology for optoelectronic integrated devices 1655

CONTENTS direct This journal is part of **ContentsDirect**, the ***free*** alerting service which sends tables of contents by e-mail for Elsevier Science books and journals. You can register for **ContentsDirect** online at: **www.elsevier.nl/locate/contentsdirect**

INDEXED IN Res. Alert, Cam. Sci. Abstr., Chem. Abstr. Serv., Curr. Cont./Phys. Chem. & Earth Sci., Curr. Cont./Eng. Tech. & Appl. Sci., Curr. Tech. Indx, Eng. Indx, INSPEC Data., PASCAL–CNRS Data., Curr. Cont. Sci. Cit. Indx, Curr. Cont. SCISEARCH Data., SSSA/CISA/ECA/ISMEC, Mater. Sci. Cit. Indx, Appl. Sci. & Tech. Indx, Wilson Appl. Sci. & Tech. Abstr.

PERGAMON

ISSN 0038-1101

PERGAMON

Solid-State Electronics 43 (1999) 1321

SOLID-STATE ELECTRONICS

Preface to special issue

The 3rd Topical Workshop on Heterostructure Microelectronics for Information Systems Applications (TWHM-ISA '98), held in Hayama-machi on the beautiful Miura Peninsula, followed the tradition set by the previous successful and internationally attended Workshops, held in Susono in 1994 and in Sapporo in 1996, by reporting the latest developments in heterostructure microelectronics. A special emphasis was placed on the applications in information systems for this workshop. This was prompted by the fact that the fast evolution of heterostructure microelectronics has already begun to show signs of solving the technological problems in information systems of today and tomorrow. The workshop title reflected this change by the addition of letters ISA.

The Workshop sessions contain a variety of contributions on devices, materials, circuits and systems. The technologies employed are based on heterostructure bipolar transistors, heterostructure field effect transistors and resonant tunneling diodes, and make use of a variety of heterostructure material systems including III-Vs (e.g. GaAs, InP and related compounds), group IV semiconductors (e.g. SiGe), and wide bandgap semiconductors (e.g. III-V Nitrides and SiC). Special emphasis was also placed on the effective insertion of these devices and circuits into systems such as mobile, fiber optic, space communications, as well as signal and data processing. This special issue includes excellent review articles on the present applications and the future impact of such technologies in information system applications.

We would like to thank all the authors for their contributions and all reviewers for their diligent efforts. Special thanks are due to the workshop committee members of TWHM-ISA '98 for soliciting papers and arranging the excellent program. The financial and organizing support made by the Asian Office of Aerospace Research and Development (AOARD), and the Office of Naval Research (ONR) are also greatly appreciated. Technical sponsorship by the IEEE Electron Device Society, the Japan Society of Applied Physics, and the Institute of Electronics, Information and Communication Engineers has also been instrumental in the organization of this Workshop.

Prof. Dimitris Pavlidis
Workshop Co-Chair
Dr. Tadao Ishibashi
Workshop Co-Chair
Prof. Takashi Mizutani
Program Co-Chair
Dr. Burhan Bayraktaroglu
Program Co-Chair

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00068-4

PERGAMON

Solid-State Electronics 43 (1999) 1323

SOLID-STATE ELECTRONICS

Committee

Workshop Co-Chairs:	D. Pavlidis (Univ. of Michigan) T. Ishibashi (NTT)
Technical Program Co-Chairs:	T. Mizutani (Nagoya Univ.) B. Bayraktaroglu (Northrop Grumman)
Financial Administrator & Local Arrangements:	Y. Miyamoto (Tokyo Inst. Tech.)
Secretary:	C. Nguyen (Hughes Labs.)
Sponsor Coordination:	S. Fujishiro (AFOSRA/AOARD) K. White (AFOSR/AOARD)
Committee members:	D. D'Avanzo (Hewlett-Packard) T. Enoki (NTT) H. Hasegawa (Hokkaido Univ.) K. Hirakawa (Univ. of Tokyo) U. König (Daimler-Benz) K. Matsumoto (ETL) S. Matsushita (Sanyo Electric) U. Mishra (UCSB) Y. Mitsui (Misubishi) K. Mizuno (Tohoku Univ.) K. Morizuka (Toshiba) T. Nakamura (Hosei Univ.) S. Nishi (Oki Electric) T. Nozaki (NEC) H. Sato (Sharp) F. Schuermeyer (Wright Labs.) S. Takamiya (Kanazawa Univ.) H. Tokuda (Toshiba) D. Ueda (Matsushita) N. Yokoyama (Fujitsu) J. Zolper (ONR)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00069-6

HIGH-SPEED & HIGH-FREQUENCY DEVICES

PERGAMON

Solid-State Electronics 43 (1999) 1325–1331

SOLID-STATE ELECTRONICS

Relationship between gate lag, power drift, and power slump of pseudomorphic high electron mobility transistors

James C.M. Hwang*

Lehigh University, 5 East Packer Avenue, Bethlehem, PA 18015, USA

Abstract

This paper attempts to clarify the confusion between different surface-related problems of PHEMTs such as gate lag, power drift, and power slump. Distinguishment is made not only in terms of electrical characteristics but also physical mechanisms. It was found that gate lag can be attributed to surface states while power drift and slump can be attributed to recoverable and irrecoverable, respectively, charge buildups in the silicon-nitride surface-passivation layer. As the result, the worsening of gate lag instead of gate lag per se can be an indicator of power drift while power drift can be a precursor for power slump. Based on such an understanding, a convenient screening test for these surface-related problems is proposed. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Compared to Si, which is by far the most common semiconductor, compound semiconductors such as GaAs have surfaces that are far more sensitive in terms of their electrical, chemical, and mechanical properties. Although amorphous hydrogenated silicon nitride has been used [1] to mitigate the surface problem thus allowing highly reliable GaAs metal-semiconductor field-effect transistors (MESFETs) to be made [2], it can be argued that the GaAs surface has never been truly passivated. Compound-semiconductor-based heterojunction devices such as pseudomorphic high electron mobility transistors (PHEMTs), having higher transconductances than MESFETs, are even more sensitive to the surface problem. As the result, certain PHEMTs are known to suffer from surface-related effects, such as gate lag [3], power drift [4], and power slump [5]. Gate lag is undesirable as it degrades the switching speed of PHEMTs. Power drift is troublesome because it makes power control difficult. Power slump shortens the lifetime of PHEMTs. These problems have confused potential users and prevented PHEMTs being more widely used in commercial applications such as RF power amplifiers for wireless communication, in spite of power PHEMT's many attractive performance characteristics [6].

This paper attempts to clarify the confusion by discussing the similarities and differences of different surface-related problems. The underlying physical mechanisms will also be discussed. However, for the sake of brevity and the emphasis on concepts that are broadly applicable to different makes of PHEMTs, the discussion may appear to be qualitative and speculative. (More definitive supports of the discussion can be found in the References.) Emphasis is also placed on transient (pulsed) instead of dc characteristics, the former being a more reliable predictor of RF performance [3]. Lastly, to emphasize that not all PHEMTs are problematic, side-by-side comparisons with PHEMTs which do not have gate-lag, power-drift, or power-slump problems are shown.

* Tel.: +1-610-758-5104; fax: +1-610-758-4244.
E-mail address: jh00@lehigh.edu (J.C.M. Hwang)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00070-2

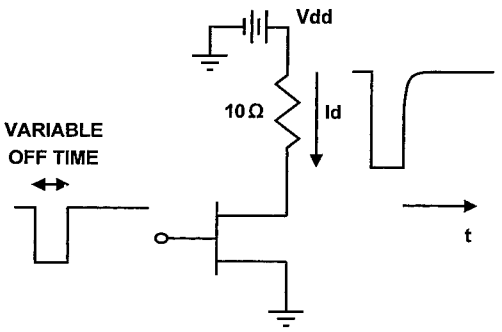


Fig. 1. Schematic illustration of a pulsed I–V stress test. A pulse generator is used as the gate stimulus and a dc power supply (V_{dd}) is connected to the drain of the PHEMT under test through a 10 Ω load resistor. Drain current, drain-source voltage, and gate-source voltage (V_{gs}) are monitored by a digital oscilloscope. During the test, V_{gs} is pulsed from on to deep pinch off repeatedly at a rate of 0.01–1 Hz. The pulse width for the off state is varied from 10 to 500 ms. The pulse at deep pinch off simulates the avalanche breakdown stress a PHEMT experiences under RF overdrive. The long holding period in the on state allows both transient and steady-state drain currents to be continuously monitored as the stress experiment progresses.

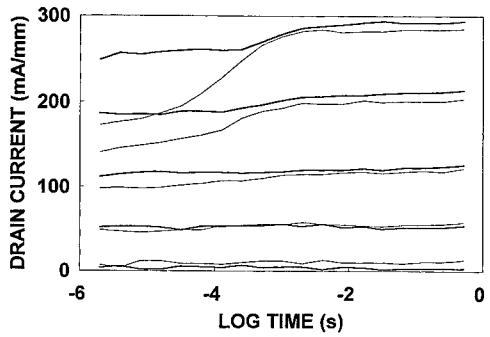


Fig. 2. Gate-lag characteristics of two PHEMTs showing (—) small and (—) large, respectively, amounts of lag due to different surface preparation processes. $V_{gs(on)} = -1.4$, -1.0 ,..., 0.2 V bottom up. $V_{gs(off)} = -4$ V. $V_{dd} = 2$ V.

Although not critical to the general concepts conveyed in this paper, the PHEMTs illustrated in this paper are of the state of the art. The PHEMTs are fabricated on 100 mm diameter semi-insulating GaAs substrates grown by using the liquid-encapsulated Czochralski method. GaAs and GaAs/AlGaAs superlattice buffer layers are next grown on the substrates by using molecular beam epitaxy. Pulse doping is used for the AlGaAs layers above and below the InGaAs channel. Each PHEMT has a 0.25 μm-long Ti–Pt–Au gate situated in a double recess which is formed by wet chemical etching of the AlGaAs layer above the channel. The recess is centrally located within a source-drain spacing of approximately 4 μm. The PHEMT is passivated by approximately 0.2 μm of amorphous hydrogenated silicon nitride which is grown by using plasma-enhanced chemical vapor deposition. Typically, the saturated and maximum drain currents are 200 and 300 mA/mm gate width, respectively. The threshold and breakdown voltages are −1 and 12 V, respectively.

2. Characteristics of gate lag, power drift, and power slump

Gate lag commonly refers to the delayed response of a PHEMT's drain current to its gate voltage change [3]. The delay time is typically of the order of ms. Using a pulsed I–V test as illustrated in Fig. 1, we found that when a PHEMT's gate is pulsed from deep pinch off to forward bias, approximately 85% of the drain current can be turned on within 1 μs, but the remaining 15% will take several ms to achieve (Fig. 2). As Fig. 3 shows, for a stable PHEMT without power drift or slump, this transient behavior is reproducible even after hours of repeated pulsing on and off.

Typical power-drift characteristics of a PHEMT are illustrated in Fig. 4. It can be seen that the output power drops by more than 0.5 dB after the RF input is applied for only 10 min. With continued RF drive, the output power will stabilize at a somewhat lower level [4]. If the RF drive is interrupted, the output power capacity of the PHEMT will recover at least partially. The recovery can be accelerated by annealing. Full recovery is usually possible within 1 h at 175°C (Fig. 5). The drift and recovery refer to the PHEMT's RF characteristics only. There is no discern-

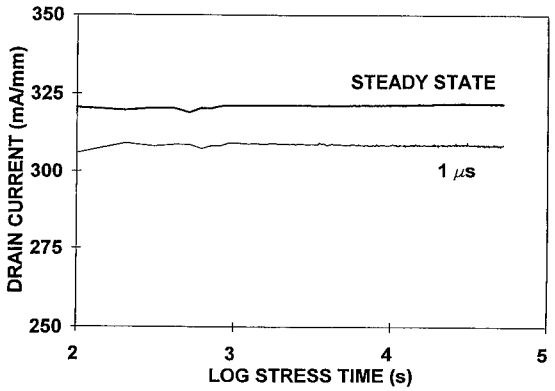


Fig. 3. With a stable PHEMT kept on most of the time but once every second pulsed off for 10 ms, transient drain current was then measured at (—) 1 μs and (—) 64 ms (steady state) after the PHEMT was pulsed back on. Notice that although the 1 μs current is always different from the steady-state current, neither current varies with the number of pulses or total stress time. $V_{gs(on)} = 0.5$ V. $V_{gs(off)} = -4$ V. $V_{dd} = 5$ V.

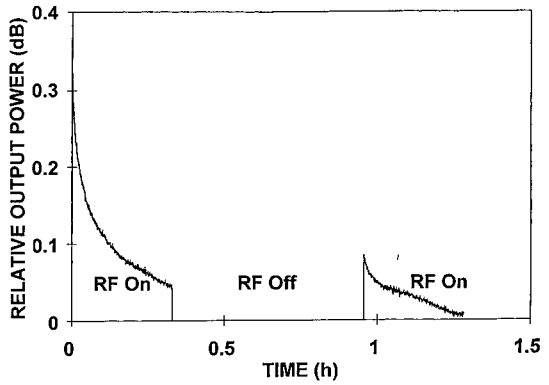


Fig. 4. Typical power-drift characteristics of a PHEMT under 5 dB RF overdrive (RF input power is increased by 5 dB beyond the 1 dB gain compression point). After the RF drive is interrupted, the output power recovers at least partially but would undergo a similar drift once the RF drive is reapplied. $V_{gs}=-0.5$ V. $V_{dd}=5$ V. $f=900$ MHz. Load impedance tuned for maximum power-added efficiency.

able change in dc. As for transient characteristics, we found power drift to be correlated with the worsening of gate lag but not gate lag per se. Using a pulsed I–V test similar to that of Fig. 3, a PHEMT which tends to drift exhibits drifting gate-lag characteristics as well (Fig. 6). Time-resolved pulsed I–V measurement similar to that of Fig. 2 indicates that only the gate-lag magnitude (deficiency in drain current right after the PHEMT is pulsed on) has drifted—the gate-lag time constant (time it takes for the drain current to reach steady state after the PHEMT is pulsed on) is unchanged. Notice that the steady-state drain current remains constant which is why no discernable change

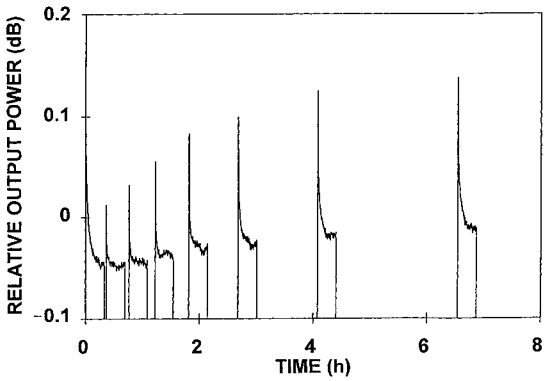


Fig. 5. Repeated drift tests of another PHEMT after RF interruptions of 2, 4, 8...128 min under an ambient temperature of 175°C. The recovery rate is increased when compared with the room temperature tests of Fig. 4. Full recovery is possible within approximately 1 h at 175°C. Bias, match and drive conditions are the same as in Fig. 4.

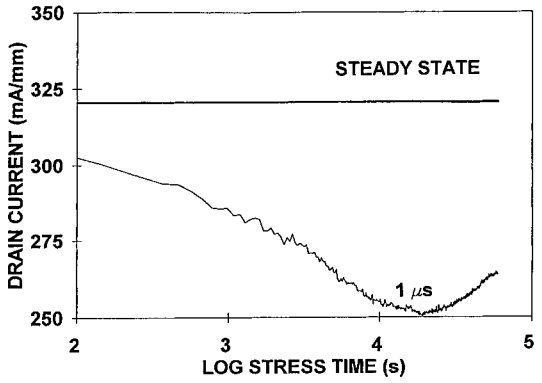


Fig. 6. Same as Fig. 3 except that a PHEMT that is known to suffer from power drift is tested. Notice that the 1 μs current decreases with increasing pulse number while the steady-state current remains the same, suggesting that power drift is correlated with the worsening of gate lag but not gate lag per se.

in dc characteristics was found. In comparison, a stable PHEMT may initially exhibit a larger or smaller magnitude of gate lag than a drifting PHEMT, but the gate-lag magnitude of the stable PHEMT will remain constant as shown in Fig. 3.

Power slump refers to gradual, yet steady drop in output power when a PHEMT is RF over-driven for hours (Fig. 7). The change in output power is associated with change in dc characteristics such as drain current, resistance, and breakdown voltage [5]. The rate of change decreases with increasing stress time but never saturates. Typically with storage time or temperature the slumped characteristics will not recover,

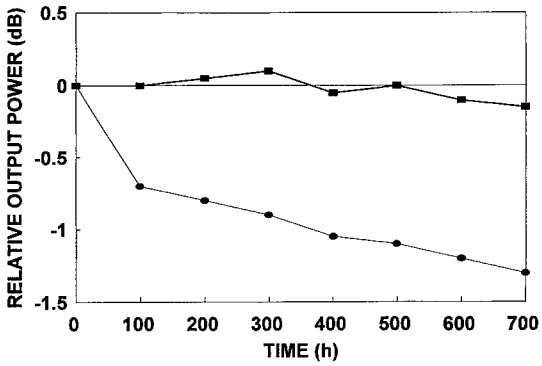


Fig. 7. Relative output power of (■) stable and (●) slumped PHEMTs under RF overdrive. The rate of slump decreases with increasing stress time but never saturates. Typically with storage time or temperature the slumped characteristics will not recover, at least not completely. When the RF drive is resumed, the output power will continue to degrade from its previously slumped value. By contrast, the stable PHEMT exhibits even a small amount of power expansion initially. Bias, match and drive conditions are the same as in Fig. 4.

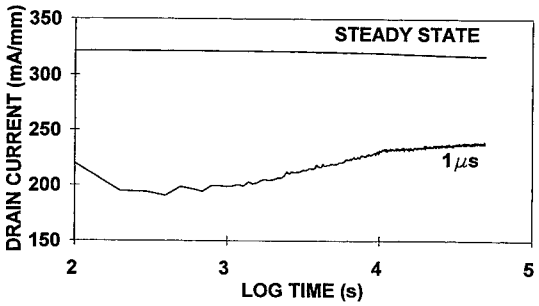


Fig. 8. Same as Fig. 6 except the off time is increased from 10 to 500 ms to accelerate power slump. In this case, the 1 μs current drifts down almost instantly but the steady-state current is stable until it starts to slump toward the end of the test. This suggests that power slump may in fact be an advanced stage of power drift.

at least not completely. When the RF drive is resumed, the output power will continue to degrade from its previous slumped value. Under a pulsed *I–V* test similar to that of Fig. 6 but with a much higher stress factor to accelerate power slump, the steady-state drain current eventually begins to decrease. This suggests that power slump may in fact be an advanced stage of power drift. Meanwhile, the initial drop in transient current (at 1 μs) due to power drift appears to be instantaneous over such an accelerated time scale (Fig. 8). Figs. 6 and 8 both show that the transient current eventually recovers slightly for an unknown reason.

3. Screening test for gate lag, power drift, and power slump

The pulsed *I–V* test illustrated in Figs. 3, 6, and 8

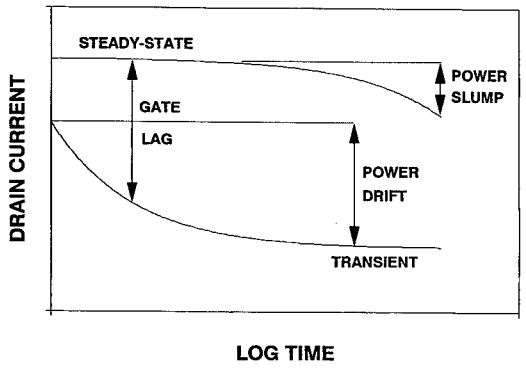


Fig. 9. Schematic illustration of the pulsed *I–V* test as a convenient screening test for gate lag, power drift, and power slump. Power drift and slump are indicated by changes in transient and steady-state drain currents, respectively, while gate lag is indicated by the difference between the transient and steady-stated drain currents.

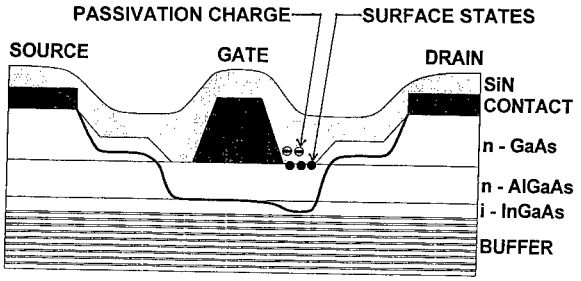


Fig. 10. PHEMT cross section showing the location of (●) surface states and (⊖) passivation traps which are potential causes for gate lag and power drift/slump, respectively.

appears to be a convenient screening test for gate lag, power drift, and power slump. Power drift and slump are indicated by changes in transient and steady-state drain currents, respectively, while gate lag is indicated by the difference between the transient and steady-stated drain currents. These quantities are indicated in Fig. 9 for a hypothetical composite case. Based on an RF waveform analysis [7], the cumulative stress under such a pulsed *I–V* test can be translated into the stress actually experienced by a PHEMT under the intended RF operating condition with its associated lifetime predicted.

For example, a PHEMT gate may be biased at −0.5 V while its drain biased at 4 V. Under an RF drive, the gate voltage swings from −2 to 1 V while the drain voltage swings from 1 to 7 V (assuming a voltage gain of 2). Since the gate and drain voltage swings are generally out of phase, the greatest stress occurs when the gate is at −2 V and the drain is at 7 V. In this case, the maximum drain-gate voltage is 9 V. Under a pulsed *I–V* test, the gate is pulsed to −4 V while the drain is held steady at 5 V, resulting in a stress level similar to that under the RF drive. However, the 'duty factor' is very different between the RF and pulsed cases. The drain-gate voltage may reach 9 V only 0.5% of the time under the RF drive but 50% of the time under the pulsed *I–V* test (as in the cases of Figs. 4 and 8, respectively). Therefore, we can use the pulsed *I–V* stress to simulate the RF overdrive stress but with an acceleration factor of 100. This implies that a PHEMT which exhibits a 10% drop in transient drain current after 10 s of pulsed *I–V* test will likely to suffer a 1 dB drift in output power after 15 min of RF overdrive. By quickly sampling a few PHEMTs on a given wafer through such a 10 s pulsed *I–V* test and by using the 10% drop in transient drain current as the pass-fail criteria, a decision can be made in the PHEMT fabrication line whether or not the wafer warrants further processing, dicing, and packaging.

Table 1
Surface-Related Problems of PHEMTs

Problem	Time	Cause	Effect	Recovery	Impact
Gate lag	Seconds	Surface states	Transient drain current	Yes	Switching speed
Power drift	Minutes	Passivation charge	Transient drain current	Yes	Power control
Power slump	Hours	Passivation charge	dc drain current	No	PHEMT lifetime

4. Mechanisms for gate lag, power drift, and power slump

Table 1 summarizes the main characteristics of gate lag, power drift, and power slump. Although they all originate from charge trapping under peak drain-gate voltage stress, gate lag is associated with surface states while power drift and slump are associated with recoverable and irrecoverable, respectively, charge trappings in the silicon-nitride passivation layer. Fig. 10 illustrates the location of surface states and passivation traps. Statically both surface or passivation charges can increase the depletion width (hence drain resistance) and their effects appear to be indistinguishable. However, dynamically they are very different because surface states typically have orders of magnitudes faster time constants than passivation traps.

Gate lag has long been attributed to surface states between the gate and drain [3]. However, we found recently that, depending on the design and growth of the buffer layer, charge carriers can also be trapped there [8]. We have since developed a pulsed dynamic transfer characteristics technique to experimentally distinguish surface and substrate effects [9]. Fig. 11 compares the transient and steady-state dynamic transfer characteristics for a hypothetical case in which the maximum channel conductance is depressed by surface states while the threshold voltage is shifted by substrate traps. Using this technique, we conclude that gate lag is caused by surface states in most PHEMTs (Fig. 12a).

The dynamic transfer characteristics of Figs. 11 and

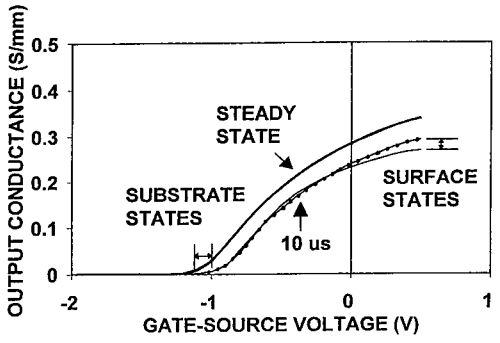


Fig. 11. Schematic illustration of the pulsed dynamic transfer characteristics of a PHEMT after its gate is pulsed from deep pinch off to forward bias. (—) Output conductance 10 μs after pulse on. (—) Output conductance when the PHEMT reaches steady state after it was pulsed on. (●) Steady-state output conductance with a hypothetical threshold voltage shift. $V_{gs(on)} = -0.5$ V. $V_{gs(off)} = -4$ V. $V_{dd} = 0.1$ V. Match and drive conditions are the same as in Fig. 4.

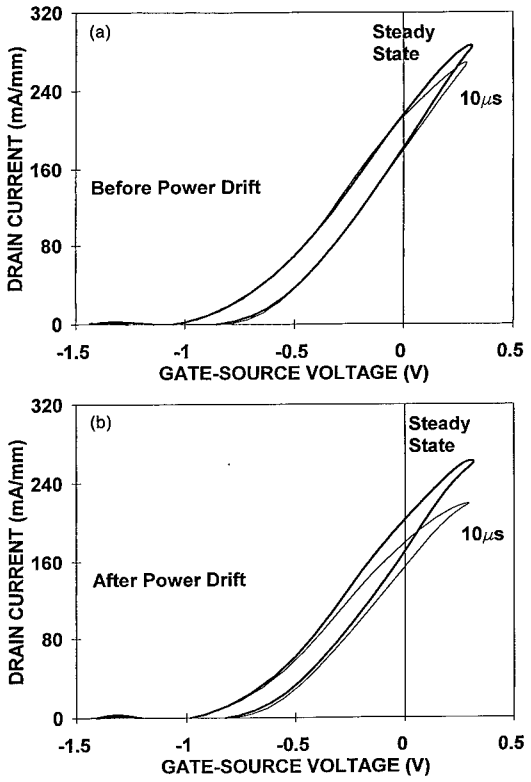


Fig. 12. Pulsed dynamic transfer characteristics of a PHEMT (a) before and (b) after power drift. The dynamic transfer characteristics were measured at (—) 10 μs and (—) steady state after V_{gs} was pulsed from −4 to −0.5 V. In (a), the peak transient current is lower than the peak steady-state current, indicating that gate lag is caused by surface states. In (b), the peak transient current is much lower than the peak steady-state current, indicating that gate lag is worse. The peak steady-state current in (b) is lower than that in (a) which indicates power drift. Bias, match, and drive conditions are similar to that of Fig. 11 except $V_{dd} = 5$ V.

12 are based on instantaneous RF drain currents and gate voltages measured at 900 MHz by using a novel pulsed RF waveform measurement technique. With the PHEMT gate biased at −4 V, the PHEMT remains pinched off even under an RF overdrive. With the gate bias pulsed to −0.5 V (for a period of 990 ms), the RF drive causes the gate voltage to swing repeatedly between approximately −1.5 and 0.5 V and the drain current to swing repeatedly between zero and approximately 300 mA/mm. In Fig. 11 a very small drain voltage is used so that capacitive loading (displacement current) is insignificant and the RF current vs voltage loci resemble a line. In Fig. 12 a more realistic drain voltage is used and the displacement current is added to the conduction current so that the RF loci resemble a loop. The loop is traced at a frequency of 900 MHz or over a period of approximately 1 ns. It should not be confused with hysteresis over a time scale of ms or longer which may be observable by using a 60 Hz curve tracer. By using the pulsed RF waveform measurement technique, hysteresis of μs or longer will not affect the magnitude of looping (if there is any), but will shift the loop up or down as shown in Fig. 12. Moreover, similar looping occurs even at steady state (with respect to gate lag). Notice that 'transient' and 'steady' are both relative to a certain time scale. To a large extent the confusion between gate lag, power drift, and power slump is caused by the very different time scales involved and the limited time span covered by some experiments.

Fig. 12b shows that power drift of a PHEMT is associated with a change in its steady-state dynamic transfer characteristics, keeping in mind that 'steady state' here continues to imply that gate lag is over but power drift is not necessarily stabilized. It can be seen that after drift the peak drain current decreases, but the transconductance and threshold voltage remain approximately the same. This can only be explained by a change in the drain access resistance since changes in the source access resistance or any other channel parameter would affect the transconductance and/or the threshold voltage. In comparison power slump results in a similar change in characteristics after a much longer period of stress and it has not been observed to be recoverable. While power slump has been attributed to the permanent charge accumulation in the passivation, the recoverable nature of power drift suggests a temporary charge trapping mechanism in the passivation.

Through measurements performed on PHEMTs and MIM capacitors we propose [10] that, driven by high electric fields associated with peak drain-gate voltages under RF overdrive, a gradual accumulation of trapped charge in the passivation relaxes the peak electric field at the drain edge of the gate. This reduced field in turn increases the surface-state occupation (when the PHEMT is swung to the off state) through decreased electron emission and/or hole capture. With an increased surface-state occupation, the peak RF current of the PHEMT (when it is swung back on) is reduced due to increased drain resistance hence power drifts. Such an indirect mechanism is different from that of power slump. In power slump there is usually a greater amount of passivation charge that, independent of surface-state occupation, the drain access resistance is increased directly under both dc and RF conditions.

In the case of power drift, when the high electric fields are removed by interrupting the RF drive, the trapped charge in the passivation is emitted thermally and power capacity gradually recovers. The temperature dependence of the recovery of drifted PHEMT's exhibits an Arrhenius behavior with an activation energy in the range of 1.4 ± 0.4 eV which is consistent with the reported energies for the K_0 center in amorphous hydrogenated silicon nitride [11] and is much higher than 0.4 ± 0.1 eV reported for the surface states [12]. This is further evidence that power drift is caused by passivation charges. The K_0 center has been associated with a dangling Si bond. Plasma-deposited silicon nitride tends to be Si rich hence can contain many dangling Si bonds except most of them are passivated by hydrogen. Under high electric field hot electrons are generated which may have sufficient energy to break a Si–H bond thus creating a new K_0 center. Once dissociated, hydrogen is probably in the form of a proton. The proton, being a fast diffuser, may outdiffuse to the ambient. Otherwise, two protons may aggregate and form a hydrogen molecule which then becomes immobile. Once the protons outdiffuse or aggregate, it will be difficult to recover power slump.

5. Conclusion

Once the physical mechanisms are understood, it is not surprising that the worsening of gate lag can be an indicator of power drift, while power drift can be a precursor for power slump. However, the correlation between gate lag, power drift, and power slump is complicated. This is because, while gate lag depends mainly on the sensitivity of channel conductance to surface-state occupation, power drift depends on both the channel-to-surface sensitivity and the rate of charge buildup in the passivation. Therefore, a PHEMT with a larger gate lag will not necessarily show a higher tendency to drift. Similarly, while power drift depends mainly on hydrogen dissociation, power slump depends on both hydrogen dissociation and hydrogen diffusion. Therefore, a PHEMT with higher tendency to drift will not necessarily slump faster.

Acknowledgements

This work was supported in part by US Air Force Research Laboratory, Sensors Directorate, Aerospace Components and Subsystems Technology Division, under Contract No. F33615-95-C-1755.

References

[1] DiLorenzo JV, Hwang JCM, Niehaus WC, Schlosser WOW, Wemple SH. US Patent 1984:4426,656.
[2] Wemple SH, Niehaus WC, Fukui H, Irvin JC, Cox HM, Hwang JCM, DiLorenzo JV, Schlosser WO. IEEE Trans Electron Devices 1981;28:834.
[3] Huang JC, Jackson S, Shanfield S, Platzker A, Salesdas PK, Weichert C. IEEE Trans Microwave Theory Techniques 1993;41:752.
[4] Leoni RE III, Bao JW, Bu JK, Hwang JCM. Digest of International Conference GaAs Manufacturing Technology. 1998. p. 147.
[5] Leoni RE, Hwang JCM. Effects of reverse gate-drain breakdown on gradual degradation of power PHEMTs. In: Technical Digest of IEEE GaAs IC Symposium. 1996. p. 31.
[6] Smith PM, Mishra UK, Chao PC, Palmateer SC, Hwang JCM. IEEE Electron Device Lett 1985;6:86.
[7] Hwang JCM. In: Technical Digest of IEEE GaAs IC Symposium. 1995. p. 81.
[8] Wei CJ, Hwang JCM. Submitted to IEEE Trans Microwave Theory Techniques.
[9] Bao JW, Leoni RE, Du X, Shirokov MS, Hwang JCM. Digest of International Conference GaAs Manufacturing Technology, 1998. p. 199.
[10] Leoni RE III, Bao JW, Bu JK, Du X, Shirokov MS, Hwang JCM. Submitted for publication to IEEE Trans Microwave Theory Techniques.
[11] Sinitsa SP. Electrophysical Properties. In: Belyi VI, et al., editors. Materials science monographs, Silicon nitride in electronics, vol. 34. New York: Elsevier, 1988. p. 188.
[12] Hasegawa H, Sawada T. J Vac Sci Technol 1979;16:1478.

PERGAMON

Solid-State Electronics 43 (1999) 1333–1338

SOLID-STATE ELECTRONICS

Novel HEMT processing technologies and their circuit applications

I. Adesida[a,*], A. Mahajan[a], G. Cueva[a], P. Fay[b]

[a]*Microelectronics Laboratory and Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, IL 61801, USA*

[b]*Department of Electrical Engineering University of Notre Dame, Notre Dame, IN 46556, USA*

Abstract

The monolithic integration of enhancement- and depletion-mode high-electron mobility transistors (E- and D-HEMTs) suitable for high-speed and low power circuit applications in the lattice-matched InP material system is examined. E-HEMT devices with gate-lengths of 0.25, 0.5 and 1.0 μm fabricated using a buried-Pt gate process demonstrate threshold voltages (V_T) ranging from +200 to +258 mV and maximum extrinsic transconductances (g_{mext}) as high as 800 mS mm^{-1}, while D-HEMT devices of identical gate-lengths exhibited a V_T ranging from −599 to −405 mV, and a g_{mext} as high as 578 mS mm^{-1}. The devices showed excellent rf characteristics, exhibiting unity current-gain cutoff frequencies (f_t) as high as 106 GHz. Based on these results, 11, 23, and 59 stage ring oscillators using direct-coupled FET logic (DCFL) technology were fabricated and characterized. Room temperature propagation delays of 9.27 ps/stage with a power-delay product of 2.37 fJ/stage were achieved. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The monolithic integration of enhancement- and depletion-mode high-electron mobility transistors (E- and D-HEMTs) lattice-matched to InP is of considerable interest in the area of low power and high speed communication systems. Compared to circuits employing a traditional D-HEMT only technology, circuits utilizing both E- and D-HEMTs have several advantages including the elimination of d.c.-level shifting stages which consume both power and area. Additionally, an E/D technology allows for single voltage supply operation which can be quite cumbersome to implement using only D-HEMTs. Heretofore, a process for the monolithic integration of E- and D-HEMTs was not viable due to the modest success of E-HEMT devices in the lattice-matched InP material system [1]. However, over the last few years, several groups have achieved E-HEMT performance by using a buried-Pt gate process [2–6], with Mahajan et al. [7,8] demonstrating a process for the fabrication of high-speed InAlAs/InGaAs/InP E-HEMTs suitable for circuit applications. Using a novel heterostructure design coupled with a buried-Pt process, E-HEMT devices with a threshold voltage (V_T) of +255 mV and a unity current-gain cutoff frequency (f_t) of 160 GHz were demonstrated. In addition, these devices exhibited a maximum frequency of oscillation (f_{max}) of over 200 GHz as well as excellent threshold voltage uniformity.

In this work, we demonstrate the feasibility of direct-coupled FET logic (DCFL) circuits made by the

* Corresponding author.
E-mail address: adesida@capone.ccsm.uiuc.edu (I. Adesida)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00071-4

monolithic integration of E- and D-HEMTs in the lattice-matched InP material system. Discrete devices with gate-lengths ranging from 0.25 to 1.0 μm are fabricated and characterized. Using this technology, DCFL inverters and ring oscillators with 11, 23 and 59 stages are fabricated and characterized at room temperature.

2. Fabrication and characterization of monolithic E- and D-HEMTs

2.1. Heterostructure material

The devices were fabricated on a heterostructure grown by molecular beam epitaxy (MBE) on a semi-insulating InP substrate. The buffer consists of a 200 nm layer of undoped InAlAs followed by a 20 nm undoped InGaAs channel. A 3.5 nm spacer layer of InAlAs lies on top of the channel followed by a Si d-doping plane, a 12 nm undoped InAlAs Schottky barrier layer, and a pseudomorphic 1.5 nm AlAs etch-stop layer. Another 10 nm-thick InAlAs layer is grown on the etch-stop layer followed sequentially by a 1.5 nm AlAs layer and an n^+ InGaAs cap layer. A 2-DEG concentration of 1.2×10^{12} cm^{-2} and an electron mobility of 10,000 cm^2/V s were obtained from room-temperature Hall measurements.

2.2. Device fabrication

The entire monolithic fabrication process was implemented using electron-beam (e-beam) technology. For the fabrication process, both the mesa isolation and the ohmic levels were patterned simultaneously for the E- and D-HEMTs. Using a citric acid-based etch, device isolation was achieved by etching to the InAlAs buffer layer. Ohmic level was then formed by evaporating AuGe/Ni/Au metallization and alloying the sample at 480°C for 20 s in a nitrogen ambient with a resulting contact resistance (R_c) of 0.15 Ω mm. Mushroom-shaped gates for the E-HEMTs were then patterned using a three-layer resist pattern. The same highly selective citric acid-based etch was used to etch to the bottom AlAs etch-stop layer. Following a dip in a diluted solution containing HCl to remove the etch-stop layer, a Pt/Ti/Pt/Au gate metallization was then evaporated and lifted off. The metallization scheme chosen is described in detail in [8]. A 60 s buried-Pt gate-anneal was then performed in a nitrogen ambient at 350°C to enhance the threshold voltage [2–8], completing the E-HEMT fabrication process. D-HEMT gates were then fabricated by etching to the top AlAs etch-stop layer using the citric acid-based etch. A Ti/Au gate metallization was deposited following the removal of the etch-stop layer. A cross-sectional view

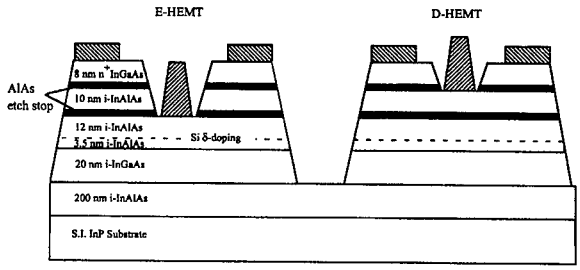


Fig. 1. Cross-sectional diagram of integrated E- and D-mode InP-based HEMTs.

of the completed E-and D-HEMT devices is shown in Fig. 1. Using this method, devices with gate lengths of 0.25, 0.5 and 1.0 μm were fabricated for both E- and

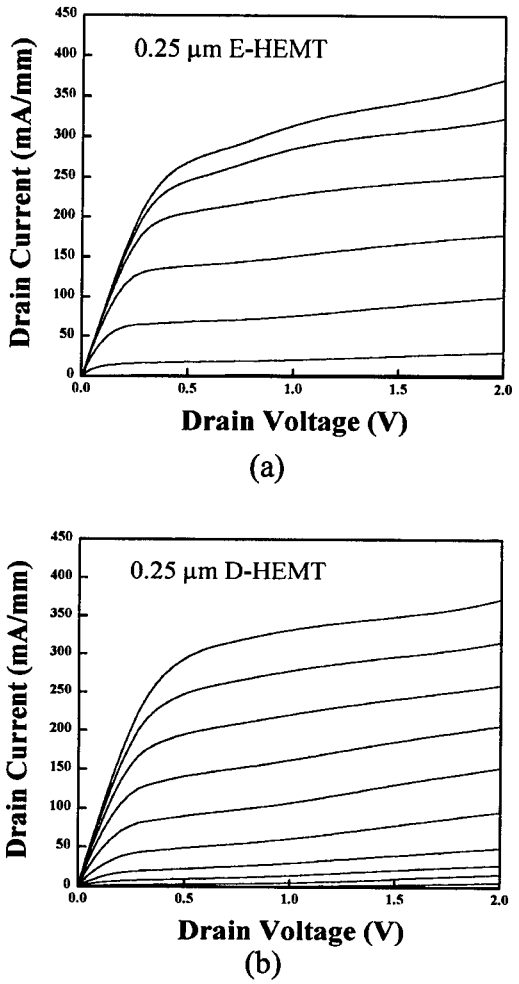


Fig. 2. Common source characteristics for 0.25 μm gate-length (a) E-HEMT where the gate potential is swept from 0 to 0.7 V in steps of 0.1 V and (b) a D-HEMT where the gate voltage is swept from −1 to 0 V in increments of 0.1 V.

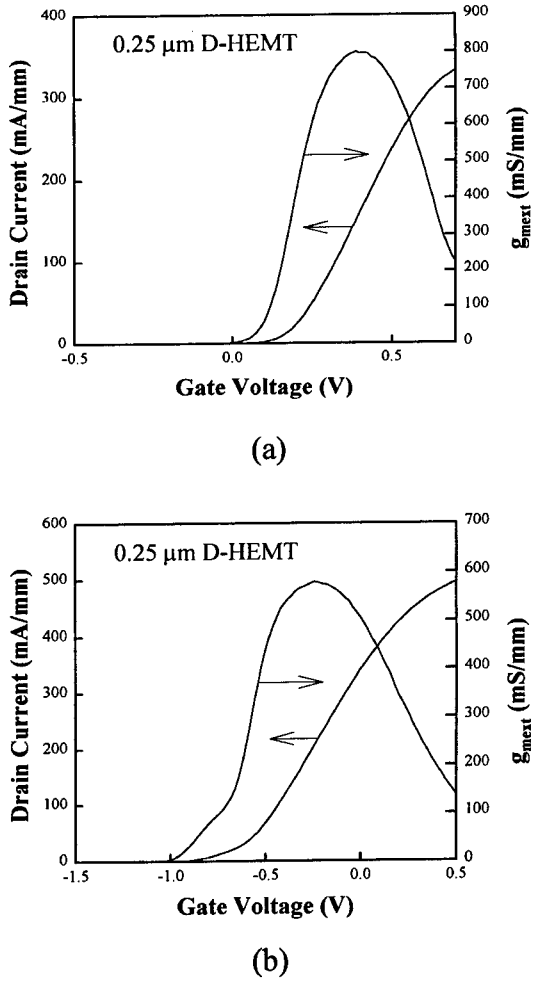


Fig. 3. Transfer chacteristics for 0.25 μm gate-length (a) E-HEMT and (b) D-HEMT.

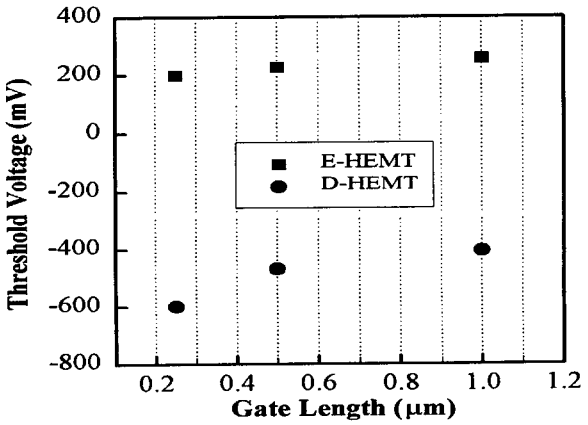


Fig. 4. Threshold voltage as a function of gate-length for both E- and D-HEMTs.

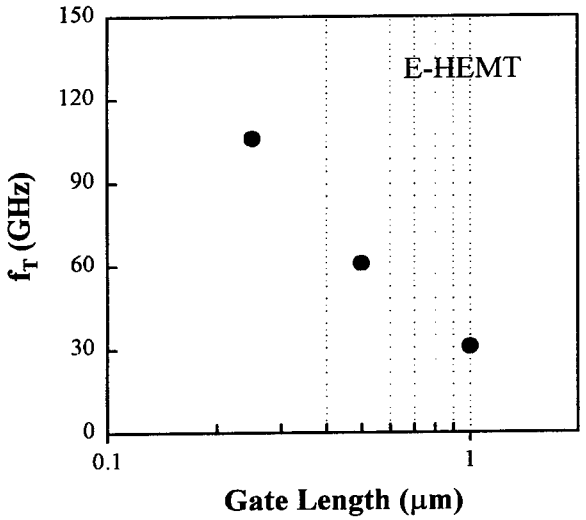


Fig. 5. Unity current-gain cutoff frequency (f_t) as a function of gate-length for the E-HEMTs.

D-HEMTs. The drain-to-source (D-S) spacing for all sub-micron gate-length devices was 2 μm, while the D-S spacing for the 1.0 μm gate-length device was 3 μm.

The common-source characteristics of a typical 0.25 μm gate-length E- and D-HEMT device are shown in Fig. 2(a) and (b), respectively. As can be seen, both devices exhibit excellent pinchoff characteristics and no significant kink effect is evident. Fig. 3(a) and (b) shows the transfer characteristics of both the 0.25 μm gate-length E and D-HEMT, respectively. Looking at Fig. 3, the 0.25 μm gate-length D-HEMT device exhibits a threshold voltage (V_T) of −599 mV and a maximum extrinsic transconductance (g_{mext}) of 577 mS mm^{-1}, while the 0.25 μm gate-length E-HEMT demonstrates a V_T of +200 mV and a g_{mext} of 800 mS mm^{-1}. A lower g_{mext} for the D-HEMT devices is due to the larger gate-to-channel distance of the D-HEMTs as compared to the E-HEMTs. Fig. 4 shows V_T as a function of gate-length for both the D- and E-HEMTs. For the D-HEMTs, V_T ranges from −599 mV for the 0.25 μm gate-length device to −405 mV for the 1.0 μm device, while for the E-HEMTs, the V_T ranges from +200 mV for the 0.25 μm gate-length device to +258 mV for the 1.0 μm device.

The unity current-gain cutoff frequency (f_t) was measured for both sets of devices using on-wafer s-parameter measurements, and the results for the E-HEMTs are displayed in Fig. 5 as a function of gate-length. Examination of Fig. 5 shows the f_t for the E-HEMT devices ranging from 31 GHz for the 1.0 μm gate-length device to 106 GHz for the 0.25 μm gate-length device. Results for the D-HEMT devices are nearly identical.

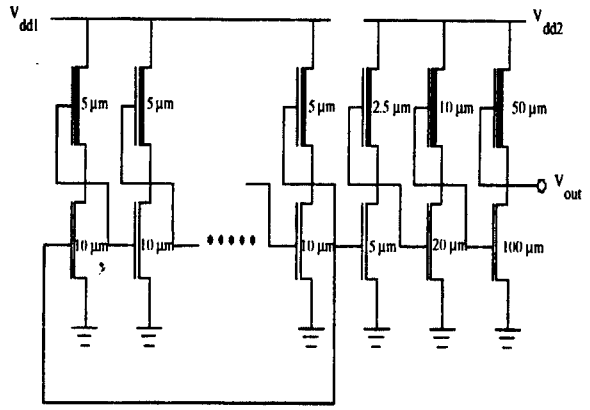


Fig. 6. Schematic for the ring oscillators.

3. DCFL ring oscillator fabrication and characterization

3.1. Ring oscillator design and fabrication

A direct-coupled FET logic (DCFL) circuit design was employed using gate lengths of 0.25, 0.5 and 1.0 μm. DCFL inverters were used as the building blocks for the ring oscillators and the D-HEMT load devices and E-HEMT driver devices were fabricated with widths of 5 and 10 μm, respectively. Both load and driver devices were of the same gate-length. The ring oscillators consisted of 11, 23 and 59 stages of inverters and a three stage buffer amplifier. The buffer amplifier was used so as not to load the output signal of the ring oscillator with the measurement equipment. The circuit schematic of the complete ring oscillator with the buffer is shown in Fig. 6. As can be seen from Fig. 6, two separate power supplies (V_{dd1} and V_{dd2}) are used in the circuit to enable the direct measurement of the dissipated power of the ring oscillator. Identical voltages are applied to both supplies, but one supply, V_{dd1}, is used for the oscillator, while the other, V_{dd2}, is used for the buffer amplifier. Total transistor count for the 11, 23, and 59 stage ring oscillator circuits are 28, 52, and 124 devices, respectively.

The process for the fabrication of the ring oscillator is similar to the fabrication process described earlier for the E- and D-HEMTs. Mesa and ohmic levels were fabricated simultaneously for both the E-HEMT drivers and D-HEMT loads. Following the ohmic process, E-HEMT gates and the first level of interconnects were patterned using the Pt/Ti/Pt/Au metallization sequence described earlier. The sample was then annealed at 350°C for 60 s in a nitrogen ambient to increase the threshold voltage of the E-HEMTs. D-HEMT gates were then fabricated using a Ti/Au gate metallization sequence. The PECVD deposition of 200 nm of Si_3N_x followed by a contact-hole etching and a second level of interconnect metal completed the ring oscillator fabrication. A photograph of a completed 23 stage ring oscillator is shown in Fig. 7.

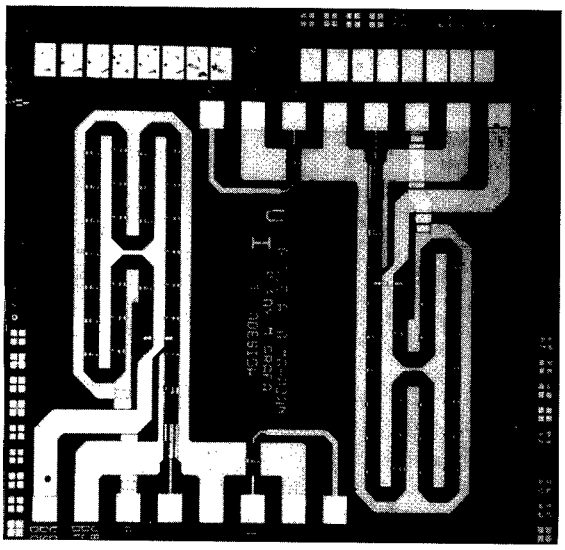


Fig. 7. Die photo of the completed 23 stage ring oscillator.

3.2. Ring oscillator results and discussion

Fig. 8 shows the output spectrum of a 23 stage ring oscillator fabricated with 0.25 μm gate-length devices, operating at room temperature, with a supply voltage of 0.4 V. From Fig. 8 a fundamental frequency of oscillation (f_0) of 2.37 GHz is observed. The propagation delay per stage (t_{pd}) can be related to f_0 by the simple expression $t_{pd} = (2nf_0)^{-1}$, where n is the number of stages, 23 in this case. Using this relationship, a t_{pd} of

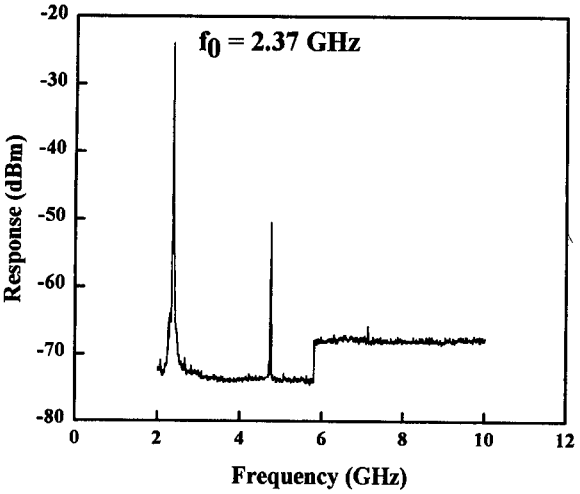


Fig. 8. Spectrum of output of 23 stage 0.25 μm gate-length ring oscillator with $V_{dd} = 0.4$ V.

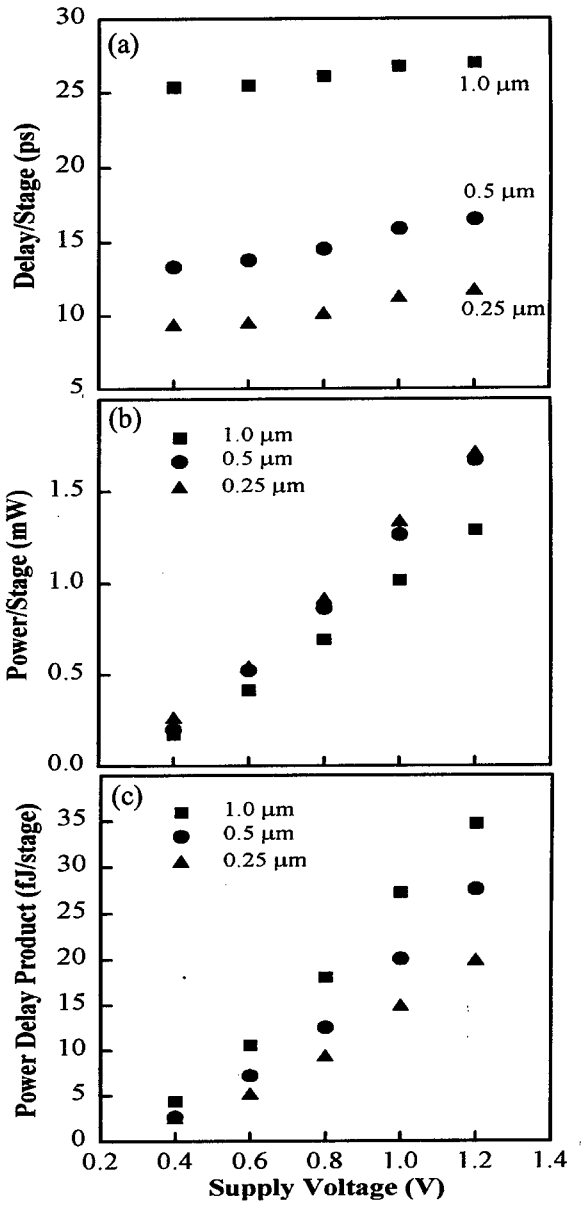


Fig. 9. (a) t_{pd}, (b) P_D and (c) PDP as a function of supply voltage for 1.0, 0.5 and 0.25 μm gate-length inverters.

9.27 ps/stage is observed. Fig. 9(a–c) shows delay/stage (t_{pd}), power dissipation/stage (P_D), and power-delay product (PDP) as a function of supply voltage each of the three gate-lengths, respectively. From Fig. 9 it is observed that t_{pd} for the DCFL inverters ranges from 27, 16.5, and 11.6 ps at 1.2 V of supply voltage to 25.4, 13.3 and 9.3 ps at a supply voltage of 0.4 V, for the 1.0, 0.5 and 0.25 μm gate-length inverters, respectively. These values compare well when compared to t_{pd}'s from other DCFL technologies, in which a room temperature t_{pd} of 23 and 22.3 ps/stage were demonstrated for a 1.2 μm gate-length InGaAs/InAlAs/InP HIGFET process [9] and a 0.15 μm gate-length InGaP/InGaAs/GaAs p-HEMT process [10], respectively. From Fig. 9(b) it is noted that P_D increases linearly with supply voltage for all three gate-length oscillators. At a supply voltage of 0.4 V, P_D is 0.17, 0.20, and 0.26 mW, and increases to 1.29, 1.67 and 1.70 mW at a supply voltage of 1.2 V, for the 1.0, 0.5 and 0.25 μm gate-length inverters, respectively. PDP, shown in Fig. 9(c), increases with supply voltage. At 0.4 V, PDP's of 4.36, 2.66 and 2.37 fJ/stage are observed for the 1.0, 0.5 and 0.25 μm gate-length inverters.

4. Conclusion

The monolithic integration of enhancement- and depletion-mode high-electron mobility transistors (E- and D-HEMTs) in the lattice-matched InAlAs/InGaAs/InP material system has been accomplished. Discrete HEMTs with gate-lengths ranging from 0.25 to 1.0 μm were fabricated and characterized. Using a buried-Pt gate process, E-HEMTs exhibited threshold voltages (V_T) ranging from +200 to +258 mV and maximum extrinsic transconductances (g_{mext}) as high as 800 mS mm^{-1}, while D-HEMT devices of identical gate-lengths exhibited a V_T ranging from −599 to −405 mV, and a g_{mext} as high as 578 mS mm^{-1}. The devices showed excellent rf characteristics, exhibiting unity current-gain cutoff frequencies (f_t) of 106 GHz for the 0.25 μm E-HEMT device. Based on these results, 11, 23 and 59 stage ring oscillators using direct-coupled FET logic (DCFL) technology were fabricated and characterized. At room temperature, with a supply voltage of 0.4 V, propagation delays of 9.27 ps/stage and a power-delay product of 2.37 fJ/stage were achieved. These results approach the speed of D-mode only circuits, but at much lower power.

Acknowledgements

This work was supported by ONR Grant N00014-98-1-0376 (program managers: J. Zolper and C. Wood) and by NSF Grant ECS 96-33535.

References

[1] Itoh T, Brown A, Camnitz AL, Wicks G, Berry J, Eastman L. In: Proc. Int. Symp. GaAs Relat. Compounds, 1985; 571–576.
[2] Harada N, Kuroda S, Ktakami T, Hikosaka K, Mimura T, Abe M. In: Proc. 3rd Int. Conf. InP Relat. Mater., 1991; 377–380.
[3] Harada N, Kuroda S, Hikosaka K. IEICE Trans 1992;10:1165–71.

[4] Harada N, Kuroda S, Watanabe Y, Hikosaka K. Electron Lett 1993;29:2100–1.
[5] Chen K, Enoki T, Maezawa K, Arai K, Yamamoto M. IEEE Trans Electron Devices 1996;43:252–2576.
[6] Mahajan A, Fay P, Caneau C, Adesida I. Electron Lett 1996;32:1037–8.
[7] Mahajan A, Arafa M, Fay P,. Adesida I. 54th Annual Device Research Conference Digest 1996;131–3.
[8] Mahajan A, Arafa M, Fay P, Caneau C, Adesida I. IEEE Electron Device Lett 1997;18:284–6.
[9] Feuer M, He Y, Shunk J-H, Vang T, Brown-Goebeler K, Chang T-Y. IEEE Electron Device Lett 1991;12: 98–100.
[10] Suehiro H, Shima M, Shimura T, Hara N. In: IEEE Gallium Arsenide Integrated Circuit Symposium, Technical Digest 1996;197–200.

PERGAMON

Solid-State Electronics 43 (1999) 1339–1345

SOLID-STATE ELECTRONICS

Suppression of G_{DS} frequency dispersion in heterojunction FETs with a partially depleted p-type buffer layer

Yasuo Ohno*, Satoru Ohkubo, Kensuke Kasahara, Kazuaki Kunihiro, Yuji Takahashi

Optoelectronics and High Frequency Device Research Laboratories, NEC Corporation, Miyukigaoka 34, Tsukuba, Ibaraki, Japan, 305-8501

Received 28 October 1998

Abstract

Drain-conductance frequency dispersion in GaAs/AlGaAs heterojunction FETs is suppressed by a partially depleted p-layer between the channel and the semi-insulating substrate. Though the p-layer is electrically floating, the electrical potential of the layer is stabilized to the source voltage by minimized capacitive coupling with the drain electrode, while at low frequencies the potential is fixed to the source voltage by the resistive balance of two pn junctions at each end of the layer. Thus, the p-layer acts as a grounded shielding layer. Though the effective drain voltage region is limited by drain p-layer depletion at low voltage and by avalanche breakdown (kink effect) at high voltage, this structure will enable to suppress the substrate deep trap problems in some GaAs FET circuits. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Heterojunction FETs have excellent high-speed performance compared with ion implanted MESFETs, and are promising for use in integrated circuits for high-speed optical communication systems. However, the frequency dispersion and the drain-lag, which appear to have a common origin, have created serious problems that prevent the digital application of heterojunction FETs, in that they cause what is called the mark-ratio effect, which degrades the phase margins in high bit-rate operations. This is due to a small variation in the rise and fall times of the switching circuit caused by the charged state in deep traps within the crystal. The variation may be small, but the effect causes erroneous signals in high-speed circuits when the clock cycle is close to the device transit time.

Deep traps located in the epitaxial layer and substrate interface or in the substrate itself cause the mark-ratio effect. In power FETs, the main origin of frequency dispersion may be surface states [1] in the ungated region which is located between the gate and drain to enable a high breakdown voltage. For digital applications, however, self-aligned structures are used, to avoid parasitic resistance, so there is no similar ungated region. However, contamination from oxygen and carbon is commonly observed at the substrate/epitaxial-layer interface of an MBE-grown layer, and this contamination can also cause frequency dispersion or drain-lag. Even if interface contamination is avoided, deep traps in the semi-insulating substrate will cause the same problems [2].

For MESFETs, a buried p-layer has been introduced to suppress the short-channel effects [3,4]. The buried p-layer does not necessarily suppress the frequency dis-

* Corresponding author. Tel.: +81-298-50-1521; fax: +81-298-50-1108.

E-mail address: ohno@uhl.cl.nec.co.jp (Y. Ohno)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00072-6

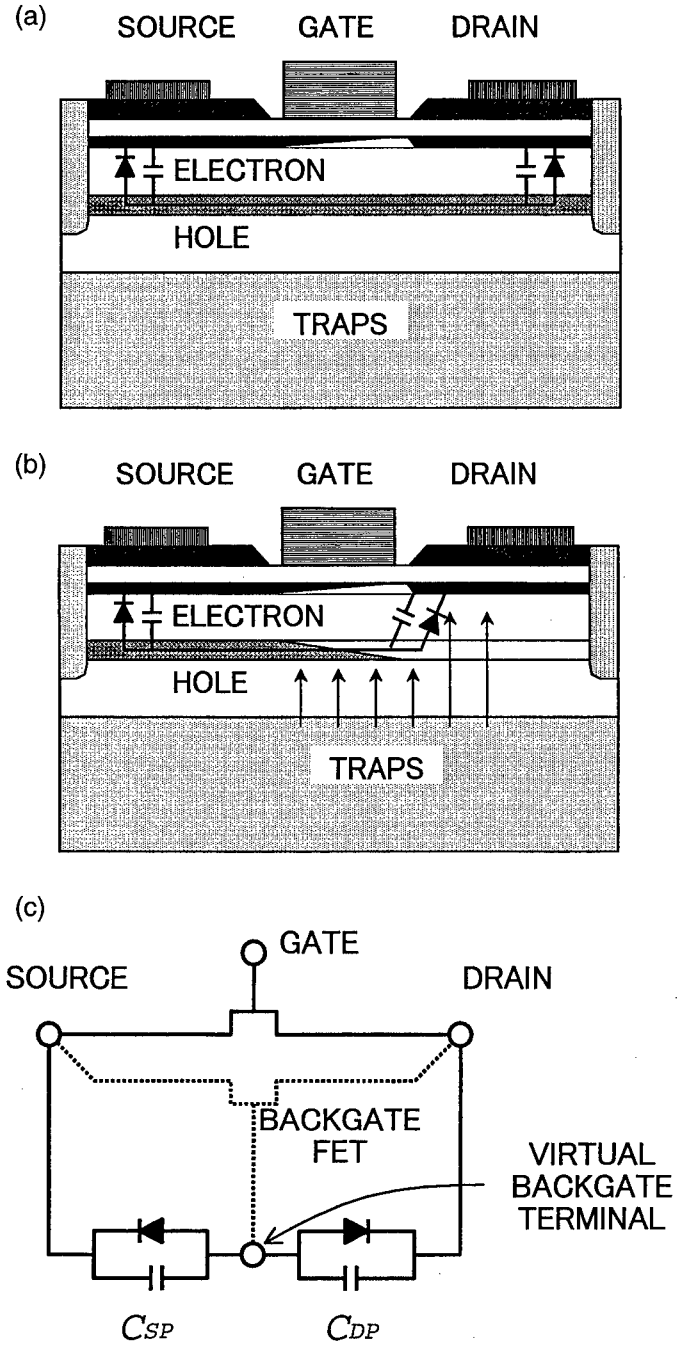


Fig. 1. Schematic explanation of trap shield model with a floating p-layer. (a) Drain voltage is low and the p-layer is not depleted. The p-region voltage fluctuates with the drain voltage; (b) Drain voltage is high enough to enable p-layer depletion. The p-layer voltage is stable at the source voltage used in this experiment; (c) Circuit model representation of the p-region voltage fluctuation.

persion [5], but such anomalies can be suppressed by a p-well structure on a GaAs MESFET [6], where the conductive p-well, whose potential is grounded through a p-well contact, screens the deep traps under the well. The disadvantage of this structure is that it requires additional process steps and a large chip area to form a contact to the well. Even if such disadvantages are tolerable, the same structure is extremely difficult to use in a heterojunction FET because selective doping by ion implantation cannot be applied since the ther-

mal annealing would degrade the modulation-doped layer.

In this paper, we propose an FET structure where the substrate deep-trap effect in a heterojunction FET is suppressed by a floating p-layer. This structure requires no additional process steps or chip area, and does not lower the high-frequency performance. We have experimentally verified the validity of the underlying theory by testing FETs with a theoretically designed p-type buffer layer.

2. Problems in trap shielding with a p-layer

When a conducting p-layer is placed between the channel and trap regions and the electric potential of the p-layer is fixed, the channel is shielded from the trap-charge effect. However, this structure requires an additional fabrication process to make an electric contact to the p-layer. In addition, the overlap between the drain and the p-layer forms an additional parasitic capacitance between the drain and the source, which will degrade the high-speed performances of the FET.

To avoid these disadvantages, the p-region is normally left electrically floating. This configuration is popular in silicon SOI MOSFETs. In terms of the DC characteristics, the floating p-layer improves the FET performance, for example, by suppressing the short channel effect and enabling good saturation characteristics. This is because the p-region potential is fixed to the source voltage even in such a symmetrical structure by the resistive balance between two pn-diodes: the forward-biased source diode and the reverse-biased drain diode (Fig. 1 a).

However, if the drain voltage is modulated at a high frequency, the p-region voltage will also vary with the drain electrode voltage due to the capacitive coupling. Expressing the capacitance between the source and the p-region as C_{PS}, that between the drain and the p-region as C_{PD}, and the drain voltage variation as δV_D, the p-region voltage fluctuation δV_P can be expressed as,

$$\delta V_P = \frac{C_{PD}}{C_{PS} + C_{PD}} \delta V_D. \quad (1)$$

Then, the drain current will be modulated according to the variation of the p-region voltage as,

$$\delta I_D = G_{MB} \delta V_P, \quad (2)$$

where G_{MB} is the virtual backgate transconductance which is mainly determined by the distance between the channel and the p-region. Fig. 1(c) is a circuit representation of this model.

When the modulation frequency is low enough, the voltage change due to the capacitive coupling is stabilized by the leakage current through the source and drain junction diodes. The critical frequency is determined by the p-region capacitance and the resistance of the diode leakage. For a frequency above the critical frequency, the drain voltage determines the p-region voltage variation, which modulates the drain current.

Table 1
Epitaxial layer structures used in the experiment

Layer	Thickness (nm)	Dopant	Concentration (cm^{-3})
(a) Conventional buffer (i-buffer)			
n-GaAs	300	Si	1×10^{18}
n-AlAs	2	Si	2×10^{18}
n-$Al_{0.2}GaAs$	35	Si	2×10^{18}
i-$In_{0.15}GaAs$	15		—
i-GaAs	500		
Semi-insulating substrate (LEC)			
(b) New structure with p-layer (p-buffer)			
n-GaAs	300	Si	1×10^{18}
n-AlAs	2	Si	2×10^{18}
n-$Al_{0.2}GaAs$	35	Si	2×10^{18}
i-$In_{0.15}GaAs$	15		—
n-GaAs	13	Si	1×10^{18}
i-GaAs	40		—
p-GaAs	20	Be	1×10^{18}
i-GaAs	500		
Semi-insulating substrate (LEC)			

The difference of drain conductance ΔG_D between the high and the low frequencies is given as,

$$\Delta G_D = \frac{\delta I_D}{\delta V_D} = \frac{\delta I_D}{\delta V_P} \frac{\delta V_P}{\delta V_D} = G_{MB} \frac{C_{PD}}{C_{PS} + C_{PD}}. \quad (3)$$

Since C_{PS} and C_{PD} are nearly the same due to the symmetrical structure of conventional FETs, C_{PD} / (C_{PS}+ C_{PD}) will be close to 0.5. In addition, the p-region should be located relatively close to the channel due to the epitaxial layer thickness limitation, so the floating p-region will cause a large dispersion [7].

However, if we can assume that C_{PD} is zero in Eq. (3), the dispersion will become zero. This concept is reasonable, because the p-region voltage is not modulated by drain-voltage variation, but is fixed to the source voltage by the stronger capacitive coupling with the source electrode. Thus, the p-region voltage is fixed to the source voltage at high frequencies as well as under DC operation.

3. Design and fabrication of the partially depleted p-layer

Such a structure is almost impossible to realize for FETs with epitaxial channel layers. Selective formation of a buried p-region under the drain region is impossible

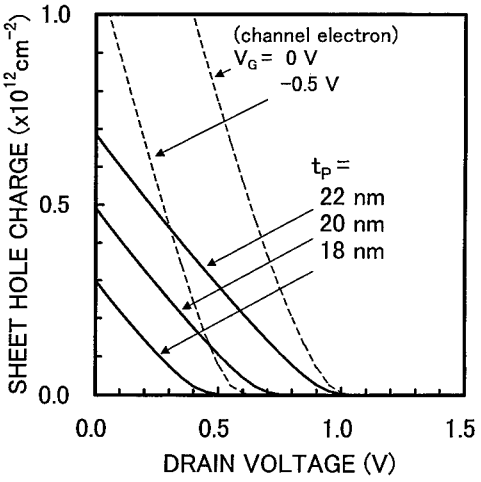


Fig. 2. Calculated sheet hole concentration in the new epitaxial structure as a function of drain voltage. The channel electron concentration is also shown.

ible because of the defect formation that will result from ion implantation. Channel layer regrowth after selective p-region formation is possible, but the alignment of the drain edge and the p-region will still be a problem.

We have designed a p-layer which will self-deplete when drain voltage is applied. When the drain voltage is increased, both the p-region and the drain n^+-region are depleted of carriers. If the p-region concentration, thickness, and depth are properly designed, the p-region under the drain n^+-region will be completely depleted at higher than a certain drain voltage (Fig. 1(b)). We call this critical drain voltage V_{DEP}. Above this voltage, the capacitance between the drain and the p-region, except for the edge capacitance eventually disappears.

Table 1 shows the epitaxial layer structures we used for the experimental heterojunction FETs. To maintain a threshold voltage close to that of a conventional

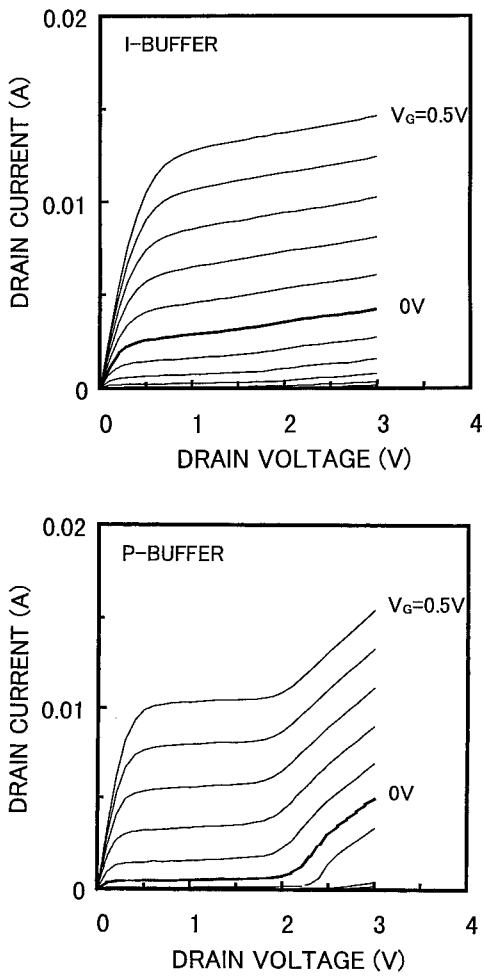


Fig. 3. I_D-V_D characteristics of the FETs. (a) FET with an i-type buffer layer; (b) FET with a p-type buffer layer.

FET, an n-type layer was inserted below the InGaAs channel.

Fig. 2 shows our one-dimensional simulation results for the sheet hole concentration in the p-buffer layer as

Table 2
FET fabrication process

Step	Conditions
1 Isolation implantation	Boron: $1 \times 10^{13}cm^{-2}$ (50 keV) + $1 \times 10^{13}cm^{-2}$(160 keV)
2 Gate recess etching	Citric Acid + H_2O_2 (5°C)
3 Gate metal	Ti/Pt/Au (lift-off)
4 Ohmic contact	AuGe/Ni/Au (Lift-off)
5 Alloy annealing	N_2, 450°C, spike
6 Passivation	CVD SiO_2
7 Contact etching	Buffered HF
8 Wiring	Ti/Pt/Au (lift-off)

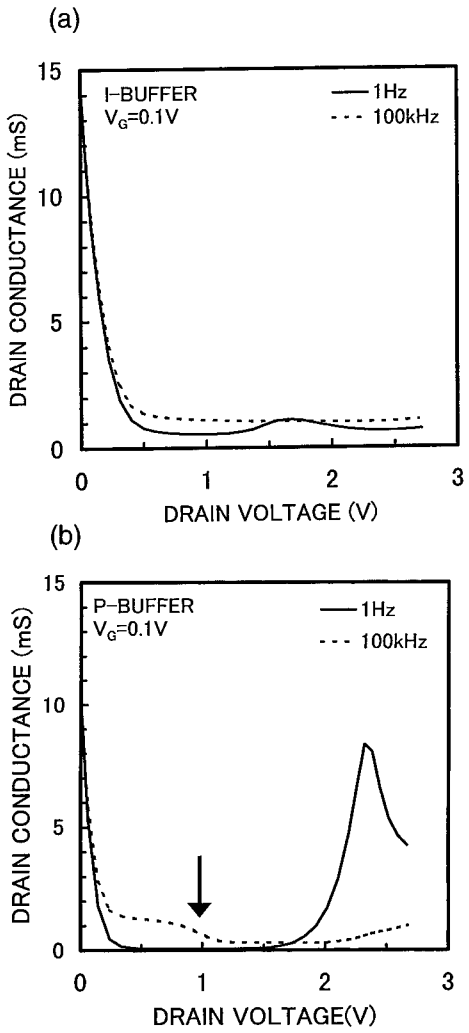


Fig. 4. Comparison of G_{DS} at 1 and 100 kHz as function of drain voltage. (a) FET with an *i*-type buffer layer; (b) FET with a *p*-type buffer layer. An arrow indicates the voltage at which the high-frequency G_{DS} falls drastically.

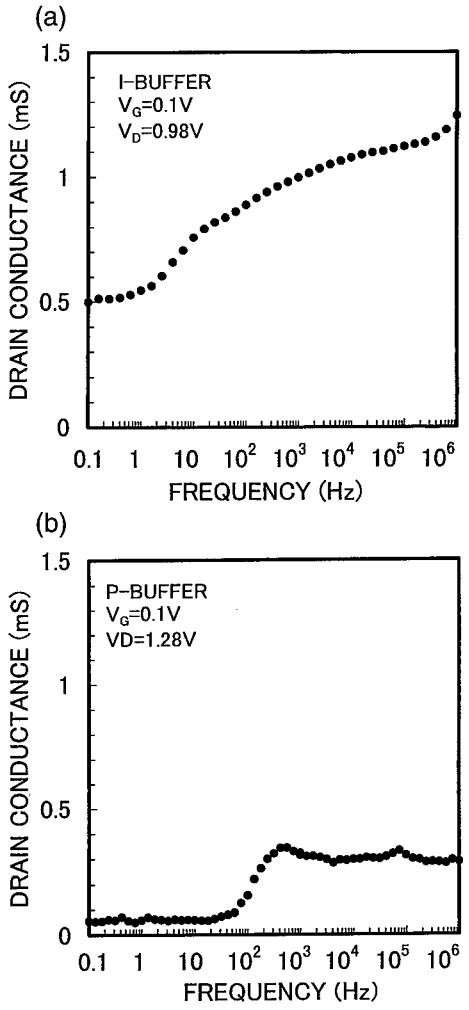


Fig. 5. Frequency dependency of G_{DS}. Since the bias was applied through a 50Ω resistor, V_D was determined by subtracting the voltage drop at the resistor. (a) FET with an *i*-type buffer layer; (b) FET with a *p*-type buffer layer.

a function of the surface *n*-layer voltage, which corresponds to the drain electrode voltage. The *p*-type layer voltage was kept at zero by making the hole quasi-Fermi level zero volts. Curves are shown for three different *p*-layer thicknesses, which give different values V_{DEP}. We chose 0.7 V for V_{DEP} (a thickness of 20 nm) in this design. For practical applications, V_{DEP} should be as low as possible to lower operation the voltage and expand the dynamic range. However, an excessively low V_{DEP} may lead to insufficient shielding under the channel. In addition, to maintain a finite C_{PS}, the *p*-layer under the source should not be deplete, so control of the *p*-region concentration and thickness will become difficult at very low values of V_{DEP}.

The wafers were prepared by MBE with Be as the *p*-layer dopant. The FETs were made by wet-etching and lift-off processes only to avoid damage during the plasma processes. For the selective etching of the cap layer, a thin AlAs layer was formed on the AlGaAs layer. The details of the fabrication process are listed in Table 2. The shortest gate length where the FETs could operate was 0.7 μm, and FETs with a channel length of 1.0 μm were measured. The total width of two FETs in combination was 100 μm.

4. Measurement results

Fig. 3 shows the DC I-V characteristics of the FETs. The G_M of both FETs was 300 mS/mm. Compared

with the conventional *i*-buffer FET, the *p*-buffer FET showed better saturation characteristics, but also showed a severe kink effect at high drain voltages. In both cases, this was due to the extremely shallow *p*-layer. The shallower the *p*-layer is, the more effectively the short channel effects are suppressed, but once the *p*-layer voltage drifts, a large drain current variation occurs due to the high virtual-backgate G_M.

Fig. 4(a) shows the drain conductance measured with small AC signals at 1 Hz and 100 kHz. There is a constant difference between the two curves in the saturation region (0.5–1.5 V). Such a frequency-dispersion is common for heterojunction FETs on semi-insulating substrates. The small hump at about 1.6 V was due to the kink effect. The kink effect also causes frequency dispersion, leading to higher conductance at low frequencies, because the current increases after avalanche holes accumulate near the source region. This effect may be slow since the hole generation is not intensive. Fig. 5(a) shows the frequency dependence of G_{DS} for the *i*-buffer FET. The transition occurs through the entire 1 Hz–100 kHz range, which indicates that different types of traps are related to the dispersion.

Fig. 4(b) compares the 1-Hz and 100-kHz drain conductance for the *p*-buffer FET. There was a large dispersion below 1 V due to the high G_M caused by the shallow *p*-layer acting as a backgate. However, the dispersion fell drastically at around 1 V, and the dispersion continues to be low up to the voltage where the kink effect begins. Fig. 5(b) shows the frequency dependence of G_{DS} in the low dispersion region. The curve shows a sharp transition corresponding to a single time constant. As explained above, the critical frequency is determined by the capacitance C_{PD} and the leakage through the pn diodes. Due to the small leakage of the pn reverse diode, the frequency is relatively low. The drain voltage of 1.0 V, at which the dispersion fell sharply, corresponded to the V_{DEP} measured from the C-V data of a large-area FET on the same wafer.

In this experiment, the drain voltage region where the frequency dispersion was suppressed was relatively small. In an SCFL circuit used for digital applications, the operating voltage can be between 0.5–1.5 V, which is attainable if the *p*-layer is properly grown. From Fig. 4(b), the suppression of the dispersion seems very effective, but comparing Fig. 5(a) and (b) suggests that the absolute value of the dispersion is still at a finite value. We need to further optimize the position and concentration of the *p*-layer. In the present FETs, we implanted boron into the *p*-layer to separate the FETs. Preserving the connection to the *p*-regions under adjacent FETs will help stabilize the *p*-region potential.

As for the high-frequency characteristics, the additional parasitic capacitance C_{PD} will be eliminated above V_{DEP}, so we do not expect any performance degradation. Fig. 6 shows that both f_T and f_{max} increased as the drain voltage increased for the *p*-buffer FET. This was due to the variation in C_{PD} caused by the drain voltage. For a sufficiently high drain voltage (2 V or higher), the high-frequency characteristics appear unlikely to be degraded by the insertion of the *p*-layer.

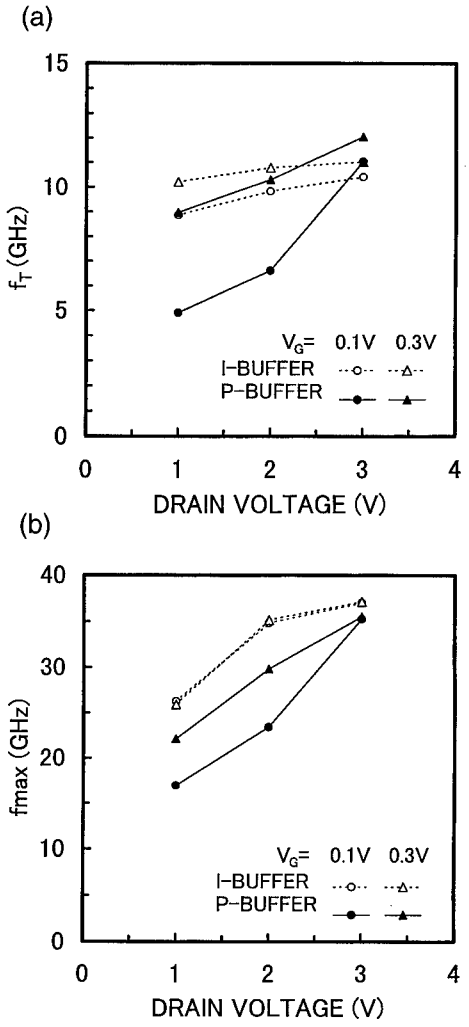


Fig. 6. Cutoff frequency and maximum oscillation frequency as functions of drain voltage for the two FETs. (a) cut-off frequency; (b) maximum oscillation frequency.

5. Conclusion

We have confirmed that the deep trap effect from the substrate can be blocked by a floating *p*-region if the holes are depleted from the drain region. In this situation, the capacitive coupling between the *p*-region and the drain is minimized and the *p*-region voltage is

stabilized to the source voltage by the p-region and the source capacitance at high frequencies. For DC and low-frequency signals, the p-region voltage is stabilized at the source voltage because the resistance of the source pn diode is lower than that of the diode at the drain, since the source-side diode bias is a forward bias while that of the drain-side diode is a reverse bias.

This structure requires no extra process steps or additional chip area. Only the precisely controlled p-layer is needed in the epitaxial layer. In this experiment, we used a very shallow p-layer to provide a shield for the channel. However, this caused a severe kink effect, which is another source of frequency dispersion. Further optimization is necessary.

In the development of a high-speed digital IC, we found that MESFET ICs with a buried p-layer showed no mark-ratio effect. The mechanism of the suppression is not yet clear, but we speculate that a similar capacitance configuration may be beneficial in MESFETs. Using this structure, stable high-speed digital ICs on heterojunction FETs will be attainable.

Acknowledgements

We thank Dr Kazuhiko Honjo for his continuous encouragement during this work.

References

[1] Canfield PC, Medinger J, Forbes L. IEEE Electron Device Lett 1987;ED-8: 88.
[2] Kasahara K, Kumihiro K, Nishizawa H, Ohmo Y. Solid-State Electronics 1995;38:1221.
[3] Yamasaki K, Kato N, Hirayama M. IEEE Trans Electron Devices 1985;ED-32: 2420.
[4] Onodera K, Tokumitsu M, Tomizawa M, Asai K. IEEE Trans Electron Devices 1991;ED-38: 429.
[5] Choi S, Das MB. IEEE Trans Electron Devices 1994;ED-41: 1725.
[6] Canfield P, Allstot D. IEEE J Solid-State Circuit 1990;25:1544.
[7] Kunihiro K, Nogome M, Ohno Y. Proc. 1996 IEEE GaAs IC Symposium, 1996. p. 179.



PERGAMON

Solid-State Electronics 43 (1999) 1347–1353

SOLID-STATE ELECTRONICS

A new method for evaluation of surface recombination in heterojunction bipolar transistors by magnetotransport

Tetsuro Nozu*, Tohru Sugiyama, Sadato Hongo, Kouhei Morizuka

Advanced Semiconductor Devices Laboratories, R&D Center, Toshiba Corporation, 1 Komukai Toshiba-cho, Saiwai-ku, Kawasaki, 210-8582, Japan

Abstract

A new method for evaluation of surface recombination in heterojunction bipolar transistors (HBTs) employing magnetotransport is reported. A new model which describes the motion of electrons in the base was developed on the basis of a diffusion equation in a magnetic field. It was shown experimentally and analyzed by this model that the magnetic force vertical to the emitter/base junction modified the electron path and suppressed the diffusion of electron to the extrinsic base surface where surface recombination took place. As a result, the surface recombination current was decreased and the current gain was increased. The present method was applied to an AlGaAs/GaAs HBT with a base doped with C at 5×10^{19} cm^{-3}. The surface recombination velocity, relaxation time, and mobility of electrons in the base were estimated without any reference samples to be 1×10^7 cm/s, 40 ps, and 2130 cm^2/Vs, respectively, which are in a good agreement with recent data. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Heterojunction bipolar transistors (HBTs) with III–V compound semiconductors have attracted considerable attention for their high-frequency performance. Among them, Npn HBTs with a GaAs base and an AlGaAs emitter grown on a GaAs substrate have been the most widely studied and developed. However, it is well known that GaAs has a high surface recombination velocity and that surface recombination at the extrinsic base region close to the emitter mesa significantly degrades the current gain. This degradation is pronounced for small devices with a large perimeter-to-area ratio of the emitter since the surface recombination current, which dominates in the base current, is proportional to the emitter perimeter length [1]. It has been recognized that the ideality factor of the surface recombination current is closer to unity than 2 if the HBTs operate in a high collector current density and the surface recombination velocity is sufficiently high [2,3]. These conditions break down the assumptions by Henry et al. that the ratio of holes to electrons at the surface is constant and the quasi-Fermi levels between the bulk and the surface are flat, which result in an ideality factor of 2 [4]. Therefore, the surface recombination current cannot be neglected even in a high collector current density where the HBTs prove their high-speed performance. The introduction of a thin depleted layer with a wide band-gap and a low surface recombination velocity onto the extrinsic base is a promising solution [5–7]. A variety of methods have also been developed to evaluate surface recombination at the GaAs surface. Most of these methods are based on optical measurement or electrical measurement. The former requires a specially designed structure which is apparently different from the HBT to obtain a sufficient intensity of photoluminescence signals [8], as in

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00073-8

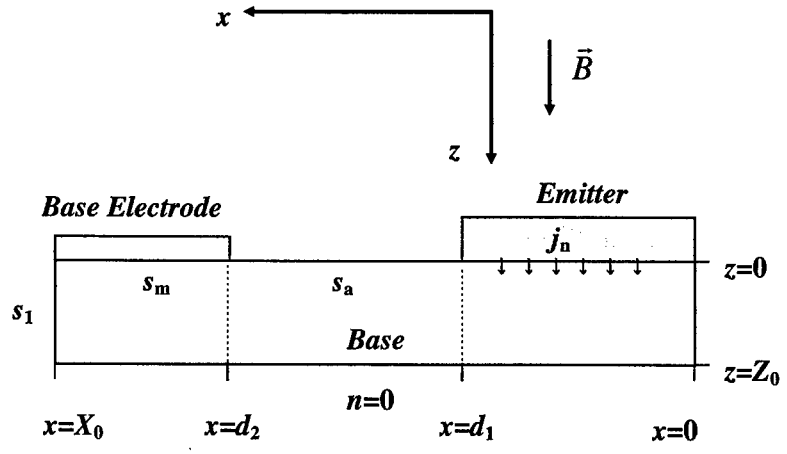


Fig. 1. Schematic illustration of one half of base region.

the case for the study of other minority carrier characteristics [9,10]. The latter is applicable to the HBT structure; however, reference samples with different surface structures or with the same layer structure but different perimeter-to-area ratios are needed [1,3,5,7]. In this sense, it has been difficult to evaluate surface recombination for an HBT by itself.

Magnetotransport in HBTs was first employed by Nottenburg et al. for the study of electron transport in InAlAs/InGaAs HBTs which exploit nonequilibrium transport due to a discontinuity in the conduction band at the emitter/base junction [11]. Betser et al. has successfully utilized magnetotransport to measure the mobility of electrons in the base of an InP/InGaAs HBT [12]. In their work, the magnetic field was applied parallel to the emitter/base junction to modify the effective diffusion constant for electrons which related the excess base current to the mobility. In the present study, we show that the surface recombination velocity can be estimated by taking advantage of a magnetic field vertical to the junction. It is also shown that the vertical field suppresses the diffusion of electrons from the emitter to the extrinsic base region and that the surface recombination velocity can be estimated by measuring the dependence of the current gain on the vertical field with the aid of a two-dimensional calculation. It is worth noting that the present method requires no reference sample as described above. Therefore, this method could be a powerful tool for studying the degradation of HBTs during a bias stress test where a reduction in current gain occurs stochastically, that is, the same degree of degradation is not expected for individual samples.

2. Model

A two-dimensional diffusion equation in the presence of a magnetic field is employed to describe the behavior of electrons in the base region. Apart from the presence of the magnetic field, the treatment of the equation is basically the same as that by Kennedy and Murley [13] in cylindrical coordinates and Liu et al. [14] in Cartesian coordinates. Fig. 1 shows a schematic illustration of one half of the base region. The emitter/base junction is in the xy-plane. The base thickness is Z_0, and the intrinsic base beneath the emitter is in the region $0 < x < d_1$ and the extrinsic base is in $d_1 < x < X_0$. The surface recombination velocities are s_a for a free surface, s_m for under the electrode metal, and s_1 for the mesa wall, respectively. The electron current density $\mathbf{j}$ in the base is given by

$$\mathbf{j} = eD\nabla n + e\mu n\mathbf{E} - \mu\mathbf{j} \times \mathbf{B}, \tag{1}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields, and n, e, μ, and D are density, charge, mobility, and diffusion constant (at $\mathbf{B}=0$) of the electron, respectively. An electric field $\mathbf{E}=0$ is assumed for simplicity. However, the case $\mathbf{E} \neq 0$ can easily be included in the model. Substituting Eq. (1) into the continuity equation, the diffusion equation in the magnetic field is obtained as follows:

$$\frac{D}{1+(\mu B_z)^2}\left(\frac{\partial^2 n}{\partial x^2}+\frac{\partial^2 n}{\partial y^2}\right)+D\frac{\partial^2 n}{\partial z^2}-\frac{n}{\tau}=0, \tag{2}$$

where τ is the relaxation time. The above expression indicates that the diffusion constant along the direction vertical to the magnetic field is reduced by the magnetic field, in other words, the diffusion in this direction is suppressed due to modification of the electron path by the magnetic force. As will be seen, the electron distribution apparently decreases along the z-direction and along the x-direction in the vicinity of the emitter mesa, $x=d_1$. Therefore, $\partial^2 n/\partial y^2$ is negligible in Eq. (2) and we finally obtain a familiar expression, except for scaling along the x-direction, as

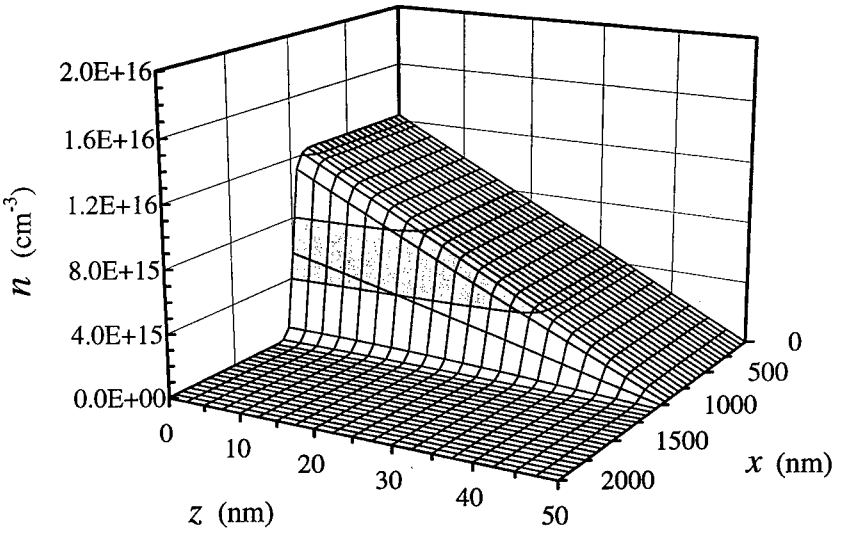


Fig. 2. Electron distribution with $B_z=0$.

follows:

$$\frac{\partial^2 n}{\partial \tilde{x}^2}+\frac{\partial^2 n}{\partial z^2}-\frac{n}{D\tau}=0, \tag{3}$$

where

$$\tilde{x}=\sqrt{1+(\mu B_z)^2}x. \tag{4}$$

Current densities are given as

$$j_x=\frac{eD}{\sqrt{1+(\mu B_z)^2}}\frac{\partial n}{\partial \tilde{x}}, \tag{5}$$

$$j_y=\mu B_z j_x, \tag{6}$$

$$j_z=eD\frac{\partial n}{\partial z}. \tag{7}$$

Eq. (3) can be solved analytically in a series expansion form. The solution and the boundary conditions are given in the Appendix. The electron mobility μ can be estimated experimentally with the parallel magnetic field, B_x [12]. The diffusion constant is derived from μ through Einstein's relation. The relaxation time τ and the surface recombination velocity are evaluated by comparing the calculated results and the experimental results.

3. Calculated results

In this section, we show the calculated results according to the model described in Section 2. Fig. 2 shows the electron distribution with $B_z=0$. The distribution with $B_z>0$ shows no difference at the scale used in Fig. 2. It is seen in Fig. 2 that apparent variation along the x-direction occurs around the mesa $x=d_1$. Therefore, the assumption $\partial^2 n/\partial y^2=0$ in Section 2 corresponds to neglecting the two-dimensional distribution over the range of 0.1 μm in the xy-plane just around the four corners of the mesa and can be justified plausibly for ordinary HBTs. We define a current I_L as follows to assess the lateral diffusion of electrons

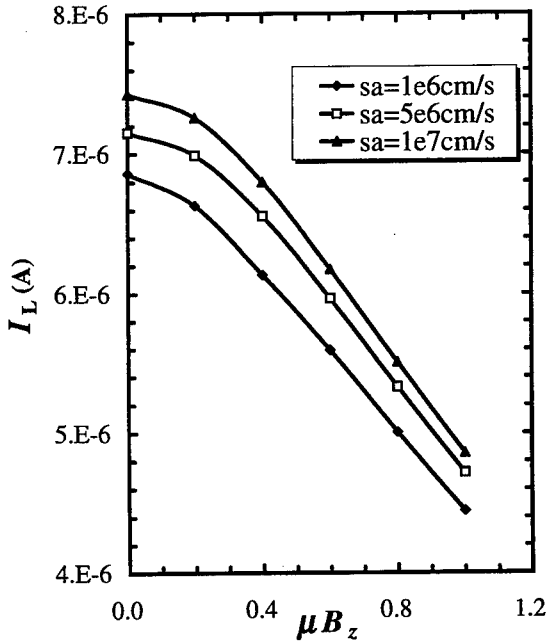


Fig. 3. Lateral diffusion current I_L as a function of μB_z with several surface recombination velocities s_a.

Table 1
Parameters for two-dimensional calculations

d_1 (μm) 1.00	τ (ps) 40
d_2 (μm) 1.20	D (cm^2/s) 50
X_0 (μm) 1.00	s_a (cm/s) 1, 5, 10 × 10^6
Z_0 (nm) 50	s_m (cm/s) 2 × 10^7
W_y (μm) 4.0	s_1 (cm/s) 2 × 10^6
	j_n (A/cm^2) 1 × 10^4

from the intrinsic to the extrinsic base region,

$$I_L = W_y \int_0^{Z_0} j_x(x = d_1, z)\, dz, \tag{8}$$

where W_y is the emitter length along the y-direction. Fig. 3 shows I_L for several values of s_a. All parameters used for the calculation are listed in Table 1. Surface recombination, except for the region around the mesa, is negligibly small, so the choices for s_1 and s_m do not result in significant differences. It is clearly shown that the vertical magnetic field suppresses the lateral diffusion of electrons. Fig. 4 shows the field dependence of the current gain h_{FE}, the surface recombination current I_{sa}, and the bulk recombination I_{bulk}. The expressions for I_{sa} and I_{bulk} are given as follows:

$$I_{sa} = e s_a W_y \int_{d_2}^{X_0} n(x,0)\, dx, \tag{9}$$

$$I_{bulk} = \frac{e W_y}{\tau} \int_0^{Z_0} dz \int_0^{X_0} dx n(x,z). \tag{10}$$

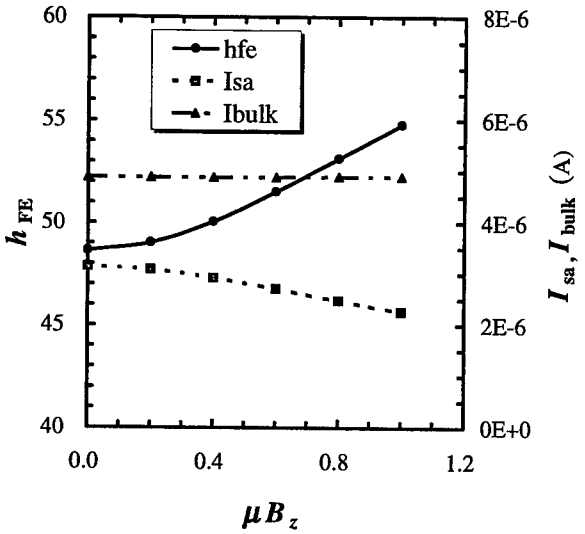


Fig. 4. Current gain h_{FE}, surface recombination current I_{sa}, and bulk recombination current I_{bulk} as a function of μB_z.

Table 2
HBT structure

Layer	Type	Doping (cm^{-3})	Thickness (nm)
$In_{0.5}Ga_{0.5}As$	n	3 × 10^{19}	50
$In_yGA_{1-y}As$ (y = 0–0.5)	n	3 × 10^{19}	50
GaAs	n	5 × 10^{18}	30
$Al_xGa_{1-x}As$ (x = 0–0.25)	n	1 × 10^{18}	30
$Al_{0.25}Ga_{0.75}As$	n	1 × 10^{17}	30
GaAs	p	5 × 10^{19}	50
GaAs	n	1 × 10^{16}	500
GaAs	n	5 × 10^{18}	500
GaAs	S.I.	–	–

The surface recombination current I_{sa} decreases with increasing field due to the suppression of the lateral diffusion. On the other hand, the bulk recombination current I_{bulk} remains constant since the electron distribution in the intrinsic base region is not affected significantly by the field and τ is treated as a constant. As a result, the current gain h_{FE} increases with the vertical field.

4. Experiment and discussion

The present method was applied to an AlGaAs/GaAs HBT. The HBT structure is shown in Table 2. The layer structure was grown by metalorganic chemical vapor deposition on a semi-insulating (100) GaAs substrate. The base layer was 50 nm thick and doped to 5 × 10^{19} cm^{-3} with carbon. Compositional grading was not employed, so no built-in electric field was present in the base layer. A standard photolithographic and lift-off process was used to fabricate the HBT with an emitter area of 2 × 4.5 μm. Fig. 5 is a schematic cross section of the HBT. The emitter mesa was formed by chemical etching to the surface of the base layer with an $H_3PO_4/H_2O_2/H_2O$ etchant. The base electrode, Pt/Ti/Pt/Au, was self-aligned to the emitter and was sintered at 350°C for 20 min in an N_2 atmosphere to obtain a low-resistance contact [15,16]. The spacing between the base electrode and the emitter mesa was approximately 200 nm. The extrinsic base region was covered with polyimide. I–V measurements were obtained at 300 K. A magnetic field up to 5 T was applied by a superconducting magnet. The HBT was operated in an emitter grounded configuration with V_{CE} = 2 V. Fig. 6 shows h_{FE} as a function of B_x and B_z. The mobility and the diffusion constant were estimated from the B_x dependence of h_{FE} according to Ref. [12] to be 2130 cm^2/Vs and 55 cm^2/s, respectively. They were in a good agreement with recent data obtained by optical methods [9,10]. We carried out a two-dimensional calculation with these parameters.

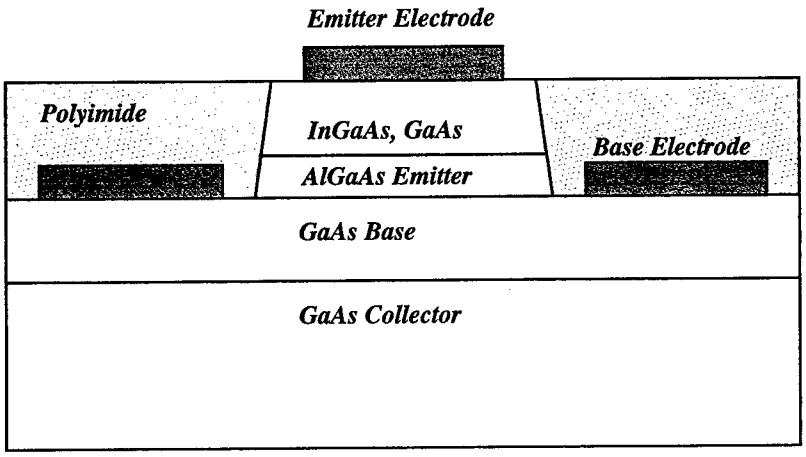


Fig. 5. Schematic cross section of HBT used in magnetotransport measurement.

Fig. 7 shows h_{FE} as a function of B_z with the calculated curve. The best fit was obtained with $\tau = 40$ ps and $s_a = 1 \times 10^7$ cm/s. The value of s_a might be larger than that previously reported in literature [14,17]. However, it should be noted that the surface recombination velocity is sensitive to the surface treatments which differ for various fabrication processes. Furthermore, previously reported data appear to have been based on the μ or D that was obtained at that time [18], and they are smaller than recent data. In Fig. 7, a small deviation between the measured data and the calculated curve is seen for large values of B_z. A possible reason is the magnetic field dependence of the surface recombination velocity. It is well known that the electron trajectory is quantized in the presence of strong magnetic field which satisfies $\mu B_z \geq 1$. In our experiment, $\mu B_z = 1$ was realized when $B_z = 4.7$ T. We treated an electron as a classical particle; however, more precise treatment based on the quantization of the electron trajectory may be needed when $\mu B_z \geq 1$.

5. Conclusion

A novel method for evaluation of surface recombination in HBTs was reported. A new model which describes the motion of the electrons in a magnetic field was developed. It was shown by the model and observed experimentally that the surface recombination current decreased with the increasing magnetic field vertical to the emitter/base junction due to the suppression of the lateral diffusion of electrons. The present method was applied to an AlGaAs/GaAs HBT and the surface recombination velocity at 300 K was estimated

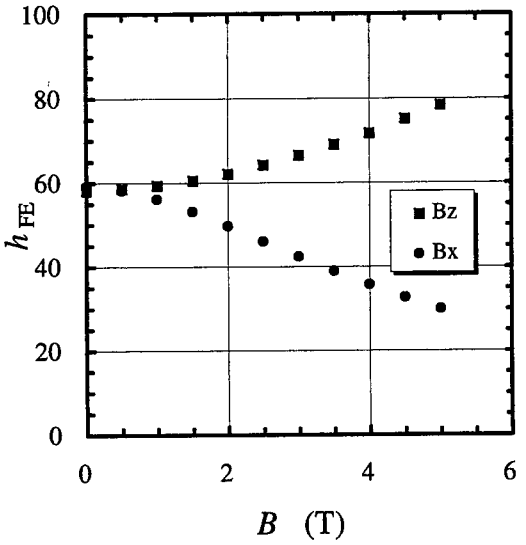


Fig. 6. Current gain h_{FE} as a function of parallel magnetic field B_x and vertical magnetic field B_z.

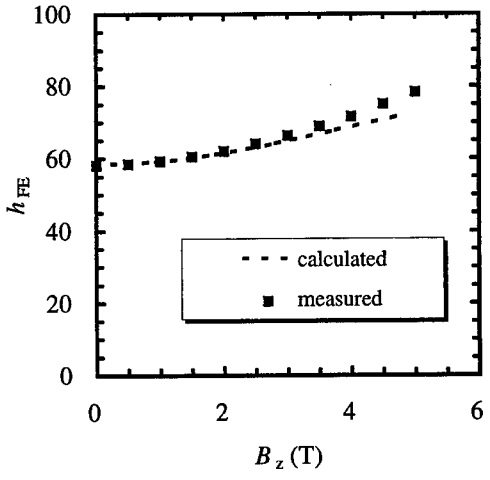


Fig. 7. Measured current gain and calculated current gain as a function of vertical magnetic field B_z.

to be 1×10^7 cm/s with the aid of a two-dimensional calculation. The mobility and the relaxation time were also estimated to be 2130 cm^2/Vs and 40 ps, respectively. They are in good agreement with recent data by optical methods. It is worth noting that we estimated the surface recombination velocity experimentally in a single HBT for the first time. The present method could be a powerful tool for studying the degradation of HBTs during a bias stress test where the same degree of degradation is not expected for individual samples.

Acknowledgements

The authors would like to thank T. Tezuka and Dr A Kurobe for measurements with a superconducting magnet and K. Tsuda for giving us a chance to make apparatus.

Appendix

In this appendix, we show how the solution of Eq. (3) in Section 2 is obtained briefly with the boundary conditions used. Similar notations as in Ref. [14] are used for convenience.

The boundary conditions are:

(a) $z = 0$,

$$\frac{\partial n}{\partial z} = -\frac{j_n}{eD}, \quad 0<x<d_1, \tag{A1}$$

$$\frac{\partial n}{\partial z} = \frac{ns_a}{D}, \quad d_1<x<d_2, \tag{A2}$$

$$\frac{\partial n}{\partial z} = \frac{ns_m}{D}, \quad d_2<x<X_0, \tag{A3}$$

(b) $x = 0$,

$$\frac{\partial n}{\partial \tilde{x}} = 0, \tag{A4}$$

(c) $x = X_0$,

$$\frac{\partial n}{\partial \tilde{x}} = -\sqrt{1+(\mu B_z)^2}\frac{ns_1}{D}, \tag{A5}$$

(d) $z = Z_0$,

$$n = 0. \tag{A6}$$

In Eq. (A1), j_n is the emitter current density. The solution of Eq. (3) in Section 2 is expressed in a series expansion as

$$n(\tilde{x},z) = \sum_{m=1}^{\infty} A_m \cos(\lambda_m \tilde{x}) \sinh\left(\frac{z-Z_0}{L_m}\right), \tag{A7}$$

where A_m is an expansion coefficient determined by the boundary conditions and λ_m is a discrete solution of the following equation which comes from (A5),

$$\lambda_m \tilde{X}_0 \tan \lambda_m \tilde{X}_0 = \frac{s_1}{D}\sqrt{1+(\mu B_z)^2}\tilde{X}_0, \quad m = 1,2\ldots, \tag{A8}$$

and L_m is defined as

$$\frac{1}{L_m^2} = \lambda_m^2 + \frac{1}{D\tau}. \tag{A9}$$

Notations with $\tilde{}$ represents a similar meaning as in Eq. (4) Section 2, that is,

$$\tilde{A} = \sqrt{1+(\mu B_z)^2}A. \tag{A10}$$

The boundary conditions (a) are rewritten as

$$\Sigma\varphi_m \cos(\lambda_m \tilde{x}) = -\frac{j_n}{eD}, \quad 0<\tilde{x}<\tilde{d}_1, \tag{A11}$$

$$\frac{n(\tilde{x},0)s_a}{D}, \quad \tilde{d}_1<\tilde{x}<\tilde{d}_2, \tag{A12}$$

$$\frac{n(\tilde{x},0)s_m}{D}, \quad \tilde{d}_2<\tilde{x}<\tilde{X}_0, \tag{A13}$$

where

$$\varphi_m = \frac{A_m}{L_m}\cosh\left(\frac{Z_0}{L_m}\right).$$

To obtain a solution which satisfies the above complicated boundary conditions at $z = 0$, we divide φ_m into three components, $\varphi_m = \varphi_m^{(1)} + \varphi_m^{(2)} + \varphi_m^{(3)}$ and add artificial conditions to determine each component. Eq. (A11) is written as

$$\Sigma\varphi_m^{(1)} \cos(\lambda_m \tilde{x}) = \left[-\frac{j_n}{eD} - \Sigma\varphi_m^{(2)} \cos(\lambda_m \tilde{x}) - \Sigma\varphi_m^{(3)} \cos(\lambda_m \tilde{x})\right] \tag{A14}.$$

As additional conditions, we required that the left-hand-side equals the right-hand-side for the region $0 < \tilde{x} < \tilde{d}_1$, and that the left-hand-side equals zero for other regions. These conditions can be written in a compact form:

$$\Sigma\varphi_m^{(1)} \cos(\lambda_m \tilde{x}) = \left[-\frac{j_n}{eD} - \Sigma\varphi_m^{(2)} \cos(\lambda_m \tilde{x}) - \Sigma\varphi_m^{(3)} \cos(\lambda_m \tilde{x})\right]\theta(\tilde{x}:0,\tilde{d}_1), \tag{A15}$$

where $\theta(\tilde{x}: 0, \tilde{d}_1)$ is a function defined as

$$\theta(x{:}a,b) = 1, \quad a<x<b, \quad 0, \text{otherwise.} \tag{A16}$$

Similar conditions are added to Eqs. (A11) and (A12). With the orthogonality of $\{\cos \lambda_m x\}$ at $0 < x < X_0$, and the normalizing factor as

$$I_0(n,m) = \int_0^{\tilde{X}_0} \cos(\lambda_n x)\cos(\lambda_m x)\mathrm{d}x = \begin{cases} \dfrac{\sin(2\lambda_n \tilde{X}_0) + 2\lambda_n \tilde{X}_0}{4\lambda_n}, & m = n, \\ 0, & m \neq n, \end{cases} \tag{A17}$$

we finally obtain

$$\varphi_n^{(1)} + \sum_m P_{nm}\varphi_m^{(2)} + \frac{s_m}{s_a}\sum_m P_{nm}\varphi_m^{(3)} = X_n, \tag{A18}$$

$$\frac{s_a}{s_a - s_m}\sum_m s_{nm}\varphi_m^{(1)} + \left[1 + \frac{s_a}{D}L_n \tanh\left(\frac{Z_0}{L_n}\right)\right]\varphi_n^{(2)} + \sum_m s_{nm}\varphi_m^{(3)} = 0, \tag{A19}$$

$$\frac{s_m}{s_m - s_a}\sum_m T_{nm}\varphi_m^{(1)} + \sum_m T_{nm}\varphi_m^{(2)} + \left[1 + \frac{s_m}{D}L_n \tanh\left(\frac{Z_0}{L_n}\right)\right]\varphi_n^{(3)} = 0, \tag{A20}$$

where

$$X_n = -\frac{1}{I_0(n,n)}\frac{j_n}{eD}\int_0^{\tilde{d}_1} \cos(\lambda_n x)\mathrm{d}x, \tag{A21}$$

$$P_{nm} = -\frac{1}{I_0(n,n)}\frac{s_a}{D}L_m \tanh\left(\frac{Z_0}{L_m}\right)\int_0^{\tilde{d}_1} \cos(\lambda_n x)\cos(\lambda_m x)\mathrm{d}x, \tag{A22}$$

$$S_{nm} = \frac{1}{I_0(n,n)}\frac{s_a - s_m}{D}L_m \tanh\left(\frac{Z_0}{L_m}\right)\int_{\tilde{d}_1}^{\tilde{d}_2} \cos(\lambda_n x)\cos(\lambda_m x)\mathrm{d}x, \tag{A23}$$

and

$$T_{nm} = \frac{1}{I_0(n,n)}\frac{s_m - s_a}{D}L_m \tanh\left(\frac{Z_0}{L_m}\right)\int_{\tilde{d}_2}^{\tilde{X}_2} \cos(\lambda_n x)\cos(\lambda_m x)\mathrm{d}x. \tag{A24}$$

The above Eqs. (A18)–(A20) are a matrix form for a $2N$ component vector $(\varphi_m^{(1)}, \varphi_m^{(2)}, \varphi_m^{(3)})$ ($m = 1, 2, \ldots, N$: N is maximum number used in series expansion), and can be solved by ordinary procedures. This calculation is easily carried out with a personal computer. The analytical method described here is suitable for our purpose since the differential method commonly used in numerical simulations requires a large amount of memory to obtain the precise distribution of electrons, which varies drastically around the emitter mesa and is the most important for our problem.

References

[1] Nakajima O, Nagata K, Ito H, Ishibashi T, Sugeta T. Jpn J Appl Phys 1985;24:L596.
[2] Tiwaria S, Frank DJ, Wright SL. J of Appl Phys 1988;64:5009.
[3] Liu W, Harris JS. IEEE Trans Electron Devices 1992;39:2726.
[4] Henry CH, Logan RA, Merritt FR. J of Appl Phys 1978;49:3530.
[5] Lee WS, Ueda D, Ma T, Pao YC, Harris J. IEEE Electron Dev Lett 1992;10:200.
[6] Malik RJ, Lunardi LM, Ryan RW, Shunk SC, Feuer MD. Electron Lett 1989;25:1175.
[7] Hayama N, Honjo K. IEEE Electron Dev Lett 1990;11:388.
[8] Ito H, Ishibashi T. Jpn J Appl Phys 1994;33:88.
[9] Harmon ES, Lovejoy ML, Melloch MR, Lundstrom MS, de Lyon TJ, Woodall JM. Appl Phys Lett 1993;63:536.
[10] Colomb CM, Stockman SA, Gardner NF, Curtis AP, Stillman GE, Low TS, Mars DE, Davito DB. J of Appl Phys 1993;73:7371.
[11] Nottenburg RN, Levi AF, Jalali B, Sivco D, Humphrey DA, Cho Y. Appl Phys Lett 1990;56:2660.
[12] Betser Y, Litter D, Bahir G, Cohen S, Sperling J. Appl Phys Lett 1995;67:1883.
[13] Kennedy D, Murley P. Solid-St Electron 1961;3:215.
[14] Liu W, Costa D, Harris J. Solid-St Electron 1991;34:1119.
[15] Sugiyama T, Kuriyama Y, Asaka M, Iizuka N, Kobayashi T, Obara M. Jpn J Appl Phys 1994;33:786.
[16] Sugiyama T, Kuriyama Y, Iizuka N, Tsuda K, Morizuka K, Obara M. IEICE Trans Electron 1995;E78-C:944.
[17] Tiwari S, Frank DJ. IEEE Trans Electron Devices 1989;36:2105.
[18] Tiwari S, Wright S. Appl Phys Lett 1990;56:563.

NOVEL DEVICES

PERGAMON

Solid-State Electronics 43 (1999) 1355–1365

SOLID-STATE ELECTRONICS

Resonant-tunneling mixed-signal circuit technology

A. Seabaugh*, B. Brar, T. Broekaert, F. Morris, P. van der Wagt[1], G. Frazier

Raytheon Systems Company, P.O. Box 660246, MS35, Dallas, TX 75266, USA

Abstract

A large-scale integration (LSI) InP-based technology is described for high-speed mixed-signal circuits. The monolithic 75-mm wafer process uses molecular beam epitaxy, InP etch stop layers, an electron-beam-defined gate, non-alloyed ohmic contacts, and 10 mask levels to provide resonant tunneling diodes (RTD's), 0.25- or 0.5-μm gate-length high electron mobility transistors (HEMT's), Schottky diodes, resistors, capacitors and two and a half levels of interconnect. Resonant tunneling circuits described here for the first time include a 2.5-GHz, ten stage, tapped shift register, a 6.5-GHz clock generator and a multivalued-to-binary converter. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Recently, the first high-speed and low-power resonant-tunneling-diode/high-electron-mobility-transistor (RTD/HEMT) circuits have been demonstrated, including a 35-Gb/s flip flop [1], a 34-GHz static frequency divider [2], a 3-GHz (40-dB spur free dynamic range) clocked quantizer [3], a 3-GHz sample & hold (> 50-dB linearity) [4], a 4-bit 2-GHz analog-to-digital converter (ADC)) [4] and an ultralow power SRAM (50 nW/bit) [5,6].

In this paper we report for the first time an LSI process for RTD/HEMT circuits. In addition, we provide new circuit results on a ten-stage 2.5-GHz shift register, a 6.5-GHz RTD clock generator, and the core circuit of a 3-bit folded quantizer (multivalued-to-binary converter) designed for use in a ripple-carry-free adder [7].

With measured switching time-constants as short as 1.5 ps [8], the RTD is the fastest large-signal semiconductor switch, and can be the enabler for ultrawideband electronic systems. Resonant tunneling diodes enhance the performance of *any* transistor technology because RTD/transistor circuits can typically be designed with fewer components, lower power dissipation and higher circuit speed than transistor-only circuits. This is illustrated in the comparison of (a) RTD/HEMT and (b) conventional HEMT latched-comparator designs (Fig. 1) operating at 25-GHz clock. The RTD's bistability enables circuit latching without the need for regenerative feedback. This saves area (6×) and power (3×) due to a reduction in the number of components. With fewer components, wiring delay is reduced and signal slew is increased. For circuit topologies which use the HEMT in the source follower configuration and the RTD as the gain element, speeds can exceed the transistor-only f_T limit.

Large-signal SPICE models of RTD's and HEMT's are now available and show good agreement with experiment when fitted to both dc and *S*-parameter measurements [4]. The RTD model is physics-based [9] and includes a dc analytic expression for the tunneling current–voltage behavior; recently an ac model which

* Corresponding author. Tel.: +1-972-344-3635, Fax: +1-972-344-3638.

[1] Present address: Rockwell International

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00074-X

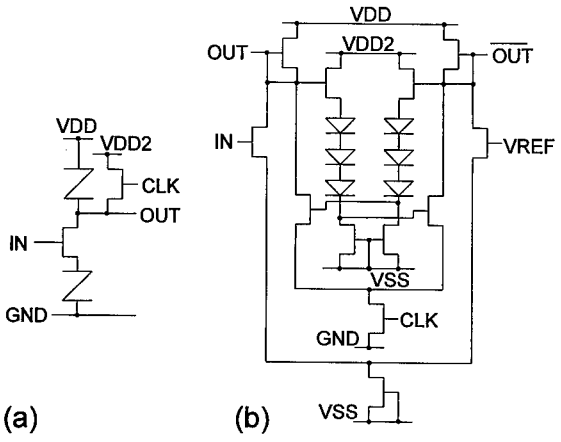


Fig. 1. Circuit schematic diagrams for (a) RTD/HEMT and (b) HEMT comparators designed for 25-GHz operation.

accounts for the capacitance–voltage dependence has been described [4]. Close agreement between measurement and RTD SPICE model is obtained not only in the prepeak and valley regions of the RTD, but also in the negative differential resistance region [4]. The HEMT model of Lee et al. [10] gives good fits to measurements, from subthreshold to saturation and from dc to 50 GHz [4]. In addition, these models are continuous for all bias conditions, an important feature for guaranteeing convergence in RTD/transistor circuit simulations.

2. Process technology

An integrated circuit (IC) process has been demonstrated for both 50- and 75-mm InP substrates. The device heterojunctions are grown by molecular beam epitaxy in a single growth, approximately 0.5 μm in thickness. As shown in Fig. 2, the RTD is integrated vertically above the HEMT. Indium phosphide etch-stop layers are used for precise placement of source/drain and gate metallizations in the heterostructure. With the RTD epitaxially integrated with the HEMT, layout is compact with the RTD located in the source and/or drain of the HEMT.

Shown in Fig. 3 is a scaled drawing, accurate in the vertical dimension, and showing the fabrication sequence for the RTD and HEMT. In the first step the emitter metal is formed using a lift-off process and used as an etch mask for reactive ion etching of the RTD mesa using a methane/hydrogen plasma. The etch is completed using a wet sulfuric acid/peroxide etch to a thin (3–5 nm) InP etch stop layer which precisely locates the level of the source/drain contact. The device mesa is formed by selective wet-chemical etching to the underlying semi-insulating InP substrate and the non-alloyed source/drain metallization is defined after removing the InP etch stop layer in a phosphoric/hydrochloric acid etch. The gate is patterned and recessed to a second InP etch stop layer which provides precise positioning of the gate with respect to the channel. Note that only three mask levels and one electron-beam (e-beam) lithography step are needed to form

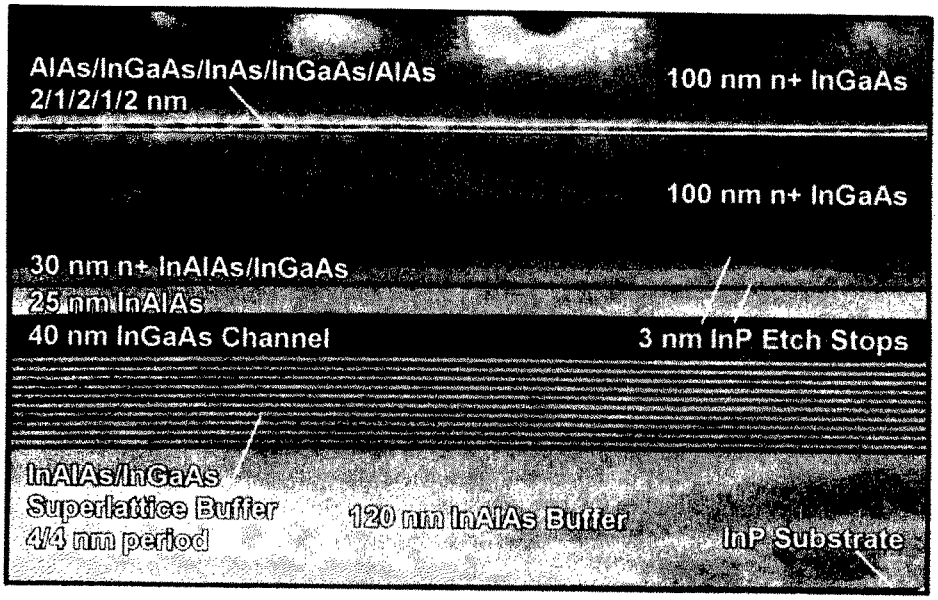


Fig. 2. Transmission electron micrograph of RTD/HEMT heterojunction material grown by molecular beam epitaxy on Fe-doped InP. The total epitaxial layer thickness for both the RTD and HEMT devices layers is 509 nm.

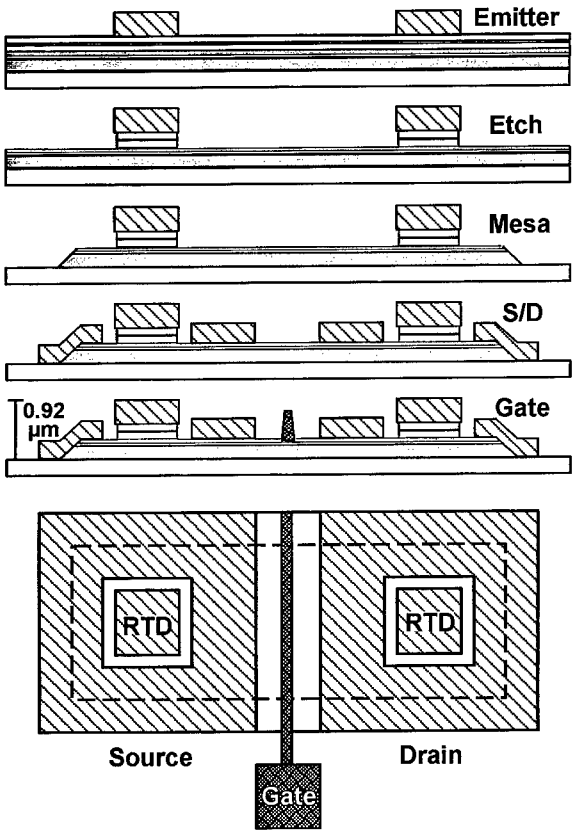


Fig. 3. Schematic diagram of the RTD/HEMT process flow with the vertical dimension of the epitaxial layers and metallizations drawn to scale. The lower drawing shows a top of view of the HEMT, where the RTD can be placed in the source, in the drain, or in both as shown. The upper drawings show cross-sections of the four-pattern, three-metallization process which forms both the RTD and HEMT in approximately the footprint of a single transistor.

both active devices, HEMT and RTD. The topography including contact metallizations is approximately half of the topography of a heterojunction bipolar transistor process.

In the next four mask levels shown in Fig. 4, resistors, capacitors and two wiring metal levels are formed. After planarizing the components with a dielectric layer, Fig. 5, vias are opened and a final wiring layer is created. In summary, 10 mask levels and 1 e-beam step are used to form RTD's, HEMT's, Schottky diodes (formed in the same step as the HEMT gate), resistors, capacitors and 2+ levels of interconnect. The extra + or "half" level of metallization is because the top plate of the capacitor (Metal2 in Fig. 4) can be used effectively for a limited amount of wiring. Process time is approximately 80 h.

A top–down scanning electron microscope (SEM) image of the process is shown in Fig. 6.Here the three-dimensional aspect and the device and interconnect economy are apparent (see also the lower portion of Fig. 3). The RTD's are located immediately below the top level via contact. The RTD/HEMT combination has approximately the same footprint as the HEMT alone.

The process described above uses only a single type of RTD over the HEMT, but it can be extended to allow the connection of two RTD types by the addition of a single mask level. An example of this process is shown in the cross sections of Fig. 7; in this case single- and four-peak RTDs were formed on the HEMT drain and source regions, respectively. A gold airbridge was used as the final wiring level.

3. Device characteristics: RTD

A representative four-peak RTD consisting of the series epitaxial integration of four independent RTD's is shown in Fig. 8(a) with the corresponding schematic conduction band diagram for the RTD's shown in Fig. 8(b). The RTD's are typically interconnected by 60 nm of n+ InGaAs layers. The AlAs/InGaAs double-barrier structure is uniformly doped with Si to minimize the peak voltage and hysteresis in the current–voltage (I–V) characteristic of series RTD's. Generally, doping decreases the peak voltage as more of the voltage drop appears across the double barrier and increases the valley current due to an increase of the inelastic scattering rate in the quantum well [11] (Fig. 9) Typical RTD's produced in this process have a peak current density of 20 kA/cm^2 with a speed index (peak-current/capacitance) of 100 mA/pF (or 0.1 V/ps).

4. Device characteristics: HEMT

The HEMT layer diagram and computed energy band diagram are shown in Figs. 10(a) and (b), respectively. The heterostructure consists of a modulation-doped InAlAs/InGaAs two-dimensional electron gas above a superlattice buffer layer. The superlattice buffer layer is intended to improve electron confinement by increasing the effective barrier height of the InGaAs/InAlAs heterostructure.

S-parameter measurements shown in Fig. 11 show a current gain cut-off frequency, f_T, of 140 GHz with a power gain cut-off frequency, f_{MAX}, of 180 GHz for a 0.25-μm gate. The dependence of f_T on inverse gate length is plotted in the inset of Fig. 11 showing the linear scaling of the technology on inverse gate length as expected based on reduction of the carrier transit time. In Fig. 12, the dependence of f_T on the gate-to-source voltage is plotted showing good frequency response over a wide range of operating voltage.

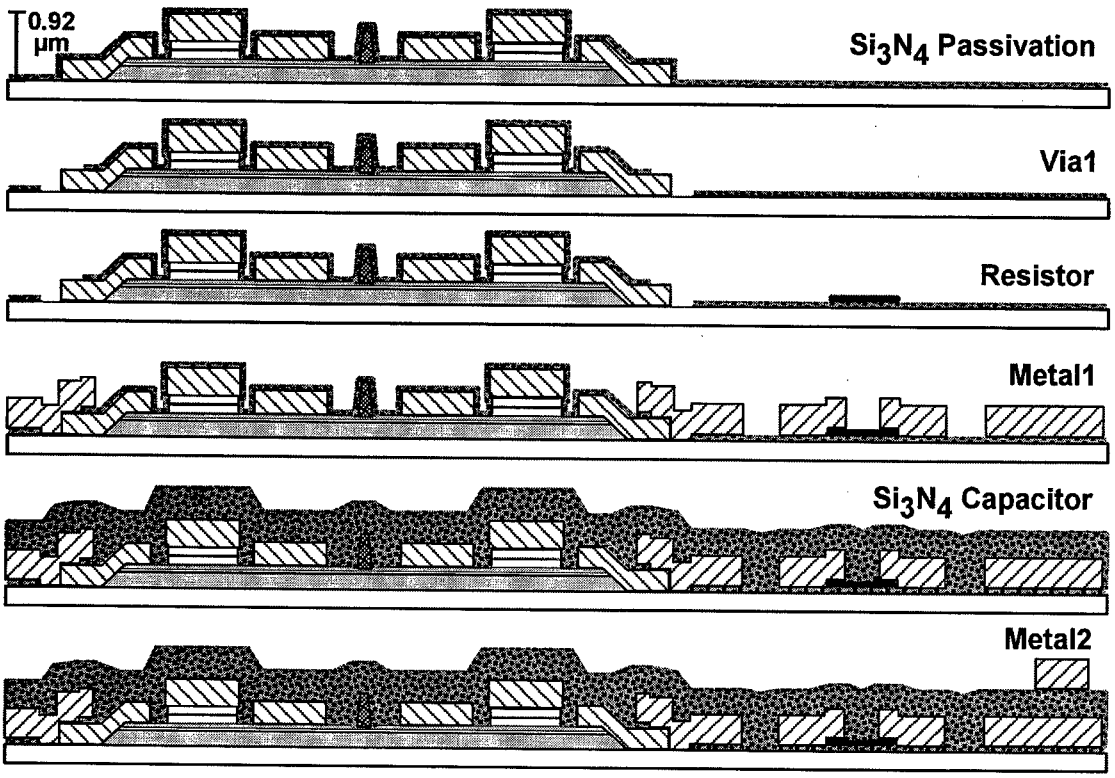


Fig. 4. Continued process flow (with vertical dimension drawn to scale) showing the next four process patterns which provide resistors, capacitors, and two interconnect levels. The first metal layer serves as both an interconnect layer and as the bottom plate of a capacitor; the second metal layer is used to form the top capacitor plate and for a limited amount of wiring.

5. Device and process characteristics: RTD/HEMT

Common source characteristics of the RTD/HEMT combination with the RTD in the source are shown in Fig. 13 along with the corresponding gate current. The gate current in this technology is typically about 50 nA/μm. When the extrinsic gate-to-source voltage increases and the channel current exceeds the peak current in the source RTD that RTD switches to a high voltage state. On switching, the internal gate-to-source voltage is forced to a lower value because the extrinsic gate-to-source voltage is held constant. This causes the sudden current decrease whenever the HEMT current exceeds the RTD peak current.

Good local and global uniformity of RTD *I–V* characteristics is shown in Figs. 14 and 15. Fig. 14, shows the *I–V* obtained from the series combination of 500 RTD's, each of the 500 RTD's switches independently in sequence from the lowest peak current to the highest peak current and when the voltage retraces the switching sequence proceeds from highest to lowest valley current. The uniformity of the peak and valley currents across 500 RTD's is shown to be better than approximately 3%. In Fig. 15, the uniformity across a 50-mm wafer is shown with the dashed curves corresponding to the wafer edges. This RTD with peak current density exceeding 10^5 A/cm^2 is highly sensitive to monolayer thickness fluctuations [12], but in this process shows remarkable uniformity across the full wafer, an indication of monolayer control of the barrier thickness.

Sizing of the RTD was used in an RTD/HEMT analog-to-digital converter [4] to set the trip points of a flash quantizer. In this application precise control the RTD area is required so as not to limit the linearity of the quantizer. For devices formed using the methane/hydrogen reactive ion etch process described earlier, the dependence of RTD peak current has been plotted against the RTD mesa area for two separate wafers which were grown consecutively but processed separately (Fig. 16). Both the growth and the etch process show good uniformity and reproducibility. The linearity in area obtained in this process is better than one part in 128, or 7 bits.

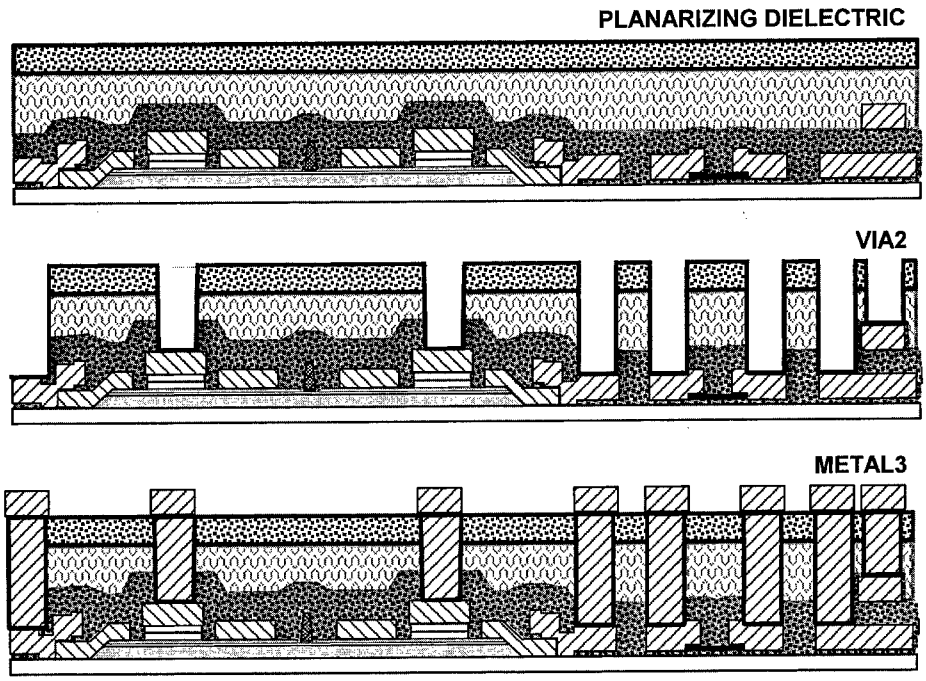


Fig. 5. Continued process flow shows the planarization and final two mask patterns to complete the interconnect process. Following this sequence the devices are passivated with silicon nitride and vias are opened to the bonding pads. In total 10 optical patterns are used; the gate is defined by electron beam lithography.

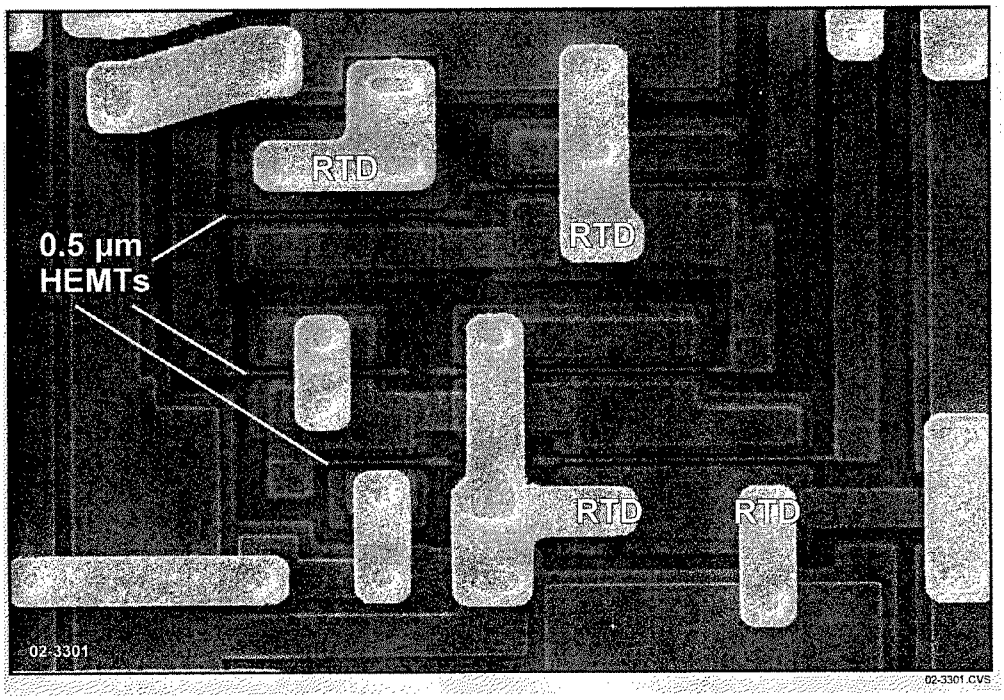


Fig. 6. Scanning electron micrograph of a completed RTD/HEMT quantizer circuit showing the three-dimensional integration of RTD's on HEMT source and drain contact regions.

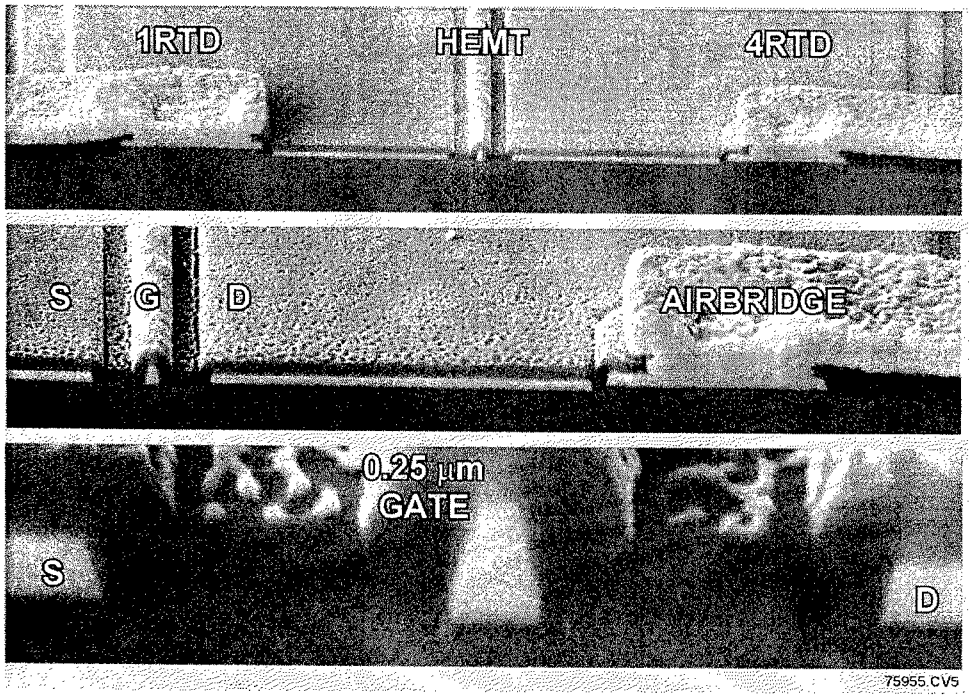


Fig. 7. Three cross-sectional scanning-electron-microscope views of an integrated RTD/HEMT structure; each view with a different magnification. As a reference the 0.25-μm gate is visible in each micrograph. The device cross section is formed by focussed-ion beam milling and the image is taken at a glancing angle to the wafer surface. In this structure two different RTD's are formed: a single RTD (1RTD) and an epitaxial stack of four RTD's (4RTD); this is readily accomplished by the addition of one additional mask step and the 4RTD's are grown in the epitaxy. The textured surface above the gate metallization is an artifact of the ion milling used to cross-section the wafer.

6. Circuits: RTD/HEMT shift register

As digital RF receivers are extended into the X-band (8–12 GHz), tapped shift-registers with input data rates exceeding 25 GHz will be needed to provide demultiplexed, parallel outputs at clock rates compatible with silicon signal processors. A single-stage of an RTD/HEMT static shift register is shown in the inset of Fig. 17. Data is clocked into the bistable RTD pair on one clock half-cycle and transferred out on the next half-cycle. Average power dissipation per stage is approximately 9 mW.

The output eye diagram shown for a pseudo-random input bit stream in Fig. 17 gives a qualitative measure of the operation of the shift register. The open-eye diagram is an indication of reliable data transmission at a clock rate of 2.5 GHz.

A direct characterization of the shift register is to measure the delay per tap. Four taps (1, 2, 5 and 10) were connected to 50-Ω pin-drivers. Fig. 18 (lower plot) shows the outputs from the four taps, for an input signal of 156.25 MHz and a clock frequency of 2.5 GHz. The data advances every half clock-cycle. The output swing from the pin-drivers is limited to less than 50 mV due to a mask layout error on the pin drivers, however clear latching and shifting of the input signal is observed. Fig. 18 (upper plot) shows the (relative) position of the falling edge of the signal from each tap plotted against tap number for both 1- and 2.5-GHz clocks. As expected, the slope of the best-fit line corresponds to half the period of the shift register clock, demonstrating circuit functionality. The tap delay extracted from the 1-GHz data is more accurate as the time scales are less sensitive to unaccounted for delays in the circuit layout and the test configuration. Power dissipation is 650 mW including pin drivers and clock shapers.

A good quantitative characterization of the high-speed operation of the shift register is a bit-error rate test. The input signal is a pseudo-random bit stream of word-lengths varying from 2^7-1 to $2^{31}-1$. Error-free operation of the shift register (tap number 2) was measured up to a clock frequency of 2.5 GHz for a 2^7-1 word-length. Changing the word-length to $2^{31}-1$ resulted in a bit-error-rate of 5×10^{-10}.

7. Circuits: RTD clock generator

On chip high-speed clocks are needed for digital sys-

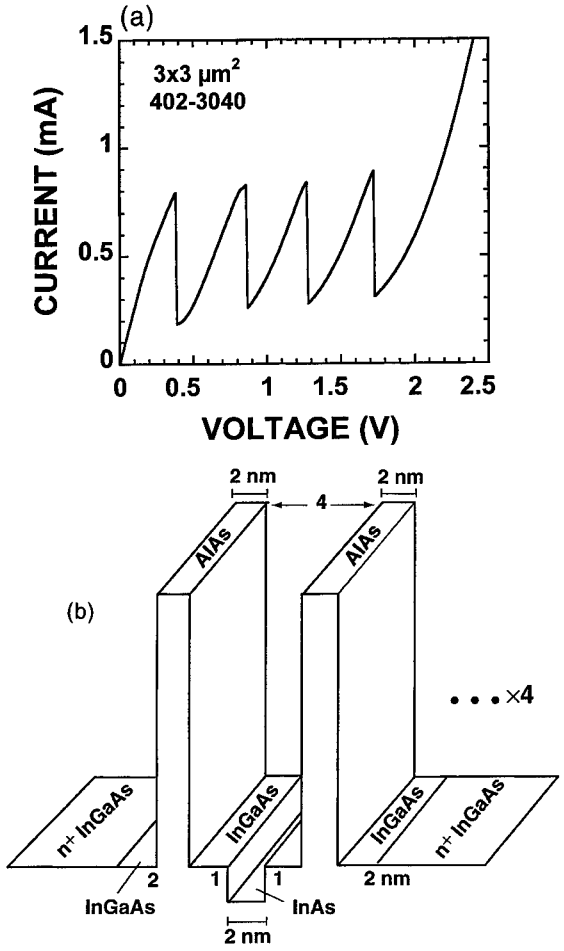


Fig. 8. Vertically-integrated four-peak resonant tunneling diode grown by molecular beam epitaxy: (a) measured current–voltage characteristics and (b) schematic energy band diagram and layer thicknesses. This particular RTD is used in the multivalued-to-binary converter circuit shown in Fig. 22.

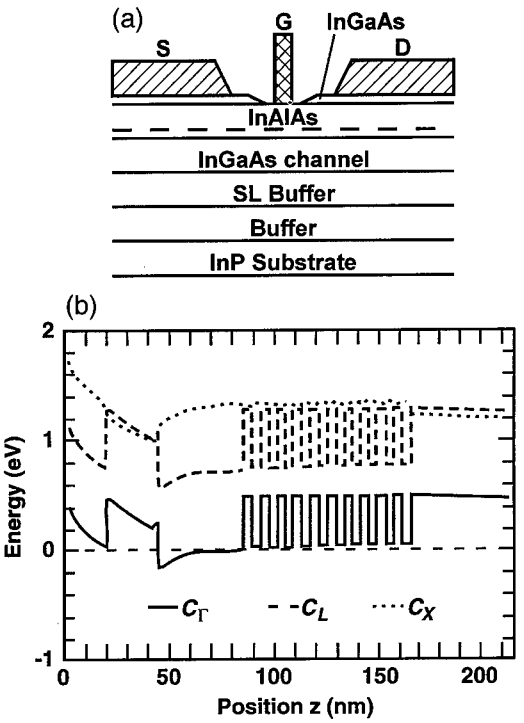


Fig. 10. A schematic layer diagram of the $In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As$ HEMT is shown in (a) where the dashed line in the drawing indicates the presence of a Si modulation-doping layer and 'SL' stands for superlattice. In (b) a computed band diagram (BandProf, Poisson solver of W.R. Frensley) of the HEMT is shown including the bands associated with the satellite valleys located along the L and X crystallographic directions.

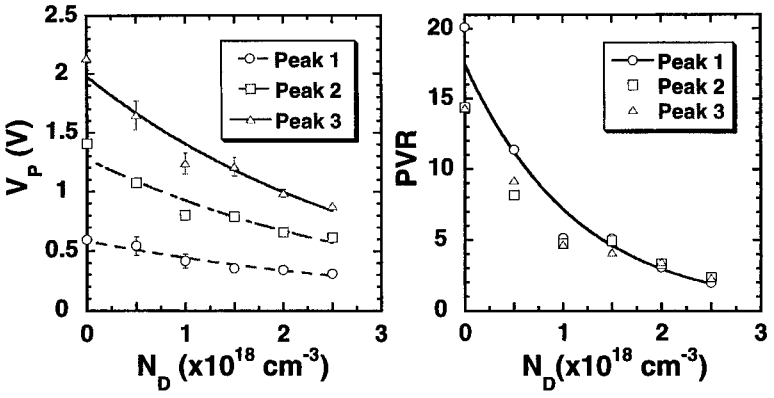


Fig. 9. Dependence of resonant tunneling diode peak voltage (V_P) and peak-to-valley current ratio (PVR) on doping density through the double barrier for an epitaxially-integrated 3RTD stack with RTD structure equivalent to the structure of Fig. 8(b). A doping density of zero corresponds to the nominally-undoped case in which dopants are shuttered off during the growth of the double-barrier structure including the 2-nm InGaAs cladding layers. The other doping densities correspond to a uniform dopant density throughout the double barrier and cladding layers; the n+ InGaAs layers have the same doping density between samples.

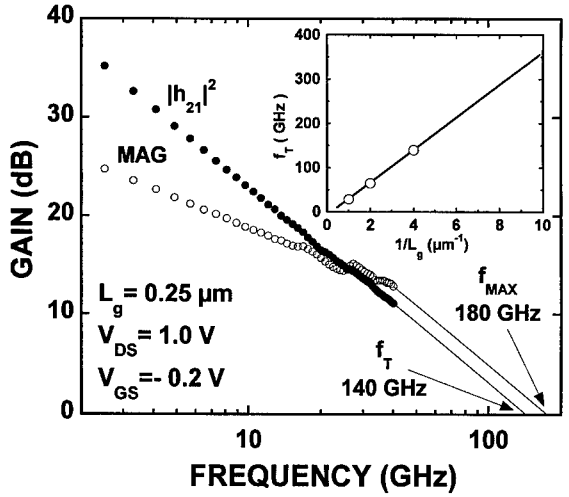


Fig. 11. Dependence of the common-source HEMT small-signal current gain (h_{21}) and maximum available power gain (MAG) on input signal frequency and (inset) the dependence of the current gain cut-off frequency (f_T) on gate length.

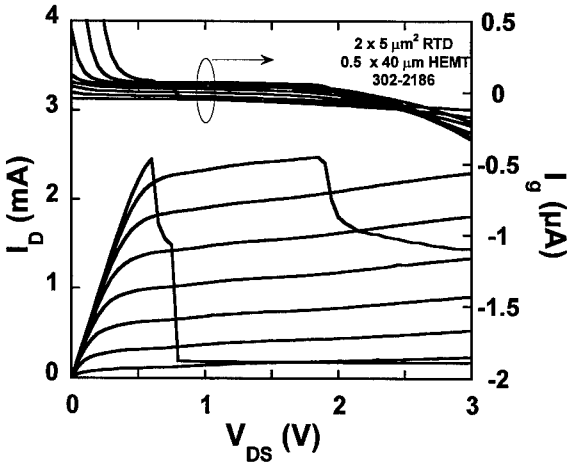


Fig. 13. Common source HEMT transfer and gate leakage characteristics with an integrated RTD in the source. The HEMT has a gate length of 0.5 μm and gate width of 40 μm; the RTD area is 2 × 5 μm^2.

tems. Traveling-wave RTD pulse generators have already been shown to operate up to 40 GHz [13] and with timing jitter as low as ~200 fs [14]. In Fig. 19, the results of a new differential clock generation circuit [15], fabricated in the RTD/HEMT process, are shown. The differential design of the RTD bridge is capable of driving 50-ohm loads. With V_{DD} = 1.4 V a 6.5 GHz output signal is generated. Once the oscillation develops, the external clock inputs can be removed and the circuit will clock freely with a speed determined by the round trip transit time in the resonant line. A close-up of the output signal is shown in Fig. 20, demonstrating a rise time of 25 ps.

8. Circuits: RTD/HEMT redundant digit adder

High-speed ripple-carry-free addition can be achieved using redundant number systems [7,16]. These number systems typically require multivalued transfer characteristics like that shown in the upper row of Fig. 21. Here, an alternative digital output signal is obtained for a monotonically increasing input signal.

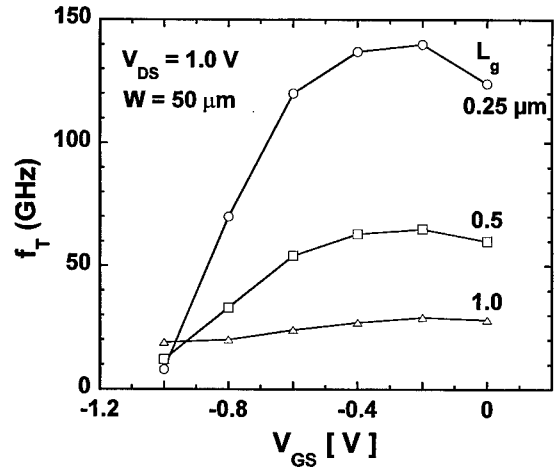


Fig. 12. Dependence of the HEMT current gain cut-off frequency (f_T) on gate-to-source voltage (V_{GS}) and gate length (L_g).

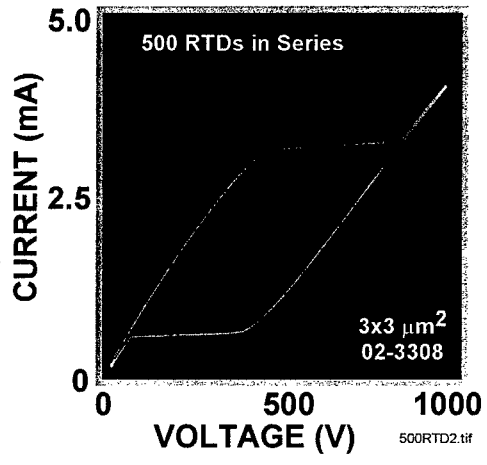


Fig. 14. Five hundred resonant tunneling diodes in series showing only a few percent variation of peak and valley currents over the entire device array: array dimension 500 × 200 μm^2, 50 RTD's × 20 RTD's.

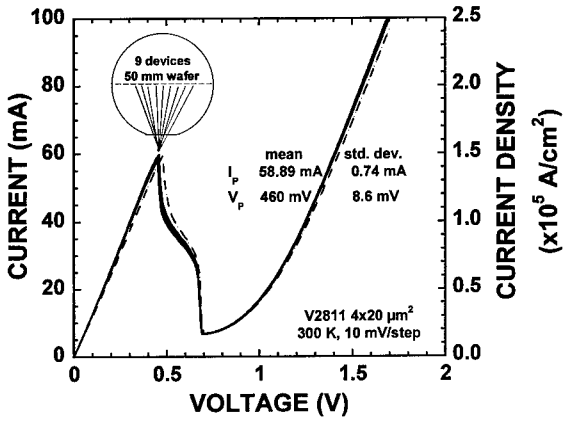


Fig. 15. Dependence of current–voltage behavior versus position across a 500-mm InP wafer.

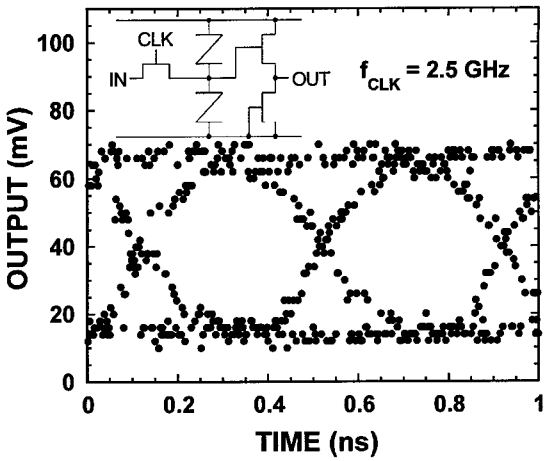


Fig. 17. Eye diagram and circuit schematic of one stage of a 10-stage RTD/HEMT shift register clocked at 2.5 GHz. The bit-error-rate is zero for a word length of 2^7-1 and 5×10^{-10} for a word length of $2^{31}-1$.

These transfer characteristics can be readily generated with the RTD/HEMT circuit in Fig. 21. The symbol labeled 3RTD refers to the series combination of 3RTD's and 1RTD refers to a single RTD. The data of Fig. 21 is from a SPICE simulation, which allows a close inspection of the internal node voltages and currents for a ramped input voltage from −0.5 to 2 V in 3 ns.

As the input voltage is increased from −0.5 V, the input transistor begins to turn on when V_{GS} exceeds its threshold voltage −0.42 V. Current then increases through the load RTD until it exceeds the RTD peak-current at which point the voltage across the load RTD jumps out by approximately 0.2 V giving rise to the first transition in the output voltage, V_{OUT}. As the input voltage continues to increase, the voltage across the 3RTD stack reaches its first peak and the 3RTD stack switches to a higher voltage. This causes V_{GS} to decrease, turning the input transistor off and resetting the RTD load to its low-voltage, low-current state. The process then repeats twice as the input voltage

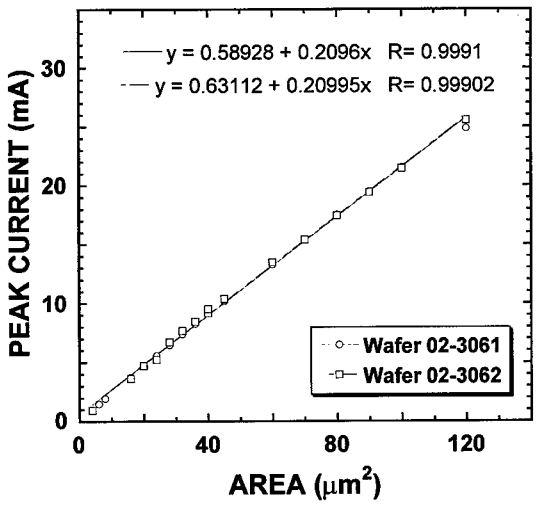


Fig. 16. Areal dependence of the resonant tunneling diode peak current for two separately processed wafers showing good linearity of the emitter mesa reactive ion etch process.

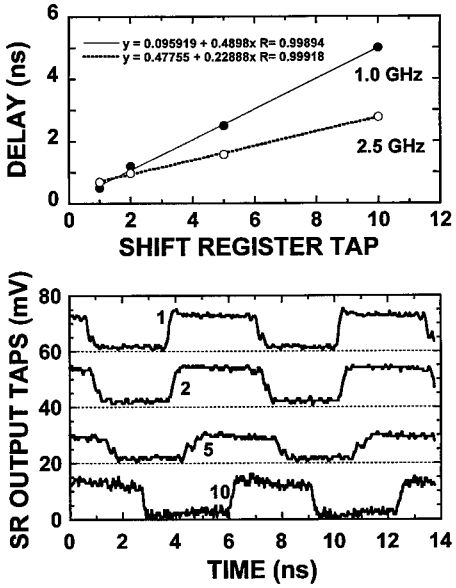


Fig. 18. Ten stage RTD/HEMT shift register with output taps after 1, 2, 5 and 10 stages. The lower figure shows the measured voltages as a function of tap points with the signals displaced along the y-axis for clarity. The upper figure shows that the delay increases linearly with the tap position.

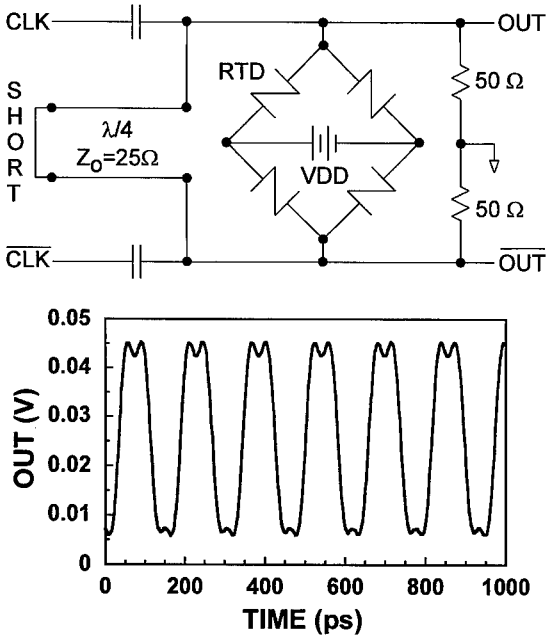


Fig. 19. Resonant tunneling diode clock generator: (a) schematic diagram and (b) 6.5-GHz output signal waveform.

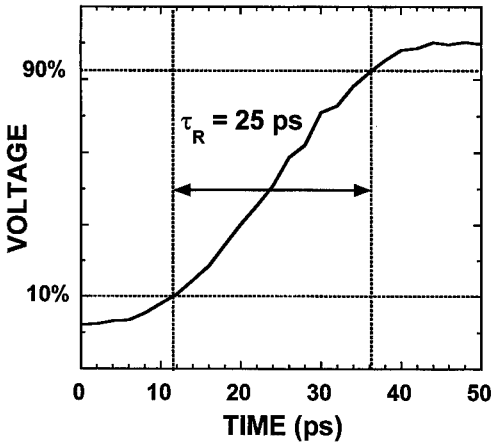


Fig. 20. An expanded view of the output signal for the resonant tunneling diode clock generator of Fig. 19.

continues to advance toward 2 V and through the three switching transitions of the 3RTD stack.

The circuit was demonstrated (Fig. 22) using 0.5-μm gate length RTD/HEMT technology with input HEMT width of 75 μm and load HEMT width of 7 μm. In practice, to yield the circuit with good uniformity of peak current in the stacked RTD pairs it is preferred to use a 4-peak vertical integration. In this way three diodes can be defined in the same etch process so as to provide closely matched peak currents. The circuit function only requires the 3-peak resonant tunneling diode to obtain the needed transfer characteristic.

9. Conclusion

An LSI process for combining 0.25- or 0.5-μm gate-length HEMT's and RTD's was described for ultrahigh speed mixed signal circuits. The 75 mm wafer

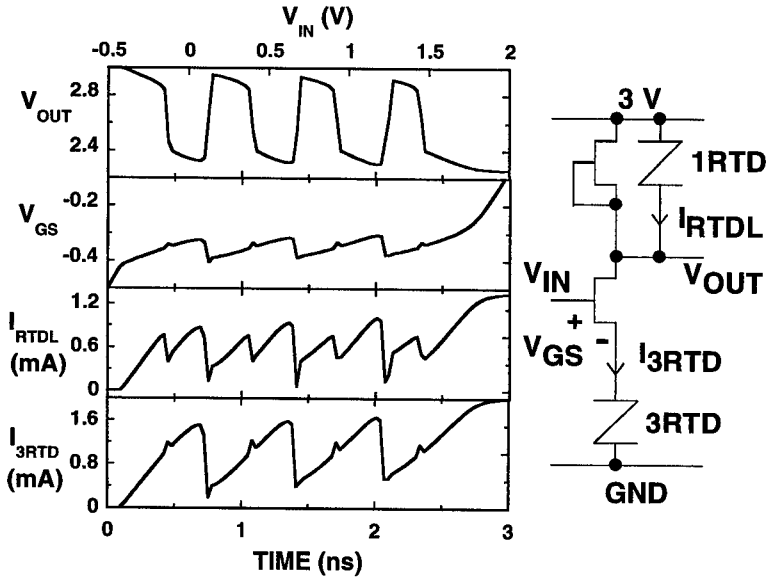


Fig. 21. SPICE simulation of the voltage and current waveforms of an RTD/HEMT multivalued-to binary-converter (quantizer).

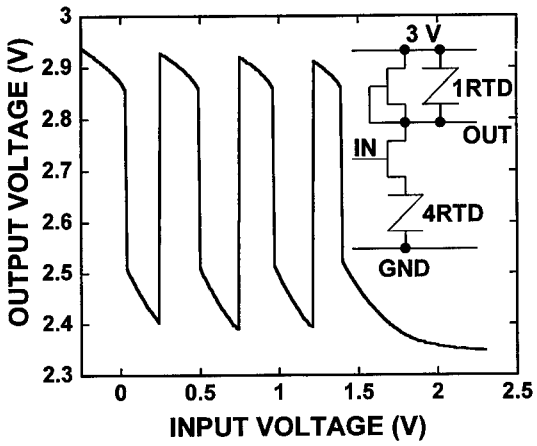


Fig. 22. Measured dc characteristics of an RTD/HEMT multivalued-to-binary converter, shown schematically in the inset. The symbol labeled 4RTD refers to the series combination of four epitaxially-integratedresonant tunneling diodes; 1RTD refers to a single resonant tunneling diode

process uses molecular beam epitaxy, InP etch stop layers, an electron-beam-defined gate, non-alloyed ohmic contacts, and 10 mask levels to provide HEMT's, RTD's, Schottky diodes, resistors, and capacitors, with two and a half levels of interconnect. Resonant tunneling diode and RTD/HEMT circuits described here for the first time included: a 2.5-GHz ten-stage tapped shift-register, a 6.5 GHz clock generator, and a multi-valued-to-binary converter. The 0.25-μm RTD/HEMT technology has the necessary bandwidth for signal processing at 10–100-GHz clock rates.

We would like to acknowledge the contributions of R. Aldert, S. Balch, E. Beam III, D. Chasse, W. Dunn, J. Hellums, Y.-C. Kao, A. Ketterson, T. Moise, B. Newell, I. Obeid, E. Pijan, J. Randall, M. Robinson, H. Schindler, T. Smith, P. Stickney, A. Taddiken, R. Turner, and P. Williams. This work was funded in part by AFOSR/DARPA, contract No. F33615-93-C-1256.

References

[1] Yamamoto M, Maezawa K, Matuszaki H, Arai K, Osaka J. In: 3rd Int Wkshop Quantum Functional Devices Assoc Future Electron Dev Japan, 1997. p. 33.
[2] Maezawa K, Matsuzaki H, Yamamoto M, Otsuji T. IEEE Electron Dev Lett 1998;19:80.
[3] Brar B, Broekaert TPE, van der Wagt JPA, Seabaugh AC, Moise TS, Morris FJ, Beam III EA, Frazier GA. In: IEEE Cornell Conference on Adv Conc High Speed Semicond DevCirc, 1997. p. 28.
[4] Broekaert TPE, Brar B, van der Wagt JPA, Seabaugh AC, Moise TS, Morris FJ, Beam III EA, Frazier GA. IEEE J Solid State Circ 1998;33:1342.
[5] van der Wagt P, Seabaugh A, Beam III E. IEEE Electron Dev Lett 1998;19:7.
[6] Van der Wagt P. Proc IEEE 1999;87:596.
[7] Micheel LJ, Taddiken AH, Seabaugh AC. In: Int. Symp. Multiple-Valued Logic, 1993. p. 164.
[8] Shimizu N, Nagatsuma T, Waho T, Shinagawa M, Yaita M, Yamamoto M. Electronics Lett 1995;31:1695.
[9] Schulman JN, De Los HJ, antos S, Chow DH. Electron Dev Lett 1996;17:220.
[10] Lee K, Shur M, Fjeldly TA, Ytterdal T. Semiconductor device modeling for VLSI. Englewood Cliffs: Prentice Hall, 1993.
[11] Wolak E, Pitner PM, Hellman ES, Park BG, Weil T, Harris JS, Thomas D. Appl Phys Lett 1988:53.
[12] Moise TS, Kao Y-C, Kao Y-C, Katz AJ, Broekaert TPE, Celii FG. J Appl Phys 1995;78:6305.
[13] Yu RY, Konishi Y, Allen ST, Reddy M, Rodwell MJ. IEEE Microwave and Guided Wave Lett 1994;4:220.
[14] Verghese S, Parker CD, Brown ER. Appl Phys Lett 1998;72:2550.
[15] Broekaert T. In: Proc. IEEE Cornell Conf. on Adv. Concepts in High Speed Semicond. Dev. and Circ., 1997. p. 132.
[16] Kawahito S, Mizuno SK, Nakamura T. In: Int Symp Multiple-Valued Logic, 1991. p. 330.

PERGAMON

Solid-State Electronics 43 (1999) 1367–1371

SOLID-STATE ELECTRONICS

Analysis of Sb-based resonant interband tunnel diodes for circuit modeling

J.N. Schulman

HRL Laboratories, Malibu, CA 90265, USA

Abstract

Sb-based resonant interband tunneling diodes are an alternative to the more commonly used InP-based resonant tunneling diodes. Potential advantages include higher speed and lower power operation due to its lower characteristic voltages. This fundamental difference is caused by the Type II band alignment of the emitter and quantum well semiconductors as compared to the Type I band alignment in the InP-based system. A simple theoretical framework is provided that exhibits and provides qualitative understanding of the influence on device behavior. A SPICE compatible analytic form for the current versus voltage characteristic is derived. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

High speed circuits utilizing resonant tunneling diodes (RTDs) have recently demonstrated rapid progress towards digital and mixed signal applications [1]. The negative resistance characteristic, similar to an Esaki diode, has allowed the design of circuits with smaller device count, higher speed, smaller area, and lower power consumption than similar circuits without RTDs. The most successful semiconductor system to date for implementing these circuits has been the InP-based system, i.e. InP substrate with InGaAs and InAlAs lattice matched alloy compositions. In addition, very thin layers of AlAs and InAs are commonly used in the RTD to enhance its properties. RTDs in this system have excellent peak-to-valley current ratios, and, most importantly, are compatible with existing HBT and HEMT growth and fabrication technologies.

To fully realize the low power consumption potential of RTD circuits, it is important to have the characteristic voltages as low as possible. The most important of these voltages is usually the peak voltage, i.e. the voltage where the current is maximum. These voltages set the voltage scale for the RTD logic states, and thus ultimately the transistor and power supply voltages.

InP-based RTDs are limited in the degree to which the voltages can be reduced. Smaller band gap InAs is commonly inserted at the center of the otherwise InGaAs quantum well part of the RTD to lower the voltages, but this tends to lower the peak current density also. Widening the well also lowers the voltage, but degrades the peak-to-valley ratio. A heterostructure system which naturally has lower characteristic voltages is the resonant interband tunnel diode (RITD) based on InAs, AlSb, and GaSb. The Type II band discontinuities are quite different in this system as compared with the InP system. Fig. 1 shows the band diagrams of the conduction and valence band edges for the two cases. Fig. 1(a) is for the Type I InGaAs/AlAs structure in which conduction electrons incoming from the InGaAs emitter tunnel through the resonance, E_R, in the central quantum well. The valence band is not involved in this case. By contrast, Fig. 1(b) shows the Type II InAs/AlSb/GaSb case in which electrons incoming from the InAs conduction band

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00075-1

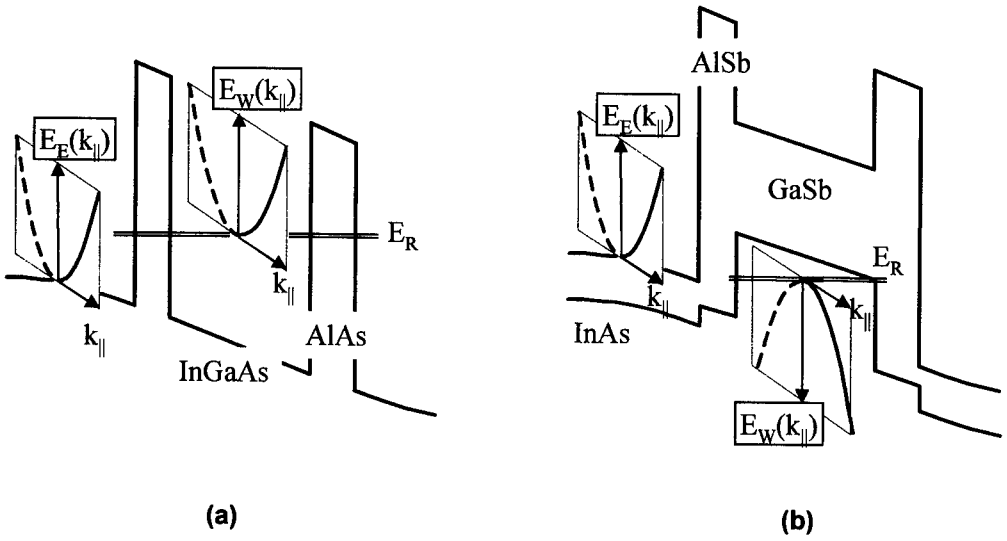


Fig. 1. Band diagrams of (a) Type I, InGaAs/AlAs, and (b) Type II, InAs/AlSb/GaSb double barrier resonant tunneling structures. $E_E(k_\parallel)$ and $E_W(k_\parallel)$ are the in-plane energy dispersion curves for the emitter and well electrons, respectively.

undergo interband tunneling through the light hole resonance in the GaSb valence band [2–4].

2. Simple analysis

The key feature which distinguishes the interband tunneling is the negative effective mass of the GaSb light hole band. Fig. 1 shows the parabolic $E(k_\parallel)$ dispersion of the emitter and quantum well subband states, where $k_\parallel$ is the wave vector in the in-plane directions. The relative lineup as a function of applied voltage of these emitter and well parabolas determine the number of electrons that tunnel through the resonance and thus the current [5]. A simple zero temperature model with a single resonance is useful for illustrating

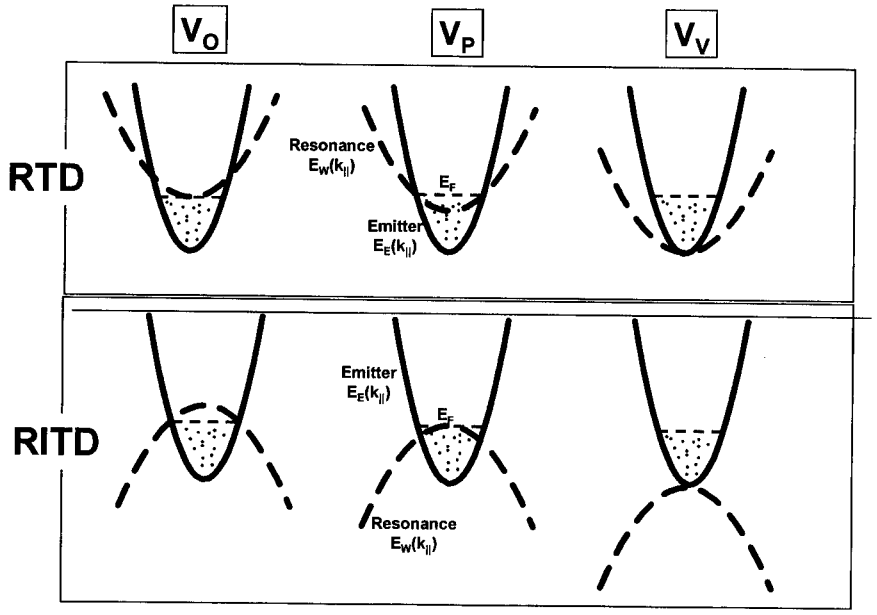


Fig. 2. Relative position of the emitter (solid) and well (dashed) $E(k_\parallel)$ dispersion curves for three voltages for Type I RTDs and Type II RITDs. E_F is the emitter Fermi level. V_O, V_P, and V_V, are the voltages at which the current turns on, peaks, and goes to zero, respectively, in a simple $T = 0$ model.

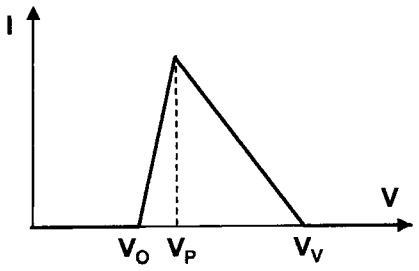


Fig. 3. Definition of V_O, V_P, and V_V as the onset, peak, and valley voltages in a simple $T = 0$ model of resonant tunneling.

the distinction between the two cases [6,10]. Fig. 2 shows the relative line-up of the emitter and well parabolas for the three voltages defined by Fig. 3. V_O (onset) is the voltage such that the current just begins to flow. V_P (peak) is the voltage at the current maximum, and V_V (valley) denotes the voltage where the current shuts off. A simple analysis based on Fig. 2 provides equations for V_O, V_P, and V_V. For example, V_O for the RTD case is the voltage required to lift the emitter conduction band edge just enough so that the Fermi level coincides with the bottom of the resonance subband. The emitter electrons that tunnel have $k_{\|}=0$ but momentum in the growth direction (not shown) such that their energy is at the Fermi level. By contrast, in the negative mass RITD case a smaller voltage is required, as the emitter Fermi level reaches the resonance subband at a nonzero $k_{\|}$ value. The effective masses determine the curvature of the two parabolas and thus the intersection points. In particular the ratio $\alpha = m_W/(m_W - m_E)$, where m_E is the emitter effective mass and m_W is the mass of the resonance in the well, is the combination that enters the formulas. α_I is the ratio for the RTD, with $m_W > m_E$ being the usual case [5], while α_{II} is for the RITD, with $m_W < 0$. (The rarer $0 < m_W < m_E$ case is also discussed in Ref. [5]. It is slightly more complicated and must be dealt with separately.) Ignoring band bending and assuming half the voltage drop occurs at the center of the well, the formulae for the RTD are as follows:

$$V_O = 2(E_R - E_F)$$

$$V_P = 2(E_R - E_F/\alpha_I)$$

$$V_V = 2E_R \tag{1}$$

E_R is the energy of the resonance above the emitter conduction band edge at zero bias. This becomes the following in the Type II case (assuming $E_R > E_F$):

$$V_O = 2(E_R - E_F/\alpha_{II})$$

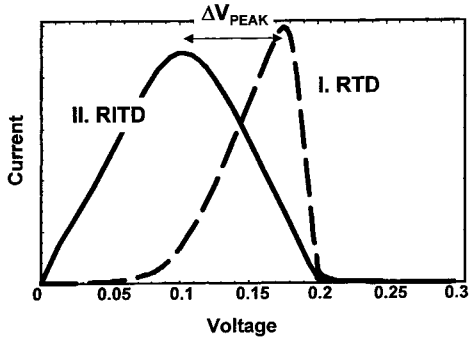


Fig. 4. Current versus voltage for Type I RTD (dashed) and Type II RTD (solid) for $m_E = 0.04$, $|m_W| = 0.05$, $T = 77$ K, $E_F = 0.05$, $E_R = 0.1$ in Tsu–Esaki model extended for unequal masses.

$$V_P = 2(E_R - E_F)$$

$$V_V = 2E_R \tag{2}$$

It is interesting to note that what is usually called the resonance condition, where the emitter conduction band edge lines up with the resonance at $k_{\|}=0$, $V = 2E_R$, corresponds to the *valley* voltage, not the peak voltage as is commonly and incorrectly assumed. The formulae for V_O and V_P are reversed in order between the two cases because $\alpha_I > 1$, while $\alpha_{II} < 1$ due to the change in sign of m_W. Assuming only the sign of m_W differs, the Type II case has a peak voltage which is $2E_F\, m_E|/m_W|$ less than for the Type I case. The actual measured voltage includes the band bending outside the double barriers which amplifies this difference.

The simplest model for resonant tunneling including temperature was developed by Tsu and Esaki [7]. It assumes that the emitter and well masses are equal. Recently, this was extended to include unequal masses thus incorporating the Type II case [5]. The resulting formula for the $I(V)$ (due to elastic tunneling only) is the following:

$$J = \frac{m_W kTe}{2\pi^2 H^3}\int_0^\infty dU t(U)\log\left(\frac{1+e^{(E_F-U)/kT}}{1+e^{(E_F-U-eV)/kT}}\right.$$

$$\left.\frac{1+e^{(E_F-\alpha U-eV)/kT}}{1+e^{(E_F-\alpha U)/kT}}\right) \tag{3}$$

Substituting a Lorentzian lineshape for the transmission coefficient, $t(U)$, allows a simple comparison of the Type I and Type II cases. Fig. 4 shows an example with $m_E = 0.04$, $|m_W| = 0.05$, $T = 77$ K, $E_F = 0.05$, and $E_R = 0.1$, with energy in eV and mass in units of the free electron mass. ΔV_{PEAK} indicates the

decrease in the peak voltage for the RITD as compared with the RTD.

3. SPICE formulation

Eq. (3) can be simplified by assuming that the transmission resonance width is much less than the thermal energy, kT, and the Fermi energy, E_F. The integral of the Lorentzian transmission produces a step-like arctangent function which indicates whether the emitter conduction band edge is above or below the resonance energy.[1] The width of the step is on the order of the width of the resonance. An earlier version of a SPICE compatible formula for the RTD $I(V)$ characteristic assumed that the emitter and well masses are equal [8]. This is not even approximately true for the RITD due to the change in sign of the masses. It is also inadequate for the RTD, because the well quantization effect usually increases the mass in the well significantly [5]. In this earlier formula the width of the resonance directly contributed empirically to determining the voltage width of the negative resistance region. However, resonance widths are typically on the order of a millivolt or less, much less than the observed negative resistance region widths. It is now understood from the above that the effective mass difference and the emitter Fermi level, not the resonance width, is more appropriate in determining the negative resistance region width, as seen, for example, in Eqs. (1) and (2) for $V_V - V_P$. Thus, the arctangent can be replaced by a truly abrupt step function and the following analytic form for the RITD $I(V)$ results (including a diode-like exponential term for the $I(V)$ beyond the resonance):

$$I(V) = -\pi a \log\left(\frac{1+e^{(b-c+n_1V)/kT}}{1+e^{(b-c-n_1V)/kT}}\right.$$

$$\left.\frac{1+e^{(b-(c-n_1V)\alpha-2n_1V)/kT}}{1+e^{(b-(c-n_1V)\alpha)/kT}}\right) + h(e^{n_2V/kT}-1) \qquad (4)$$

Here, the physical quantities E_F and E_R have been replaced by the empirical parameters b and c, respectively. n_1 helps to represent band bending effects. It represents the fraction of the applied voltage at the center of the well, 1/2 in the simplest approximation. This replacement of the physical quantities by empirical ones acknowledges that this extended Tsu–Esaki framework is too simple to model actual resonant tunneling devices, but that it accurately captures the essential physics and thus the form of the $I(V)$ relationship. The explicit negative sign preceding the first term comes from the negative value of m_W whose magnitude is otherwise included in the prefactor a (positive by convention) [5]. The log is also negative and thus the overall contribution is positive. The sign should be changed to be explicitly positive for the $m_W > m_E$ Type I RTD case, and the log is also then positive. Again, the $0 < m_W < m_E$ case requires modifications which will not be discussed here.

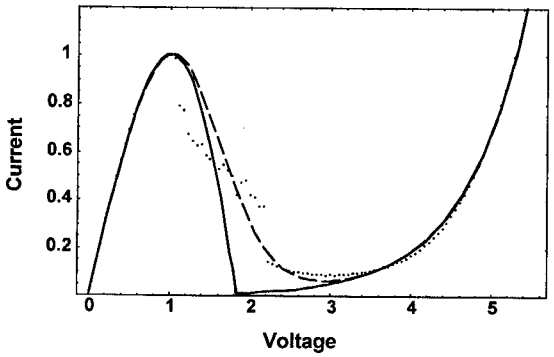


Fig. 5. Empirical analytic RITD $I(V)$ curve without (solid) and with (dashed) resonance broadening term. Data are the dots. The curve is normalized to $V=I$ at the $I(V)$ peak.

The solid line in Fig. 5 plots an example of Eq. (4) for a typical Type II RITD case. The $I(V)$ is normalized to peak voltage and current of 1.0. A shortcoming can be seen in the valley current region. The log term in Eq. (4) goes strictly to zero when the voltage approaches the resonance, $V = V_R = c/n_1$ in this notation. This is correct but incomplete, in that various mechanisms such as interface roughness and scattering broaden the resonance voltage. The following function can be substituted for V to include the broadening empirically:

$$V_s = V_R - s \log(1 + e^{-(V-V_R)/s}) \qquad (5)$$

V_s has the simple property that it is very close to V for $V_s < V_R$, and asymtotically approaches V_R for

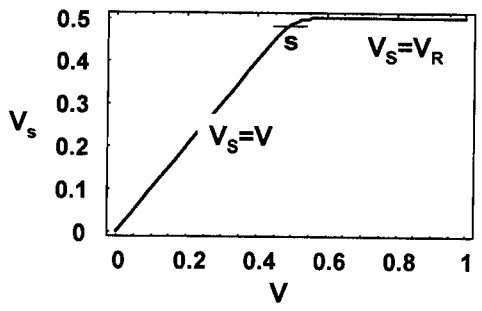


Fig. 6. Broadening function $V_s = V_R - s \log(1 + e^{-(V-V_R)/s})$.

[1] Eq. (6) in Ref. [5]. There is a typographical error: '−1/2' should be '+1/2' before the arctangent term.

$V > V_R$, as shown in Fig. 6. The transition region is on the order of s. Any similarly behaved function can also be used. The effect of this function is to broaden the resonance and avoid the condition of zero current. The dashed curve in Fig. 5 illustrates the resulting more realistic behavior, with $s = 0.308$ for this case. This approach is also effective for RTDs with $m_W > m_E$, including the case where m_W approaches m_E (large positive α).

An application of this formula to measured RITD $I(V)$ data is shown in Fig. 5. The parameters for Fig. 5 were fit using a least square routine, with the peak and valley voltages and currents emphasized, as well as the slope near $V = 0$. The $T = 0$ limit of Eq. (4) (ignoring the diode term) provides useful initial guesses for the parameter values. Recognizing that the $T = 0$ limit is derived by replacing the $\log(1 + e^x)$ terms with $x\theta(x)$, (θ is the step function) in Eq. (4) produces the triangular form seen in Fig. 3. The following simple relationships in this limit are then easily derived from the RITD case:

$$V_O = (c\alpha - b)/\alpha n_1$$

$$V_P = (c - b)/n_1$$

$$V_V = c/n_1$$

$$I_P = ab\pi(\alpha - 1)/kT \tag{6}$$

I_P is the peak current. Initial guesses for the diode term parameters, h and n_2, can be obtained by fitting to the $I(V)$ beyond the valley voltage separately. Further refinement of all the parameters is then accomplished using the least squares software.

4. Conclusions

A simple framework is provided that demonstrates the lower characteristic voltages of Type II RITDs as compared with Type I RTDs. An extension of the Tsu–Esaki model for differing emitter and quantum well masses is used to derive an analytic formula for a description of both the RTD and RITD $I(V)$ curves. This formula (Eqs. (4) and (5)) improves on a previous version [8] in that the negative resistance region is empirically related to the effective mass ratio and the Fermi level instead of to the width of the resonance. The previous version assumed equal effective masses in the emitter and well which is not realistic, especially for the Type II case. Even though measured $I(V)$ data can be adequately fit using the previous version, this new more physically correct form allows for an empirical fit in which the parameters, derived from resonance energies, Fermi levels, etc., have more realistic physical values and a clearer relationship to the form of the measured $I(V)$. This provides a natural initial guess for the parameter values which leads to quicker convergence and less variability in the results of the least square fit for the refined values. At this stage the resulting parameter values must still be regarded as purely empirical. However, it is hoped that this improved realism will eventually provide a closer connection between the physical values and the empirical parameter values. This might be achieved, for example by using a full quantum transport calculation [9] to develop semi-empirical rules for relating the physical and empirical values.

References

[1] Seabaugh A, Brar B, Broekaert T, Morris F, Frazier G Solid State Electronics, in press.
[2] Sweeny M, Xu Jingming. Appl Phys Lett 1989;54:546.
[3] Söderström JR, Chow DH, McGill TC. Appl Phys Lett 1989;55:1094.
[4] Luo LF, Beresford R, Wang WI. Appl Phys Lett 1989;55:551.
[5] Schulman JN. Appl Phys Lett 1998;72:2829.
[6] Ohno H, Mendez EE, Wang WI. Appl Phys Lett 1990;56:1793.
[7] Tsu R, Esaki L. Appl Phys Lett 1973;22:562.
[8] Schulman JN, De Los Santos HJ, Chow DH. IEEE Elect Dev Lett 1996;17:220.
[9] Klimeck G, Blanks D, Lake R, Bowen RC, Fernando CL, Leng M, Frensley WR, Jovanovic D, Sotirelis P VLSI Design. In: Fifth International Workshop on Computational Electronics, Notre Dame, May 1997. Gordon and Breach, 1998. 8. p. 79.
[10] Mendez EE, Nocera J, Wang WI. Phys Rev B 1992;45:3910.

PERGAMON

Solid-State Electronics 43 (1999) 1373–1381

SOLID-STATE ELECTRONICS

SiGe-HBTs for mobile communication

Andreas Schüppen* [1]

TEMIC Semiconductor GmbH, Theresienstr. 2, D-74072, Heilbronn, Germany

Abstract

SiGe-HBTs are the most favored candidates for mobile communication systems. They demonstrated excellent high frequency and noise performances on research level, but now the production era of SiGe technologies is starting. First production-ready SiGe-DECT front-end will be presented including an LNA with a typical noise figure of 1.6 dB and a gain of 20 dB, and a power amplifier with 27.5 dBm output power and up to 47% on packaged devices at 1.9 GHz. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Silicon Germanium heterobipolar transistors (SiGe-HBTs) offer the opportunity to fabricate high performance ICs for mobile communication systems in the 1–10 GHz range, e.g. GSM, DECT, DCS, PCS1800, WLANs and SAT TV as depicted in Fig. 1. The greatest progress with SiGe for rf circuits are made in analog and mixed signal, e.g. in low noise amplifiers (LNAs), power amplifiers (PA), mixers, frequency dividers, voltage controlled oscillators (VCOs) and phase locked loops (PLLs). In addition, the fabrication of high frequency analog to digital and digital to analog converters (ADCs and DACs) are also reasonable circuits, concerning also direct conversion.

The physical advantages of SiGe over pure silicon for these applications are mainly the extremely high cutoff frequencies with record values from research HBTs of $f_T = 130$ GHz [1] and $f_{max} = 160$ GHz [2], leading to ECL gate-delays of ring oscillators down to 9.3 ps [3]. The good high frequency noise performance with rf-noise figures of 0.9 dB at 10 GHz [4] emphasizes the potential of SiGe-HBTs for mobile communication systems. In addition, the good low frequency noise behaviour of SiGe-HBTs with corner frequencies below 1 kHz, due to the high quality surface passivation on the pn-junctions by thermal SiO_2; which does not exist on III/V semiconductors (Fig. 2). This is very important for large scale integration and for high quality oscillators.

Furthermore, an important argument for using HBTs in wireless applications is the high power added efficiency (PAE) at low DC-voltages, which allows a low voltage battery power supply and a long up-time for the hand held systems. From the economic point of view it should be pointed out, that a standard Si production line and the most of the standard bipolar process modules can be used for SiGe device fabrication. This results in low cost production with high yield and excellent reliability.

2. Two HBT concepts

In principle, there exist two types of SiGe-Bipolar Transistors: type 1 the box-shaped real SiGe-HBT, favored by, for example, Daimler–Benz [5], TEMIC, NEC and Philips and type 2 the triangular shaped SiGe-Drift transistor, favored by, for example, IBM [6] and Siemens. In the literature both types are called

* Tel.: +49 7131 672796; fax: +49 7131 672933.
[1] andreas.schueppen@temic-semi.de

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00076-3

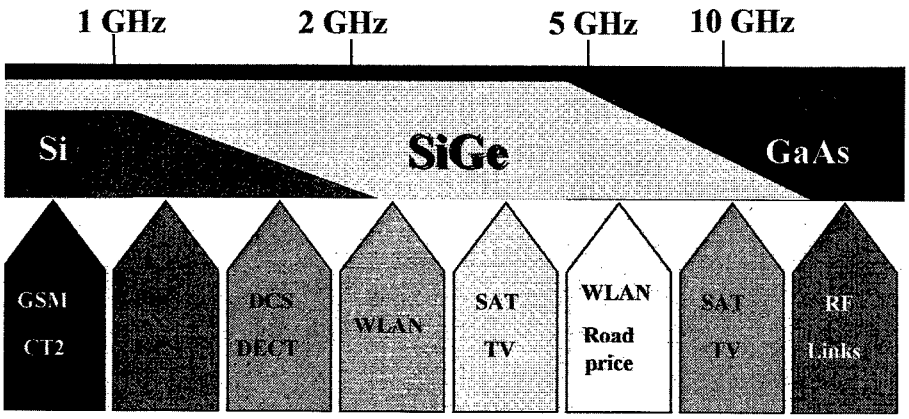


Fig. 1. Application domains of Silicon Germanium.

SiGe-HBT. However, the box-shaped type has a lot of advantages over the transistor with triangular Germanium profile. Due to the inversion of the doping between base and emitter, i.e. the base doping is higher than the emitter doping in contrast to a standard homo junction bipolar transistor and also in contrast to the drift field HBT, the base sheet resistance of the real HBT is up to 10 times smaller than in a rf-BJT, as shown in Fig. 3. For that reason the total base resistance is lower in real SiGe-HBTs leading to higher f_{max}, higher power gain and a lower high frequency noise figure comparing transistors with the same emitter geometry. This yields a more independent f_T and f_{max} behavior for different emitter widths (Fig. 4). In addition, the position of the emitter-base junction is defined by the epitaxy and not by a critical high temperature outdiffusion process from a poly emitter. In the box-shaped version the emitter-base diode can be easily used as varactor diode for oscillators owing to the low tunneling leakage currents and the higher breakdown voltage.

3. The SiGel-technology

TEMIC's production technology originates from the Daimler–Benz differential SiGe research technology with box-shaped SiGe profiles. TEMIC's SiGel is in

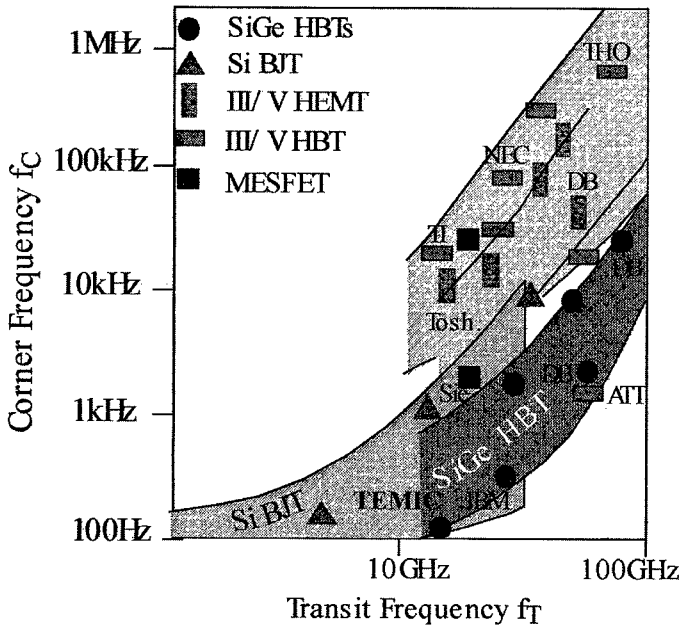


Fig. 2. Low frequency noise: corner frequencies vs transit frequencies for Si, SiGe and III/V devices.

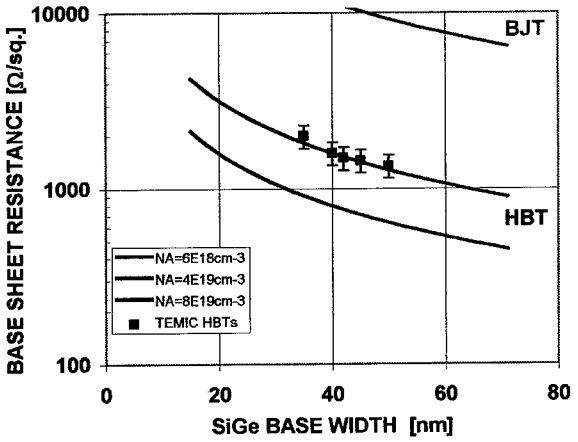


Fig. 3. Low base sheet of SiGe-HBTs.

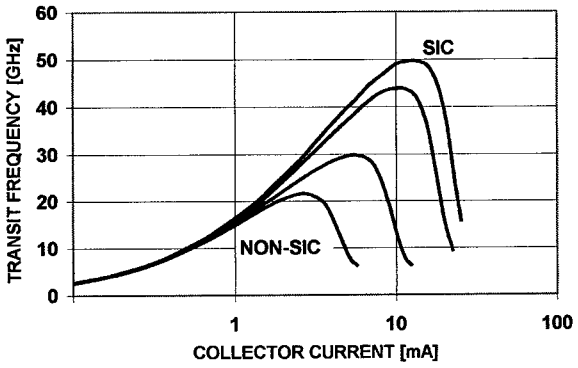


Fig. 5. SiGe-HBT with 0.8×20 μm^2 emitter size with and without selectivity implanted collector (SIC). $v_{CB}=0$ and 1.8 V.

principle a double poly technology including as special feature *npn* HBTs with and without selectivity implanted collectors (SIC) on the same wafer [7]. The SiGe-HBTs reveal transit frequencies of 30 GHz with $BV_{CEO}=6$ V and 50 GHz with $BV_{CEO}=3$ V, respectively (Figs. 4 and 5). The latter HBTs includes a SIC with a doping concentration of 1×10^{17} cm^{-3}.

In addition, spiral inductors, nitride capacitors, three types of poly resistors, a LPNP, rf- and dc-ESD protection and varactor diodes are incorporated in the present technology. In order to grow the monocrystalline SiGe epitaxial layers, TEMIC uses a single wafer reactor from Applied Materials, which was well proofed by homo silicon epitaxy. The homogenity of the epitaxial films over the wafer and the lots is in both cases, thickness and Ge content, better than ±5% as shown in Fig. 6, measured by X-ray diffraction. The technology is planar and comparable in terms of masks and costs to a standard double poly Si BJT process. SiGel is well suited for large scale integration (LSI), as nicely demonstrated in Fig. 7, which shows a wafer mapping of 10 k transistor arrays over a typical 6 in. wafer. To demonstrate the performance of multi finger power devices, HBTs with different emitter numbers were investigated. The effective emitter size of each finger is 1.6×30 μm^2. On each side of the emitter is a base contact in order to achieve a low base resistance and a collector finger contact is placed between a group of base and emitter contacts. To achieve homogenious current distribution for all emitters, a 4 Ω ballast resistance is inserted in series to each emitter. The large signal analysis was carried out by an on-wafer

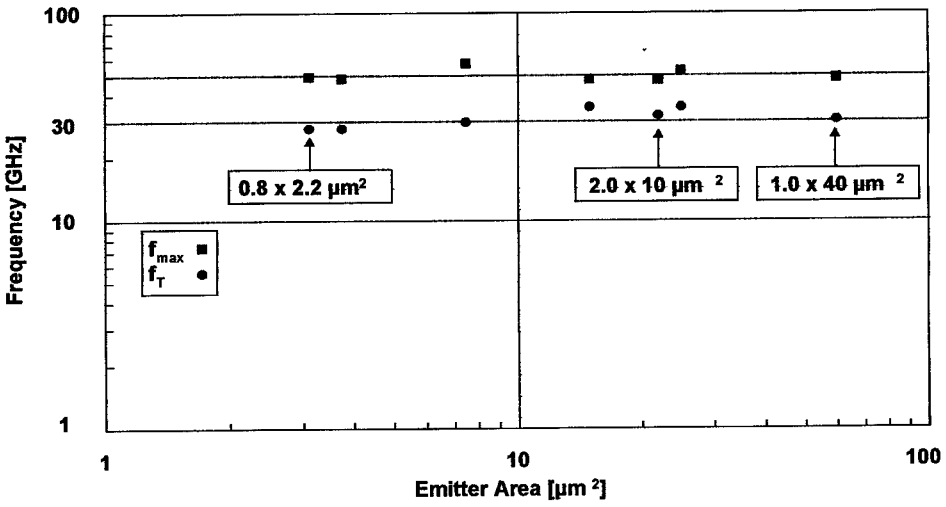


Fig. 4. SiGe-HBTs with highly doped base are nearly independent on the emitter width for both f_T and f_{max}.

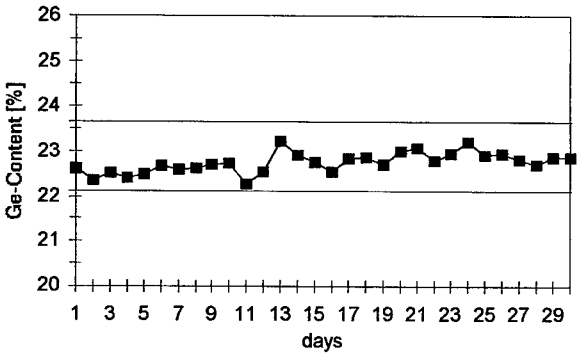


Fig. 6. Statistical process control (SPC) of the SiGe LPCVD epitaxy by monitoring the base germanium content over 3 months (each third day one wafer).

loadpull measurement setup from ATN. P_{out}, power added efficiency (PAE) and gain of a 10 emitter device with output matching for the fundamental frequency 1.9 GHz and 2nd harmonic is shown in Fig. 8. The 2nd harmonic tuning leads to a 8% higher efficiency achieving a value of 62%. The results of the 10, 20, 40 and 80 emitter devices show that the power scales with the number of emitters, i.e. 21, 24, 27 and 30 dBm. The parameters of the SiGel-technology are summarized in Table 1.

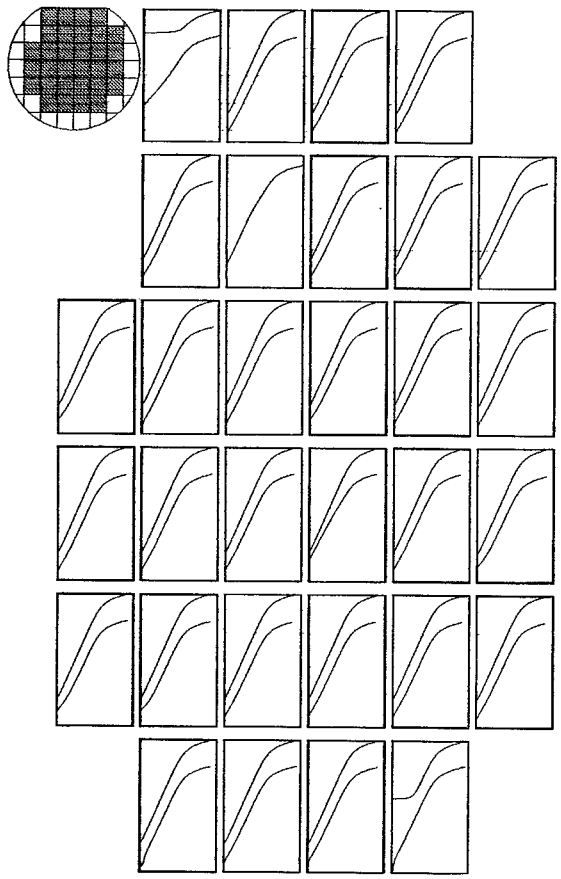


Fig. 7. Gummel plot wafer mapping of 10 k transistor arrays with 0.8 × 1.6 μm^2 SiGe-HBTs: y-axis: 1 pA–100 mA, x-axis: 0.2–1 V.

4. Research circuits with SiGe-HBTs

On researh level a couple of circuits were investigated in the last few years since the technology breakthrough of SiGe-HBTs in 1989 [8] and 1990 [6], respectively, e.g. digital to analog converters (DACs) [9,10], optical transmitter circuits, transimpedance amplifiers for optical fibre links [11–14] and demultiplexer [15,16]. An example of a wideband amplifier fabricated by Daimler–Benz's research technology is shown in Fig. 9. This two stage amplifier reached a 3 dB bandwidth of 18 GHz with a power gain of 10 dB [17].

For wireless high frequency systems VCOs at 11, 26 and 40 GHz were investigated [18,19]. A hybrid 10 GHz DRO using TEMICs SiGe-HBTs reached a very low phase noise of −115 dBc/Hz at 10 kHz off-carrier, as depicted in Fig. 10, owing from the low trap density of the bulk Si/SiGe/Si and the good surface passivation by thermal oxide [20]. Additionally, Gilbert cell mixers [21], a return-to-zero comparator [22] and a couple of dividers [23,24] were fabricated with SiGe technologies. Until now, the fastest circuit in SiGe is a 42 Gb/s static frequency divider, which was fabricated by Siemens' technology with a design from Rein et al. [24]. A more exotic SiGe application is the SPDT T/R antenna switch in [25]. Very important rf-circuits are LNAs and PAs, i.e. the rf front-ends, which were only fabricated by TEMIC until now. Erben et al. (private communication) have designed three types of LNA-ICs for 1–2 GHz, for 5.8 GHz and for 10 GHz using the SiGel technology. The 2 GHz LNA achieved an associated gain of 12–18 dB and a noise figure of 0.9 dB for the whole circuit. The 5.8 GHz circuit revealed a record noise figure of 1.6 dB at 5.8 GHz and a gain as high as 26 dB by the two stage approach. The 10 GHz LNA has to be optimized in the future (Fig. 11).

5. Commercial SiGe-ICs

Nevertheless, at the moment the biggest market share for SiGe is to be seen in wireless communication systems in the 1–6 GHz range. Hence, mixers, GSM power modules, dual band frontends for GSM and PCS1800 and DECT front-ends are in the focus. DECT front-ends including LNAs with 1.5 dB noise figure and 20 dB gain combined with a 28 dBm power

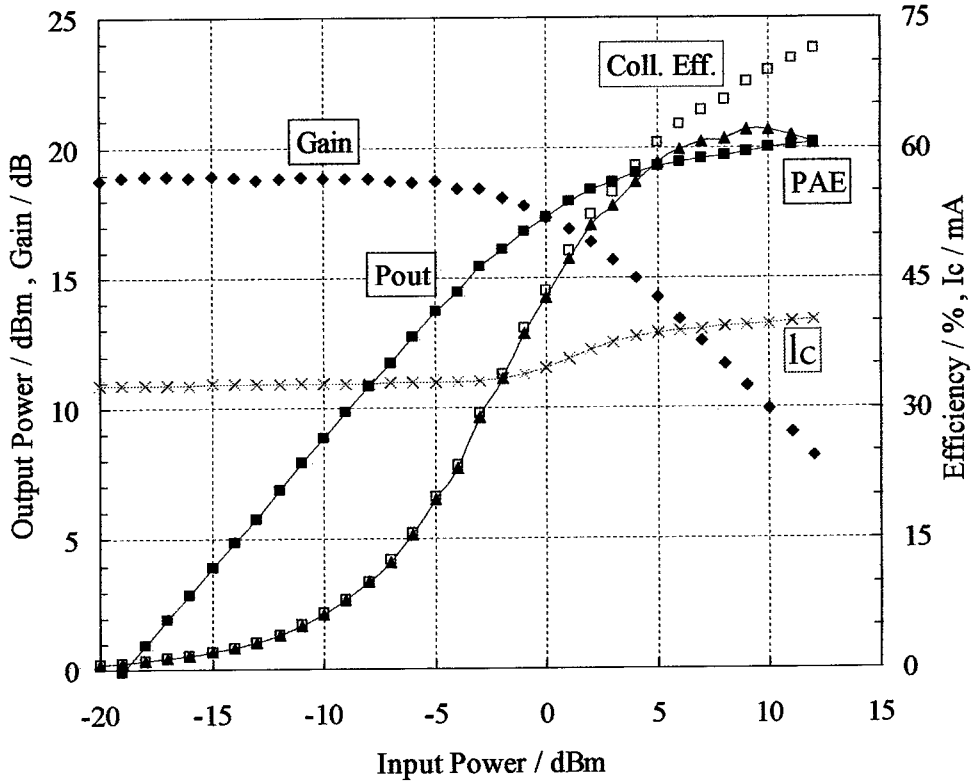


Fig. 8. Load pull measurement at 1.9 GHz of a SiGe power HBT with $10 \times 1.6 \times 30$ μm^2 emitter area at $v_{CE}=3.6$ V.

Table 1
Summary of the essential parameters of the SiGel technology

Parameter (unit)	non-SIC	SIC
NPN transistor		
Transit frequency, f_T (GHz)	30	50
Max. frequency of oscillation, f_{max} (GHz)	50	50
Current gain, h_{FE} (I)	180	180
Early voltage, V_A (V)	>40	>40
Collector emitter breakdown voltage, BV_{CEO} (V)	6.0	3.0
Collector base breakdown voltage, BV_{CBO} (V)	15	12
Noise figure at 2 GHz, F_{min} (dB)	1.0	1.0
LPNP		
Collector emitter breakdown voltage, BV_{CEO} (V)	7	
Current gain, h_{FE} (I)	5	
Typical collector current, I_{CP} (μA)	40	
Passive devices		
High ohmic poly resistor (poly1), R_H (Ω/sq)	400	
Medium ohmic poly resistor (poly2), R_M (Ω/sq)	110	
Low ohmic poly resistor (poly1-$TiSi_2$), R_L (Ω/sq)	4.5	
Precision MIM capacitor, C_{SPEC} (fF/μm^2)	1.1	
Spiral inductor 4 nH, Q-value at 2 GHz, O (I)	7	
ESD Zener diode, parasitic capacitance, (pF)	5.5	
RF-ESD diode, parasitic capacitance, (pF)	0.3	

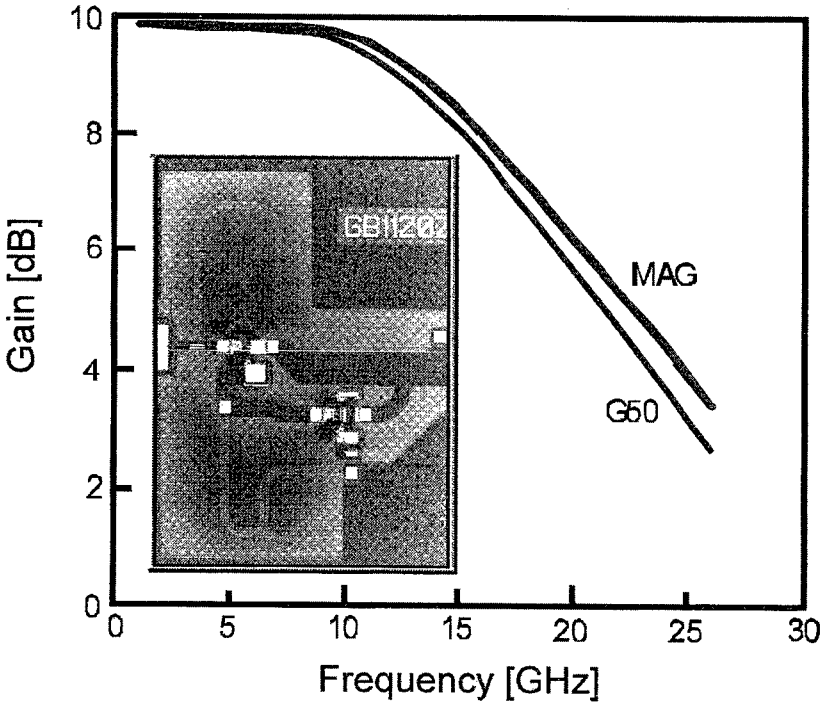


Fig. 9. Wideband amplifier with 18 GHz bandwidth and 10 dB gain (cowork from Daimler–Benz Research Center and University of Ulm).

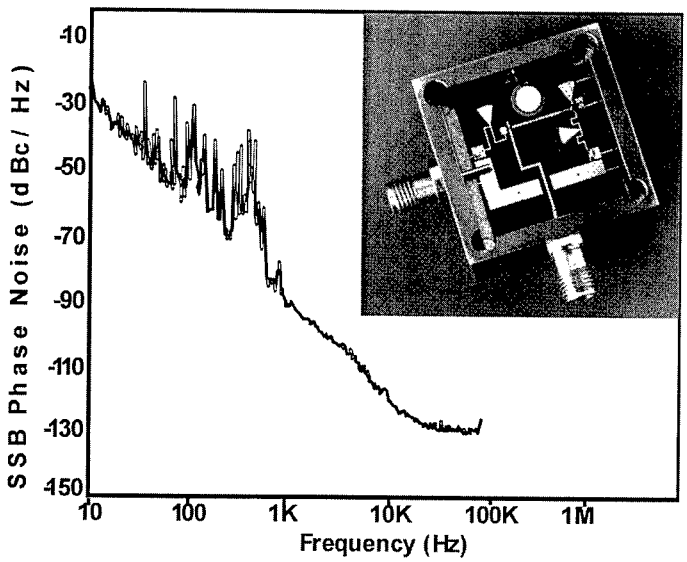


Fig. 10. 10 GHz Dielectric Resonat Oscillator (DRO) revealing—115 dBc phase noise at 10 kHz off-carrier [20].

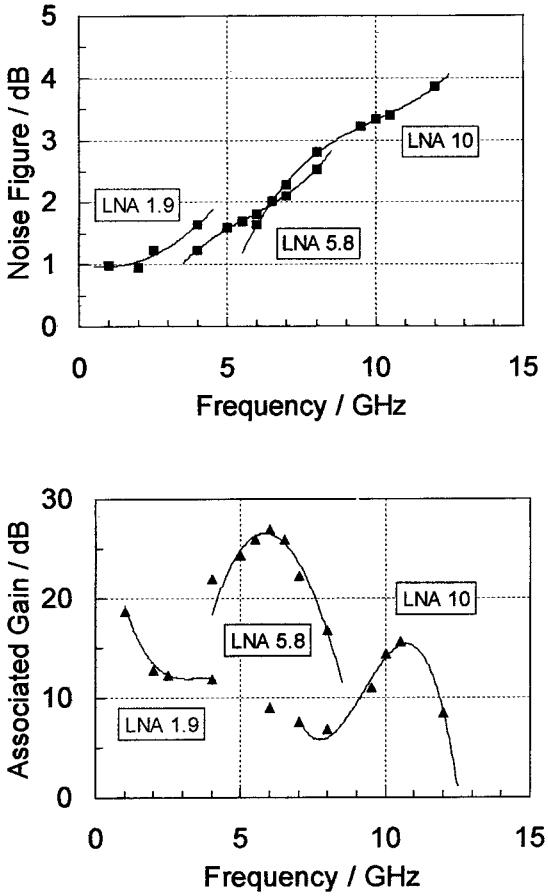


Fig. 11. rf-Noise figures and associated gain of high performance LNAs at 1.8, 5.8 and 10 GHz.

amplifier with 47% PAE, measured over the whole packaged device, are production ready (Figs. 12–15). The PA works between 2.5 and 4.0 V with nearly constant output power. The demoboard for the complete

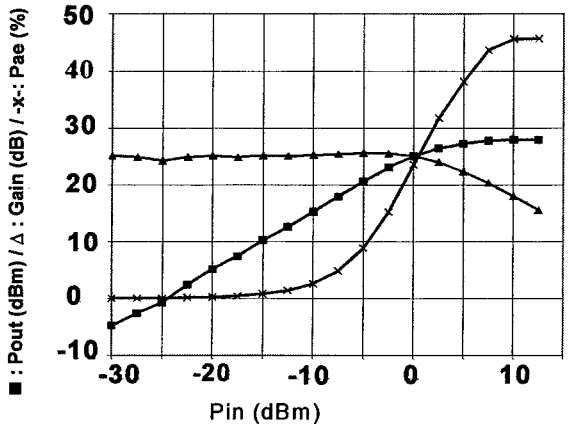


Fig. 13. Three complete stages DECT-PA for 2.7–3.6 V operation in a SSO20 package (loadpull measurement), reaching a gain of 25 dB, an output power of 28 dBm and a PAE of 47% at 1.9 GHz.

DECT chip-set, including two Si chips and one SiGe front-end IC are shown in Fig. 15. In addition, there exist some functionable higher integrated customer ICs using SiGel.

6. Conclusions

At the moment two HBT concepts exists. Both show good circuit results. However, it seems that the graded SiGe profile already reached its optimum performance, whereas the box-shaped HBT can be improved by using carbon in the base to suppress the boron diffusion. The first commercial SiGe ICs are ready and the first customers trust them to buy. The frequency limit for the commercial applications of SiGe is not f_T or f_{max}, the challenge is the handling of the package para-

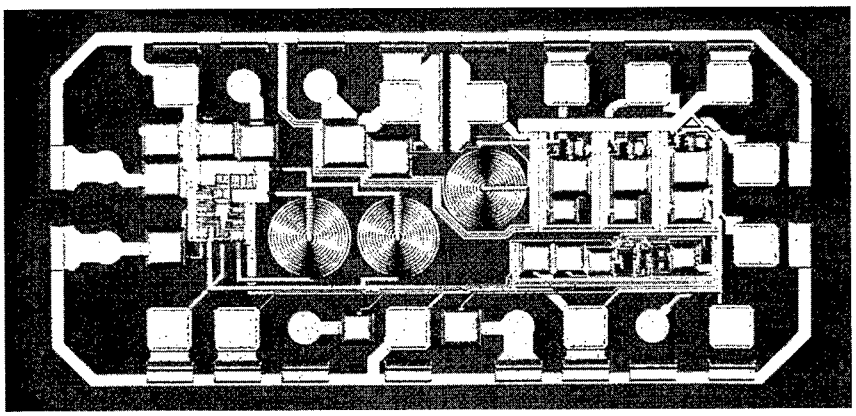


Fig. 12. Chip-Foto of the SiGel DECT front-end, including LNA, PA and T/R switch control.

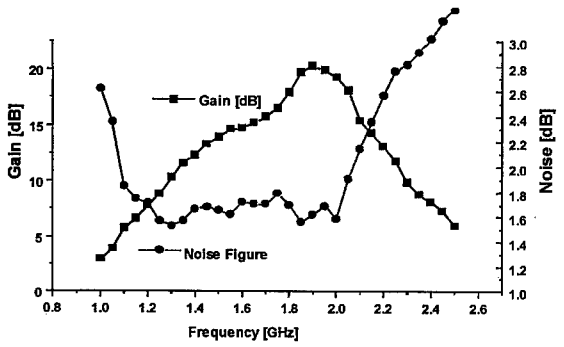


Fig. 14. LNA of the SiGel DECT front-end with 1.6 dB noise figure and 20 dB associated gain.

sitics, hence the next SiGe-ICs above 2.4 GHz will be mounted by flip-chip or chip scale packaging.

However, now, there is a frequency and time slot for SiGe, but CMOS does not wait.

Acknowledgements

The author would like to thank J. Arndt, H. Dietrich, U. Seiler, H.v.d. Ropp, M. Lentmaier, N. Gellrich, K. Burger and all the other colleagues for the support by transferring the SiGel process into production. In addition, I would like to acknowledge the circuit design from U. Erben and H. Schumacher and the fruitful discussions and friendly words of U. König and A. Gruhle. This work was supported by the BMBF within the SIGMA project and by the European Community within the BETA project.

References

[1] Oda K, Ohue E, Tanabe M, Shimamoto H, Onai T, Washio K, 130 GHz f_T SiGe HBT technology, IEDM Tech. Dig 1997;791–794.

[2] Schüppen A, Erben U, Gruhle A, König U, Enhanced SiGe heterojunction bipolar transistors, IEDM Tech. Dig 1995;743–746.

[3] Washio K, Ohne E, Oda K, Tanabe M, Shimamoto H, Onai T. A selective-epitaxial SiGe HBT with SMI electrodes, featuring 9.3 ps ECL gate-delay. In: Tech. Dig. Int. Electron Device Meeting, 1997. p. 795–8.

[4] Shumacher H, Erben U, Gruhle A. Low-noise performance of SiGe heterojunction bipolar transistors. In: Tech. Dig. IEEE MTT-S Int. Microwave Symp, 1994. p. 1167–70.

[5] Kasper E, Gruhle A, Kibbel H. High speed SiGe-HBT with very low base sheet resistivity. In: Tech. Dig. Int. Electron Device Meeting, 1993. p. 79–81.

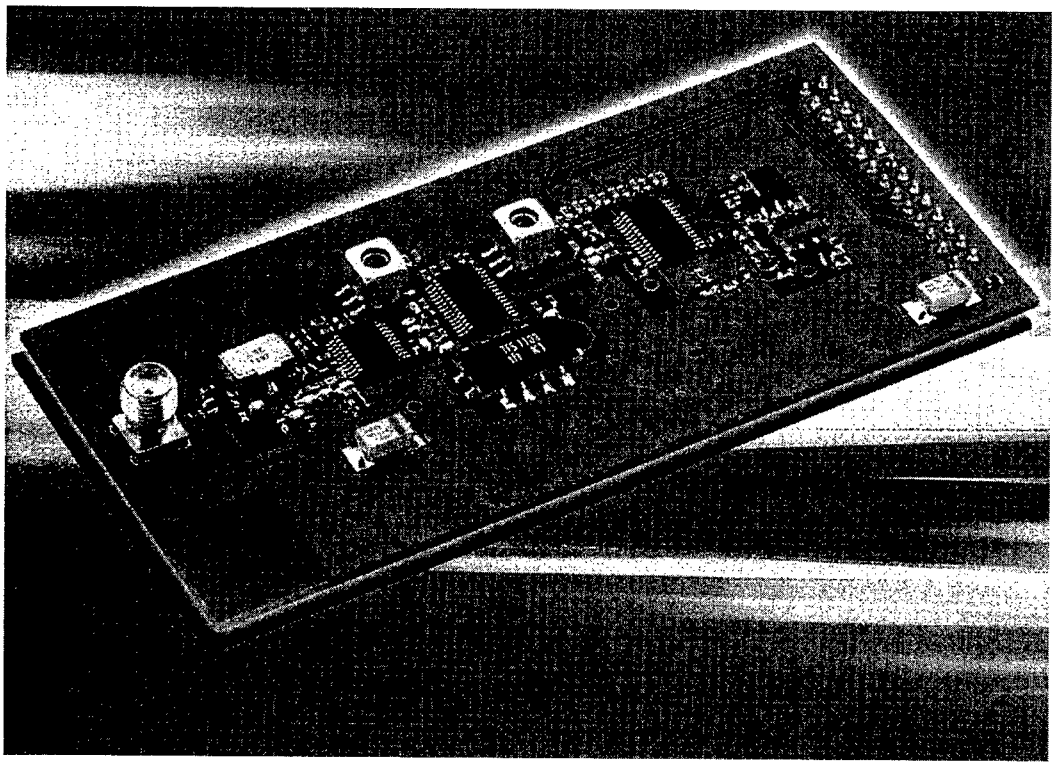


Fig. 15. Demo board of the DECT chip set.

[6] Patton GL, Comfort JH, Meyerson BS, Crabbe EF, Scilla GJ, de Fresart E, Stork JMC, Sun JY-C, Harame DL, Burghartz J. 75 GHz fT SiGe base heterojunction bipolar transistors. IEEE Electron Device Lett 1990;1(1):171–3.

[7] Schüppen A, Dietrich H, Gerlach S, Höhnemann H, Arndt J, Seiler U. SiGe technology and components for mobile communications systems. In: IEEE Bipolar/BiCMOS Circuits Technol. Meeting, 1996. p. 130–3.

[8] Iyer S, Platton GL, Stork JMC, Meyerson BS, Harame DL. Heterojunction bipolar transistors using Si-Ge alloys. IEEE Trans on Elect Dev 1989;36(10):2043–64.

[9] Harame DL, Stork JMC, Meyerson BS, Hsu KY-J, Cotte J, Jenkins KA, Cressler JD, Restle P, Crabbe EF, Subbanna S, Tice TE, Scharf BW, Yasaitis JA. Optimization of SiGe HBT technology for high speed analog and mixed-signal applications. In: Tech. Dig. Int. Electron Device Meeting, 1993. p. 71–4.

[10] Xiao P, Jenkins K, Soyeur M, Ainspan H, Burghartz J, Shin H, Dolan M, Harame D. A 4 b 8 GSample/s A/D converter in SiGe bipolar technology. In: Tech. Dig. IEEE Int. Solid-State Circuits Conf, 1997. p. 124–5.

[11] Neuhäuser M, Rein H-M, Wenz H. Design and realization of low-noise, high gain Si.bipolar transimpedance preamplifiers for 10 Gb/s optical-fibre links. In: Bipolar/BiCMOS Circuits and Technology Meeting, 1994. p. 163–6.

[12] Hashimoto T, Tezuka H, Sato F et al., SiGe bipolar ICs for 20 Gb/s optical transmitter, BCTM 1994;163–170.

[13] Masuda T, Ohhata K-l, Ohne E, Oda K, Tanabe M, Shimamoto H, Onai T, Washio K. 40 Gb/s analog IC chipset for optical receiver using SiGe HBT's. In: IEEE Int. Solid-State Circuits Conf, 1998. p. 314–5.

[14] Soda M, Morikawa T, Shioiri S, Tezuka H, Sato F, Tatsumi T, Emura K, Tashiro T. A 1 Gb/s 8-channel array OEIC with SiGe photodectors. In: Tech. Dig. IEEE Int. Solid-State Circuits Conf, 1997. p. 120–1.

[15] Sato F, Tezuka H, Soda M, Hashimoto T, Suzaki T, Tatsumi T, Morikawa T, Tashiro T, A 2.4 Gb/s receiver and a 1:16 demultiplexer in one chip using a super self-aligned selectively grown SiGe base (SSSB) bipolar transistor, IEEE J. Solid-State Circuits 1996;1451–1457.

[16] Shiori S, Soda M, Morikawa R, Hashimoto T, Sato F, Emura K. A 10 Gb/s SiGe bipolar framer/demultiplexer for SDH system. In: IEEE Int. Solid-State Circuits Conf, 1998.

[17] Schumacher H, Erben U, Gruhle A, Kibbel H, König U. A 3 V supply voltage, DC-18 GHz SiGe HBT wide-band amplifier. In: IEEE Bipolar/BiCMOS Circuits and Technol. Meeting, 1995. p. 190–3.

[18] Gruhle A, Schüppen A, König U, Erben U, Schumacher H. 40 GHz VCO's with SiGe heterojunction bipolar transistors. In: Tech. Dig. Int. Electron Device Meeting, 1995. p. 725–8.

[19] Soyuer M, Burghartz J, Ainspan H, Jenkins K, Xiao P, Shahani A, Dolan M, Harame D. An 11 GHz 3 V SiGe voltage-controlled oscillator with integrated resonator. In: IEEE Bipolar/BiCMOS Circuits Technol. Meeting, 1996. p. 169–72.

[20] Gruhle A, Mähner C. Low 1/f Noise SiGe HBTs with applications to low phase noise microwave oscillators. Electronics Letters 1997;33:2050–1.

[21] Glenn J, Poisson R, Case M, Harame D, Meyerson B. 12-GHz Gilbert mixers using a manufacturable Si/SiGe epitaxial-base bipolar technology. In: IEEE Bipolar/BiCMOS Circuits Technol. Meeting, 1995. p. 186–9.

[22] Gao W, Snelgrove W, Varelas T, Kovacic S, Harame D. A 5-GHz SiGe HBT return-to-zero comparator. In: IEEE Bipolar/BiCMOS Circuits Technol. Meeting, 1995. p. 166–9.

[23] Case M, Larson L, Rensch D, Rosenbaum S, Knorr S, Hararne D, Meyerson B. A 23 GHz static 1/128 frequency divider implemented in a manufacturable Si/SiGe HBT process. In: IEEE Bipolar/BiCMOS Circuits Technol. Meeting, 1995. p. 121–3.

[24] Wurzer M, Meister T, Schäfer H, Knapp H, Böck J, Stengl R, Aufinger K, Franosch M, Rest M, Möller M, Rein H-M, Felder A. 42 GHz static frequency divider in Si/SiGe bipolar technology. In: Tech. Dig. IEEE Int. Solid-State Circuits Conf, 1997. p. 122–3.

[25] Götzfried R, Itoh T, Luy JF, Schumacher J. Zero power consumption Si/SiGe HBT SPDT T/R antenna switch. In: Tech. Dig. IEEE M77-S Int. Microwave Symp, 1996. p. 651–3.

PERGAMON

Solid-State Electronics 43 (1999) 1383–1388

SOLID-STATE ELECTRONICS

n- and p-Type SiGe HFETs and circuits

U. König[a,*], M. Zeuner[a], G. Höck[a], T. Hackbarth[a], M. Glück[a], T. Ostermann[b], M. Saxarra[c]

[a]*Daimler-Benz Research Center Ulm, D-89081 Ulm, Germany*
[b]*University Linz, A4040 Linz, Austria*
[c]*Ruhr-University Bochum, D-44780 Bochum, Germany*

Abstract

n- and p-Type SiGe HFETs exhibit advanced performance especially favourable for RF-applications. Due to strained channels high carrier mobilities at room temperature (2700 and 1870 cm^2/V s) and large sheet carrier densities ($n_s = 6.4 \times 10^{12}$ $cm^{\times 2}$ and $p_s = 2.1 \times 10^{12}$ $cm^{\times 2}$) have been achieved. For the n-MODFET ($L_G > = 150$ nm) tensile strained Si channels embedded in SiGe layers lead to a maximal g_{me} of 476 mS/mm and to cut-off frequencies of $f_t = 43$ GHz and $f_{max} = 92$ GHz. The best results for p-type HFETs were attained for a pure Ge channel MODFET with $f_t = 32$ GHz and $f_{max} = 85$ GHz. Analog and digital circuit realizations for the n-MODFET resulted in a transimpedance amplifier yielding a Z_{21} of 56 dB Ω at a bandwidth of 1.8 GHz and an inverter with a gate delay of 25 ps. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The introduction of SiGe offers the possibility of band gap engineering in the Si system and thus the realization of n- and p-type hetero FETs, both with enhanced performance. Tensile strained Si films on strain relieved SiGe buffer layers enable the formation of electron quantum wells which exhibit enhanced electron mobilities. Compressively strained SiGe or even pure Ge layers can be used to create two-dimensional hole channels. Room temperature mobilities of up to 2900 and 1870 cm^2/V s have been achieved in 2D electron and 2D hole channels, respectively [1,2]. In contrast to III–V, SiGe offers compatibility with the well-established Si technology and its capability for low cost mass production [3]. This means especially the natural oxide, which can be grown on a Si capped SiGe hetero structure and thus enables the fabrication of MOS gated hetero FETs. The combination of high performance n- and p-type SiGe hetero-MOSFETs to a complementary circuit arrangement promises the realization of advanced analog and digital high speed circuits [4–6].

2. Technology

The basic layer structure of the n-MODFET [7] is grown by molecular beam epitaxy (MBE) on p^--substrate with $\rho \geq 1000$ Ω cm starting with a relaxed SiGe buffer layer whose Ge content is linearly graded to 40%. The core of the layer structure is the 9 nm thick biaxially strained Si channel embedded in undoped $Si_{0.6}Ge_{0.4}$ spacers which separate the following carrier supply layers from the channel. Due to a Ge content of 40–45% in the SiGe layers, a high conduction band

* Corresponding author. Tel.: +49-731-505-2050; fax: +49-731-505-4102.
E-mail address: ulf.koenig@daimlerchrysler.com (U. König)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00077-5

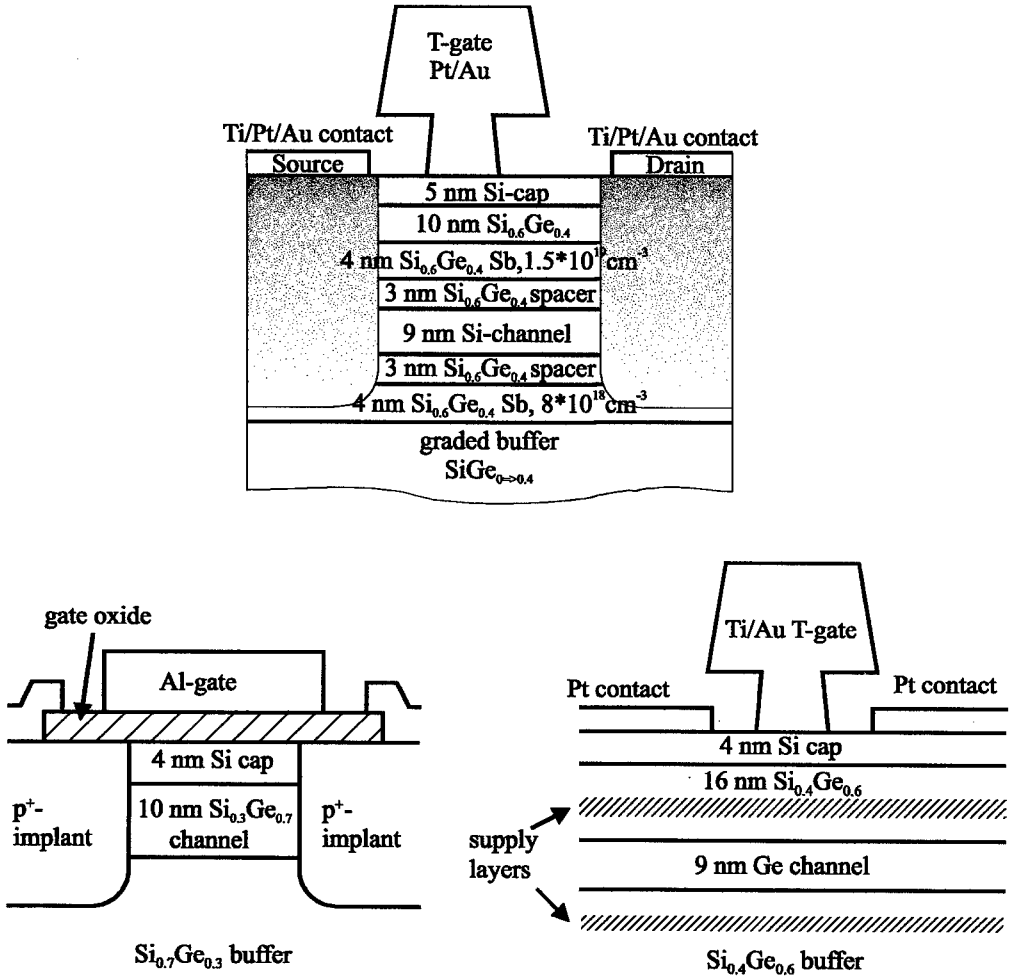


Fig. 1. Layer structure of n-type SiGe MODFETs (a), p-type $Si_{0.3}Ge_{0.7}$ channel MOSFETs (b) and p-type Ge channel MODFETs (c).

offset [8] is achieved and the therefore existing quantum well enables sheet carrier densities up to $n_s = 6.4 \times 10^{12}$ cm^{-2} and electron mobilities up to 2700 cm^2/V s. Fig. 1(a) exhibits a cross-section of the layer sequence with detailed information.

The device processing of the n-MODFET started with the formation of the mesa realized by dry etching in a SF_6/O_2–plasma. The ohmic contacts of drain and source are defined by ion implantation (P, 20 keV, 2×10^{15} cm^{-2}) followed by the electrical activation of the dopants using rapid thermal annealing (30 s, 625°C). A lift-off process of Ti/Pt/Au forms the metallisation layer of the ohmic contacts. The asymmetrically located T-shaped Schottky-gate is patterned by e-beam-lithography and consists of 50 nm Pt and 300 nm Au with a footprint of $L_G = 250$ nm. In the case of recessed gates prior to the gate metallisation a dry etch process is performed in a SF_6/O_2-plasma reducing the distance d_{GC} between gate and channel.

For the p-type HFETs two concepts have been under investigation: a modulation doped HFET with a strained Ge channel and a $Si_{0.3}Ge_{0.7}$ channel hetero MOSFET, making use of the native oxide, which can be formed on a Si capped hetero structure.

These samples were grown by molecular beam epitaxy (MBE) on high resistivity n-type (100) Si substrate. Fig. 1(b) shows the structure of a p-type SiGe MOSFET. The compressively strained $Si_{0.3}Ge_{0.7}$ channel (10 nm) with a hole mobility of 350 cm^2/V s was grown on a relaxed $Si_{0.7}Ge_{0.3}$ buffer and capped by a thin Si layer (7 nm). The 3.5 nm thick gate oxide was formed by thermal oxidation of the Si cap layer at 750°C in an H_2O atmosphere. Ohmic contacts were fabricated by a BF_2^+ implant at 50 keV. Finally, the

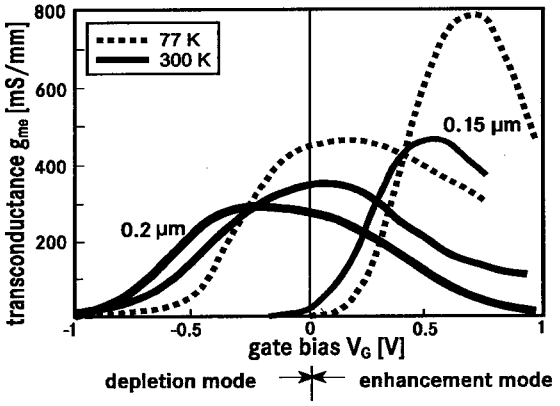


Fig. 2. Extrinsic transconductance as a function of the applied gate voltage for n-MODFETs with different gate to channel distances.

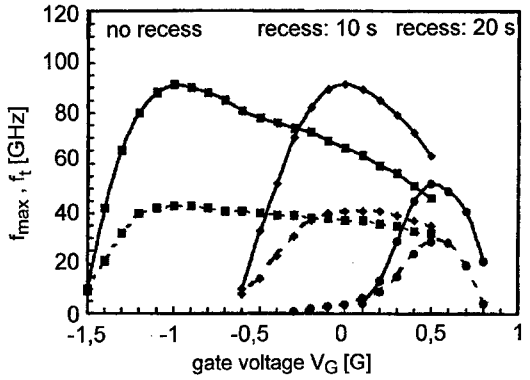


Fig. 3. Measured cut-off frequencies f_t (dotted lines) and f_{max} (solid lines) as a function of the gate voltage and recess etching time.

overlapping Al gate and the contact metallisation were deposited.

The structure of the Ge channel MODFET is shown in Fig. 1(c). The 9 nm thick strained Ge hole channel is grown on a strain relieved SiGe buffer layer with a Ge content of 60%. The channel is followed by a 16 nm $Si_{0.4}Ge_{0.6}$ layer and a 4 nm thick Si cap. Two boron doped supply layers are placed above (5 nm, 8×10^{18} cm^{-3}) and beneath (5 nm, 2×10^{18} cm^{-3}) the channel. The sample exhibits a low sheet resistance of about 1300 Ω/sq. A high 2D gas hole mobility of 1870 cm^2/V s was extracted from magnetic field dependent Hall measurements [9,10], even at the high sheet carrier density of 2.1×10^{12} cm^{-2}. The processing started with mesa definition by dry etching, then the sample was covered by 200 nm SiO_2, deposited by CVD, for passivation. In the active transistor areas this passivation oxide was removed by wet chemical etching. The T-shaped Ti/Au Schottky gates with footprints of $L_G = 0.25$ µm were structured by electron-beam-lithography, using a three layer resist system. To form ohmic contacts, 20 nm Pt was evaporated in the active areas. In the gate region the evaporated Pt was screened by the 250 nm wings of the T-shaped gates, which leads to a quasi self aligned structure [11]. Then contact pads were defined by deposition of Ti/Au.

3. n-Type Si channel HFETs

Enhanced electron mobility in strained Si layers compared to bulk Si offers new possibilities for high speed semiconductor devices realized by the n-MODFET [12,13].

Dc-characterization of transistors with different Schottky-gate to Si-channel distances d_{GC} exhibit a change in the operation mode of the MODFET from depletion to enhancement [14]. The three solid lines in Fig. 2 (from left to right) show the dependence of g_{me} on the applied gate voltage V_G. The maxima of g_{me} correspond to the recess etching times of 0, 10 ($d_{GC} \approx 12$ nm) and 20 s ($d_{GC} \approx 20$ nm) and reach 285, 345 and 476 mS/mm at room temperature, respectively. The threshold voltages of the three transistors shift with d_{GC} from $V_{th} \approx -1.6$–0 V. Under reverse bias ($V_G < V_{th}$) a low gate leakage current below 1 µA/mm was measured, indicating the good quality of the Schottky-gate even on SiGe.

Considering the RF-behavior of the n-MODFETs (Fig. 3) we found equal values for the unity current gain cut-off frequency $f_t = 43$ GHz and the maximum frequency of oscillation $f_{max} = 92$ GHz (the highest reported so far) for the unrecessed transistor and that produced with a 10 s recess process. These frequencies were extrapolated at $V_{DS} = 2.5$ V and operation points of $V_G = -1$ V and $V_G = 0$ V. For the deepest recess the cut-off frequencies of the enhancement MODFET are reduced to $f_t = 30$ GHz and $f_{max} = 52$ GHz due to the lower drain currents and the increased C_{gs}. Fig. 4

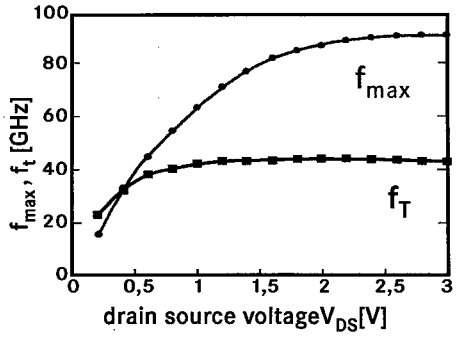


Fig. 4. f_t and f_{max} values of a 10 s recessed n-MODFET as a function of the drain-source voltage at $V_G = 0$ V.

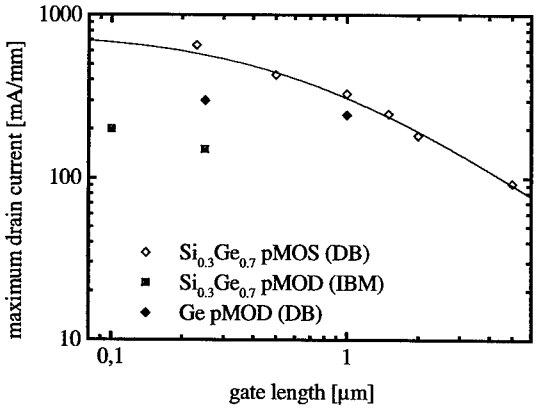


Fig. 5. Maximum drain current as a function of gate length for various p-type hetero FETs.

shows the dependence of the cut-off frequencies on the applied V_{DS} for the 10 s recessed device. f_t attains the maximum of 43 GHz at $V_G=0$ V and $V_{DS}=1$ V, exactly at the peak of g_{me} and saturates for higher V_{DS}. This indicates that the ratio g_m/C_{gs}, which has the main influence on f_t, remains constant. Since g_m decreases for $V_G>0$ V (see Fig. 2), C_{gs} has to abate equally. The saturation of f_{max} occurs at 92 GHz for $V_{DS}>2$ V. This high frequency value results from the low parasitic source resistance of $R_S=0.18$ Ω mm attributed to the asymmetric position of the e-beam-gate at $d_{SG}=200$ nm. Additionally we found an increasing R_{ds} and C_{gs}/C_{gd} up to the point of saturation at $V_{DS}=2$ V, so that f_{max} grows far above f_t. Finally the very low gate resistance ($R_G=250$ Ω mm^{-1}) of the thick T-shaped gate has a significant influence on the maximum of the frequency of oscillation.

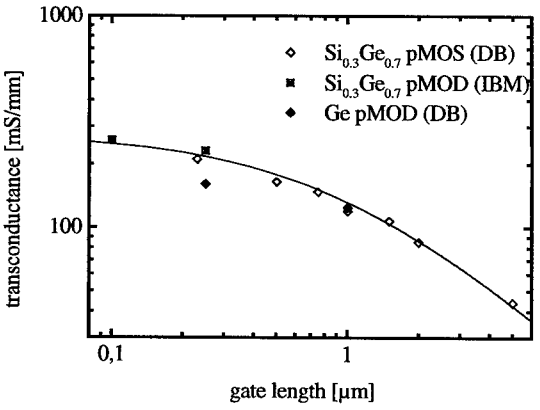


Fig. 6. Extrinsic transconductance as a function of gate length for various p-type hetero FETs.

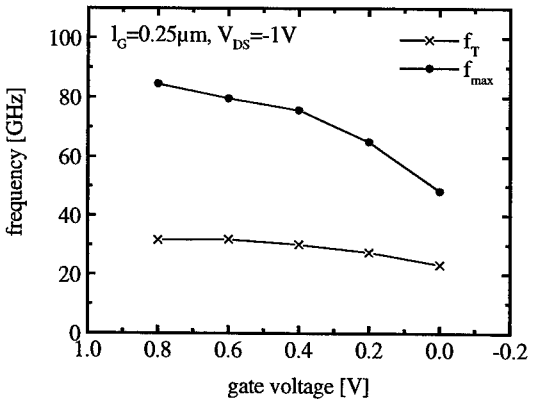


Fig. 7. Gate voltage sweep off f_T and f_{max} of the 0.25 μm gate length Ge channel MODFET at $V_{DS}=-1$ V.

4. p-Type SiGe and Ge channel HFETs

Fig. 5 shows the maximum drain saturation current of our $Si_{0.3}Ge_{0.7}$ MOSFET and Ge channel MODFET as a function of the gate length. In case of the hetero MOSFET drain current values of up to 650 mA/mm have been achieved at 0.23 μm gate length. For the Ge channel MODFET values of up to 225 mA/mm at $L_G=0.8$ μm and 300 mA/mm at $L_G=0.25$ μm were measured. Results achieved by IBM for a $Si_{0.3}Ge_{0.7}$ channel MODFET [11,15] are shown for comparison. Here the drain current reaches 200 mA/mm at a gate length of 0.1 μm. The hetero MOSFET exhibits the highest current values, even though it is operating in enhancement mode ($V_{Th}=-0.7$ V). This is due to the higher gate bias, which can be applied over the gate oxide.

Fig. 6. shows the gate length dependence of the extrinsic transconductance of the devices. The $Si_{0.3}Ge_{0.7}$ channel MOSFET exhibits values of up to 210 mS/mm at a gate length of 0.23 μm. A maximum transconductance of 160 mS/mm at $L_G=0.25$ μm and 125 mS/mm at $L_G=0.8$ μm was achieved for the Ge channel MODFET.

For RF characterization, S-parameters of the Ge channel MODFET were measured on-wafer in the range of 2 to 40 GHz. Fig. 7 shows the gate bias dependence of the unity current gain frequency f_t and the maximum frequency of oscillation f_{max} at $V_{DS}=-1$ V. f_t shows only a small variation and saturates at 32 GHz for gate voltages higher than 0.6 V. f_{max} increases monotonously for increasing V_G and shows a maximum value of 85 GHz at $V_G=0.8$ V.

5. HFET circuits

The exploitation of the main electrical advantages of

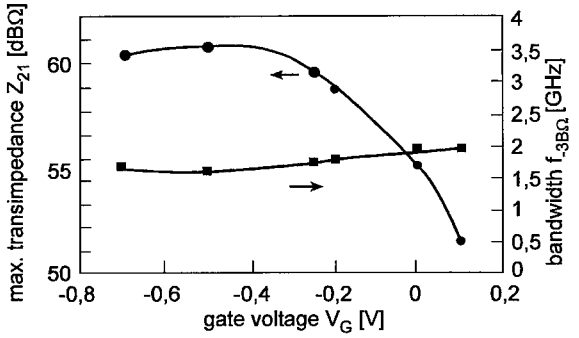


Fig. 8. Transimpedance Z_{21} and bandwidth $f_{-3\,dB\,\Omega}$ as a function of V_G for the analog circuit.

the n-MODET (i.e. high saturation current I_{Dsat}, large g_{me} and high cut-off frequencies f_t and f_{max}) have been the aim for designing analog and digital IC applications. Especially for RF-circuits in mobile- and satellite-communication and in automotive sensors the MODFET is favourable.

For the fabrication of the analog and digital circuits the standard process technology described before has to be extended by a 300 nm thick passivation-oxide and a second metallisation consisting of 100 nm TiW and 1 μm Al which was sputter-deposited and structured by wet etching.

A transimpedance amplifier for transforming current signals into amplified voltage signals was realized [16]. These kinds of amplifiers play an important role in modern communication systems as front-end receivers or low noise amplifiers. The analog circuit consists of two functional stages. The first one transforms the incoming current into a voltage using a single transistor common-source-stage with a drain-to-gate feedback realized by a series connection including a capacitance and a feedback resistor. The second stage amplifies the output voltage of the transimpedance stage by the help of a common-source-circuit with two parallel n-MODFETs. The evaluation of the amplifier performance by RF-measurements showed a high transimpedance of 56 dB Ω with a −3 dB Ω frequency of 1.8 GHz obtained at the dc operation point of $U_{DD}=5$ V and $V_G=-0.2$ V. A variation of the operation point around the previous considered peak of transconductance (see Fig. 8) led to the crest of transimpedance Z_{21} of 61 dB Ω at $V_G=-0.5$ V and the maximal bandwidth $f_{-3\,dB\,\Omega}$ of 2 GHz at $V_G=0.1$ V

In order to appraise the potential of the n-MODFET for digital IC applications inverter circuits have been realized [17]. The basic inverter consists of two n-MODFETs. The first one is the current-driver with W = 100 μm, $d_{SG}=0.5$ μm and 180 and 300 nm gate length. The second transistor is ungated and therefore works as a resistor in the load path. Fig. 9

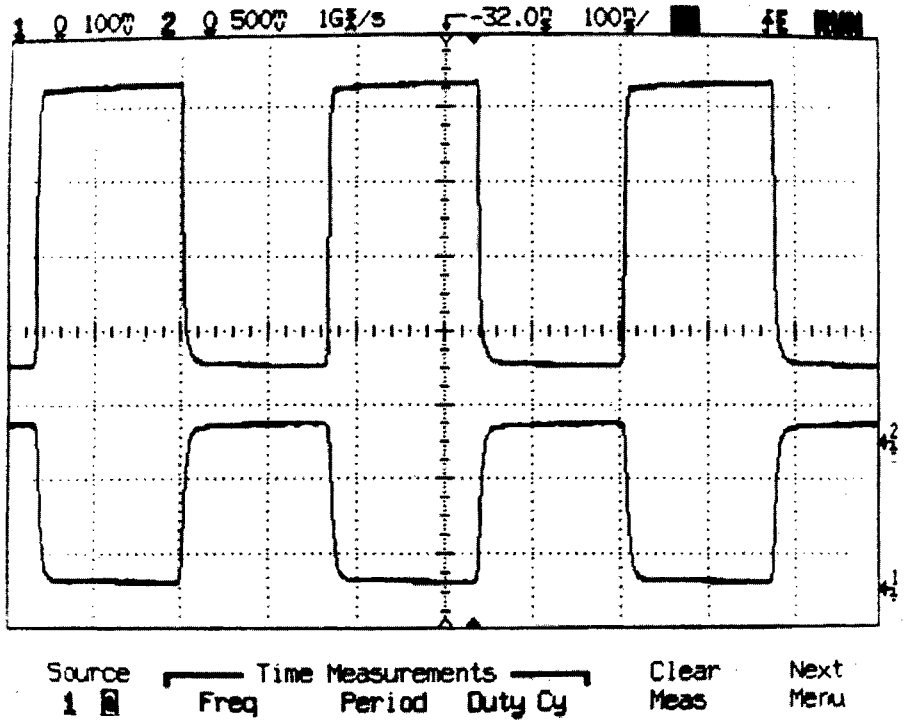


Fig. 9. Input and output signal of the inverter circuit.

shows the input and output signal of the inverter stage operating at a frequency of 1 GHz and $V_{DD}=2$ V. Applying an input signal of 600 ps rise time the circuit exhibited a delay of 70 ps for $L_G=300$ nm. Inverters with $L_G=180$ nm attained a gate delay of 25 ps and a maximum output voltage swing of 430 mV at an input rise time of 150 ps. The delay time was determined between the 50% input and output signal values, taking into account the different RC-delays of the measurement equipment. For standard n-Si MOS delays at respective gate lengths are above 30 ps.

6. Conclusion

In this paper we present n- and p-type hetero FETs based on MBE grown layer structures. Carrier mobilities and cut-off frequencies of p-SiGe based MODFETs are now comparable to those of n-type SiGe based MODFETs. This is a milestone towards high speed complementary HFET circuits. In case of n-type MODFETs first circuit applications have been demonstrated: a transimpedance amplifier with a bandwidth of 1.8 GHz and an inverter with a gate delay of 25 ps. Further improvements in the performance of n- and p-type HFETs are expected by downscaling and selfaligning the gates of the device and by a reduction of parasitic series resistances. These topics are under investigation at the moment.

Acknowledgements

The authors would like to thank R. Hagelauer (University Linz), K. Hofmann (University Hannover), E. Kohn (University Ulm) and U. Langmann (University Bochum) for stimulating discussions. The financial support of the German BMBF in the frame of the project NANOMOS, grant M 2957, is gratefully acknowledged.

References

[1] Ismail K, Nelson SE, Chu JO, Chan K, Meyerson BS. Appl Phys Lett 1993;63(5):660–2.

[2] Höck G, Hackbarth T, Erben U, Kohn E, König U. High performance 0.25 μm p-type Ge/SiGe MODFETs. IEE Electron Lett 1998;34(19):1888–9.

[3] König U, Gruhle A. Proc. of the IEEE Cornell Conf. on Advanced Concepts in High Speed Semiconductor Devices and Circuits. 1997. p. 14–23.

[4] König U, Glöck M, Gruhle A, Höck G, Kohn E, Bozon B, Nuembergk D, Ostermann T, Hagelauer R. Solid State Electron 1997;41(10):1541–7.

[5] Hagelauer R, Ostermann T, König U, Glück M, Hock G. IEE Electron Lett 1997;33(3):208–10.

[6] Sadek A, Ismail K, Armstrong MA, Antoniadis DA, Stem F. IEEE Trans Electron Dev 1996;ED-43(8):1224–32.

[7] Dämbkes H, Herzog HJ, Jorke H, Kibbel H, Kasper E. The n-channel SiGe/Si modulation-doped field-effect transistor. IEEE Trans Electron Dev 1986;ED-33(5):633–8.

[8] People R, Bean JC. Band alignments of coherently strained Ge_xSi_{1-x}/Si hetero structures on $\langle 001\rangle Ge_ySi_{1-y}$substrates. Appl Phys Lett 1986;48(8):538–40.

[9] Beck WA, Anderson JR. Determination of electrical transport properties using a novel magnetic field dependent Hall technique. J Appl Phys 1987;62(2):541–54.

[10] Höck G, Glück M, Hackbarth T, Herzog H-J, Kohn E. Carrier mobilities in modulation doped $Si_{1-x}Ge_x$ hetero structures with respect to FET applications. E-MRS Spring Meeting, June 16–19, 1998, Strasbourg, France. Thin Solid Films 1998;336:141–4.

[11] Arafa M, Ismail K, Chu JO, Meyerson BS, Adesia I. A 70-GHz f_T low operating bias self aligned p-type SiGe MODFET. IEEE Electron Dev Lett 1996;17(12):586–8.

[12] König U, Schäffler F. Modulation-doped high-mobility Si/SiGe hetero structures for device applications. In: International Conference on Solid State Devices and Materials. SSDM, 1993.

[13] Ismail K. Si/SiGe high-speed field-effect transistors. In: IEDM 95 Tech. Digest, 1995. p. 509–12.

[14] Glück M, Hackbarth T, Birk M, Haas A, Kohn E, König U. Design and fabrication of SiGe n-type MODFETs. Physica E: Low dimensional Systems and Nanostructures 1998. p. 763–7.

[15] Arafa M, Fay P, Ismail K, Chu JO, Meyerson BS, Adesia I. Dc and RF performance of 0.25 μm p-type SiGe MODFET. IEEE Electron Dev Lett 1996;17(9):449–51.

[16] Saxarra M, Glück M, Albers JN, Behammer D, Langmann U, König U. Transimpedance amplifier based on Si/SiGe MODFETs. IEE Electron Lett 1998;34(5):499–500.

[17] Ostermann T, Glück M, Hagelauer R, König U, Dobersberger M, Birk M, Oehler F. Inverter circuits with Si/SiGe n-type MODFETs. In: International Semiconductor Device Research Symposium. ISDRS, 1997.

PERGAMON

Solid-State Electronics 43 (1999) 1389–1393

SOLID-STATE ELECTRONICS

Process design for SiGe-HBTs prepared using cold-wall UHV/CVD

F. Sato[a,*], T. Hashimoto[a], T. Tatsumi[b], T. Tashiro[a]

[a]*ULSI Device Development Laboratory, NEC Corporation, 1120, Shimokuzawa, Sagamihara, Kanagawa, 229-1198, Japan*
[b]*Silicon Systems Research Laboratories, NEC Corporation, 34, Miyukigaoka, Tsukuba, Ibaraki, 305-8501, Japan*

Abstract

We have studied the influence of Si/SiGe/Si profile on electrical characteristics of a self-aligned SiGe-base heterojunction bipolar transistor (HBT) with a poly-crystalline silicon (polysilicon) emitter. Particular attention has been given to a collector-base undoped SiGe intermediate layer. This undoped layer acts as a spacer for boron out-diffusion of the p^+-type SiGe to the n^--type Si collector, during heat treatment after the SiGe base formation. The SiGe-base was selectively grown by a cold-wall ultra-high vacuum/chemical-vapor-deposition (UHV/CVD) technique. Using both low temperature oxide (LTO) as a passivation film, and in-situ phosphorus-doped emitter polysilicon for suppressing of boron out-diffusion, a 15-nm-thick undoped layer was an insufficient boron out-diffusion spacer. This was confirmed by measuring the cut-off frequency, f_T, characteristics. With a 30-nm-thick spacer layer, the f_T of the SiGe base bipolar transistor increased from 38 to 63 GHz. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

In order to obtain a higher breakdown voltage, BV_{CEO}, and/or a reduction of base resistance, R_B, an increase of the doped impurity concentration in the base of bipolar transistors is necessary. When this concentration increases, the current gain, h_{FE} decreases. This phenomenon is one of the limitations of bipolar junction transistors (BJTs). Because of this limitation, HBTs were proposed as alternative transistors in the 1950 s. In HBTs, the band gap difference $\Delta E_g(E-B)$ between the emitter and the base materials can compensate for the decrease of h_{FE} [1,2]. Also, HBTs, with their graded bandgap base, can reduce the base transit time of minority carriers τ_B, because they have an accelerating drift field built in a neutral base region [3]. Recent advancement in Si-based hetero-epitaxial growth technology [4] has opened up the application of heterostructures to Si devices. Wide-gap emitter Si-HBTs [5,6], have been studied first, because of their compatibility to state-of-the-art double polysilicon self-aligned Si-BJTs. They have succeeded in increasing h_{FE}. The emitter transit time τ_E, one of the components of the cut-off frequency f_T can be decreased due to the increase in h_{FE} [7]. However, the base transit time τ_B cannot be decreased, because it is mainly determined by three parameters, such as the base width W_B, the minority carrier diffusion coefficient D_n, and the electric field E in the neutral base. In order to form the ultra-shallow base junction, which leads to an increase in f_T, an epitaxial growth technology is necessary. A graded SiGe-base HBT has succeeded in increasing f_T [8]. In terms of the emitter formation

* Corresponding author. Tel.: +81-42-779-9926; fax: +81-42-771-0938.
E-mail address: f-sato@lsi.nec.co.jp (F. Sato)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00078-7

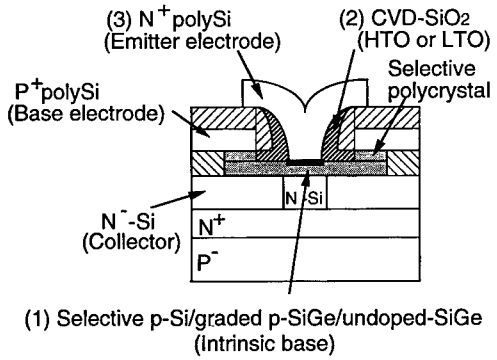


Fig. 1. Schematic cross-section of SiGe-HBT. Selective epitaxial base has triple layers, *p*-Si/*p*-graded SiGe/undoped-SiGe.

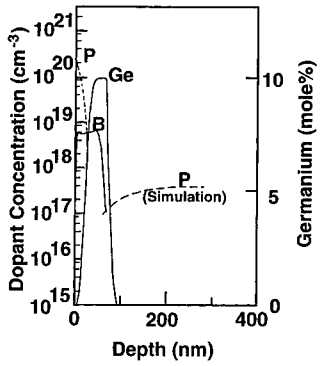


Fig. 2. Secondary ion mass spectrometry (SIMS) profiles of germanium and boron in the base, and of phosphorus in the emitter. A simulated collector profile of implanted phosphorus is also added.

method, there are two types of SiGe-base HBTs. One type, the SiGe-base HBT, like the III-V HBT, has an epiaxially-grown emitter region. That is, the collector, the base, and the emitter are sequentially grown on a substrate by a molecular beam epitaxy (MBE) technique, and useless regions are removed by mesa-etching. A low-temperature CVD film is used as the passivation film on the epitaxial layer. Consequently, as-grown profiles are preserved in the final devices [9–11]. However, it is difficult to decrease the emitter area for low-current operation. The other type of SiGe-HBT has an emitter fabricated by the diffusing of *n*-type dopant from the polysilicon. In Si bipolar transistors, polysilicon emitter technology is very attractive in terms of self-alignment and enhancement of current gain h_{FE} [12]. Furthermore, higher f_T can be simulated for a polysilicon emitter than for an epitaxial silicon emitter as is the case for the emitter in graded SiGe-base transistors [13]. However, heat treatment of the emitter drive-in simultaneously causes the diffusion of the base dopants into the collector region. Therefore, the as-grown base doping profile cannot be preserved in the final profile of polysilicon-emitter BJTs. If out-diffusion of boron into the n^--type Si region occurs, there is a conduction bandgap difference at the junctions, which has an undesirable effect for the electrons travelling through the base to the collector. Minority carrier electrons injected from the emitter into the base were prevented from travelling by the hetero-interface positioned in the base. The parasitic energy barrier at the interface of the hetero-structure causes degradation of electrical characteristics, for example, decrease in h_{FE} and f_T [14–16]. This paper describes an experimental study of the influence of the collector-base undoped intermediate layer on electrical characteristics of double polysilicon self-aligned SiGe-base HBTs. The study is aimed at eliminating the parasitic energy barrier effect.

2. Experiment

Selective epitaxial growth (SEG) technology has many attractive points which make it suitable for application to transistors [17]. The intrinsic base of the experimental self-aligned SiGe-base bipolar transistor shown in Fig. 1 was fabricated using SEG technology [18,19]. We used cold-wall ultra-high vacuum (UHV)/CVD technology. The conditions for the selective epitaxial process can be seen in detail elsewhere [20]. The Ge profile in the SEG layer is trapezoidal. The Si/SiGe base layer was composed of three layers [21]. The first layer on the Si collector was an undoped (in effect, a p^--type) SiGe layer with constant Ge content to act as a boron out-diffusion spacer between the Si collector and the p^+ base. This undoped layer was doped by selective implantation of phosphorus ions after SEG. The second layer was p^+-type graded SiGe. The upper layer was a pure Si layer in which the emitter dopant impurity had been diffused from the emitter polysilicon film, when the emitter drive-in was carried out. The profiles in the intrinsic region of the optimized transistor with 63-GHz-f_T are shown in Fig. 2. The peak Ge

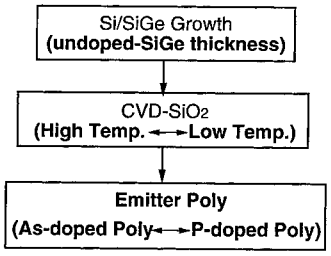


Fig. 3. Main process flow after Si/SiGe growth. Three points were studied: (1) undoped SiGe-layer thickness, (2) CVD-SiO_2 deposited at high (>800°C) or low (<700°C) temperatures, and (3) emitter polysilicon dopant (arsenic or phosphorus).

concentration is 10 mol%. A simulated phosphorus profile implanted into the Si-collector is also shown. At a high current density, we studied f_T degradation due to the parasitic barrier effect. First, we studied the influence of process temperature (CVD and emitter drive-in). Second, we studied how the outcome is affected by the thickness of the spacer undoped SiGe between the p^+-type graded SiGe and the n^--type Si collector. The process flow after the selective SiGe epitaxial base formation is illustrated in Fig. 3. After the epitaxial base was selectively grown, a SiO_2 sidewall for separating the emitter from the graft base region was formed by SiO_2 deposition and subsequent etch-back. An in-situ doped polysilicon film for an emitter electrode was deposited.

2.1. Spacer thickness

The lattice constant of a SiGe alloy is not coincident with that of Si. The strained/unstrained transition of SiGe/Si is determined by the Ge concentration, its thickness, and by the heat treatment (temperature and duration). From the point of view of suppressing collector-base junction degradation, the thinner the undoped SiGe layer used as the boron-diffusion spacer, the better the junction characteristics. A boron diffusion coefficient in SiGe has been reported [22–24], and it is about one order of magnitude smaller than that in Si. To avoid formation of the parasitic energy barrier, optimization of the thickness of the undoped SiGe layer is needed. The thickness of the SiGe spacer layer between the Si-collector and the p^+-graded SiGe intrinsic base was varied from 15 to 30 nm.

2.2. Sidewall

We used two different types of CVD SiO_2 films for electrical separation between base and emitter polysilicon, and for passivation of the epitaxial base. One was high-temperature oxide (HTO) and the other was low-temperature oxide (LTO). The deposition temperatures of HTO and LTO were >800 and <700°C, respectively. This temperature difference caused different boron profiles in the as-grown p^+-type graded-SiGe/undoped-SiGe (as a spacer)/n^--type Si-collector hetero-structure.

2.3. Emitter polysilicon impurities

In state-of-the-art double-polysilicon self-aligned Si-BJTs, arsenic atoms are usually used as the emitter dopant impurity, because a steep emitter junction can be formed by arsenic diffusing from the polysilicon film into the single-crystalline silicon region. However, from the point of view of lowering the processing temperature, phosphorus dopant is a promising emitter impurity for SiGe-HBTs, because its diffusion coefficient is higher than that of arsenic. To compare the effect of emitter dopant impurities, in-situ arsenic-doped and phosphorus-doped polysilicon were each used for the polysilicon electrode of the emitter in the experimental self-aligned SiGe-base HBTs. Emitter drive-in of rapid thermal annealing (RTA) was carried out at 980°C and 910°C for arsenic-doped and phosphorus-doped emitter polysilicon, respectively.

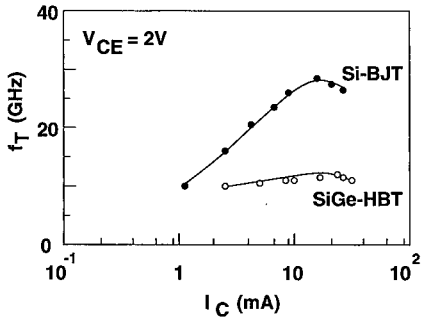


Fig. 4. f_T versus I_C characteristics of SiGe-HBT with a parasitic barrier and Si BJT using high temperature oxide and in-situ arsenic doped emitter polysilicon.

3. Results and discussion

We measured the cut-off frequency f_T versus the collector current I_C characteristics of a Si-BJT and a SiGe-HBT (30-nm undoped-SiGe layer) fabricated using high temperature oxide (HTO) film and the in-situ As-doped polysilicon film. They have the same as-grown boron-doping profile in the SEG layer. As shown in Fig. 4, the measured f_T for the Si BJT was 28 GHz, but it was as low as 12 GHz for the SiGe-HBT. This difference cannot be explained by the base width. The generalized base transit time τ_B formula is represented as [25],

$$\tau_B = \int_0^{W_B} \left[\frac{n_i^2(z)}{N_B(z)} \int_z^{W_B} \frac{N_B(y)}{D_n(y) n_i^2(y)} dy \right] dz \quad (1),$$

where notations have their usual meanings. This equation means that a shallower base produces a smaller base transit time, if other parameters, such as, W_B, D_n, are of the same order. We can expect the base widths of SiGe-HBTs to be shallower than those of Si-BJTs, because the boron diffusion coefficients in the SiGe alloy film are reported to be about one order small than those in the Si film. This f_T difference can be explained by the existence of an energy barrier for the current flow. The potential barrier blocks the electron flow injected from the emitter through the base to

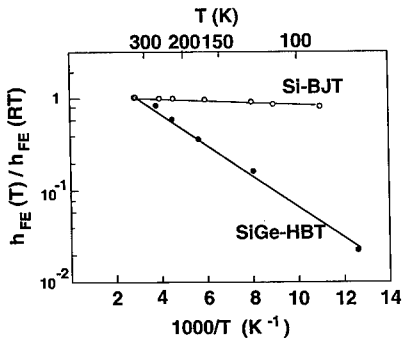


Fig. 5. Current gain versus temperature of the same SiGe-HBT and Si-BJT as shown in Fig. 4.

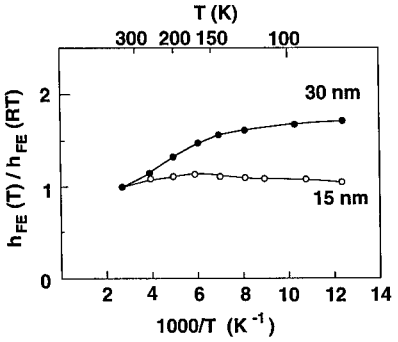


Fig. 6. Current gain h_{FE} versus temperature T for transistors with a 15-nm or a 30-nm-thick undoped-SiGe intermediate layer between a p^+-type SiGe-base and an n^--type Si collector.

the collector, and consequently the injected charge is stored in the base.

The effect of temperature on current gain h_{FE} for the SiGe-HBT and the Si-BJT is shown in Fig. 5. The h_{FE} of BJT is almost independent of temperature. The h_{FE} of usual Si-BJTs has been known to decrease with decreasing temperature because there is a bandgap difference $\Delta E_g(E-B)$ between the emitter and the base caused by the higher-doped emitter (this is known as the heavily-doped bandgap narrowing effect). The bandgap difference $\Delta E_g(E-B)$ decreases as the base doping concentration for the narrow base width increases. Base boron concentration in this Si-BJT was about 6×10^{18} cm^{-3}, which was high enough to compensate for the bandgap narrowing effect in the emitter. On the other hand, current gain in the SiGe-base HBT decreased steadily with decreasing temperature. This behavior suggests the existence of a thermionic potential barrier. The potential barrier blocks the electron flow through the base to the collector.

Next, we decreased the SiO_2 CVD temperature and used the in-situ phosphorus-doped emitter polysilicon film instead of the arsenic-doped polysilicon film. There was a meaningful difference in the breakdown voltages of transistors with 15- or 30-nm-thick interlayers within process deviation. BV_{CBO} were 7.7–7.8 for the 15-nm interlayer and 8.0–8.2 for the 30-nm interlayer. BV_{CEO} were 3.0–3.3 and 3.2–3.5 for 15 and 30-nm, respectively. The current gain (=65) of the HBT with a 15-nm-thick undoped layer was slightly smaller than that (=82) with a 30-nm-thick undoped layer. The remarkable difference in the DC characteristics was the reverse/forward current ratio, I_r/I_f. Here I_r and I_f mean collector currents for reverse and forward connection at the same bias voltage, respectively. I_r/I_f had a positive correlation with f_T, that is, it increased with increasing f_T. For example, it was 2.6–3.0 for a high f_T transistor and 0.9–1.0 for an extremely low f_T transistor. These results can be interpreted as follows; I_r decreased in the transistor which had a parasitic energy barrier, because the collector-base junction was a homojunction. The h_{FE} versus $1/T$ characteristics of the SiGe-HBTs with a 15- or 30-nm undoped-SiGe layer are shown in Fig. 6. With the 15-nm-thick undoped SiGe spacer, h_{FE} is nearly independent of temperature, as it is for the Si-BJT. This indicates that the 15-nm thick SiGe is insufficient for the boron out-diffusion spacer. In contrast, the h_{FE} of the SiGe-HBT with the 30-nm-thick spacer, increases as temperature decreases. Fig. 7 shows the f_T versus the I_C characteristics of the SiGe-HBTs. A maximum cut-off frequency of 63 GHz was obtained for the SiGe-HBT with the 30-nm undoped SiGe spacer.

4. Conclusion

We studied the process design of selectively-grown SiGe-HBTs. Heat treatment, such as the CVD and the emitter drive-in, caused boron out-diffusion from an epitaxially-grown boron-doped SiGe-base layer to the n^--type Si collector. This made the energy barrier for

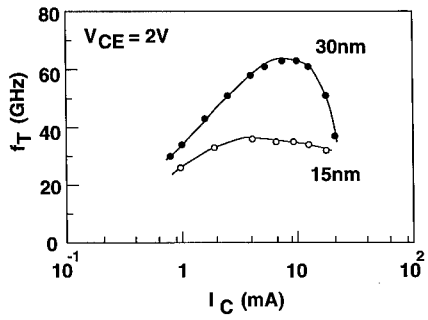


Fig. 7. f_T versus I_C characteristics. Emitter area is $0.4 \times 7.6 \times 2$ μm^2. The collector current density at the peak f_T is about 1.2×10^5 A/cm^2.

current flow. The f_T of a SiGe-HBT with the energy barrier was 12 GHz compared to an f_T of 28 GHz for a Si-BJT, in spite of the same as-grown boron doping profile. When SiO_2 deposition temperature was reduced to $<700°C$, the temperature dependence of the current gain was changed. A maximum cut-off frequency f_T of 63 GHz was obtained.

Acknowledgements

The authors would like to thank S. Shishiguchi and S. Oguro for their support in the fabrication of the transistor. The authors also thank Drs O. Kudoh, O. Mizuno, H. Abe, and M. Nakamae for their encouragement during this study.

References

[1] Shockley W US Patent 2569347 1951.
[2] Kroemer H. Proc IRE 1957;45:1535.
[3] Kroemer H. RCA Review 1957;18:332.
[4] Kasper E, Bean JC, editors. Silicon-molecular beam epitaxy. Florida: CRC Press Inc, 1988.
[5] Sugii T, Ito T, Furumura Y, Doki M, Mieno F, Maeda M. J Electrochem Soc 1987;134:2545.
[6] Sasaki K, Fukazawa T, Furukawa S Tech. Dig. of International Electron Devices Meeting, 1987. p. 2889.
[7] van den Biesen JJH. Solid-State Electron 1986;29:529.
[8] Patton GL, Comfort JH, Meyerson BS, Crabbé EF, Scilla DJ, Frésart E, de Stork JMC, Sun JY-C, Harame DL, Burghartz J Tech. Dig. 1990 Symposium on VLSI Technology, 1990. p. 49.
[9] Gruhle A, Kibble H, Konig U, Erben U, Kasper E. IEEE Electron Device Lett 1992;13:206.
[10] Tatsumi T, Hirayama H, Aizaki N. Appl Phys Lett 1988;52:895.
[11] Temkin H, Bean JC, Antreasyan A, Leibenguth R. Appl Phys Lett 1988;52:1089.
[12] Kapoor AK, Roulston DJ, editors. Polysilicon emitter bipolar transistors. IEEE Press, 1989.
[13] Patton GL, Stork JMC, Comfort JH, Crabbé EF, Meyerson BS, Harame DL, Sun JY-C Tech. Dig. of 1990 International Electron Devices Meeting, 1990. p. 13.
[14] Prinz EJ, Garone PM, Schwartz PV, Xiao X, Strurm JC Tech. Dig. of 1989 International Electron Devices Meeting, 1989. p. 639.
[15] Slotboom JW, Streutker G, Pruijmboom A, Gravesteikin DJ. IEEE Electron Device Lett 1991;12:486.
[16] Pruijmboom A, Slotboom JW, Gravesteijm DJ, Fredriksz CW, van Gorkum AA, van de Heuvel RA, van Rooij-Mulder JML, Streutker G, van de Walle GFA. IEEE Electron Device Lett 1991;12:357.
[17] Ginsberg BJ, Burghartz J, Bronner GB, Mader SR. IBM J Res Develop 1990;34:816.
[18] Sato F, Hashimoto T, Tashiro T, Tatsumi T, Hiroi M, Niino T Proceedings of 1992 Symposium on VLSI Technology, 1992. p. 62.
[19] Meister TF, Stengl R, Meul HW, Weyl R, Packan P, Felder A, Klose H, Schreiter R, Popp J, Rein HM, Treitinger L Tech. Dig. of International Electron Devices Meeting, 1992. p. 401.
[20] Sato F, Tatsumi T, Hashimoto T, Tashiro T. IEEE Trans Electron Devices 1994;41:1373.
[21] Sato F, Hashimoto T, Tatsumi T, Tashiro T. IEEE Trans Electron Devices 1995;42:483.
[22] Hu SM, Ahlgren DC, Ronsheim PA, Chu JO. Phys Rev Lett 1991;67:1450.
[23] Kuo P, Hoyt JL, Gibbons JF, Turner JE, Jacowitz RD, Kamins TI. Appl Phys Lett 1993;62:612.
[24] Loechelt GH, Tam G, Steele JW, Knoch LK, Klein KM. J Appl Phys 1993;74:5520.
[25] Kroemer H. Solid-State Electronics 1985;28:1101.

PERGAMON

Solid-State Electronics 43 (1999) 1395–1398

SOLID-STATE ELECTRONICS

Barrier thickness dependence of peak current density in GaInAs/AlAs/InP resonant tunneling diodes by MOVPE

Y. Miyamoto*, H. Tobita, K. Oshima, K. Furuya

Department of Electrical and Electronic Engineering, Tokyo Institute of Technology, 2-12-1 O-okayama, Meguro-ku, Tokyo, 152, Japan

Abstract

GaInAs/AlAs/InP resonant tunnelling diodes with three different barrier thicknesses (3.5, 5.3, and 7 nm) were fabricated by metalorganic vapor phase epitaxy and the barrier thickness dependence of the peak current density was measured. The range of peak current was from 100 to 0.1 A/cm^2. In the measurement of peak current density distribution, the deviations of peak current density became larger when the barrier became thicker. This fluctuation of peak current density can be explained by the thickness fluctuation of the barrier in the wafer's millimeter range. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Recently, integrated circuits with a combination of resonant tunneling diodes (RTDs) and other electron devices such as an HEMT have exhibited high-speed capabilities, low power consumption, and/or multi-functionality in logic circuits [1–4]. A GaInAs/AlAs heterostructure on an InP substrate is the most promising for RTD materials because InP-based electron devices show the highest speed as electron devices [5], while AlAs barriers on the InP show a high peak to valley current ratio (PVR) at room temperature [6–11]. Metalorganic vapor phase epitaxy (MOVPE) has superior characteristics in productivity and uniformity as a method for fabricating InP related devices [5]. However, RTD structures are usually grown by molecular beam epitaxy (MBE) for abrupt heterointerfaces. The successive growth of the RTD structure by MBE after the growth of HEMT structures by MOVPE was reported for HEMT/RTDs circuits [3].

On the other hand, we reported GaInAs/InP RTDs [12] and GaInAs/GaInP/InP RTDs [13] with good PVR and an atomically flat InP/GaInAs surface [14] by using MOVPE. Currently, there has been only one report of GaInAs/AlAs RTDs by MOVPE [8] with no report of peak current distribution, which is an important characteristic for the application of integrated circuits [11] and an advantage in using MOVPE.

In this paper, we report GaInAs/AlAs/InP RTDs grown by MOVPE and investigate barrier thickness dependence of peak current density and uniformity. We also discuss methods for achieving better characteristics.

2. Fabrication

The MOVPE apparatus has a horizontal reactor under reduced pressure (76 Torr). Samples are rotated

* Corresponding author. Tel.: +81-3-5734-2572; fax: +81-3-5734-2907.

E-mail address: miya@pe.titech.ac.jp (Y. Miyamoto)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00079-9

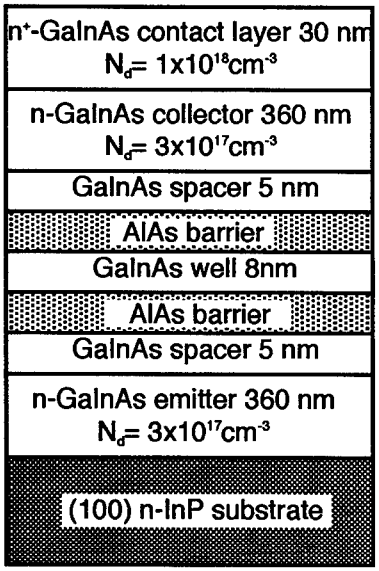


Fig. 1. Growth structure for GaInAs/AlAs RTDs with four different AlAs barrier thicknesses (3.5, 5.3, 7, and 10 nm).

with 12 rpms for uniformity. There is a load-lock system with exhaust system from a turbo molecular pump to prevent contamination from the atmosphere. The MOVPE sources are trimethylindium, triethylgallium and trimethylaluminium for III groups, arsine and phosphine for V groups, and disilane for the *n*-type dopant. The growth temperature was 600°C. The growth rate and the V:III ratio of GaInAs and AlAs were 1.7 μm/h, 1.5 μm/h, 70 and 40, respectively. The partial pressure of As was 0.76 Torr. The RTD structures are shown in Fig. 1. We made four different samples with different AlAs barrier thicknesses: 3.5, 5.3, 7, and 10 nm. The AlAs thickness was estimated from the growth rate of the AlInAs lattice-matched to InP. The contact layer was a 30 nm-thick n^+-GaInAs layer with a doping of 1×10^{18} cm^{-3}. The emitter and the collector were 360 nm-thick *n*-GaInAs layers with a doping concentration of 3×10^{17} cm^{-3}. For the spacer layers, 5 nm-thick undoped GaInAs layers were inserted between the barriers and *n*-GaInAs layers.

Diodes were fabricated by the following process after growth. Cr/Au square electrodes with a 100×100 μm^2 area were evaporated with the metal mask on the wafer. The pitch of electrode was 500 μm. Electrodes were separated into mesas by Br-methanol etching down to the RTD structure. Cr/Au was evaporated on the reverse side of the wafer after etching. The current density-voltage (J–V) characteristics were measured at room temperature.

3. Measurement results

When the barrier thicknesses were 3.5, 5.3, and 7 nm, we observed negative differential resistance (NDR). Typical J–V characteristics of the RTD with 3.5 nm-thick barriers are shown in Fig. 2. Here, we defined the forward bias as the bias for current injection from the top electrode. The highest PVRs were 2.5 at the forward bias and 3.6 at the reverse bias, respectively. The highest peak current densities were 172 A/cm^2 with 0.28 V applied voltage for the forward bias and -300 A/cm^2 with 0.66 V applied voltage for the reverse bias. In three samples with different barrier thicknesses, we observed that the peak current and peak voltage at the reverse bias became larger than at the forward bias. Fig. 3 shows a semi-log plot of |current|–voltage characteristics of diodes with three different barrier thicknesses. The peak current density change was almost a three order magnitude as shown in Fig. 3. When the barrier thickness was 7 nm, the highest PVR was 2.0 when the applied voltage was 0.3 V for the forward bias and 3.3 at -0.81 V for the reverse bias. No dependence was observed between the PVR and the barrier thickness.

To evaluate the distribution of peak current density, we measured J–V characteristics of 50 samples with a 3.5 nm-thick barrier and a 5.3 nm thick barrier and 14 samples with a 7 nm thick barrier. Since the area of the RTD electrode was relatively larger, we ignored electrode area fluctuation. When the barrier thickness was 3.5 nm, the peak current density average and the standard deviation were 132 A/cm^2 and ±22% at forward bias and 230 A/cm^2 and ±20% at the reverse

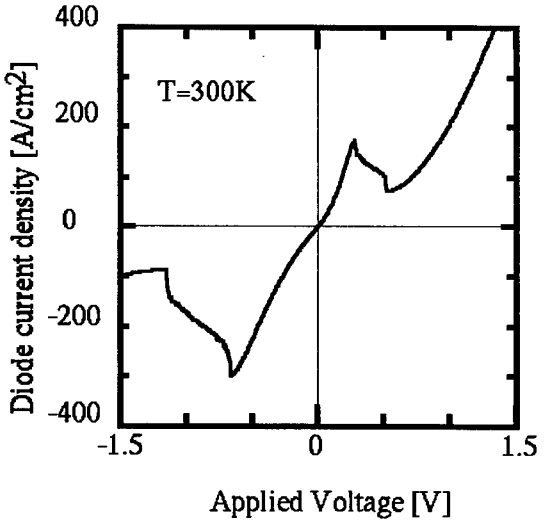


Fig. 2. Measured J–V characteristics of GaInAs/AlAs RTD at 300 K (3.5 nm barrier thickness).

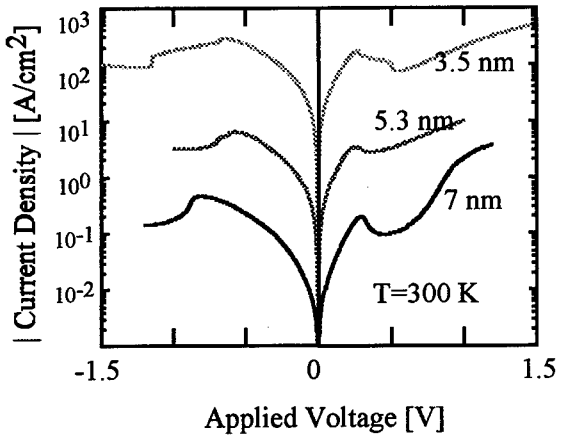


Fig. 3. Semi-log plot of |current|–voltage characteristics of diodes with three different barriers.

bias, respectively. When the barrier thickness was 5.3 nm, the peak current density average and the standard deviation were 1.74 A/cm^2 and ±43% at forward bias and 3.9 A/cm^2 and ±38% at the reverse bias, respectively. When the barrier thickness was 7 nm, the peak current density average and the standard deviation were 0.099 A/cm^2 and ±59% at the forward bias and 0.25 A/cm^2 and ±53% at the reverse bias, respectively. These results show that the deviations of peak current density became larger when the barrier became thicker. Fig. 4 shows the barrier thickness relation to average peak current density. The peak current is almost proportional to the barrier thickness exponent. We could not observe NDR when the barrier thickness was 10.5 nm.

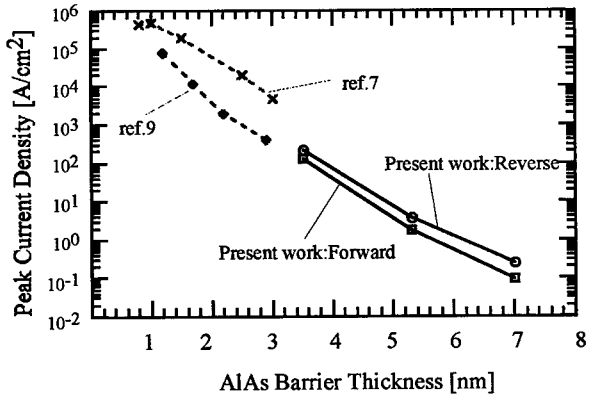


Fig. 4. Barrier thickness dependence of average peak current density (reported barrier thickness dependence is also plotted).

4. Discussion

Since a lower peak current density in RTDs is important for low power consumption of memory circuits [4], RTDs with a thicker barrier are important. We could explain the lack of NDR at the 10.5 nm thick AlAs barrier from the estimated critical thickness by strain (10 nm) [15]. On the other hand, the observed critical thickness for NDR was approximately 4 nm in our previous GaInAs/GaInP/InP RTD despite the similarity of the lattice mismatching to the present one [13]. The thickest reported AlAs barrier for RTDs by using MBE is 4.5 nm to our knowledge [10] and pre-barriers by AlInAs were inserted to reduce peak current density for memory applications [4]. Thus our critical thickness for observing NDR was quite larger than previous reports. One possibility for explaining such characteristics is the incorporation of indium. However, the incorporation of indium causes a lowering of barrier height and weaker barrier thickness dependence. We plotted the reported barrier thickness dependence [7,9] in Fig. 4. No significant difference between our tendency and the reported tendency was observed. Moreover, the same apparatus exhibited a good heterostructure for InP related optical/electron devices. The threshold current density of strained quantum well laser is 300 A/cm^2, and the current gain of the heterojunction bipolar transistors is 50 with 600 Ω/□ as base sheet resistance.

The smallest reported fluctuation of peak current [11] was 7.2% as 3 σ and our obtained standard deviation is not sufficient for logic applications [16]. We estimated the corresponding fluctuation of barrier thickness from Fig. 4. The standard deviation of RTDs with 3.5, 5.3, and 7 nm-thick AlAs barrier correspond to 0.08, 0.15, and 0.2 nm as barrier thickness fluctuations. When the thickness uniformity is due to fluctuations in growth, we can estimate the fluctuation using a thicker layer. The measured thickness fluctuation in a 1 μm-thick GaInAs layer is ±1.5% in the several millimeter range. This fluctuation corresponds to 0.053 nm of a 3.5 nm-thick barrier, 0.8 nm of a 5.3 nm-thick barrier, and 0.11 nm of a 7 nm-thick barrier, respectively. Thus, the estimated barrier thickness fluctuation from the growth rate fluctuation in the millimeter range and the estimated barrier thickness fluctuation from the peak current are in the same order. Of course, some relation between the barrier thickness fluctuation and the peak current fluctuation must be calculated for a precise comparison. A thinner barrier will provide a smaller fluctuation as noted from previous observation. Moreover, our thickness fluctuation is poorer than the one reported in modern MOVPE. A good thickness fluctuation (±0.2% in 3″ wafer) [17] was reported by MOVPE and was seven

times better than our results. Thus, MOVPE growth improvement will provide a reduction of fluctuations.

The obtained PVR is lower than the reported value in GaInAs/AlAs RTDs. Since no dependence was observed between the peak current and the PVR when the barrier thickness was changed, the low PVR could not be explained by a current leakage at the mesa side surface. The PVR at 77 K was around 6 and is lower than the previous PVR of our GaInAs/InP RTDs grown by MOVPE. We used a different MOVPE from the former experiment to introduce Al-sources and substrate rotation. We grew the GaInAs/InP RTD structure by the present MOVPE to compare apparatuses. The epitaxial structure is the same in Fig. 1 except that the barrier became a 8 nm-thick InP. The observed highest PVR at 77 K was 1.9. This is less than a quarter of the previous PVR. One reason for the lower PVR is the non-homogeneous heterointerface. We could observe islands or valleys and these lateral sizes were several hundred nm in an observation of AFM image of the GaInAs surface. No parallel step was observed. On the other hand, an AFM image of the GaInAs surface by the previously used MOVPE showed parallel equidistant monolayer steps [14]. Neither an island nor a valley was observed. We think the observed islands and valleys made many parallel small RTDs with different well thicknesses, and this resulted in a lower PVR. The carrier gas flow speed in the reactor is different between the former apparatus and the present apparatus. The flow speed was 300 cm/s in the previous apparatus and 50 cm/s in the present apparatus. A faster speed flow provides a thinner stagnant layer resulting in enhancement of migration and the step-flow growth mode. Thus, increasing the flow speed to obtain atomically flat surface will provide a higher PVR.

5. Summary

GaInAs/AlAs/InP resonant tunneling diodes (RTDs) grown by MOVPE were studied. The peak current density dependence on barrier thickness was measured in three different barrier thicknesses (3.5, 5.3, and 7 nm). The range of peak current was from 100 to 0.1 A/cm^2. The results suggest that OMVPE has the possibility to provide a wider peak current range than that by MBE. From a measurement of the distribution of peak current density, the peak current density deviations became larger when the barrier became thicker. This fluctuation of peak current density can be explained by the thickness fluctuation of the barrier in the wafer's millimeter range. We have discussed the reasons for peak current density fluctuation and PVR degradation using millimeter range thickness fluctuation and surface AFM images.

Acknowledgements

We would like to thank Professors S. Arai, M. Asada, M. Watanabe, and Dr M. Suhara for fruitful discussion and T. Kojima, T. Arai, K. Sato, and H. Inazawa for technical support. This work was supported by the Ministry of Education, Science, Culture and Sports through a Scientific Grant-In-Aid, and by the 'Research for the Future' Program JSPS-RFTF96P00101 from the Japan Society for the Promotion of Science (JSPS).

References

[1] Broekaert TPE, Brar B, van der Wagt JPA, Seabaugh AC, Morris FJ, Moise TS, Beam III EA, Frazier GA. IEEE J Solid-St Circ 1998;33:1342.
[2] Maezawa K, Matsuzaki H, Yamamoto M, Otsuji T. IEEE Electron Dev Lett 1998;19:80.
[3] Chen KJ, Maezawa K, Yamamoto M. IEEE Electron Device Lett 1996;17:127.
[4] van der Wagt P, Seabaugh A, Beam E. III IEDM Tech Dig 1996;425.
[5] Enoki T, Yokoyama H, Umeda Y, Otsuji T. Int Conf on Solid State Devices and Materials, D10-1 1987.
[6] Inata T, Muto S, Nakata Y, Sasa S, Fujii T, Hiyamizu S. Jpn J Appl Phys 1987;26:L1332.
[7] Broekaert TPE, Fonstad CG. J Appl Phys 1990;68:4310.
[8] Keller BP, Yen JC, Denbaars SP, Mishra UK. Appl Phys Lett 1994;65:2159.
[9] Moise TS, Kao Y-C, Katz AJ, Broekaert TPE, Celii FG. J Appl Phys 1995;78:6305.
[10] Mattia JP, Brown ER, Calawa AR, Manfra MJ. Appl Phys Lett 1993;63:521.
[11] Seabaugh A, Lake R. Encyclopedia Appl Phys 1998;22:335.
[12] Sekiguchi T, Miyamoto Y, Furuya K. J Crystal Growth 1992;124:807.
[13] Oobo T, Takemura R, Suhara M, Miyamoto Y, Furuya K. Jpn J Appl Phys 1997;36:5079–80.
[14] Suhara M, Nagao C, Honji H, Miyamoto Y, Furuya K, Takemura R. J Cryst Growth 1997;179:18.
[15] Matthews JW. J Cryst Growth 1974;27:118.
[16] Huber JL, Chen J, McCormack JA, Zhou CW, Reed MA. IEEE Trans Electron Dev 1997;44:2149.
[17] Hou HQ, Chui HC, Choquette KD, Hammons BE, Breiland WG, Geib KM. IEEE Photon Tech Lett 1996;8:1285.

MICROWAVE/MILLIMETER WAVE DEVICES

PERGAMON

Solid-State Electronics 43 (1999) 1399–1403

SOLID-STATE ELECTRONICS

Advanced heterostructure transistor technologies for wireless communications

N-L. Larry Wang*, Barry Lin, Frank H-F. Chau, Gordon Jackson, Zhengming Chen, C.P. Lee [1]

EiC Corp., 45738 Northport Loop West, Fremont, CA 94538, USA

Abstract

Wireless communication has enjoyed tremendous growth in the last five years. Most of the market is below the 3 GHz. Recently, millimeter wave frequency band was also opened up to commercial applications, such as the Local Multipoint Distribution System. The rapid growth of the market demands cost effective RF circuitry with ever better performance. Thus, the heterostructure transistors are pursued to meeting the market needs. This article will first analyze the technical demand on RF transistor circuitry for wireless application. Existing and emerging transistor technologies will be discussed for its strength. A general comparison will be made. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

Wireless application started in the mid 80s. However, the cellular market did not really take off until the 90 s. The first generation standard, Advanced Mobile Phone System and other FM systems, also gave way to the second-generation digital standards, including Global System Mobile Phone, Personal Digital Cellular, IS54 Time Division Multiple Access, and IS95 Code Division Multiple Access. In recent years, the third generation standard was setup and the direction was the wideband CDMA to handle data as well.

In pace with the advance of the phone generation, the size, weight, and performance of the handheld phone are also constantly improved. The frequency of operation has moved up to the 2 GHz range of PCS application. In Industrial, Scientific, Medical application, it even goes up to the 5 GHz band. It poses a constant demand for higher performance RF components. When the wireless application started, the phone designer had to be satisfied with any available parts to cover the RF circuitry. Now designers are asking for much improved RF semiconductors to satisfy market demands. Therefore, heterostructure transistors are becoming common in the wireless market.

Below 3 GHz, GaAs HBT (Heterojunction Bipolar Transistor) and GaAs PHEMT (Pseudomorphic High Electron Mobility Transistor) are commonly used in the wireless RFICs, complementing the workhorse transistor technologies of Si bipolar, Si BiCMOS, and GaAs FET. The acceptance of the heterostructure transistors is ever increasing! As the market demands higher level integration, performance, and low price, the production related issues are as important as the RF performance in the value of the RFICs, and heterostructure transistors would be an attractive solution.

* Corresponding author.

[1] Sabbatical leave from National Chao-Tung University, Hsinchu, Taiwan.

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00080-5

Table 1
Commonly encountered wireless systems

800 to 1000 MHz	1700 to 2000 MHz	2400 MHz	Millimeter wave
AMPS	DCS1800	WLAN[b]	LMDS
GSM	DCS1900	WLL[d]	Auto-radar
IS-54 TDMA	CDMA		
IS-95 CDMA	W-CDMA		
PDC	TDMA		
ISM	PHS[a]		
	DECT[c]		

[a] PHS Personal Handyphone System.
[b] WLAN Wireless Local Area Network.
[c] DECT Digital European Cordless Telephone.
[d] WLL Wide Area Local Loop.

2. Wireless systems

There are many different wireless systems in use. Table 1 outlines several systems at different frequency bands. The first-generation phone system refers to the analog phone, such as AMPS, using the FM scheme. The second-generation phone refers to the digital modulation scheme, including GSM, IS-54 TDMA, IS-95 CDMA, and PDC. These systems have improved security over the first generation and more capacity. The next (third) generation will handle data as well as voice, and the preferred approach is wide-bandwidth CDMA (5 MHz channel bandwidth).

These systems have different frequency band, channel bandwidth, modulation scheme, power level, etc. As a result the architecture of the RF transceiver will be very different. Figs. 1 and 2 are the generic function block diagram of two commonly encountered digital modulation systems: CDMA and TDMA.

The commonly shared RFICs include LNA (low noise amplifier), Downconverter, variable gain amplifier, power amplifier, driver amplifier [1]. However, different systems have quite different specifications on each function block. For example, the CDMA requires much higher IIP3 (third-order input intercept point) in LNA than the TDMA's.

The RF transceiver architecture is also quite different. The CDMA system requires almost 80 dB transmit power control (versus less than 30 dB in the TDMA systems). This presents a great challenge in the RF circuit design due to the signal leakage among plastic package pins. Therefore, the gain control is implemented in the IF frequency, and the signal must be up-converted from IF to RF before amplification.

Different modulation schemes are used in the second generation phone system. GSM uses GMSK modulation which is a constant envelope signal. Therefore, the spectrum regrowth of the signal is not a concern. A $\pi/4$ differential quadrature phase shift keying is used in both IS-54 TDMA and PDC [2]; an offset quadrature phase shift keying is used in IS-95. Both schemes have amplitude and phase modulation and are sensitive to the spectrum regrowth caused by the circuit non-linearity (AM-to-AM and AM-to-PM distortion) [3]. Therefore, the circuits along the transmitter chain must operate in the backed-off state to guarantee a clean spectrum (low adjacent channel power).

In the transmitter chain, the key specifications include output power, ACPR (adjacent channel power rejection), spurs, rms vector errors. Therefore, the RFICs must maintain good linearity to keep ACPR low. In the receiver chain, the interference tolerance

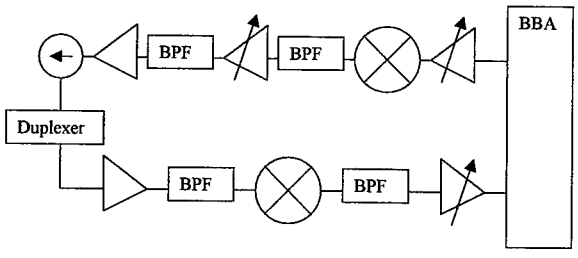


Fig. 1. A generic function block diagram of a CDMA RF transceiver.

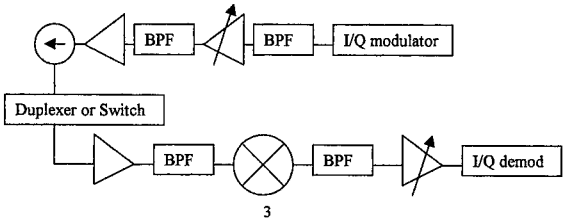


Fig. 2. A generic TDMA (GSM, IS-54, PDC) RF Transceiver block diagram.

level demands high IP3, and signal sensitivity demands low noise figure.

The TDMA systems run in the Time Division Multiplex mode: the TX is turned on while the RX is off, and vice versa. The circuit is only turned on over a portion of the full frame. This is essentially a pulse modulation. In AMPS and CDMA systems, TX and RX are turned on simultaneously. The interaction between the TX and RX, as well as the heating from the transmitter must be taken into design consideration.

With such a wide variety of circuits and different specifications, the RFIC design is a challenging task. The frequency band is also shifting to 2 GHz which further demands higher performance transistor technologies.

3. RFIC circuits

The RFIC can be classified into three categories: power amplifier, low noise amplifier, and analog-like RFICs.

3.1. Power amplifier

In AMPS, the saturated power is all one needs. In IS-54, IS-95, PDC systems, the amplifier must be linear enough to maintain the ACPR within specifications [3]. Since the peak power of the modulated signal is a few dB higher than the average power, the PA runs at a backed off situation to maintain the linearity. Thus, the efficiency of the linear PA is lower than the saturated power version.

To achieve high efficiency, the power transistor must be power matched. Therefore, the output matching circuit of the PA is designed with the reactive matching method.

The power amplifier must deliver the rated RF power; therefore, it is usually the most important component deciding the talk time in a handheld phone. As the operation voltage goes lower, the load resistance value is reduced. This situation presents a challenge to the low loss of the output matching circuit (due to the limited Q factor of the matching components), as well as the resistive power loss of the larger current flowing from battery to the PA. Contrary to the baseband VLSI circuit, PA prefers to run at higher voltage for better performance.

3.2. Low noise amplifier

As mentioned in the previous section, wireless systems are 'interference dominant'. The third-order intercept point is as important as the noise figure. For standby time consideration, the LNA should run at the lowest current. The noise figure can be maintained by scaling the LNA design with current and transistor size. But the IP3 demands a corresponding amount of dc power, and set the limit for the minimum amount of current consumption in LNA.

To achieve the lowest possible noise figure, the input impedance of the LNA must be correctly chosen and is designed with the reactive matching approach.

3.3. Other RFIC

Other function blocks, such as a down/up converter, variable gain amplifier, driver amplifiers, are often designed by a mixture of design approaches [4]. Within the semiconductor chip, the circuit is usually designed with the analog-like approach, i.e., with a transistor, resistor, and capacitor (no inductors). No impedance matching is used (which requires an inductor) on chip to save chip size. Outside the package, surface mount components are used to provide the reactive impedance match.

As a wireless system moves up in operation frequency, a transistor may lack the performance. An on-chip inductor becomes popular to provide impedance matching and recover the lost gain through the mismatch of analog-like approach. The negative effect of an on-chip inductor is its large size, the cross talk, and limited bandwidth.

All these components must meet gain, IP3, power consumption, and noise figure requirements. The higher performance transistor extends the specification envelope, and allows better performance at lower current.

4. Heterostructure transistor technologies

The most commonly used transistor technologies include: Si bipolar, Si BiCMOS, GaAs MESFET, GaAs PHEMT, and GaAs HBT. Table 2 compares the key transistor parameters.

Although all the transistors have good f_T and f_{max} for 2 GHz operation, the intrinsic property makes a difference. For example, the low BV_{ceo} limits the Si BJT and Si BiCMOS from power amplifier application. Table 3 outlines the most common applications for each transistor technologies.

4.1. Si bipolar transistor

Si BJT is mainly used in the small signal circuit, including LNA. Its high transconductance is most useful in analog-like circuit approach. The uniformity of the base-emitter junction voltage is also important for product yield and statistic control. However, most bipolar transistors have a limited performance at 2

Table 2
Comparison of transistor parameters[a]

	Si BJT	Si BiCMOS	GaAs FET	GaAs HBT
Min feature	0.5 μm	0.65 μm	0.5 μm	2 μm
f_T	26 GHz	20 GHz	30 GHz	40 GHz
f_{max}	47 GHz	27 GHz	65 GHz	65 GHz
Breakdown	4 V	4 V	12 V	12 V

[a] BV_{cco} is used in the breakdown voltage for a bipolar transistor.

Table 3
Application of transistor technologies

	Si BJT	Si BiCMOS	GaAs FET	GaAs HBT
Power amplifier	No	No	Yes	Yes
Low noise	Yes	Yes	Yes	Yes
Other RFICs	Yes	Yes	Some	Yes

GHz band and an on-chip inductor becomes a necessity.

4.2. Si BiCMOS

Si BiCMOS has the addition of the CMOS which can be used as a switch and other bias circuitry in the RFIC, complementing the bipolar circuit. The tradeoff is the complexity of the process. With the CMOS, the baseband circuit can be implemented as well.

Can the RF circuit be integrated with the baseband circuit using BiCMOS? From a technology view point, certainly yes. The technical challenge lies in the fundamental difference between the analog and the digital circuit. The digital circuit switching noise can be easily coupled into the RF circuit which has an extremely wide dynamic range to handle. A typical receiver sensitivity is in the −110 dBm range. Until the in-package cross-coupling is controlled, the integration is limited to systems that have loose requirements.

4.3. GaAs FET

GaAs FET is widely used in power amplifiers for its performance. There are two main deficiencies: the gate bias often requires a negative voltage, and a *p*-channel FET along the drain bias is needed for power down. These two deficiencies add to the cost and PCB area.

GaAs FET has lower noise than the Si Bipolar transistor, with the advantage most noticeable at 2 GHz band. With its high performance, GaAs FET can also be used in the analog-like RFICs. However, the lower transconductance and the fluctuation of the threshold voltage limit the yield and usage in such a role.

4.4. GaAs HBT

GaAs HBT is now widely used in power amplifiers, low noise amplifiers, and many other RFICs. The key improvement of GaAs HBT comes from two features: low base sheet resistance and semi-insulating substrate [5]. The low base sheet resistance (300 Ω/sq) allows wide emitter fingers without the current crowding effect: 2 μm emitter finger width makes the processing easy and the uniformity high. The semi-insulating substrate eliminates the substrate capacitance which is a limiting factor in the Si bipolar circuit performance [6].

In the power amplifier application, the HBT does not have the two deficiencies of GaAs FET. In the low noise amplifier case, HBT is similar to Si BJT in noise figure but higher gain is due to higher f_{max}. Since HBT is a bipolar transistor, the Si BJT circuit schematic can also be implemented with HBT. However, GaAs HBT has a higher Vbe of 1.35 V, which is higher than the 0.8 to 0.9 V of Si BJT. In the low voltage (3 V) circuit, a GaAs HBT circuit needs to be finely adjusted for dc level.

4.5. GaAs PHEMT

Pseudomorphic HEMT offers an improvement of performance over GaAs FET. From the user's view point, there is no difference between PHEMT and MESFET except the performance. PHEMT also inherits the MESFET's deficiencies in the PA application. Some PHEMT are designed in the enhancement mode to eliminate the need of negative gate bias. This presents a challenge in the gate etch process control.

PHEMT with short gatelength (0.25 μm or shorter) offers the highest f_T and f_{max}. It is the obvious candidate for millimeter wave application, such as LMDS. The MMIC format is often used to avoid the hybrid assembly difficulty at such a high frequency. PHEMT also offers the lowest noise figure. The low noise figure of PHEMT makes it the number one choice for a satellite communication receiver. In the wireless application, the PHEMT low noise advantage is often smeared by the loss of duplexer which precedes LNA in the receiver chain.

4.6. Si CMOS on Sapphire

SOS CMOS has become commercially available in recent years. The claimed advantage is the simplified process (compared with standard CMOS) and improved performance from the non-conducting Sapphire substrate. The SOS version is different from the SOI: SOI has only 1 μm or so oxide to separate the transistor from the conducting substrate, where SOS has a non-conducting substrate. At RF frequency, the SOS approach guarantees the elimination of the substrate signal leakage.

Is SOS CMOS an improvement more attractive for VLSI application, or for RFIC, or for both? It will be decided by cost, performance, and production yield.

4.7. SiGe HBT

SiGe HBT is a heterojunction bipolar transistor on Si substrate. The base material is SiGe. Like GaAs HBT, it can have a low base sheet resistance. However, most reported results still have a sheet resistance of 7 to 9 kΩ/sq [7]. The transistor characteristic is very similar to the Si BJT. Therefore, SiGe HBT can be considered as an improved Si BJT. In applications, SiGe HBT also is subject to the same constraints of Si BJT.

4.8. Fundamental difference between Si and GaAs

Beyond the transistor technology difference, the substrate also plays an important role in the RFIC. Si substrate is conducting, giving rise to the parasitic substrate capacitance to the transistor which reduces the performance of the circuit. It also provides a path of RF signal leakage. GaAs and Sapphire are semi- or non-conducting, which allows the full exploit of the transistor performance.

5. Trend in the below-3 GHz wireless market

New systems in the 2 to 3 GHz range demands more performance out of the transistor. As the second generation phone is popular, and moving toward the third generation phone, the demanding system specification asks for high performance. High level integration of RF function blocks is a solution to the price, size and part count reduction [8]. This approach requires mature technology in production and sufficient transistor performance margin. All the trends favor a higher performance transistor, therefore, more and more RFICs were made with heterostructure transistors in recent years.

6. Conclusion

Heterostructure transistors are an improvement over existing homostructure transistors. Although the device physics is very different, the users see only an evolution of performance.

In this paper, an objective comparison of the heterostructure transistors is attempted. However, in the industry each technology is supported by a handful of companies only. Therefore, business execution to explore the full technical advantage of each transistor technology is more important than the perceived potential of each technology. Even so, we can be sure that heterostructure transistors will be much more widely used in the wireless market to answer the price, performance, and production demands.

Acknowledgements

The authors would like to thank colleagues at EiC Corp., as well as many others in the industry who bring the heterostructure transistors from the R&D into the real applications!

References

[1] Razavi B. Recent advances in RF integrated circuits. IEEE Communications Magazine. 1997. p. 36–43.
[2] Ariyavisitakul S, Liu TP. Characterizing the effects of nonlinear amplifiers on linear modulation for digital portable radio communications. IEEE Trans Vehicular Tech 1990;39(4):383–9.
[3] Struble W, McGrath F, Harrington K, Nagle P. Understanding linearity in wireless communication amplifiers. IEEE J Solid State Circuits 1997;32(9):1310–8.
[4] Meyer RG, Mack WD, Hageraats J. A 2.5 GHz BiCMOS transceiver for wireless LAN's. IEEE J Solid State Circuits 1997;32(12):2097–104.
[5] Wang NL. Transistor technologies for RFICs in wireless applications. Microwave J. 1998. p. 98–110.
[6] Gray PR, Meyer RB. Analysis and Design of Analog Integrated Circuits. John Wiley and Sons, 1977.
[7] Kovacic S Performance Consideration of Si, SiGe and GaAs Bipolar Technologies. White paper by SiGe Microsystems Inc., www.sige.com.
[8] Antes T, Conkling C. In: RF Chip Set Fits Multimode Cellular/PCS Handsets. MW&RF, 1996. p. 177–86.

PERGAMON

Solid-State Electronics 43 (1999) 1405–1411

SOLID-STATE ELECTRONICS

A novel GaAs flip-chip power MODFET with high gain and efficiency

Tsuyoshi Tanaka*, Hidetoshi Furukawa, Haruto Nagata, Daisuke Ueda

Electronics Research Laboratory, Matsushita Electronics Corporation, 3-1-1 Yagumo-Nakamachi, Moriguchi, Osaka, 570-8501, Japan

Abstract

A novel GaAs flip-chip power MODFET with a spike-gate structure has been developed for high-gain and high-efficiency operation under extremely low supply voltages. The spike-gate MODFET is featured by a unique gate structure that has almost zero effective gate length. The spike gate provides both a low on-resistance of 1.5 Ω mm and a high transconductance of 280 mS/mm without increasing the drain conductance. The flip-chip bonding, with an Au–Sn alloy system, has been used for mounting of the spike-gate power MODFET on an AlN substrate. The thermal resistance of the flip-chip power MODFET is 12°C/W which is lower than that of a conventional wire-bonded MODFET. The power-added efficiency (PAE) of 71% at the output power of 30.0 dBm can be obtained even when the supply voltage is reduced to 1.5 V. By eliminating parasitic wire inductance, the linear gain of the MODFET is 2 dB higher and the PAE is 5% higher than that of the conventional wire-bonded MODFET. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

GaAs power FETs have been widely used as key devices in mobile communication systems. In these systems, both low-voltage and high-efficiency operations are of major concern for the power FETs because their performances give long battery life and small volume of the hand sets. However, it is difficult to maintain high output power and high efficiency as the supply voltage is reduced, because the available drain voltage swing is narrowed. This narrowed operating area is determined by the knee voltage of GaAs FETs which is determined by the on-resistance.

Recently, several authors have reported on GaAs MESFET [1–3] and heterojunction FETs [4–6] operating under low voltage while maintaining high output power. We have reported on GaAs power MESFETs with the spike-gate structure for both analog and digital cellular phones [7–9]. These FETs exhibited high PAE under an extremely low supply voltage of 1.5 V. It is expected that the spike-gate structure will have a great potential for extension to double-doped MODFETs.

Moreover, how we can get high power gain under the low voltage operation is also a crucial problem. There have been few reports of high power gain performance obtained under low supply voltage while maintaining high output power and efficiency. Flip-chip technology [10–13] has the advantage of high gain in the high-frequency characteristics. Since the FET operates free from parasitic inductances caused by bonding wires, the negative feedback can be eliminated. However, in flip-chip bonding, it is difficult to transfer the heat from the channel to the heat sink. Since SiN passivation has poor thermal conductivity, the heat path from the channel becomes narrower than that of conventional wire bonding. Thus, the heat gen-

* Corresponding author. Tel.: +81-6-6906-4922; fax: +81-6-6907-2640.

E-mail address: tanaka@erl.mec.mei.co.jp (T. Tanaka)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00081-7

erated at the channel is confined around the gate causing thermal runaway. In order to harness the superior characteristics of flip-chip FETs, it becomes necessary to cope with the extreme brittleness as well as low thermal conductivity of GaAs and SiN. These thermophysical properties become particularly important in GaAs flip-chip bonded power devices.

In this paper, we demonstrate a flip-chip bonded GaAs power MODFET which achieves high gain and efficiency by using the spike-gate structure. Because the depletion widening from the fringe of the spike gate alleviates the electric field between the gate and the drain, this structure can suppress the increase of the drain conductance which reduces the power gain and thereby the PAE. The present power MODFET achieved the extremely low on-resistance by using the unique gate structure where the effective gate length was almost zero. The attained on-resistance of 1.5 Ω mm was less than a half of that without spike. The obtained thermal resistance of the flip-chip power FET is 12°C/W which is even lower than that of conventional wire-bonded one. Fabricated GaAs MODFET showed the output power of 30.0 dBm with 71% PAE at the drain bias of 1.5 V. By eliminating parasitic wire inductance, the linear gain of the MODFET is improved 2 dB resulting in increasing PAE of 5% compared to that of conventional wire-bonded one.

2. Device structure

A schematic cross section of the spike-gate MODFET is shown in Fig. 1. The epitaxial structure is based on a pseudomorphic double-doped heterostructure, which consists of a 15 nm undoped $In_{0.2}Ga_{0.8}As$ channel layer sandwiched between a 10 nm thick upper Si-doped $Al_{0.2}Ga_{0.8}As$ layer with a donor concentration of 6×10^{18} cm^{-3} and a 5 nm thick lower Si-doped $Al_{0.2}Ga_{0.8}As$ layer with 3×10^{18} cm^{-3}. The surface undoped layer plays a role in reducing the gate capacitance and reducing the gate leakage, which contribute to superior power performance.

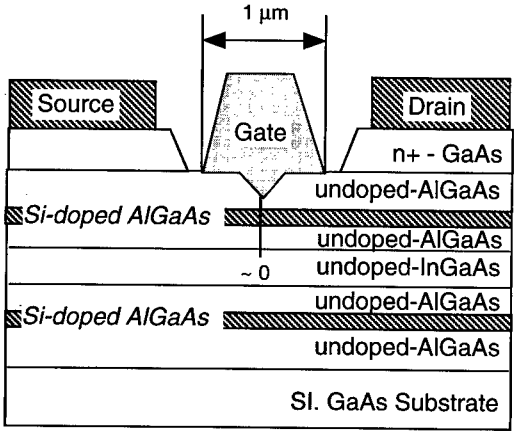


Fig. 1. Schematic cross section of the spike-gate power MODFET.

It is difficult to maintain high efficiency as the supply voltage is reduced, because the available drain voltage swing is narrowed. The operating area is determined by the knee voltage which is determined by the on-resistance. In order to reduce the on-resistance, it is important to reduce parasitic resistances such as contact, source/drain and channel resistances. In particular, reduction of the channel resistance is most effective to reduce the on-resistance. The short channel length makes it possible to reduce the channel resistance. However, shorter channel results in higher drain conductance due to the short channel effect. Higher drain conductance in turn lower power gain and lower efficiency. In order to achieve both low on-resistance and low drain conductance, we proposed the spike-gate structure [7–9].

The spike-gate MODFET consists of two gate regions. One is the extremely short gate length region formed on the spike recess, the other is the fringing gate region. The spike gate gives the lowest channel resistance, because the effective gate length is almost zero. While the fringing gate suppresses the increase of drain conductance.

3. Device fabrication

The process flow of the spike-gate MODFET is shown in Fig. 2. Device processing began with standard mesa formation to define the active region, followed by a dummy gate formation. A 0.2 μm wide dummy gate was defined by the pattern-edge-line method [14]. Then the dummy gate was ashed in O_2 plasma to reduce the pattern width to 0.15 μm. The dummy gate defines the footprint of the gates (Fig. 2a). The obtained variance of 3σ of the line width was 0.023 μm. The cause of this variance is the etch rate difference of the resist in a wafer and the difference between wafers. This value is sufficient for the large-volume production. The dummy gate line was inverted into the gate footprint space by SiO_2 deposition followed by the lift-off. The SiO_2 deposition was performed at room temperature by electron-beam evaporation. The SiO_2 layer is 400 nm thick. Conventional Au/AuGeNi metallurgy was used for source and drain electrodes which were defined by lift off using the SiO_2 as the spacer and were subsequently alloyed with GaAs (Fig. 2b). The first narrow gate recess was formed by wet etching in a mixture of tartaric acid and hydrogen peroxide (Fig. 2c). This etch gives (111)A and (111)B sidewalls with low etching

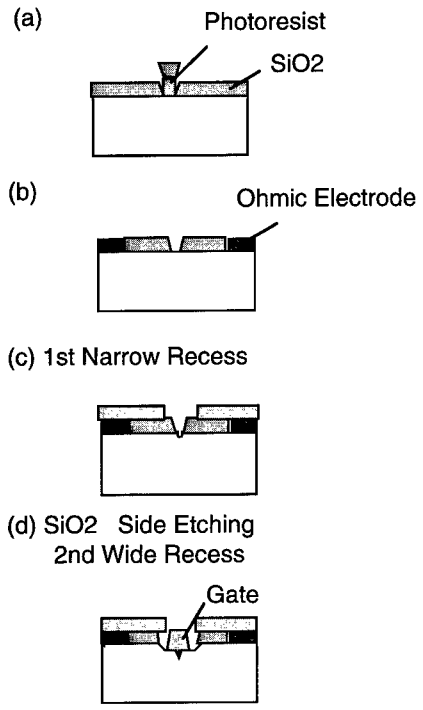


Fig. 2. Process flow of the spike-gate power MODFET.

rate. After side etching of SiO_2, a second wide etching was performed which gives a spike-gate structure. If the etchant is chosen with preference, it is possible to etch only the bottom of the recess selectively. This is because the insoluble (111)A and (111)B sidewalls become a barrier to the second etching. Thus, the spike shaped recess was formed by the second etching as shown in Fig. 2d. The Al/Ti gate electrode was formed by lift off. After interconnecting of the gate fingers by Au electro-plating, the MODFET was passivated using 1.0 μm-thick SiN film for high reliability.

Fig. 3a shows the cross-sectional SEM micrograph of the photoresist pattern defined by pattern-edge-line method. The line width of the dummy gate was 0.15 μm and the height was approximately 1.0 μm which was sufficiently high for lift off. Fig. 3b shows the cross-sectional SEM micrograph of the fabricated spike gate. The tip of the spike formed by (111)A and (111)B sidewalls was 80 nm long and the base of the spike was 100 nm wide. The upper gate was 1.0 μm long and the effective gate length is close to zero. Because the insoluble (111)A and (111)B sidewalls become a barrier to spike gate formation, the length and width of the spike is mainly determined by the footprint of the controllable dummy gate. For this reason, the spike-gate formation is quite reproducible.

Conventional interdigitated source-gate-drain pattern was used for flip-chip bonding of the MODFET. After MODFET was passivated, all source fingers were connected over the active area by a thermal bump as shown in Fig. 4. To reduce the drain-gate feedback capacitance, we use thick SiN passivation. The total thickness of the passivation over the gate metal was 1.5 μm. Au electro-plating was used as the thermal bump by using 50 μm photoresist process. The height of the bumps are 30 μm.

The GaAs chip was mounted on an AlN substrate patterned with Sn/Au/Cu. Au–Sn was alloyed at 280°C under a 200 g load. The reason for using Au–Sn alloy bonding is because GaAs is too brittle to tolerate heavy pressure or ultrasonic vibration. Au–Sn alloy bonding

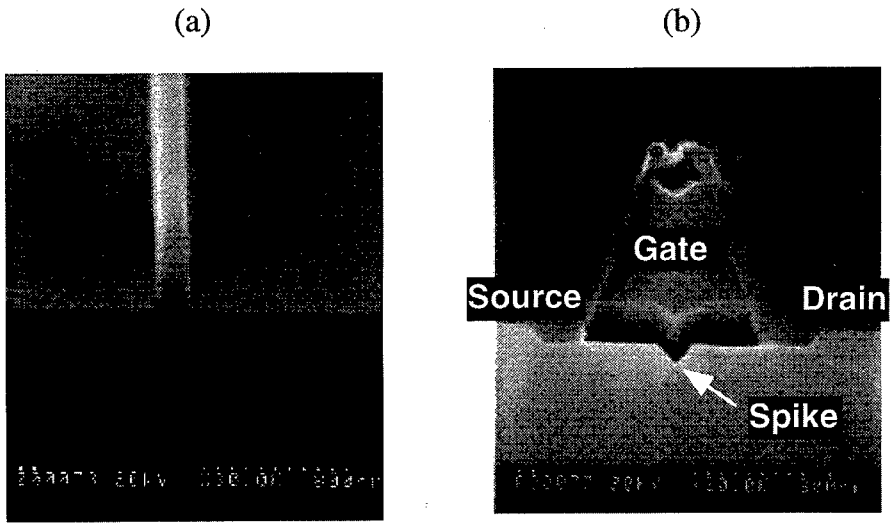


Fig. 3. (a) SEM cross section of the dummy gate pattern defined by PEL; (b) SEM cross section of the spike gate.

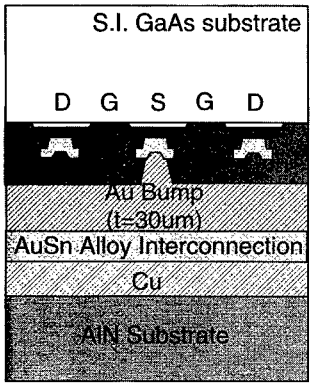


Fig. 4. Schematic cross section of the flip-chip spike-gate power MODFET.

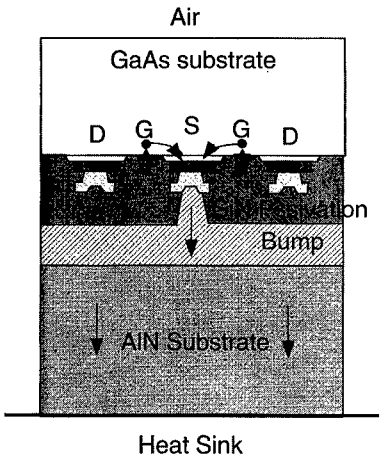


Fig. 6. Schematic illustration of the heat transfer path in the flip-chip power MODFET.

also has an advantage of relieving the mechanical stress and absorbing the substructure roughness. Fig. 5 shows the cross-sectional SEM micrograph of a flip-chip mounted power MODFET. There is usually no void in the alloy or crack in the GaAs chip. X-ray-micro-analysis indicates that Au is uniformly distributed between Au bump and Cu of the substrate, implying all Au–Sn is in the 80/20 eutectic phase.

4. Device performance

4.1. Thermal properties

The essential elements of the flip-chip bonding scheme are: (a) the bonding metallurgy and technique; (b) the substrate; and (c) the bumps. In designing the novel flip-chip interconnection, these elements were considered with a view of minimizing the thermal resistance. In flip-chip bonding, it is difficult to transfer the heat from the channel to the heat sink. Fig. 6 illustrates the heat flow path in a flip-chip power MODFET. Since SiN passivation has poor thermal conductivity, the flow path from the channel is narrower than that of conventional wire-bonding. So, the heat generated at the channel is confined around the gate, causing thermal runaway. Fig. 7 shows a simulated temperature contour map of the flip-chip power MODFET. The heat generated at the channel is transferred only via the source contact. And heat confinement is occurred around drain. To reduce the thermal resistance, there are mainly two points to consider. One is the heat transfer through substrate, the other is the finger pitch of the power MODFET, i.e., the spacing between heat sources. Fig. 8 shows the simulated channel temperature shift as a function of the finger pitch and the AlN thickness.

The channel temperature increases drastically as the finger pitch narrows. Moreover, a thin AlN substrate can be useful to suppress the temperature shift. It is noted that if we choose the thickness of the AlN to be 250 μm, the channel temperature can be lower than that of conventional wire bonding. The measured temperature shift of the flip-chip MODFET is compared with that of the wire-bonded one in Fig. 9. The thickness of GaAs substrate was 150 μm for both devices. This shows that a thermal resistance as low as 12°C/W can be achieved by flip-chip bonding with a thin AlN substrate.

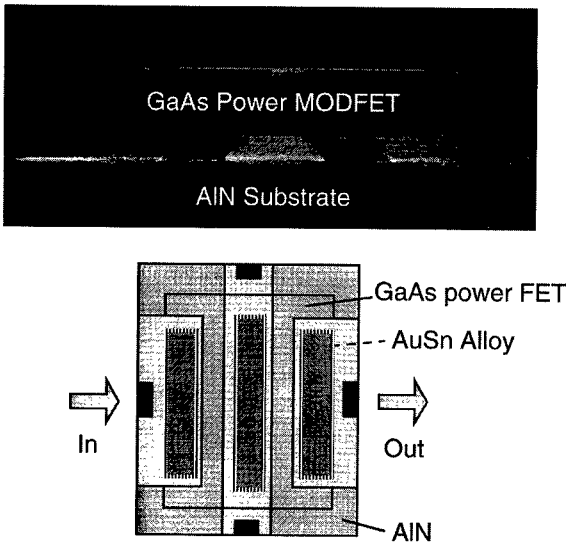


Fig. 5. Cross-sectional SEM of the flip-chip power MODFET.

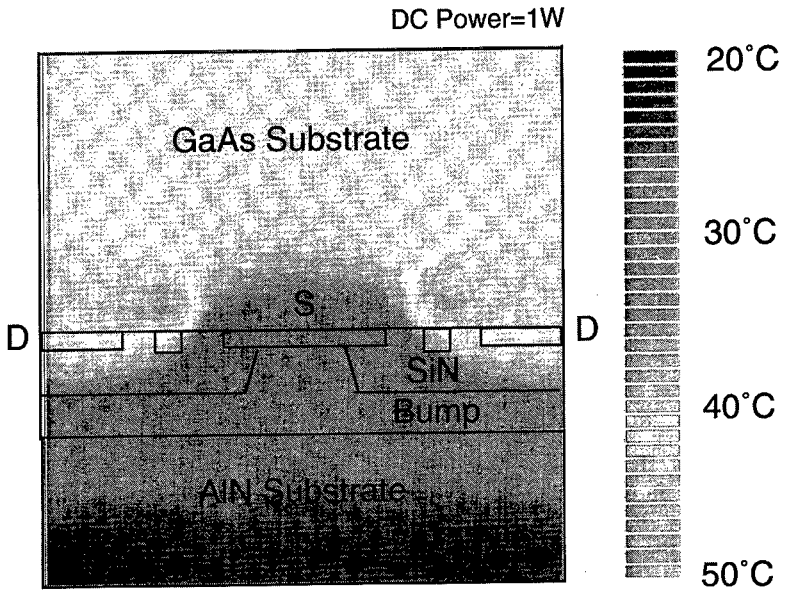


Fig. 7. Simulated temperature contour map of the flip-chip power MODFET.

4.2. DC performance

Fig. 10 shows typical DC current–voltage characteristics of a power MODFET (W_g=200 μm) with and without the spike gate. The gate length of the device without spike was 0.6 μm with a single recess. The fabricated devices, with and without the spike gate, have pinch-off voltages, the maximum drain currents (Imax) of about −1.1 V, 600 mA/mm and −0.8 V, 450 mA/mm, respectively. A lower knee voltage is seen for the device with the spike gate. The on-resistance of 1.5 Ω mm of the device with the spike gate is almost half of that without it. Owing to the fringing upper gate, the drain conductance is almost the same in spite of zero effective-gate length. The gate-drain breakdown voltage (BV_{gd}), measured at a gate current density of 100 μA/mm, was 12 V. This value is sufficiently high enough for low voltage power operation less than 1.5 V. This high value is due to the fringing gate structure, because the depletion by the fringing gate alleviates the electric field between the gate and drain. Another reason of this higher breakdown voltage is the undoped surface layer which separates the surface and the channel.

4.3. RF performance

Fig. 11 shows the S-parameters of the flip-chip MODFET vs that of the conventional wire-bonded one. The S-parameters were measured under a DC supply-voltage of 1.5 V and the bias current of 100

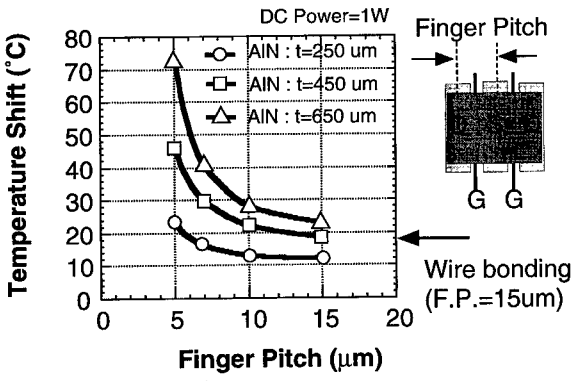


Fig. 8. Simulated channel temperature shift as a function of the finger pitch and the AlN thickness.

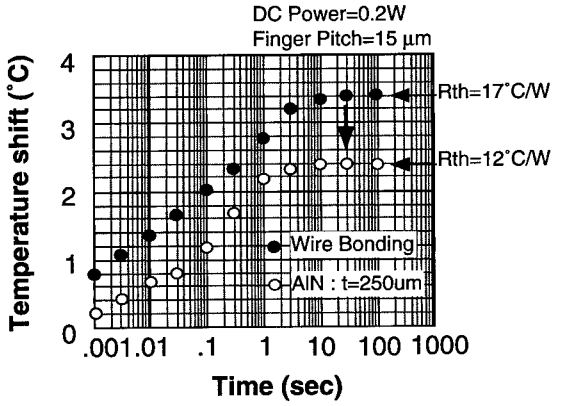


Fig. 9. Measured channel temperature shift of the flip-chip power MODFET vs that of the wire-bonded one.

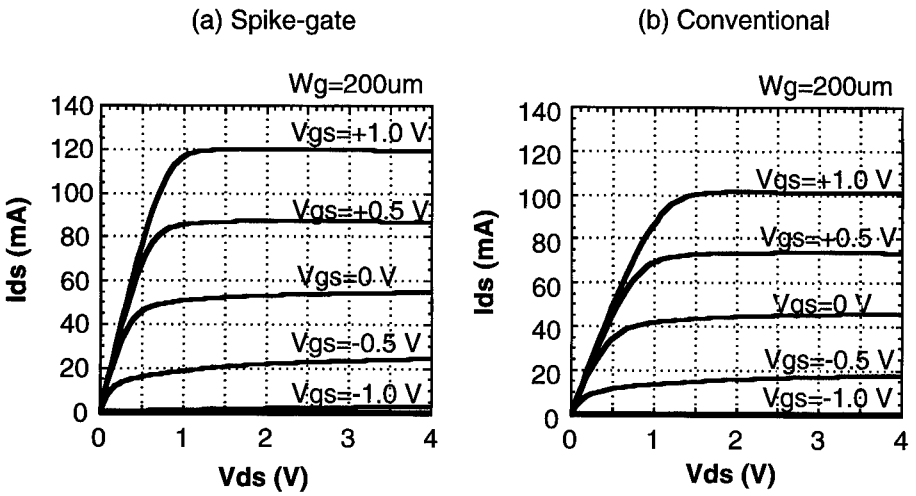


Fig. 10. Typical DC current-voltage characteristics of MODFETs (a) with and (b) without the spike-gate.

mA matched for 50 Ω. The gate width of the device was 20 mm. The gain of the flip-chip MODFET is 2 dB higher. This result indicates that the bump interconnection of the present flip-chip power MODFET is successful in reducing the ground inductance. On the other hand, the isolation of the flip-chip MODFET is almost the same as that of a wire-bonded one. This means that the thick passivation effectively suppresses the gate-drain feedback capacitance. It is noted that the obtained rather low gain is caused by the mismatch of the impedance, because the power device has quite low impedance compared to 50 Ω.

The power performance of the device was measured by using an automated tuner system at 900 MHz under a DC supply-voltage of 1.5 V. The device under test has a total gate width of 30 mm with a unit gate width of 250 μm. The device was operated under Class AB condition with a quiescent bias current of 0.8 A. The resultant output power and power-added efficiency (PAE) are shown in Fig. 12 as a function of the input power for flip-chip bonded and wire bonded MODFETs. The input and output circuits were matched for the maximum efficiency. The maximum PAE of 71% was obtained even when the supply voltage was reduced to 1.5 V. This high performance with low supply voltage is due to the unique spike-gate structure. By matched for the device impedance, linear power was measured to be 16.0 dB. By eliminating parasitic wire inductance, the linear gain of the MODFET is improved by 2 dB resulting in a 5% higher PAE than that of the conventional wire-bonded one.

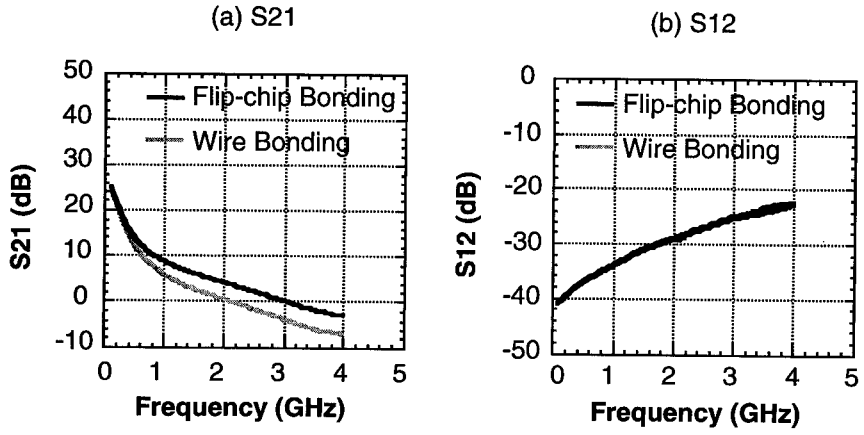


Fig. 11. S-parameters of the flip-chip power MODFET vs that of the conventional wire-bonded one: (a) S21; (b) S12.

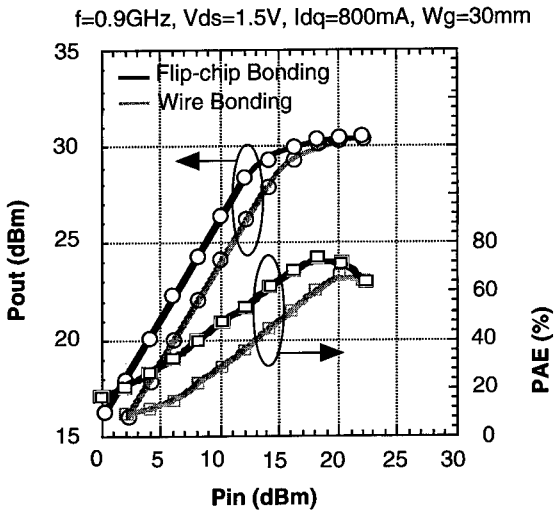


Fig. 12. Output power and power-added efficiency as a function of input power at 900 MHz.

5. Conclusions

A GaAs flip-chip power MODFET with a spike-gate structure has been developed for low-voltage operation over 70% PAE and high gain. The present power MODFET achieved a low on-resistance of 1.5 Ω mm by using the unique gate structure for which the effective gate length was almost zero. The on-resistance was almost half of that without spike. The flip-chip bonded spike-gate MODFET shows a low thermal resistance of 12°C/W. A 20 mm device gave 30.0 dBm output power with a PAE of 71% and a drain bias 1.5 V. The linear power gain was measured to be 16.0 dB. By eliminating parasitic wire inductance, the linear gain is 2 dB higher and the PAE is 5% higher than that of conventional wire-bonded MODFETs. These performance improvements were attributed to a device structure with minimum parasitics due to spike gate and flip-chip bonding. The present flip-chip MODFET has a great potential to be used in a variety of mobile communication handy sets.

Acknowledgements

The authors would like to thank G. Kano and M. Kazumura for their support of this work. We would like to thank also K. Kanazawa, K. Tateoka and K. Miyatsuji for their technical discussions.

References

[1] Ngo D, Beckwith D, O'Neil P, Camilleri N IEEE MTT-S Digest 1993 1461–64.
[2] Tateoka K, Sugimura A, Furukawa H, Yuri M, Yoshikawa N, Kanazawa K IEEE MTT-S Digest 1994 569–72.
[3] Lee J-L, Mun JK, Kim H, Lee JJ, Park H-M, Park S-C GaAs IC Symposium Tech Dig, 1994. p. 133–6.
[4] Iwata N, Inosako K, Kuzuhara M IEEE MTT-S Digest 1993 1465–8.
[5] Smith P, Lester L, Chao P, Lee B, Smith R, Ballingall J, Duh K IEDM Tech Dig 1987 854–6.
[6] Iwata N, Yamaguchi K, Kuzuhara M. Solid-State Electronics 1997;41:1587–90.
[7] Tanaka T, Furukawa H, Takenaka H, Ueda T, Noma A, Fukui T, Tateoka K, Ueda D IEDM Tech Dig 1995 181–4.
[8] Furukawa H, Tanaka T, Takenaka H, Ueda T, Fukui T, Ueda D. Solid-State Electronics 1997;41:1599–604.
[9] Tanaka T, Furukawa H, Takenaka H, Ueda T, Fukui T, Ueda D. IEEE Trans Electron Devices 1997;ED-44:354–9.
[10] Hang J et al. Digest Paper of IEEE MTT-S International Microwave Symposium, 1994. p. 1711–14.
[11] Krems T et al. Digest Paper of IEEE MTT-S International Microwave Symposium, 1996. p. 247–50.
[12] Sturdivant R et al. Digest Paper of IEEE MTT-S International Microwave Symposium, 1997. p. 505–7.
[13] Feng Z et al. Digest Paper of IEEE MTT-S International Microwave Symposium, 1998. p. 1823–6.
[14] Takenaka H, Ueda D. IEEE Trans Electron Devices 1996;ED-33:238–44.

PERGAMON

Solid-State Electronics 43 (1999) 1413–1417

SOLID-STATE ELECTRONICS

A low power dissipation 0.4 ~ 7 GHz transimpedance amplifier IC for SCM optical communication system

Kazuhisa Fujimoto[a,*], Hiroyuki Masato[a], Katsuhiko Kawashima[a], Mitsuru Nishitsuji[a], Masato Takata[b], Yoshiharu Kudo[c], Osamu Ishikawa[a]

[a]*Electronics Research Laboratory, Matsushita Electronics Corporation, 3-1-1 Yagumo-Nakamachi, Moriguchi, Osaka, 570-8501, 2-1-45 Hikoso-cho, Kanazawa-city, Ishikawa, 920-0901, Japan*

[b]*Matsushita Communication Kanazawa R&D Laboratories Co Ltd, 2-1-45 Hikoso-cho, Kanazawa-city, Ishikawa, 920-0901, Japan*

[c]*Osaka Development Center, Matsushita Communication Ind Co Ltd, 1006, Kadoma, Kadoma-city, Osaka, 570-8501, Japan*

Abstract

A low power dissipation and broadband transimpedance amplifier IC has been developed for sub-carrier multiplexing (SCM) optical communication systems such as a multi-channel video signal transmission system. A 0.25 μm pseudomorphic double-heterojunction modulation-doped FET (MODFET) is adopted for the IC to achieve the required bandwidth and noise performance. Specifically, transimpedance gain of 52 dBΩ with equivalent input noise current of 12 pA/√Hz has been obtained between 0.4–7 GHz. In order to minimize the group delay (GD) deviation, a peaking control technique is investigated. We also adopted the flip-chip bonding method to reduce the parasitic elements. As a result, group delay deviation of less than 30 ps has been obtained. A novel $SrTiO_3$ (STO) capacitor process technology is incorporated to realize low power dissipation. By using integrated STO capacitors for DC blocking, single power-supply voltage of 5 V and power dissipation of 300 mW are realized. This power dissipation is less than one half of that required by conventional transimpedance amplifiers. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Recently, demands for high speed and broadband optical communication systems have increased. In particular, sub-carrier multiplexing (SCM) optical communication systems have been generally adopted for multi-channel video signal transmission such as cable television distribution. A super wideband frequency modulation (FM) scheme, which converts multi-channel AM signals to a FM signal, is very attractive for video signal distribution [1].

For the transimpedance amplifier, broadband and low noise performances are required with low power dissipation. In addition, a super wideband frequency modulation scheme requires flat gain and small group delay deviation to achieve extreme low distortion characteristics [2]. Then, we adopted a 0.25 μm pseudomorphic double-heterojunction modulation-doped FET (MODFET) to obtain required bandwidth and low noise with its high gain and low total input capacitance. In order to minimize the group delay deviation, a peaking control technique is investigated. The novel STO capacitor process technology is also incorporated to realize low power dissipation.

* Corresponding author. Tel.: +81-6-6906-4923; fax: +81-6-6907-2640.

E-mail address: fujimoto@oerl.src.mei.co.jp (K. Fujimoto)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00082-9

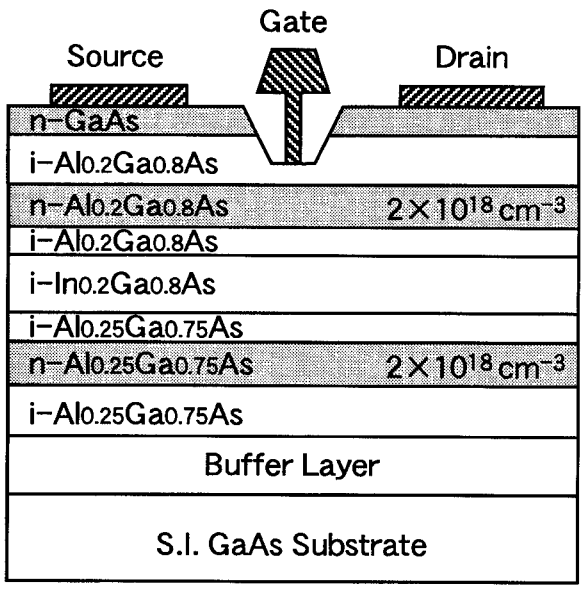


Fig. 1. Cross-sectional view of the 0.25 μm MODFET.

In this paper, we describe device structure, a novel $SrTiO_3$ (STO) capacitor process technology, circuit design and performance of the transimpedance amplifier IC.

2. Device structure

Fig. 1 shows the cross-sectional view of the 0.25 μm MODFET [3,4]. Double-doped double-heterojunction structure is adopted in order to obtain high gain and low noise with low power dissipation. The active part consists of an undoped InGaAs channel layer sandwiched between an upper *n*-AlGaAs layer and a lower *n*-AlGaAs layer. An *i*-AlGaAs layer is inserted as a schottky barrier between *n*-GaAs and *n*-AlGaAs layer to improve breakdown voltage. The gate length (L_g) of the MODFET is 0.25 μm with mushroom-shaped gate structure.

Fig. 2 shows typical transfer characteristics of the 0.25 μm MODFET. The threshold voltage (V_{th}) is chosen to be −0.6 V. The maximum transconductance (gm_{max}) and *k*-value are 450 mS/mm and 510 mS/Vmm, respectively. The breakdown voltage is −8 V. The current gain cutoff frequency (f_t) and maximum frequency of oscillation (f_{max}) are 42 GHz and 110 GHz, respectively.

3. STO capacitor process technology

Fig. 3 shows the cross-sectional view of the IC. We have developed high dielectric-constant STO thin-film

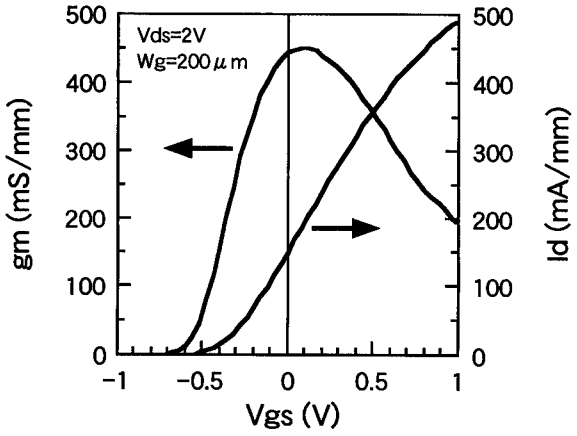


Fig. 2. Typical transfer characteristics of the 0.25 μm MODFET.

capacitor on the GaAs epitaxial substrate by using a low-temperature RF sputtering method [5–7]. RF sputtering is performed at 200°C to avoid degradation of the elaborate heterojunction or ohmic contact. The STO film has a greater than 100 relative dielectric constant (ϵ_r) and a less than 10^{-6} A/cm^2 leakage current density at 1 MV/cm. This relative dielectric constant is more than 20 times higher than that of a conventional SiNx film ($\epsilon_r \sim 6.5$). By using this STO capacitor for bypassing or DC blocking, small-size and low-cost microwave ICs can be realized. Fig. 4 shows that the STO capacitor can be used for a bandwidth of over 20 GHz. Insertion loss is as low as 0.17 dB at 10 GHz for a 100 pF capacitor.

4. Design and fabrication

4.1. Circuit design

Fig. 5 shows the proposed circuit diagram of the transimpedance amplifier IC. There is always a compromise between bandwidth and noise performance. A large-value feedback resistor (R_f) is desired for low noise performance. However, with a large feedback

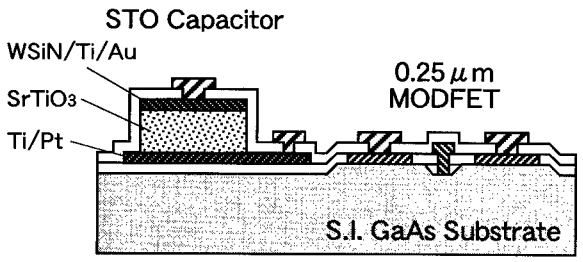


Fig. 3. Cross-sectional view of the fabricated IC.

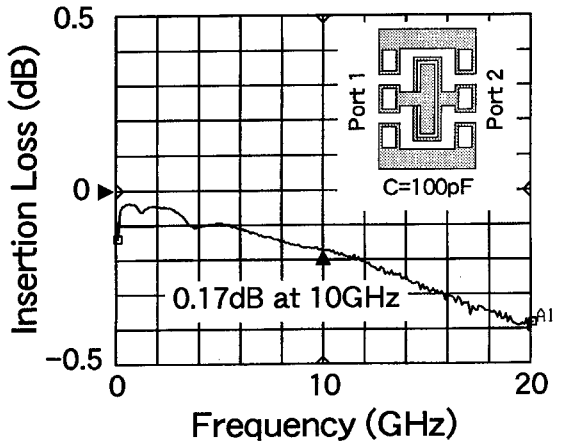


Fig. 4. Insertion loss (S21) of the STO Capacitor.

resistor, the bandwidth is limited. Then, in order to achieve the required bandwidth and low noise performance, higher gain and low input capacitance are required. For this reason, we adopted the 0.25 μm MODFET with a 270 Ω feedback resistance. The output impedance is matched to 50 Ω.

4.2. Group delay

The peaking technique that produces a high-frequency peak is popular in order to increase the bandwidth. However, too much peaking causes large group-delay deviation at nearby frequencies. Therefore, a peaking control technique is needed to keep the group delay constant. In order to control peaking, the gm R_L product (RL: load resistance) of the first-stage FET and parasitic elements such as drain and/or source series inductance are investigated. Fig. 6 shows sample frequency dependencies of transimpedance gain (Z_t) and group delay with various load resistances. Larger load resistance causes lower-frequency gain peaking and larger group delay deviation.

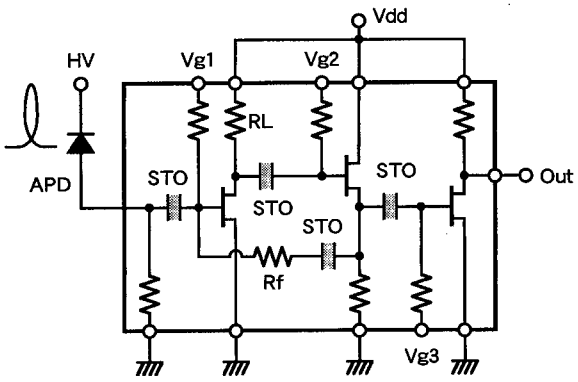


Fig. 5. Proposed circuit diagram of the transimpedance amplifier IC.

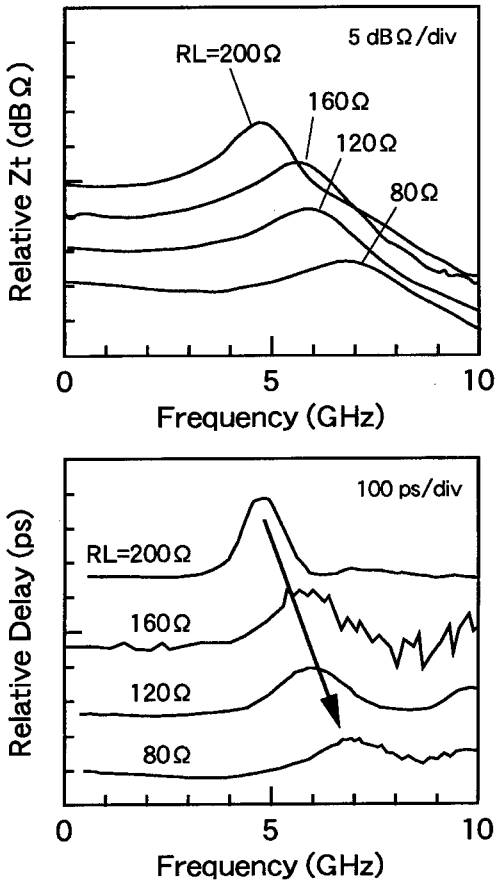


Fig. 6. Sample frequency dependencies of transimpedance gain (Z_t) and group delay (GD) with various load resistor (RL).

Fig. 7 shows the cross-sectional view of the optical receiver module. To reduce parasitic elements, we also adopted the flip-chip bonding method. The GaAs IC and an avalanche photo diode (APD) are mounted on an alumina substrate by using the stud bump bonding (SBB) technique. Finally, the load resistance (RL) of the first-stage FET is determined by optimizing gmR_L

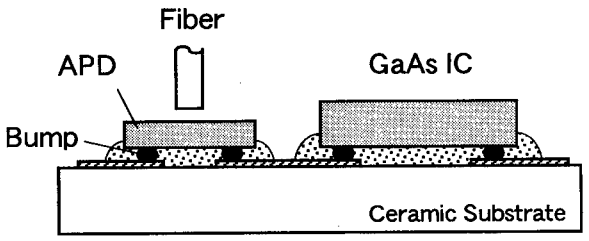


Fig. 7. Cross-sectional view of the optical receiver module.

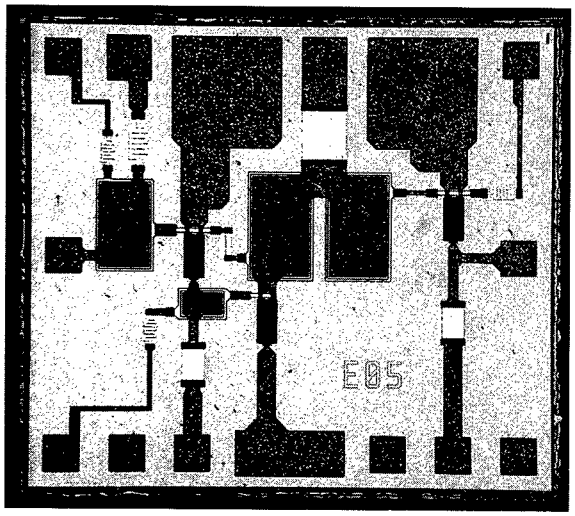


Fig. 8. Microphotograph of the fabricated transimpedance amplifier IC. (Chip size: 1.29 mm × 1.14 mm).

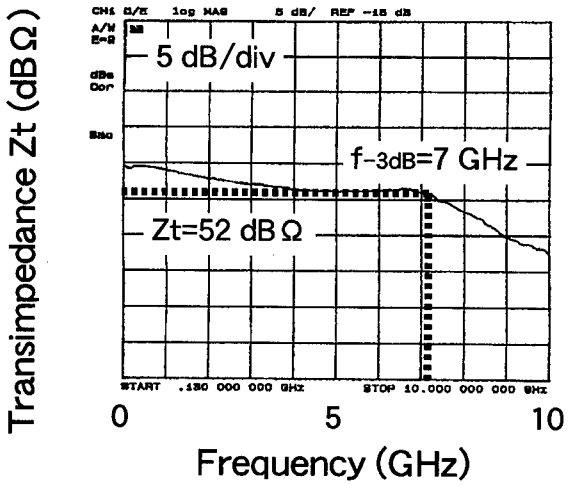


Fig. 9. Transimpedance gain of the IC measured with an avalanche photo diode.

and parasitic elements. This way, the group delay deviation has been improved over the frequency range of 0.4–7 GHz.

4.3. Power dissipation

Conventional transimpedance amplifier requires two supply voltages ($+V_{dd}$ and $-V_{ss}$). However, the 0.25 μm MODFET can operate with a drain voltage of only 2–3 V. Moreover, level-shifting diodes are sensitive to temperature. Therefore, in order to realize low power dissipation and wide-temperature-range operation, level-shifting diodes are replaced by DC blocking STO capacitors. As a result, this IC can operate with a V_{dd} of +5 V without $-V_{ss}$ (negative gate bias voltage is still needed), and less than 1/2 power dissipation is realized compared with conventional transimpedance amplifier with level-shifting diodes.

4.4. Fabrication

Fig. 8 shows the microphotograph of the fabricated transimpedance amplifier IC. The chip size of the transimpedance amplifier IC is 1.29 × 1.14 mm. Resistors with a sheet resistance of 50 Ω/square are made by using a thin-film NiCr process.

5. Performance

Fig. 9 shows the transimpedance gain of the IC measured with an avalanche photo-diode. A transimpedance (Z_t) of 52 dBΩ with a −3 dB bandwidth of 7 GHz is obtained. Fig. 10 shows that the equivalent

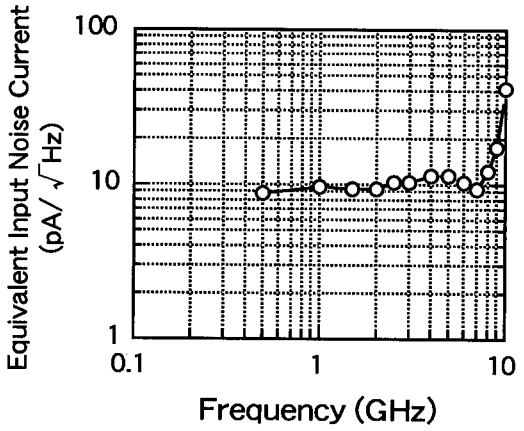


Fig. 10. Noise performance of the IC.

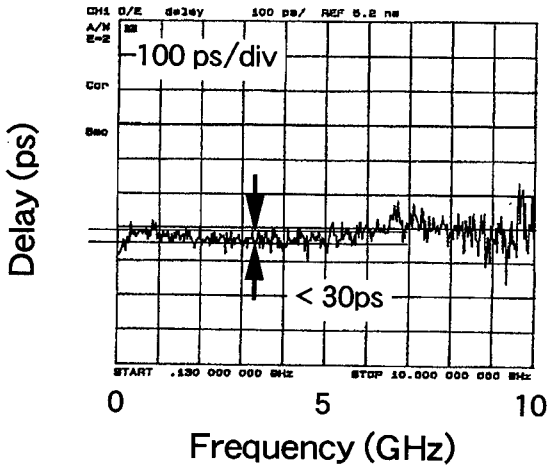


Fig. 11. Group delay (GD) of the IC.

input noise current is less than 12 pA/√Hz. Fig. 11 shows that the group delay deviation is less than 30 ps between 0.4–7 GHz. The output return loss (S22) is less than −8 dB over the same frequency range. This IC operates from a single +5 V supply at less than 60 mA. Thus, the power dissipation is less than 300 mW. This is less than one half of that required by conventional ICs with level-shifting diodes. These results indicate that the transimpedance amplifier IC developed in this work is suitable for the SCM optical communication system.

6. Conclusion

We have successfully developed a low power dissipation and broadband transimpedance amplifier for the SCM optical communication system. Using a novel STO capacitor process, we have integrated large-value capacitors (~200 pF) on GaAs without degrading the 0.25 μm MODFET. By using integrated STO capacitors for DC blocking, single low power supply voltage and low power dissipation is realized.

Acknowledgements

The authors would like to thank Dr Masaru Kazumura and Dr Daisuke Ueda for their continuous encouragement throughout this work. They are also grateful to the colleagues at Matsushita Electronics Corporation.

References

[1] Kikushima K, Kishimoto C, Yoshinaga H, Kumozaki K, Mawatari H, Shibata N. ECOC'95, Th.L.3.1 1995.
[2] Ishii Y, Nojima K, Tougo H, Iwamoto M, Takata M, Kawashima S, Negishi H, Kitaji S. OAN'98, AN-31 1998.
[3] Masato H, Nakatsuka T, Matsuno T, Inoue K. Extended Abstracts of the 1993 International Conference on Solid State Devices and Materials, 1993, p. 715.
[4] Kunihisa T, Yokoyama T, Nishijima M, Yamamoto S, Nishitsuji M, Nishii K, Nakayama M, Ishikawa O. GaAs IC Symposium, 1997. p. 37.
[5] Nishitsuji M, Tamura A, Kunihisa T, Yahata K, Shibuya M, Kitagawa M, Hirao T. GaAs IC Symposium, 1993. p. 329.
[6] Nishitsuji M, Uda T, Yokoyama T, Fujimoto K, Nishii K, Shibuya M, Kitagawa M, Tamura A. Semicond Sci Technol 1995;10:1534.
[7] Nishimura T, Iwata N, Takemura K, Kuzuhara M, Miyasaka Y. Jpn J Appl Phys 1996;35:L1683.

PERGAMON

Solid-State Electronics 43 (1999) 1419–1424

SOLID-STATE ELECTRONICS

44% efficiency operation of power heterojunction FET at near pinch-off for 3.5 V wide-band CDMA cellular phones

T.B. Nishimura*, N. Iwata, Y. Bito, G. Hau

Kansai Electronics Research Laboratories, NEC Corporation 9-1, Seiran 2-Chome, Otsu, Shiga, 520-0833, Japan

Abstract

This paper describes 1.95 GHz power performance of a double-doped AlGaAs/InGaAs/AlGaAs heterojunction FET (HJFET) operated at 3.5 V drain bias voltage for Wide-band CDMA (W-CDMA) cellular phone systems. An HJFET with a novel multilayer cap and a narrow recess structure showed a low on-resistance (R_{on}) of 1.4 ohm mm. With an optimum output impedance matching at a reduced quiescent drain current of 80 mA (less than 1% of the maximum drain current), the HJFET exhibited a high power added efficiency (PAE) of 44.2% with an output power of 600 mW (28.0 dBm) and an adjacent channel leakage power ratio (ACPR) of −43 dBc at 5 MHz off-center frequency. This high PAE obtained was ascribed to an ACPR dip behavior with respect to the input power at around the W-CDMA criteria. Through an ACPR simulation with measured AM–AM and AM–PM conversion characteristics, the AM–AM conversion characteristic was found to dominate the W-CDMA ACPR rather than the AM–PM conversion. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Much attention has been paid to Wide-band CDMA (W-CDMA) cellular phone system as a candidate for the next generation cellular phones standard (IMT-2000) to further increase data stream rate and expand system capacity [1]. Since W-CDMA system requires lower distortion characteristic than the IS-95 system due to the strict adjacent channel leakage power ratio (ACPR) criteria (−40 dBc at 5 MHz off-center frequency with a 4.096 Mcps chip rate), transmitter amplifiers for the W-CDMA system must be operated at high output back-off level from the saturation power under near class A operation. This results in significantly poor power added efficiency (PAE). In addition, a low voltage operation of less than 3.6 V is required for the light weight handsets to employ Li-ion battery, which has a maximum power density per volume (also weight) among rechargeable batteries. To achieve low voltage operation with high PAE, low on-resistance (R_{on}) is a key issue [2]. In this work, a low R_{on} of 1.4 ohm mm has been accomplished with a novel power HJFET structure and low quiescent drain current (I_q) biasing has been investigated for high PAE operation. These result in a remarkably high PAE of 44.2% satisfying the W-CDMA criteria.

2. Device structure

Fig. 1 shows the cross section of the developed HJFET. The active part of the FET consists of a 15 nm-thick undoped $In_{0.2}Ga_{0.8}As$ channel layer sandwiched between upper and lower Si-doped

* Corresponding author. Tel.: +81-77-537-7688; fax: +81-77-537-7689.

E-mail address: nisimura@kel.cl.nec.co.jp (T.B. Nishimura)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00083-0

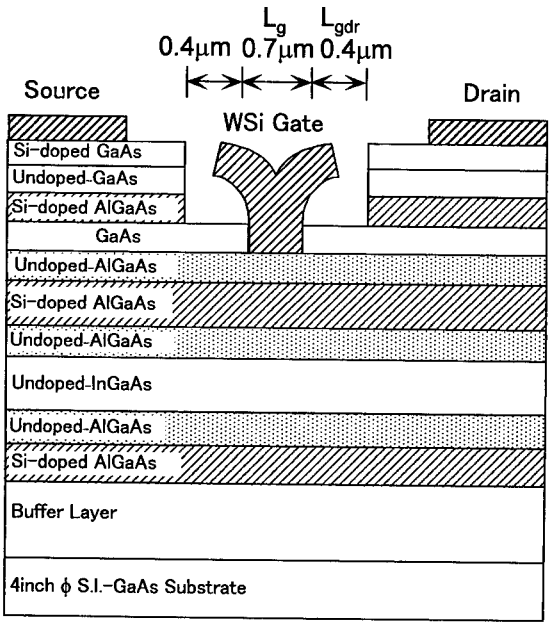


Fig. 1. Cross section of the HJFET.

$Al_{0.22}Ga_{0.78}As$ layers. A 15 nm-thick undoped $Al_{0.22}Ga_{0.78}As$ Schottky layer was incorporated on the upper Si-doped $Al_{0.22}Ga_{0.78}As$ layer to achieve a high gate-to-drain breakdown voltage (BV_{gd}) [3]. A double recess structure was fabricated by electron cyclotron resonance plasma dry-etching with SF_6 and BCl_3 gases, using $Al_{0.22}Ga_{0.78}As$ layers as etch stops [4], and the resulting standard deviation of the threshold voltage (V_T) was as small as 20 mV on a 4″ wafer. WSi metal was sputter-deposited onto the narrow recess to form a 0.7 μm long gate. With this recessed gate structure, parallel conduction and surface trapping effects are suppressed, thus resulting in a low R_{on} as well as preventing from pre-mature power saturation [3].

We investigated two novel designs for the HJFET. The first is a multilayer cap consisting of highly Si-doped GaAs, undoped GaAs and highly Si-doped AlGaAs layers. The multilayer cap was designed to reduce the potential barrier height of the AlGaAs layer, which is indispensable as an etch stop for the wide recess formation, as shown above. Two dimensional electron gases are formed at the upper and lower interface of the AlGaAs etch stop layer, thus the potential barrier height of the AlGaAs is reduced. With this structure, 0.1 ohm mm reduction of R_{on} was achieved due to a reduction of the contact resistance between the cap and the channel layers. The second design is a narrow recess structure. In this investigation, total recess width (L_w) from 1.0 to 2.3 μm were

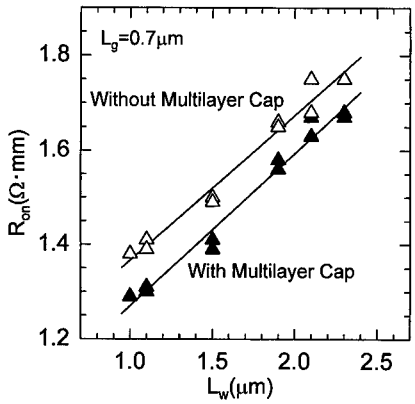


Fig. 2. R_{on} vs L_w for the fabricated HJFET.

employed with a 0.7 μm long gate. Fig. 2 shows the dependence of R_{on} on L_w of HJFETs with and without the multilayer cap. We found that 1.0 μm shrinkage of L_w reduces R_{on} by 0.3 ohm mm. Though a lower R_{on} can be obtained by reducing L_w, further reduction in L_w results in the degradation in BV_{gd}. Fig. 3 shows the dependence of BV_{gd} on the gate-to-drain spacing (L_{gdr}) of HJFETs. An HJFET with L_{gdr} of 0.4 μm achieved a BV_{gd} of more than 13.5 V, which is sufficiently high for Li-ion battery operation [5]. Thus the optimized recess structure was determined to be L_w of 1.5 μm with L_{gdr} of 0.4 μm, resulting in a low R_{on} of 1.4 ohm mm. This R_{on} is 0.7 ohm mm lower than the previously reported HJFET [6].

3. Results and discussions

3.1. DC characteristics

Fig. 4 shows the drain I–V characteristics of the HJFET. The fabricated FET exhibited a maximum

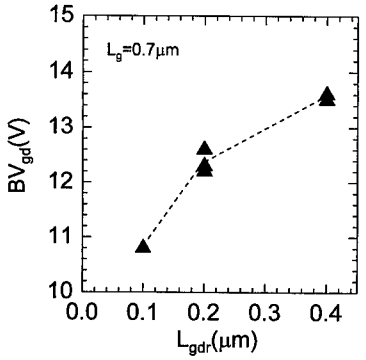


Fig. 3. BV_{gd} vs L_{gdr} for the fabricated HJFET

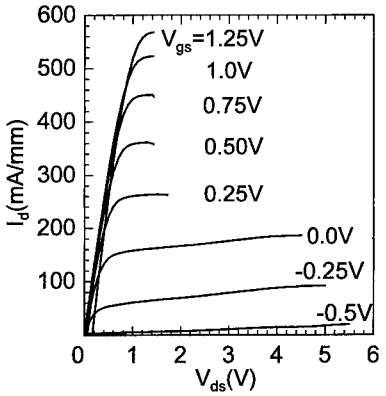


Fig. 4. *I–V* characteristics of the HJFET.

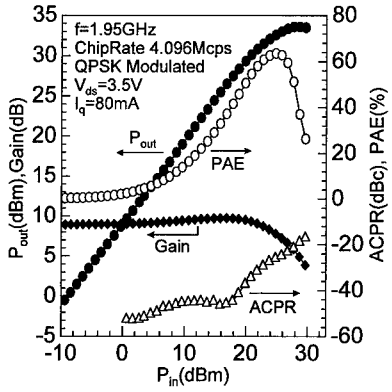


Fig. 6. P_{out}, PAE, gain and ACPR vs P_{in} at I_q of 80 mA.

drain current (I_{max}) of 580 mA mm^{-1} at a gate-to-source voltage (V_{gs}) of 1.25 V. The BV_{gd} was 14.2 V. A maximum g_m of 400 mS mm^{-1} was achieved at around V_{gs} of 0 V. The threshold voltage was −0.6 V.

3.2. Power performance

A 25.6 mm gate width HJFET was evaluated at 1.95 GHz with QPSK signal of a 4.096 Mcps chip rate. The device was mounted on the heat sink of a ceramic package of 9 mm × 6 mm in size, and was grounded by plated gold through the chip side. Through load-pull measurements with computer controlled tuners, the optimum load impedance (Z_L) and source impedance (Z_s) were evaluated. Fig. 5 shows the PAE and ACPR dependence on Z_L evaluated at P_{out} = 28 dBm. The Z_L for maximum PAE was around 3.5–j8.0 ohm. This impedance is higher than that for the maximum P_{out} (2.7–j12.3 ohm). ACPR was found to be strongly affected by the imaginary part of Z_L and the ACPR becomes lower as the imaginary part of Z_L decreases.

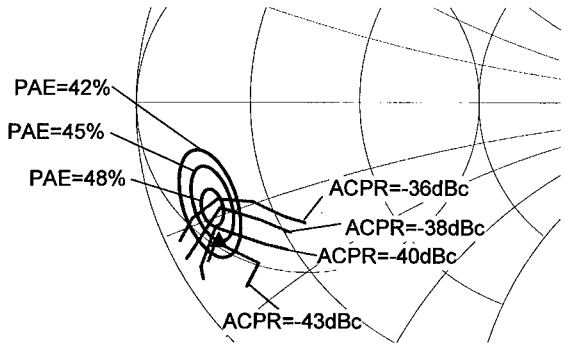


Fig. 5. PAE and ACPR vs Z_L at P_{out} = 28 dBm.

A Z_L of 4.0–j12.2 was chosen to obtain maximum PAE with an ACPR of −43 dBc. Fig. 6 shows P_{out}, PAE, gain and ACPR as a function of P_{in} for the HJFET. At a drain bias voltage (V_{ds}) of 3.5 V and a quiescent drain current (I_q) of 80 mA (< 1% I_{max}), the HJFET demonstrated a P_{out} of 600 mW (28.0 dBm) and a PAE of 44.2% with an associated gain (G_a) of 9.7 dB at ACPR = −43 dBc. This PAE is the best value among power transistors for the W-CDMA cellular phones [7]. A dip was observed in ACPR at around the W-CDMA criteria of P_{in} = 17 dBm. The dip behavior was found to be important for high PAE operation of the HJFET within the W-CDMA criteria, as will be discussed later. Fig. 7 shows P_{out}, PAE, gain and ACPR as a function of P_{in} at I_q = 300 mA (3% I_{max}) for the same HJFET. The HJFET exhibited lower PAE of 33.2% with a P_{out} of 600 mW (28.0 dBm) at ACPR = −43 dBc. Under this condition, the ACPR increases monotonically with P_{in} and no dip

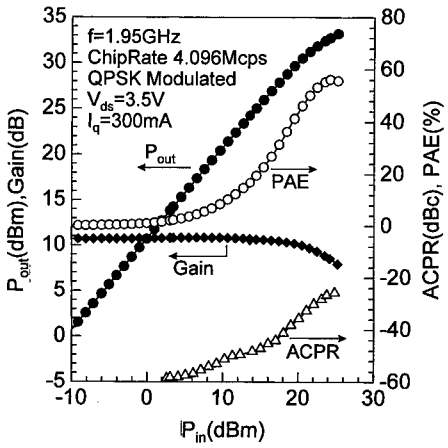


Fig. 7. P_{out}, PAE, gain and ACPR vs P_{in} at I_q of 300 mA.

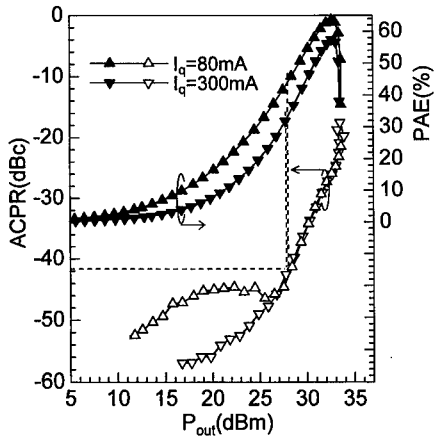


Fig. 8. ACPR and PAE vs P_{out}.

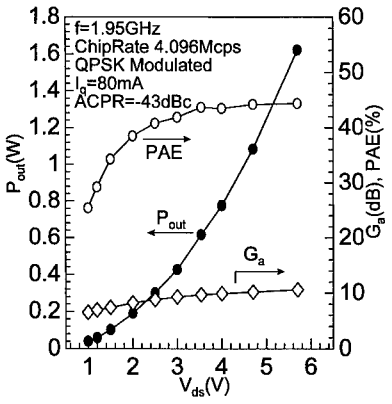


Fig. 10. P_{out}, PAE, gain and ACPR vs P_{in}.

was observed. The ACPR and PAE as a function of P_{out} for I_q values of 80 mA and 300 mA are compared in Fig. 8. Although the $I_q = 300$ mA condition provided a better ACPR at a lower P_{out} level of less than 26 dBm, the $I_q = 80$ mA condition also satisfies the criteria of −43 dBc. In the higher output power range, the ACPRs of these two conditions were similar. Since lower I_q condition provides lower operation current at an identical P_{out} level, the PAE at the criteria is significantly improved.

Fig. 9 shows P_{out}, PAE and G_a as a function of I_q at ACPR of −43 dBc. The HJFET was found to exhibit a maximum PAE of 44.2% at $I_q = 80$ mA. For an I_q range of less than 60 mA, P_{out} and PAE at the criteria drastically degraded due to the rapid increase in ACPR in the lower output power range. This shows setting of quiescent drain bias current is critical for the W-CDMA performance of the HJFET.

The power performance was also evaluated as a function of V_{ds} ranging from 1.0 to 5.7 V. Fig. 10 shows the P_{out}, PAE and G_a dependence on V_{ds} at ACPR of −43 dBc. For V_{ds} of above 2.5 V, a high PAE of more than 40% with P_{out} of more than 300 mW was achieved. Operated at a reduced V_{ds} of 1.2 V, rather high PAE of 29.2% with P_{out} of 60 mW was obtained. These results indicate that the developed HJFET is suitable for W-CDMA cellular phones that require high PAE over a wide range of output power.

3.3. ACPR simulation

An ACPR simulation was carried out using Omnisys system model simulator (HP-EESOF). A behavioral model (GCOMP7 model) of the HJFET was constructed for the analysis utilizing measured AM–AM and AM–PM conversion characteristics. Fig. 11 shows

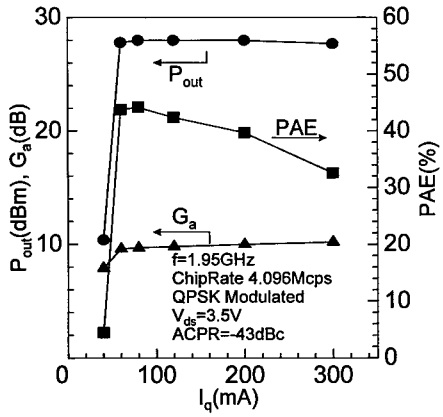


Fig. 9. P_{out}, PAE and G_a vs V_{ds} at ACPR = −43 dBc.

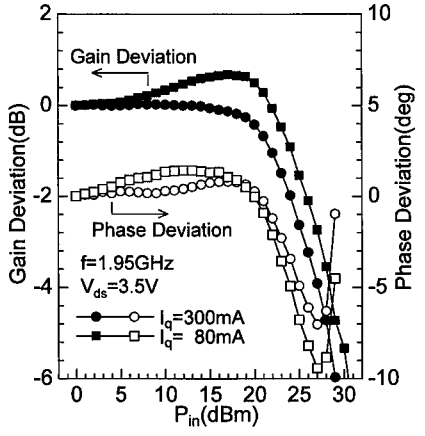


Fig. 11. AM–AM and AM–PM conversion characteristics for the HJFET.

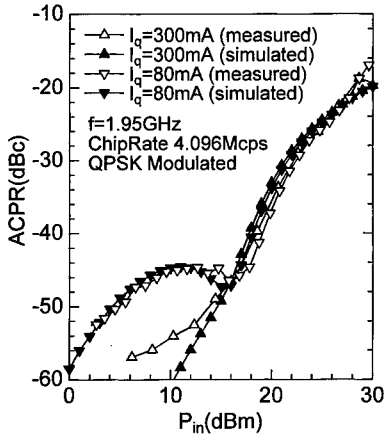


Fig. 12. Measured and calculated ACPRs vs P_{in}.

the AM–AM and AM–PM conversion characteristics of the HJFET measured at 1.95 GHz using single tone input signals. An input power range over 30 dBm was provided by a vector network analyzer with a booster amplifier. The conversion characteristics were measured under two I_q conditions at 80 mA and 300 mA. For the $I_q = 80$ mA case, gain expansion of approximately 0.6 dB was observed at P_{in} of around 17 dBm, whilst no gain expansion was observed at $I_q = 300$ mA. The phase deviations at the criteria were similar between the two cases.

Fig. 12 shows the measured and calculated ACPRs as a function of P_{out}. The calculated ACPR was in good accordance with the measured one for both I_q values. The calculated ACPR of the $I_q = 80$ mA condition also shows dip at P_{in} of around 17 dB. The disagreement observed at lower input power level is due to the noise floor of the spectrum analyzer employed, which limits the dynamic range of measurements.

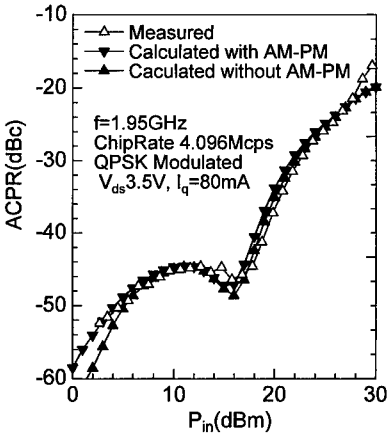


Fig. 13. Measured and calculated ACPRs vs P_{in} with and without AM–PM characteristics.

Another ACPR simulation was carried out to investigate the effect of phase distortion of the HJFET. In the simulation, the AM–AM conversion characteristic was used by assuming no phase deviation throughout the P_{in} range. Fig. 13 shows the measured and calculated ACPRs with and without AM–PM characteristic. The calculated ACPR without AM–PM conversion characteristic is still in good accordance with measured one and also shows the dip behavior. This indicates that the AM–AM conversion characteristic dominates the overall ACPR level of the W-CDMA system under low I_q operation and the AM–PM characteristic has a minimal effect. This implies a linearizer that improves an AM–AM characteristic will be effective to further enhance the PAE of the HJFET amplifier for W-CDMA application.

4. Conclusions

1.95 GHz power performance of a double-doped AlGaAs/InGaAs/AlGaAs HJFET operated at 3.5 V drain bias voltage for W-CDMA cellular phone systems was discussed. An HJFET with a novel multilayer cap and a narrow recess structure showed a low R_{on} of 1.4 ohm mm. With an optimum output matching at a reduced I_q of 80mA ($< 1\%$ I_{max}), the HJFET exhibited a high PAE of 44.2% with P_{out} of 600 mW (28.0 dBm) at a low ACPR of −43 dBc. This high PAE obtained was ascribed to an ACPR dip behavior with respect to the input power at around the W-CDMA criteria.

ACPR simulation indicated that the AM–AM conversion characteristic dominates the W-CDMA ACPR rather than the AM–PM conversion. These results indicate that the developed HJFET tuned for low I_q bias condition has a great potential for W-CDMA cellular phone applications, which can be operated with one Li-ion battery cell.

Acknowledgements

The authors would like to thank Mr M. Ichihara for useful discussions, and Mr M. Tomita for his help in measurements. Thanks also go to Drs M. Kuzuhara, I. Mito and M. Ogawa for encouragement.

References

[1] NTT Mobile Communication Network Inc, NTT DoCoMo Brochure 'W-CDMA' Nov 1997.

[2] Yamaguchi K, Iwata N, Kuzuhara M, Takayama Y. IEEE Trans Electron Devices 1998;45:361.

[3] Iwata N, Inosako K, Kuzuhara M. IEEE MTT-S Dig 1993;1465.

[4] Oikawa H, Kohno M, Mochizuki A, Nashimoto Y. Electrochem Soc Proc 1996;96(2):246.

[5] Mori K, Nakayama M, Itoh Y, Murakami S, Nakajima Y, Takagi T, Mitsui Y. IEICE Trans Electron 1995;E78-C:1229.

[6] Yamaguchi K, Iwata N, Tomita M. APMC Proc 1997;1105.

[7] Nishimura TB, Iwata N, Yamaguchi K, Takemura K, Miyasaka Y. IEEE MTT-S Dig 1998;447.

PERGAMON

Solid-State Electronics 43 (1999) 1425–1428

SOLID-STATE ELECTRONICS

Reliability investigation of heavily C-doped InGaP/GaAs HBTs operated under a very high current-density condition

K. Mochizuki*, T. Oka, K. Ouchi, T. Tanoue

Department of Electrical and Computer Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0407, USA

Abstract

The reliability of InGaP/GaAs heterojunction bipolar transistors, whose base layer was doped with C to 1×10^{20} cm^{-3}, was investigated at an emitter current density J_E of 1 and 2×10^5 A/cm^2. The bias stress with $J_E = 2 \times 10^5$ A/cm^2 increased emitter resistance R_E, but that with $J_E = 1 \times 10^5$ A/cm^2 did not change R_E even at a base-emitter junction temperature T_j of 275°C. Bias stress also increased base current I_B. An ideality factor of 1.8 for the increased I_B indicates that a generation-recombination process in the emitter depletion region dominated the increase in I_B under the low injection condition. The increase in I_B under the high injection condition indicates that the minority carrier lifetime in the neutral base region decreased with the bias stress. The activation energy for the current gain degradation due to this I_B increase under the high injection condition was 1.1 eV. A fairly long lifetime of 8.8×10^4 h was extrapolated at $T_j = 125$°C when $J_E = 1 \times 10^5$ A/cm^2. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

InGaP/GaAs heterojunction bipolar transistors (HBTs) have attracted considerable attention as a potential replacement for AlGaAs/GaAs HBTs in the field of high-speed integrated circuit applications [1–3]. The large valence band discontinuity at the InGaP/GaAs interface improves the carrier injection efficiency, enabling much higher base doping levels [4,5], and the low surface recombination velocity of InGaP can reduce the base leakage current [6,7], which improves reliability characteristics [8].

To fully utilize the excellent performance of these HBTs, their base should be as highly doped as possible (up to 1×10^{20} cm^{-3}) and they should be operated at a high emitter current density J_E ($J_E = 1$–2×10^5 A/cm^2 [9]) at which the cutoff frequency f_T and maximum oscillation frequency f_{max} peak. However, in a previous study, the base doping level and the J_E during bias testing were quite low (4×10^{19} cm^{-3} and 6×10^4 A/cm^2, respectively [8]). To investigate the degradation characteristics of HBTs that have a heavily doped base and are operated under a very high current–density condition, we have fabricated InGaP/GaAs HBTs whose base layer was doped with C to 1×10^{20} cm^{-3} and carried out bias testing at $J_E = 1$ and 2×10^5 A/cm^2.

2. Experiments

The epitaxial layer structures were grown on a semi-insulating GaAs (100) substrate by gas-source molecu-

* Corresponding author. Tel.: +1-619-822-2285; fax: +1-619-534-2486.

E-mail address: motizuki@crl.hitachi.co.jp (K. Mochizuki)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00084-2

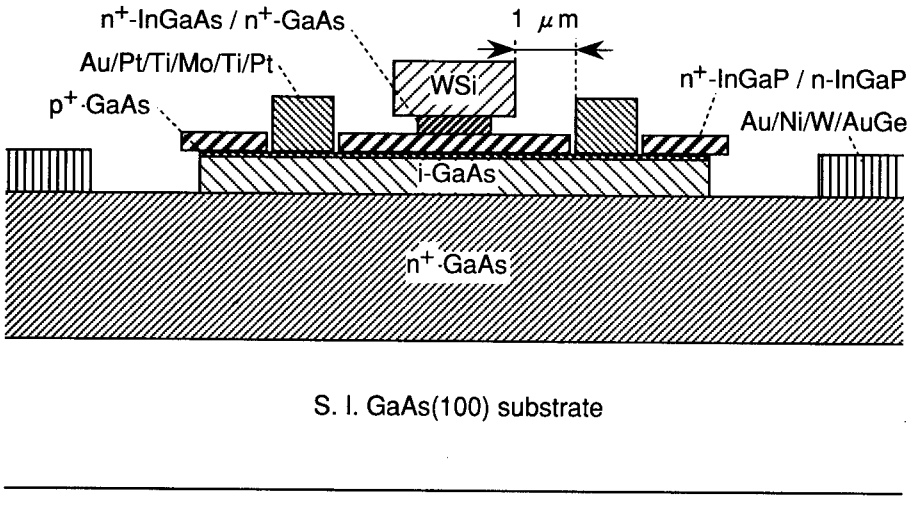


Fig. 1. Schematic cross-section of the fabricated HBTs.

lar beam epitaxy using Si and C as *n*- and *p*-type dopants, respectively. A triple emitter cap layer was used to reduce emitter resistance R_E: 50 nm-thick $In_{0.5}Ga_{0.5}As$ doped to 4×10^{19} cm^{-3}, 50 nm-thick GaAs doped to 5×10^{18} cm^{-3}, and 50 nm-thick $In_{0.5}Ga_{0.5}P$ doped to 8×10^{18} cm^{-3}. The $In_{0.5}Ga_{0.5}P$ emitter layer was 50 nm thick and doped to 1×10^{18} cm^{-3}. The GaAs base layer was 30 nm thick and doped to 1×10^{20} cm^{-3}. The H concentration in the base layer measured by secondary-ion mass spectrometry was about 5×10^{18} cm^{-3}. The collector layer was 200 nm-thick undoped GaAs. The GaAs subcollector layer was 800 nm thick and doped to 5×10^{18} cm^{-3}.

A schematic cross section of the fabricated HBTs is shown in Fig. 1. We fabricated mesa structure transistors with an emitter metal area S_E of 2×5 μm by using selective wet chemical etching and a standard photolithographic process [9]. The InGaAs and GaAs cap layers were etched with $H_3PO_4+H_2O_2+H_2O$ using the WSi emitter electrode as an etching mask. The InGaP cap and emitter layers were etched with HCl only where the Au/Pt/Ti/Mo/Ti/Pt base electrode was formed. We left the InGaP layers on the 1 μm-long extrinsic base, expecting them to act as a surface passivation layer to reduce surface recombination [10,11]. After the base mesa and the collector mesa were formed by wet chemical etching using a photoresist as an etching mask, Au/Ni/W/AuGe was evaporated onto the subcollector layer and alloyed at 350°C for 30 min. The effective emitter area, determined by the emitter size dependence of collector current I_C, was 1.2×3.4 μm. The S_E independence of current gain h_{FE} reported previously [9] indicates that surface effect is negligibly small even for such small transistors.

Bias testing was carried out using a substrate temperature T_s of 170°C, J_E of 1 and 2×10^5 A/cm^2, and a collector-base bias V_{CB} ranging from −0.5 to +4 V. The base-emitter junction temperature during bias testing, T_j, was estimated from the dependence of h_{FE} on the power consumption and T_s [12].

3. Results and discussion

Fig. 2 shows Gummel plots of the HBTs before and after 30 h-long bias stress at $J_E=2 \times 10^5$ A/cm^2 and $V_{CB}=0$ V. We observed a forward base-emitter voltage

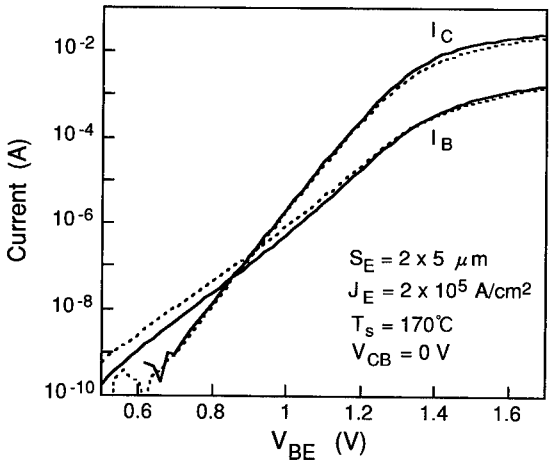


Fig. 2. Gummel plots of the fabricated HBTs before (solid line) and after (dashed line) 30-h bias stress at $J_E=2 \times 10^5$ A/cm^2 and $V_{CB}=0$ V.

(V_{BE}) shift of 4.8 mV for I_C, i.e., a 5.4% reduction of I_C when $V_{BE} < 1.2$ V. Since this reduction of I_C agrees with the ratio of H vs C concentration in the base (5%), we believe that debonding of H from C–H pairs occurred during the bias testing [13].

In Fig. 2, we also observed increases in R_E and base current I_B. The R_E, determined from emitter current I_E vs V_{BE} linear plots, was 12.7 Ω before and 17.1 Ω after the stress. There are two possible causes for this R_E increase. One is carrier depletion at the InGaAs/GaAs interface due to an increased density of misfit dislocations. The other is degradation of the WSi/InGaAs interface, possibly due to segregation of In into the WSi. Since a similar R_E increase was observed for the InGaAs cap layer of InP/InGaAs HBTs lattice-matched to InP substrates [14], this R_E increase should be attributed to increased contact resistivity of the WSi on the InGaAs cap layer. We found that when $J_E = 1 \times 10^5$ A/cm^2, R_E did not change even at $T_j = 275$°C. Therefore, we concluded that the maximum J_E should be 1×10^5 A/cm^2 for practical applications.

The ideality factor n for I_B after the stress (Fig. 2) was 1.8 when $V_{BE} < 1.3$ V. Since the base layer was heavily doped to 1×10^{20} cm^{-3}, we believe that a generation-recombination process in the emitter depletion region dominated the increase in I_B under the low injection condition. The I_B also increased under the high injection condition, as shown in Fig. 3. This indicates that the minority carrier lifetime in the neutral base region decreased with the bias stress. We defined the lifetime of the HBTs as the time taken for the peak h_{FE} to decrease by 20%. The dependence of the lifetime on T_j is shown in Fig. 4. The activation energy E_a was 1.1 eV for both $J_E = 1$ and 2×10^5 A/cm^2. This is within the range of reported values (from 0.7 to 2.0 eV) for InGaP/GaAs HBTs [15,16]. The extrapolated lifetime at $T_j = 125$°C was 1.7×10^4 h for $J_E = 2 \times 10^5$ A/cm^2 and 8.8×10^4 h for $J_E = 1 \times 10^5$ A/cm^2. Thus, these HBTs remain very stable even under a very high current-density condition.

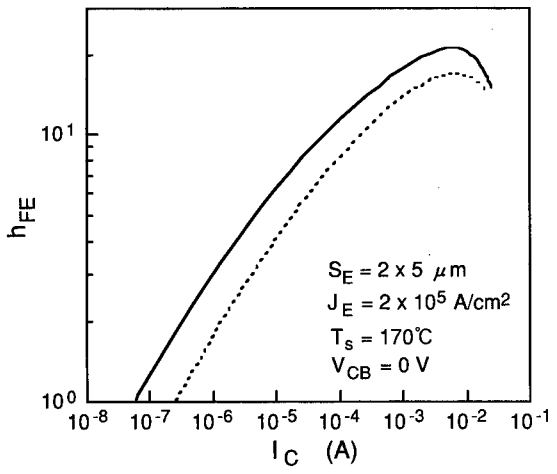


Fig. 3. Dependence of current gain on the collector current before (solid line) and after (dashed line) 30-h bias stress at $J_E = 2 \times 10^5$ A/cm^2 and $V_{CB} = 0$ V.

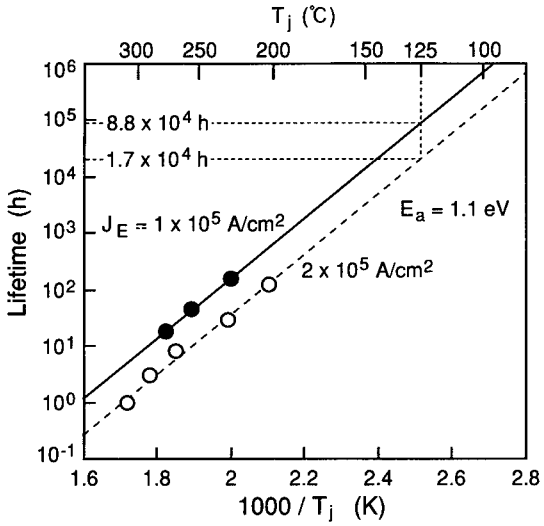


Fig. 4. Arrhenius plots of the lifetime of the fabricated HBTs.

4. Conclusions

We investigated the current-induced degradation of InGaP/GaAs HBTs, whose base layer was doped with C to 1×10^{20} cm^{-3}, with $J_E = 1$ and 2×10^5 A/cm^2. The bias stress with $J_E = 2 \times 10^5$ A/cm^2 increased R_E, which we attributed to increased contact resistivity of the WSi electrode on the InGaAs cap layer. The bias stress with $J_E = 1 \times 10^5$ A/cm^2, on the other hand, did not change R_E even at $T_j = 275$°C. Therefore, we concluded that the maximum J_E should be 1×10^5 A/cm^2 for practical applications. Bias stress also increased I_B. The increase in I_B with $n \sim 2$ indicates that a generation-recombination process in the emitter depletion region dominated under the low injection condition. The increase in I_B under the high injection condition indicates that the minority carrier lifetime in the neutral base region decreased with the bias stress. The activation energy for the h_{FE} degradation due to this I_B increase under the high injection condition was 1.1 eV. A fairly long lifetime of 8.8×10^4 h for $J_E = 1 \times 10^5$ A/cm^2 was extrapolated at $T_j = 125$°C, showing that these HBTs remain very stable even under a very high current density.

References

[1] Kren DE, Rezazadeh AA, Tothill N. Electron Lett 1994;30:825.
[2] Liu W, Fan SK. IEEE Electron Device Lett 1992;13:510.
[3] Abernathy CR, Ren F, Wisk PW, Pearton SJ, Esagui R. Appl Phys Lett 1992;61:1092.
[4] Kroemer H. J Vac Sci Technol B 1983;1:126.
[5] Biswas D, Debar N, Bhattacharya P, Razeghi M, Defour M, Omnes F. Appl Phys Lett 1990;56:833.
[6] Costa D, Harris Jr J. IEEE Trans, Electron Dev 1992;39:2383.
[7] Chen J, Sites JR, Spain IL, Hafich MJ, Robinson GY. Appl Phys Lett 1991;58:744.
[8] Takahashi T, Sasa S, Kawano A, Iwai T, Fujii T. In: IEDM Tech Digest, 1994. p. 194.
[9] Oka T, Ouchi K, Mochizuki K, Nakamura T. Solid St Electron 1997;41:1611.
[10] Hayama N, Tanaka S, Honjo K. In: Third Asia Pacific Microwave Proceedings, 1993. p. 1039.
[11] Mochizuki K, Masuda H, Kawata M, Mitani K, Kusano C. Jpn J Appl Phys 1991;30:L266.
[12] Waldrop JR, Wang KC, Asbeck PM. IEEE Trans Electron Dev 1992;39:1248.
[13] Ren F, Abernathy CR, Chu SNG, Lothian JR, Pearton SJ. Solid-St Electron 1995;38:1137.
[14] Mochizuki K. unpublished.
[15] Bahl SR, Camnitz LH, Houng D, Mierzwinski M, Turner J, Lefforge D. In: IEDM Tech Dig, 1995. p. 815.
[16] Ueda O, Kawano A, Takahashi T, Tomioka T, Fujii T, Sasa S. Solid-St Electron 1997;41:1605.



PERGAMON

Solid-State Electronics 43 (1999) 1429–1436

SOLID-STATE ELECTRONICS

Comparison of conventional and thermally-stable cascode (TSC) AlGaAs/GaAs HBTs for microwave power applications

Shawn S.H. Hsu[a,*], Burhan Bayraktaroglu[b], Dimitris Pavlidis[a]

[a]*Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109-2122, USA*
[b]*Northrop Grumman Corporation, Baltimore, MD 21203, USA*

Abstract

Conventional and thermally-stable cascode HBT (TSC-HBT) were fabricated using a self-aligned emitter-base process on MOCVD-grown wafers. The pronounced self-heating effect of conventional AlGaAs/GaAs HBT was reduced dramatically by the cascode design approach. The DC, small and large-signal characteristics of conventional common–emitter (CE) and TSC-HBTs were compared and a direct assessment of the new HBT design is provided. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

AlGaAs/GaAs heterojunction bipolar transistors (HBTs) have been implemented for a wide range of power applications such as power amplifiers and mixers. However, the low thermal conductivity of the material causes pronounced thermal effects, which are an important issue for power applications. To control these effects, various approaches have been used in the past including ballast resistors and thermal-shunt structures [1–3]. More recently, a new Thermally-Stable Cascode HBT (TSC-HBT) design was developed, which not only provides an effective solution to the thermal runaway issue, but also can improve the robustness of high power HBTs under overstressed DC or RF bias conditions [4]. The TSC-HBT employs the cascode configuration to regulate the current in each emitter finger independently. In other words, the cascode HBTs reduce the undesired thermal effects by locating the current and power generation regions into separate temperature zones.

In this paper, the performance of conventional and cascode HBTs are compared under dc, small-signal and large-signal conditions. In addition, a Gummel–Poon based model including thermal sub-circuit and behavior sources is used to simulate the junction temperature increase and to provide a clearer picture of device characteristic [5,6]. Conventional HBTs with identical characteristics to the common–emitter (CE) stage of the cascode configuration are used for comparison. The large-signal model of the cascode configuration also used two single HBT stages with identical characteristics to those of conventional HBTs. We found that TSC-HBT, in addition to being more robust, can provide higher power gain and efficiency than the conventional HBTs at microwave frequencies.

2. Device design and fabrication

A TSC-HBT is made up of a CE and a CB stage connected as shown in Fig. 1. In such a device, the CB

* Corresponding author.
E-mail address: shuohung@engin.umich.edu (S.S.H. Hsu)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00085-4

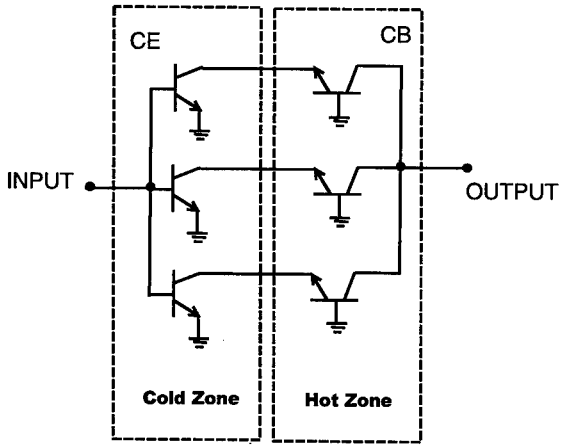


Fig. 1. Schematic drawing of a TSC-HBT cell showing individual connection between CE and CB stages.

stage that provides the output power (and therefore becomes hot) is physically separated from the part that regulates the current (CE-stage). Because the electrothermal feedback is effectively eliminated in this configuration, the collector current remains uniformly distributed across all parts of the CB stage. The net result is that a uniform temperature distribution is achieved at all DC and RF drive levels without thermal instability. As shown by the results of this paper, this increased thermal stability is possible without compromising the microwave performance of power HBT cells.

All devices studied here were fabricated on MOCVD-grown 100-mm diameter wafers with a self-aligned emitter-base process. A constant emitter geometry $2.5 \times 20\ \mu m^2$ was used in all designs. The conventional CE HBT had four emitter fingers separated by 30 μm. A similar layout configuration was used for the CE and the CB stages of the cascode cell with 100 μm separation between the stages. This separation distance was found from computer simulations to be sufficient to provide thermal isolation between the stages. To increase the thermal stability, thermal shunt structures were used for both the conventional HBT and the CE stage of the TSC-HBT. The thermal shunt structure provides stronger thermal coupling between adjacent emitter fingers and therefore can minimize the variation in the current value of each emitter finger of the cell. In the case of the conventional HBT, the thermal shunt approach alleviates the thermal runaway conditions and allows higher power operation. For a TSC-HBT, the thermal shunt serves to minimize the temperature difference between the emitter fingers of the CE stage to ensure a uniform collector current supply to the CB stage. No separate thermal-shunt structure was used for the CB stage. The final substrate thickness was 100 μm for all devices.

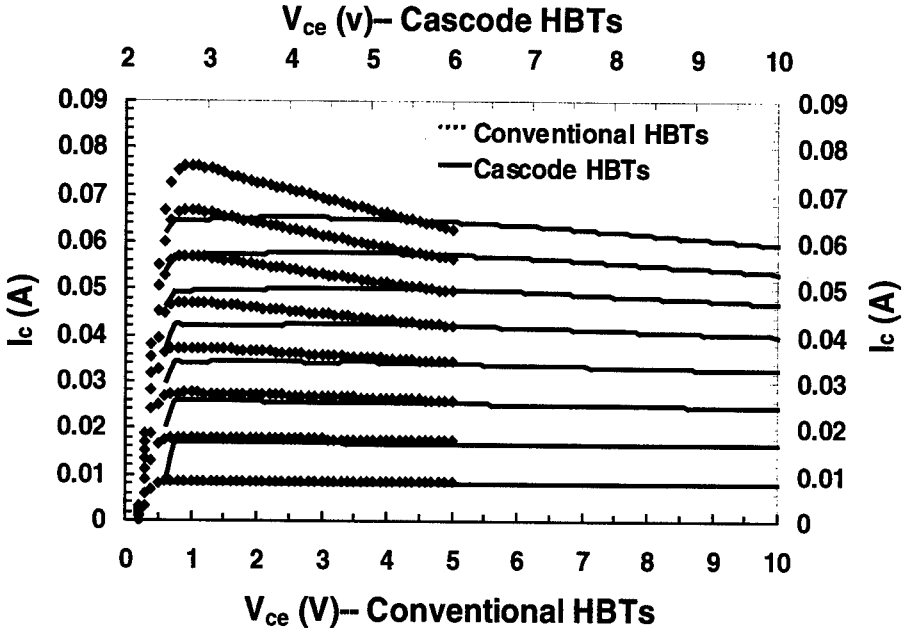


Fig. 2. Forward I_c– V_{cc} characteristics for conventional and cascode HBTs. $I_b = 0.3, 0.6, \ldots, 2.4$ mA, $V_{cc} = 0$–5 V for conventional HBT and 2.5–10 V for cascode HBTs.

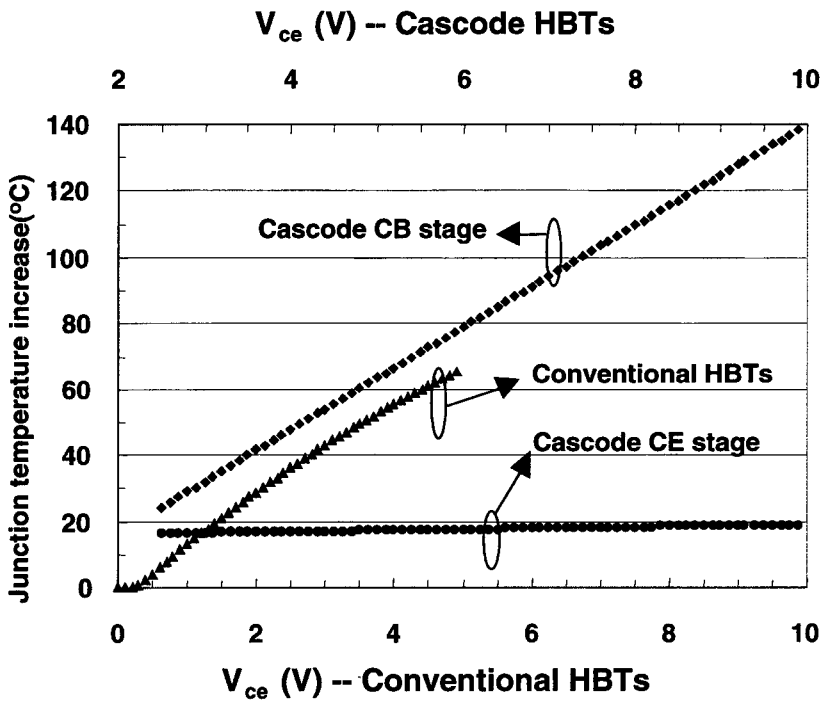


Fig. 3. Simulated junction temperature increase for conventional, cascode CE stage and CB stage HBTs. I_b=2.4 mA, V_{ce}=0–5 V for conventional HBTs and 2.5–10 V for cascode HBTs.

An identical cell layout approach was used in TSC-HBTs for both CE and CB stages. Virtually complete thermal isolation was provided between CE and CB stages by separating these cells by at least 100 μm. Thermal shunt structures were used for the CE stage to minimize the temperature variation between emitter fingers and therefore maintain a uniform collector current generation. The collector of each CE subcell was directly connected to the corresponding emitter of the CB subcell. No thermal shunt structures were used for the CB stage cell.

We compared the performance of 4-finger cells (4-finger subcells for the CE and CB stages for TSC-HBT) throughout this paper. Additional information was provided for larger devices to confirm that the superior microwave performance of TSC-HBT can be extended to higher power levels.

3. Experimental results and discussion

3.1. DC characteristics

The room temperature I_c– V_{ce} characteristic of conventional and cascode HBTs under constant I_b bias condition are shown in Fig. 2. A negative slope of the collector current in the forward I_c– V_{ce} characteristics of conventional devices indicates the strong influence of junction temperature on current gain. The reverse hole-current injection increase with the junction temperature is the main reason causing a reduced gain under constant I_b bias [5]. A similar effect was not observed in TSC-HBTs because the rise in temperature is confined mostly to the CB cell, whose current is controlled by the CE cell located at a cooler temperature zone. As can be seen, the collector current of the conventional HBT dropped by about 15 mA at I_b=2.4 mA, V_{ce}=5 V due to the thermal effect. On the other hand, the TSC-HBT suffered considerably less from the thermal effects; I_c dropped by about 6 mA at I_b=2.4 mA, even when the collector voltage was increased to 10 V.

A large-signal microwave device model including self-heating effects was employed to investigate the thermal characteristics of both devices. A temperature controlled current source was employed to model the increase in the reverse hole injection from the base to emitter, while a temperature controlled voltage source was used to describe the decrease in the emitter junction built-in potential. In addition, a thermal subcircuit including a current source and a thermal resistance were used to obtain a self-consistent thermal HBT model. In the sub-circuit, the current source describes the device power consumption and the voltage drop in

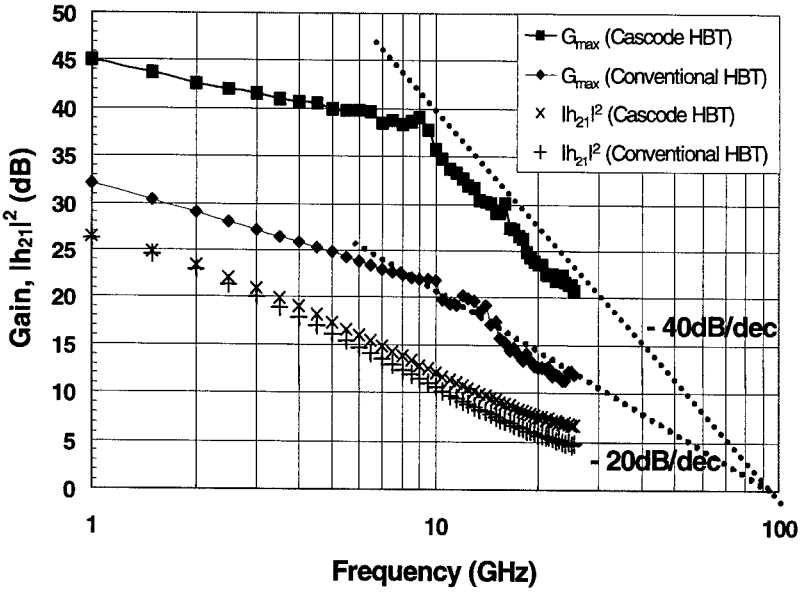


Fig. 4. G_{max} and $|h_{21}|^2$ for conventional and cascode HBTs. $I_b = 2.4$ mA, $I_c = 65.06$ mA, $V_{ce} = 4$ V for conventional HBTs; $I_b = 2.1$ mA, $V_{b2} = 2.5$ V, $I_c = 57.44$ mA, $V_{ce} = 7$ V for cascode HBTs.

the thermal resistance presents the junction temperature increase. The parameters for this physical-based model were extracted from the Gummel-plot, I_c– V_{ce} characteristics under both constant I_b and V_{be} conditions and from the small-signal S-parameters [7]. A large-signal model was developed first for conventional devices and then applied to the cascode configuration. Good agreement was obtained between measured and simulated results.

Based on the developed large-signal models, the base–emitter junction temperature increase of conventional HBTs and cascode CE, CB stages were extracted from the output of the thermal sub-circuit and the results are shown in Fig. 3. The TSC-HBT CE stage junction temperature was found to increase by ~20°C at $V_{ce} = 10$ V, while the temperature increase was ~60°C for the conventional HBT at $V_{ce} = 5$ V. As can be seen, the junction temperature of the cascode CB-stage was somewhat higher than that of the CE-stage at all power dissipation levels. This is due to the fact that the main voltage drop occurs across the CB stage, which therefore, consumes the largest amount of power.

Since most of the collector voltage dropped on the CB stage and the CE and CB stages were thermally isolated, the current control stage could be kept at lower temperature in the cascode HBTs even under high power consumption conditions. These results highlight the advantages of the TSC-HBT design to suppress electrothermal effects by maintaining a lower junction temperature increase at the CE stage.

3.2. Microwave small-signal characteristics

The small-signal S-parameters of the conventional and TSC-HBT devices were measured using an automated network analyzer (HP8510C) in the frequency range of 0.5–25.5 GHz. The MSG (maximum stable gain), MAG (maximum available gain) and $|h_{21}|^2$ versus frequency characteristics are shown in Fig. 4 for both the conventional and the TSC-HBT. The bias conditions corresponding to the highest f_{max} were $I_b = 2.4$ mA, $I_c = 65.06$ mA, $V_{ce} = 4$ V for conventional HBTs and $I_b = 2.1$ mA, $V_{b2} = 2.5$ V (base voltage for CB-stage), $I_c = 57.44$ mA, $V_{ce} = 7$ V for cascode device. The maximum stable gain for TSC-HBT devices is ~13 dB higher than for conventional HBTs at lower frequencies. Above 10 GHz, G_{max} dropped at different slopes (−20 dB/dec and −40 dB/dec for CE and TSC-HBT, respectively). Obviously, the common–base stage provided extra gain to the cascode configuration, but caused a sharper decrease in gain–frequency characteristics. The extrapolated results suggest a similar f_{max} value of ~92 GHz for both devices. On the other hand, $|h_{21}|^2$ is higher for TSC-HBTs in the measured frequency range (0.5–25.5 GHz) leading to higher f_t. The higher value of f_t for cascode devices is due to lower bias voltage applied to the CE stage compared to the bias applied to the conventional HBT.

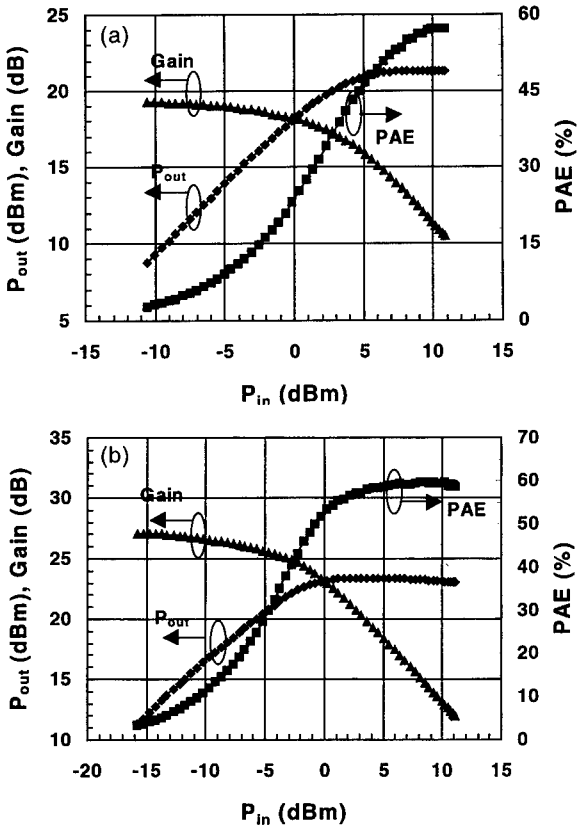


Fig. 5. Load-pull measurement of conventional and cascode HBTs (a) Gain, P_{out} and PAE for conventional HBTs (b) Gain, P_{out} and PAE for cascode HBTs.

3.3. Large-signal power performance

On-wafer power characterization was performed at 8 GHz using a load-pull measurement system with electromechanical tuners, which provide the ability to control the input and output impedances. The power performance such as gain, power-added efficiency and output power were evaluated. In addition, load-pull contour measurements were performed to provide a clearer picture of the device power characteristics. The devices were biased under the same conditions as above. The input and output impedances were optimized for maximum gain at $P_{in}=0$ dBm. The corresponding source (Γ_S) and load (Γ_L) reflection coefficients were $\Gamma_S=0.793\angle-173.3°$ and $\Gamma_L=0.254\angle105.6°$ for the conventional HBTs; $\Gamma_S=0.793\angle-173.3°$ and $\Gamma_L=0.539\angle77.7°$ for the cascode HBTs. Fig. 5 shows the measured gain, PAE and P_{out} for both devices. The results indicate that under this input power level, the optimized gain was 18.47 and 25.40 dB for conventional and TSC-HBT devices respectively. The higher gain of the cascode HBT, compared to the conventional device, appears to be due to the common–base stage providing additional power gain. The corresponding P_{out} was 18.11 and 22.98 dBm; PAE was 25.4%, 53.4% for each device. Compared with the PAE of conventional HBTs at the low input-power level, the cascode HBTs have much higher PAE due to the contribution of the CB-stage. It is important to note that TSC-HBT was able to maintain PAE values greater than 50% for a wide range of P_{in} values (−1 to 12 dBm). The peak efficiency of 60% was maintained across a wide power range also ($5 < P_{in} < 12$ dBm).

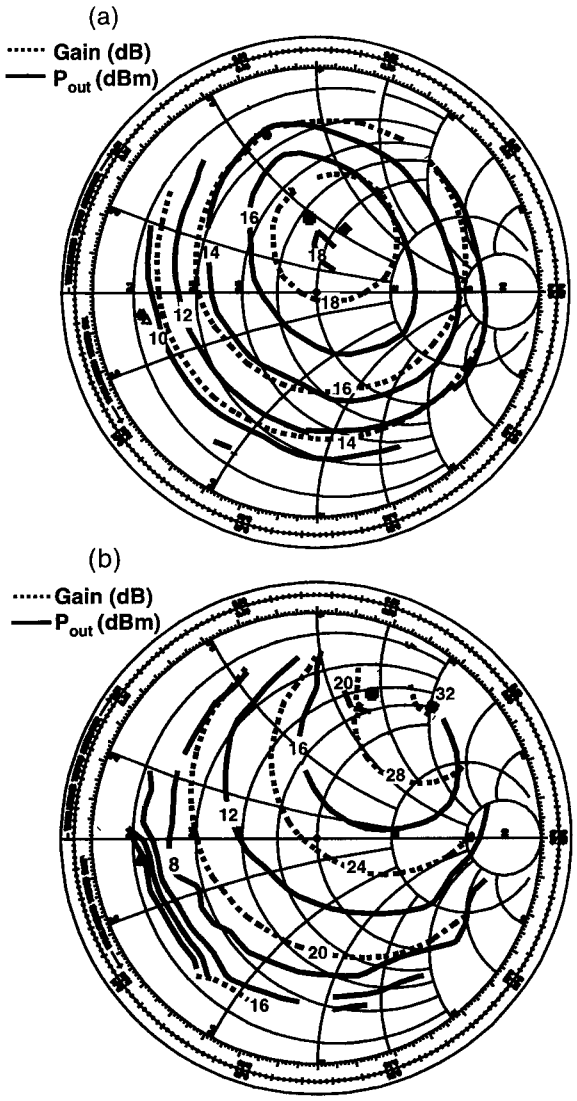


Fig. 6. 1-dB gain compressed on-wafer load pull measurement of (a) conventional HBTs: maximum gain = 18.5 dB and $P_{out}=18.2$ dBm; (b) cascode HBTs: maximum gain = 32.9 dB and $P_{out}=21.2$ dBm.

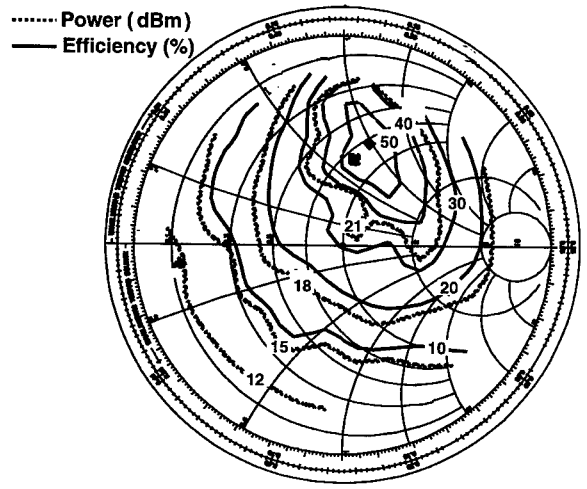


Fig. 7. P_{out} and PAE contours for cascode HBTs under large-signal conditions. Maximum $P_{out}=23.4$ dBm at $\Gamma_L=0.43\angle 81.7°$; maximum PAE = 60% at $\Gamma_L=0.52\angle 74.9°$.

The power handling capability was investigated by measuring the 1 dB-gain compression characteristics of each device type. The measured power and gain characteristics are shown in Fig. 6 for the conventional and TSC-HBTs. As can be seen, the maximum compressed gain is 18.5 dB for conventional HBTs at $\Gamma_L=0.34\angle 70.2°$, maximum $P_{out}=18.2$ dBm at $\Gamma_L=0.34\angle 94.4°$. On the other hand, the cascode device can reach a gain of 32.9 dB at $\Gamma_L=0.79\angle -129.2°$ while producing $P_{out}=21.2$ dBm at $\Gamma_L=0.79\angle -129.2°$.

Fig. 7 shows the load-pull contour measurement results of the cascode device under high input-power level conditions. The load mismatch corresponded in this case to reflection coefficient of 0.9 and is therefore very close to the edge of the Smith Chart (VSWR = 19). The input power level was ~5 dBm and the device showed ~14 dB gain compression. It should be noted that the device could still operate with a maximum 60% of PAE under this highly gain-compressed condition. The purpose of stressing the devices with a high input power was to observe their stability under such conditions. As this figure shows, the contours for constant P_{out} and PAE are no longer uniform circles as under small-signal conditions. In addition, the optimum loads for maximum P_{out}, PAE and gain shifted toward the center of the Smith Chart compared to the positions observed under small-signal conditions. The distorted contour circles and Z_{opt} shift could be due to self-biasing, as well as, the presence of a very large collector voltage sweeping under large-signal conditions. In other words, the bias dependent elements such as C_{be} (I_c, V_{be}), C_{bc} (I_c, V_{bc}) appear to change with the input-power level. The cascode device was found to maintain its thermal stability even under these high gain compression and load-mismatched conditions. On the other hand, the conventional device

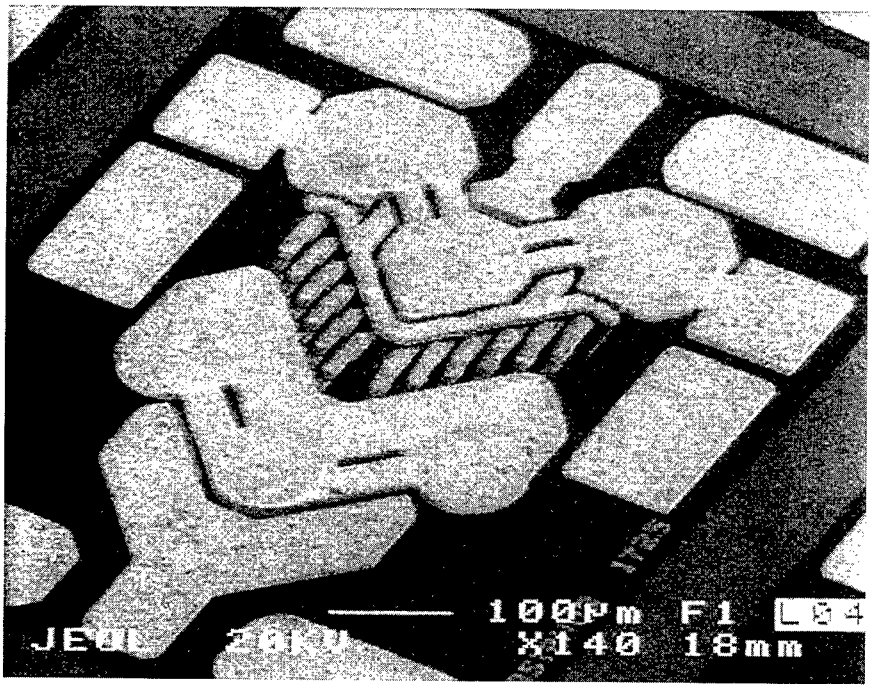


Fig. 8. SEM picture of a microwave high power TSC-HBT cell with 24 emitter fingers.

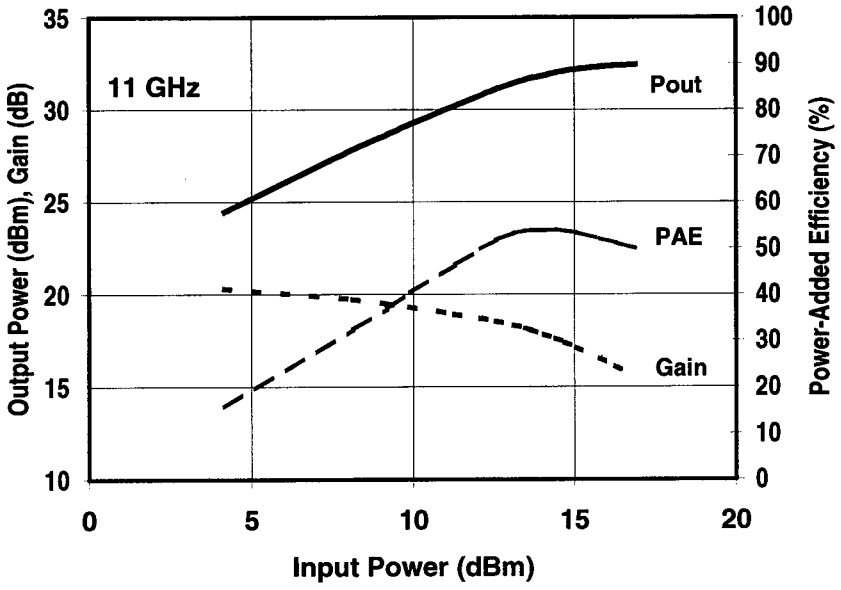


Fig. 9. Power performance of 24-finger TSC-HBT at 11 GHz with $V_{cc}=11$ V and $V_{b2}=2.5$ V.

was found to be thermally unstable under such strenuous gain compression conditions.

To identify the influence of the cascode CB-stage, load-pull measurements were performed on this stage alone. The results showed a relatively small variation of gain and P_{out} for different loads selected on the Smith Chart. This suggests that the power characteristics of the CB-stage are less sensitive to the matching condition provided by the output load. The load-pull measurement results led to the conclusion that the CB-stage of the cascode configuration improved thermal stability and relaxed the load-matching requirement of the device. Thus, we found that the TSC-HBT devices perform considerably better as microwave power amplifiers than conventional devices.

3.4. High power performance

We have employed relatively small size devices (4-fingers) to make a comparison between the conventional and TSC-HBT cells. The superior electrothermal properties demonstrated with small cells can only be useful in practice if such properties can be extended to higher power levels. At X-band frequencies, for example, typical HBT cells are expected to produce over 0.5 W and preferably over 1.0 W output power levels. To demonstrate that the TSC-HBT performance can be maintained as the devices are scaled up for higher power, we have fabricated a cell with 24 emitter fingers (in each of CE and CB stages) as shown in Fig. 8. A staggered layout approach was used in the design of this cell to minimize the thermal interaction between adjacent emitter fingers. The distance between CE and CB stages was kept constant at 100 μm, which means that identical staggering of fingers was used for both stages. Multiple bypass capacitors were used, as shown, to maintain uniform base grounding for the CB stage. The total width of the cell was only 300 μm.

The power performance of the high power cell is shown in Fig. 9. At 11 GHz, with $V_{ce}=11$ V and $V_{b2}=2.5$ V, the cell produced over 1.75 W CW output power with 17 dB gain and 55% PAE. This result, which is the highest reported for a HBT cell at this frequency with such high gain and efficiency, underscores the power capability of TSC-HBTs at microwave frequencies.

4. Conclusion

We have compared the DC, small-signal and large-signal power characteristics of conventional and TSC-HBTs and found that TSC-HBTs provide a higher power handling capability than conventional HBTs. The CB-stage in the TSC-HBTs not only leads to smaller temperature increase but also acts as an additional power amplification stage. The net result is that TSC-HBTs are eminently more suitable for high frequency power applications than conventional CE HBTs.

Acknowledgements

This work is supported by DARPA MAFET Thrust 2 Program, Contract No. N00014-95-C-6026.

References

[1] Gao G, Unlu MS, Morkoc H, Blackburn DL. IEEE Trans Electron Devices 1991;38:185.
[2] Liu W, Khatibzadeh A. IEEE Trans Electron Devices 1994;41:1698.
[3] Bayraktaroglu B, Barrette J, Kehias L, Huang CI, Fitch R, Neidhard R, Scherer R. IEEE Trans Electron Devices Lett 1993;14:493.
[4] Salib M, Bayraktaroglu B. IEEE MTT-S Int Microwave Symp Dig 1997;897.
[5] Zhang QM, Hu H, Sitch J, Surridge RK, Xu JM. IEEE Trans MTT 1996;44(11):2001.
[6] Lu K, Perry PA, Brazil TJ. IEEE Trans MTT 1995;43(7):1433.
[7] Samelis A, Pavlidis D. IEEE Trans MTT 1997;45(6): 886.

PERGAMON

Solid-State Electronics 43 (1999) 1437–1444

SOLID-STATE ELECTRONICS

InGaP HBT technology for RF and microwave instrumentation

Tom Low*, Tim Shirley, Craig Hutchinson, Gilbert Essilfie, Wes Whiteley, Bob Yeats, Don D'Avanzo

Microwave Technology Division, Hewlett Packard Co., 1400 Fountaingrove Parkway, Santa Rosa, CA 95403, USA

Abstract

An InGaP emitter HBT IC process developed for RF and microwave instrumentation is described. The process is based on MOCVD epitaxial material, 1 μm critical dimension (CD), G-line, stepper aligned lithography and $SiCl_4$ based reactive ion etching. F_t and F_{max} values of 65 and 75 GHz, respectively are achieved. The HBT technology is well suited for instrument applications in that it can simultaneously achieve both excellent reliability and high performance in terms of broad bandwidth, low phase noise, high gain and linearity. Circuits designed in the process include a Darlington feedback amplifier, which achieves 9.8 dB gain from dc to 20 GHz and dc to 16 GHz dividers. The circuits have been utilized in numerous instrument applications and have resulted in improvements in dynamic range, bandwidth and time-domain jitter. Extensive reliability testing of the InGaP emitter process indicates that MTTF values at $T_j = 150°C$ and $J_C = 6 \times 10^4$ A/cm^2 are greater than 4×10^5 h and are an order of magnitude larger than MTTF values for AlGaAs emitter devices. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

Heterojunction bipolar transistors have emerged as a competitive technology for wireless and communication markets. High efficiency and high power amplifiers have been successfully utilized for output amplifiers in cellular phones [1,2]. Mixed signal HBT technology has been used to implement 10 Gbit/s fiber optic systems [3]. Another potential application for HBT technology is in RF and microwave instrumentation. The instrumentation market tends to be a demanding application because of the simultaneous needs for broad bandwidth, high performance and excellent reliability. In order to cover multiple applications with the same instrument, bandwidths can often range from 100 kHz to 26.5 GHz. High performance is required since the test equipment must have superior performance to the device under test. The useful life of an instrument often exceeds 20 years, therefore requiring excellent component reliability.

It is well established that HBT frequency performance is optimized at relatively high current densities [4]. However, the primary reliability failure mode in HBTs, β degradation, is believed to be accelerated by increasing current density [5]. The challenge for HBTs in microwave instrument applications is to simultaneously achieve high values of F_t and F_{max} at relatively high current density while maintaining reliability.

HBTs offer an attractive complement to the existing 0.25 μm PHEMT technology for microwave instruments. The higher power and gain density provided by HBTs leads to significantly smaller chip sizes and potentially lower cost per function. For example a transconductance of 100 mS can be achieved with a 3×3 μm^2 HBT operating at 5 mA. A 250 μm wide PHEMT

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00086-6

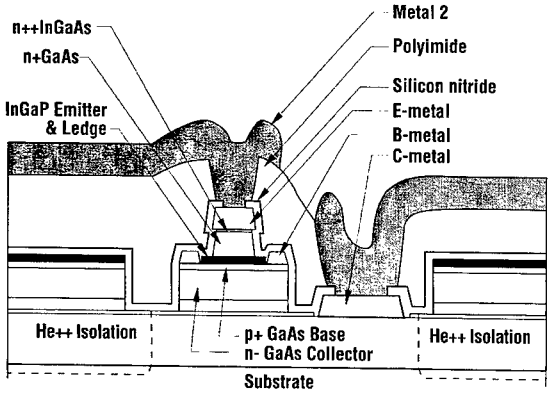


Fig. 1. Schematic cross-section of a completed InGaP emitter HBT.

operating at 100 mA is required to achieve the same transconductance. The high HBT intrinsic voltage gain, 2000 versus 20 for the PHEMT, allows extensive use of feedback to minimize circuit performance variations over process distributions. HBTs achieve comparable F_t (~65 GHz) to 0.25 μm PHEMTs with much larger (~1 μm) critical dimensions (CDs). Furthermore HBT performance variations are dominated by variations in epitaxial growth rather than sub-micron dimensions and are inherently more uniform and better controlled, enabling higher levels of integration and more functionality on a chip. In addition, the inherently lower $1/f$ noise corner frequency, 20 kHz compared to 100 MHz for a PHEMT or MESFET, is a major advantage for low phase noise oscillators and dividers. In general, HBT integrated circuits are well positioned to meet the key instrumentation requirements of high linearity, broadband gain, low phase noise, moderate levels of integration and a high degree of reliability.

This paper describes an InGaP emitter, HBT technology that has been developed to meet these requirements and address the instrumentation market. Section 2 describes the HBT process and epitaxial structure. Vehicle circuits and several instrument applications are explained in Section 3. Reliability characterization of the InGaP process is presented in Section 4 and compared to similar reliability tests on AlGaAs processes. Conclusions are presented in Section 5.

2. InGaP emitter HBT process

The HBT process is based on MOCVD epitaxial wafers and G-line, stepper aligned lithography. A schematic cross-section of the device is shown in Fig. 1. The epitaxial material is manufactured by Kopin Corp. [6] and features a 4000 Å collector drift region, an 800 Å carbon doped GaAs base, an InGaP emitter and a highly doped InGaAs contact layer. The process utilizes G-line stepper lithography that readily defines the minimum geometries, 2×2 μm^2 emitters and 1×1 μm^2 vias. $SiCl_4$ based RIE is used to define the emitter mesa, emitter ledge/base contact, and the base–collector mesa. Non-alloyed, TiPtAu contacts are used for the emitter and base while an alloyed Au/Ge/Ni contact is used for the collector. Helium implantation isolates passive and active devices. Device passivation is realized with PECVD Si_3N_4. Planarization and an intermetal dielectric layer are achieved with a 1 μm thick layer of Dupont's PI2555 polyimide.

Table 1
Key HBT parameters for a 2×2 μm^2 emitter device with $V_{ce} = 1.5$ V and $J_c = 6 \times 10^4$ A/mm^2. Measurements on HBTs from 38 wafers are included in the statistics

	Mean	Total standard deviation (%)
F_t(GHz)	65	7
F_{max} (GHz)	75	5
β	132	20
R_b (Ω)	73	14
R_e (Ω)	16	19
R_c (Ω)	8.8	7
C_{bc} (fF)	12.6	5
BV_{cco} (V)	8.3	3

The HBT transistors are integrated with a standard MMIC passive component platform which includes 22 Ω per square Ta_2N thin film resistors, MIM capacitors with 1500 Å PECVD Si_3N_4, 2 μm thick TiPtAu interconnect metal and dry etched backside vias in 100 μm thick substrates. Because of the demanding performance requirements of the instrumentation applications the HBTs are operated at a relatively high current density of 6×10^4 A/cm^2. Nominal values of key process monitors include: $F_t = 65$ GHz, $F_{max(MAG)} = 75$ GHz, $\beta = 132$, $R_e = 16$ Ω and $BV_{ceo} = 8.3$ V, see Table 1.

Relatively large critical dimensions, extensive use of dry etching and excellent control of both the epitaxial growth and wafer fabrication technology all contribute to maintaining tight parameter control and high yields. Total standard deviations of key parameters are less than 10% except for β, R_b and R_e, which are less than or equal to 20%, see Table 1. The variations in β and R_b are partially caused by intentional base doping variations which will be eliminated. The variation in R_e is expected to decrease significantly by replacing a wet etch of the emitter contact layer with dry etching. As a result of the tight parameter distributions and relatively large critical dimensions high circuit yields have been observed. For example a 1300 transistor inverter chain IC used for process monitoring typically yields greater then 70%.

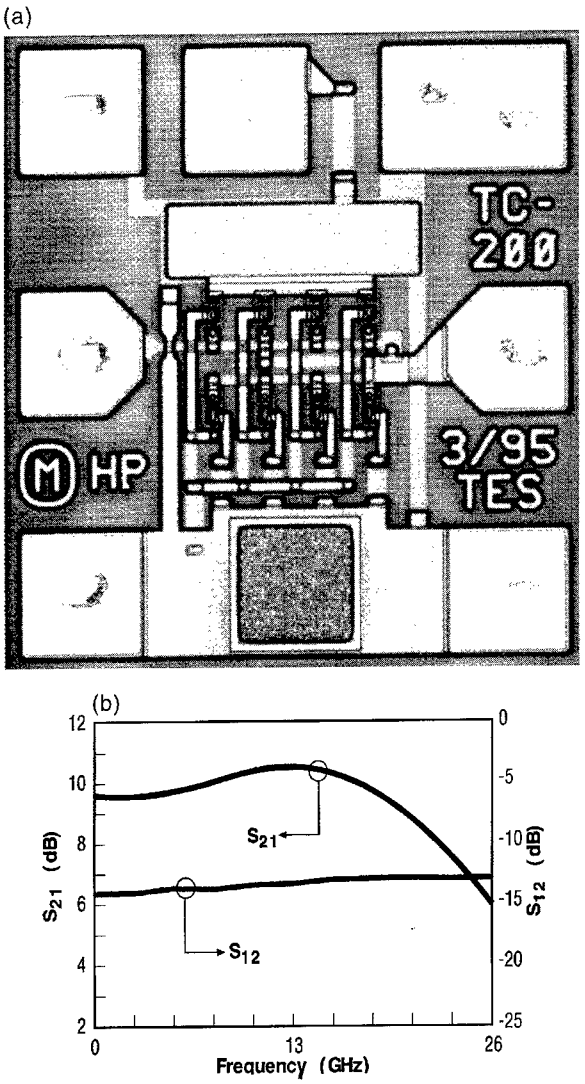


Fig. 2. dc to 20 GHz Darlington feedback amplifier: (a) chip photomicrograph and (b) S_{21} and S_{12} versus frequency.

3. HBT circuits and instrument applications

Several integrated circuits have been developed to address the instrumentation market. A typical example is a broadband Darlington feedback amplifier that provides 9.8 dB gain from dc to 20 GHz. This product is designed as a cascadable gain block which uses feedback to minimize sensitivity to process variations. It features 50 Ω input and output matches, a 0.41 × 0.46 mm^2 chip size and 177 mW dc power dissipation. A chip photomicrograph and a typical S_{21} versus frequency characteristic are diagrammed in Fig. 2. This amplifier has been used as a broadband pre-amp for receiver applications to 20 GHz, a driver amplifier for RF output stages and an LO driver amp. An example of a pre-amp application is a 12.4 GHz frequency counter, the HP53131A.

A customized version of this amplifier was developed for the specific purpose of providing improved reverse isolation. Higher levels of isolation are achieved by adding a common base gain stage to the front-end of the amplifier. The resulting isolation amp has been utilized in a microwave network analyzer, HP8753E, to improve its performance when characterizing high rejection filters. Without the presence of the isolation amplifier spurious IF mixing products can produce undesirable side-bands which appear as spikes in the stop band of the filter response. The HBT isolation amp provides 55 dB of reverse isolation enabling accurate characterization of filters with rejection bands up to 110 dB, as compared to only 85 dB without it (Fig. 3).

A second product family is comprised of a series of high performance frequency dividers. The family consists of dc to 16 GHz divide-by-2, 4 and 8 pre-scalers. On-chip pre- and postamplifiers have been incorporated in the design to improve input sensitivity and output signal swing. The divide by 8 chip photomicrograph and input sensitivity window are shown in Fig. 4. The chip operates over most of the frequency range with input powers between −20 and +20 dBm. The input sensitivity decreases to −15 to +7 dBm at 16 GHz. Low phase noise, −153 dBc/Hz at 100 kHz offset, is also achieved. The chip size is 1.33 × 0.44 mm^2.

The dividers have been designed into a wide variety of applications including phase locked loops, trigger circuits and frequency converters where bandwidth, phase noise and dynamic range are critical. For example the divide by 2 is used in a high performance spectrum analyzer, the HP8563-K35. The K35 option, which is enabled by the HBT divider, improves the dynamic range of the spectrum analyzer from roughly 65 to 75 dB for adjacent channel power rejection (ACPR) measurements in wideband CDMA systems. Other examples include the HP53131A 12.4 GHz counter and the HP83480 option 100 communications analyzer. In the latter case, the high performance HBT divider extends the frequency range of the scope trigger from 2 to 12 GHz. The time domain jitter of the trigger is limited to 0.7 ps rms over most of the band, Fig. 5, as a result of the excellent phase noise performance of the HBT divider.

4. Reliability characterization

Maintaining a high level of reliability, even while operating at relatively high current densities, is essential for components addressing the instrumentation market. The original HBT process was based on an

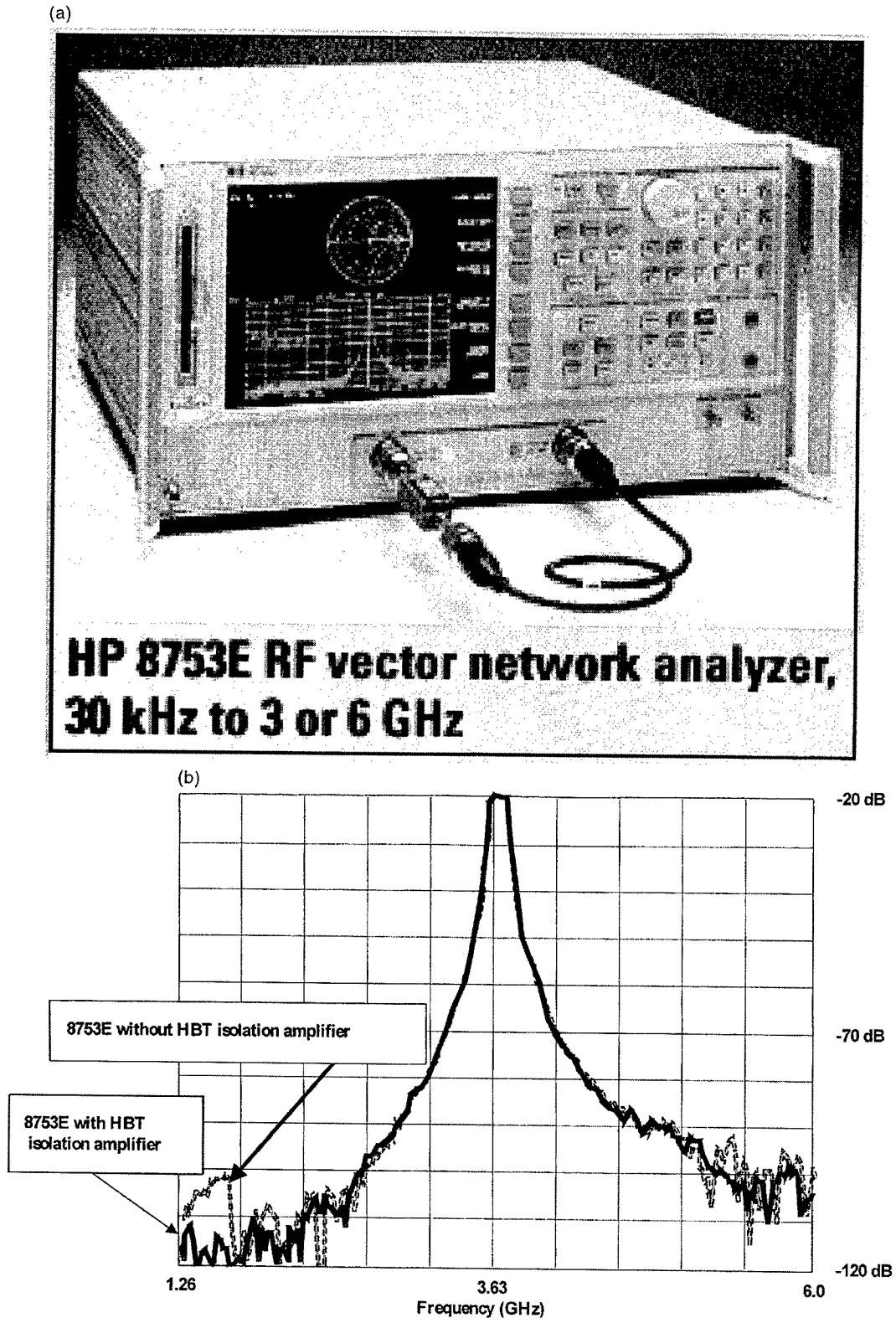


Fig. 3. (a) HP8753E network analyzer and (b) high dynamic range filter response with and without the HBT isolation amplifier.

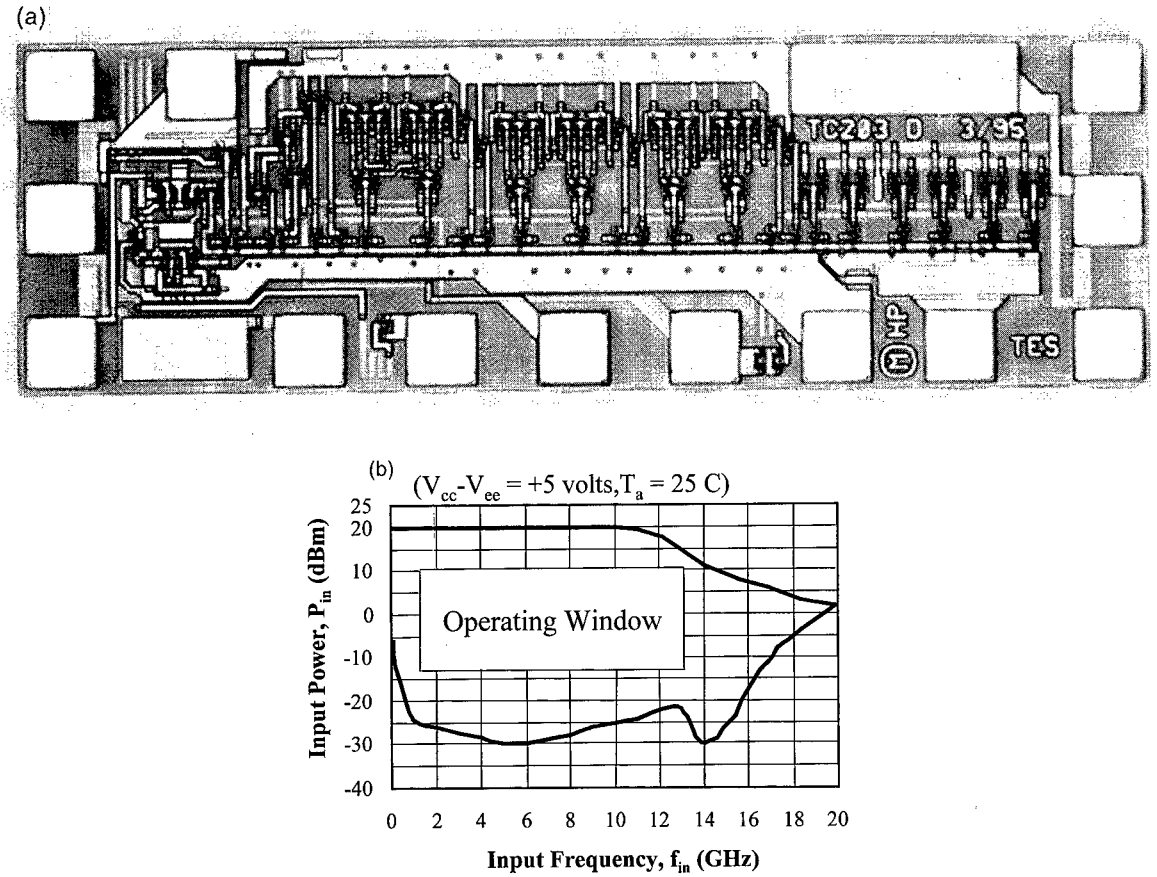


Fig. 4. dc to 16 GHz frequency divider: (a) chip photomicrograph and (b) input sensitivity curve.

AlGaAs emitter process. While performance and manufacturability of the AlGaAs process was comparable to the InGaP process the reliability was found to be unacceptable for instrument applications. To characterize the reliability of the HBT processes, several hundred Darlington amplifiers and high speed dividers were subjected to high temperature operating life (HTOL) stress tests. In all cases, collector currents were held at the maximum use condition of $J_c = 6 \times 10^4$ A/cm^2. Peak junction temperatures were varied between 230 and 363°C. Failures were defined as a 1 dB change in S_{21} for the amplifiers and a 1 dB decrease in output power for the dividers. The dominant failure mode for both the AlGaAs and InGaP processes was β degradation as illustrated in the Gummel plot in Fig. 6. Since the β degradation was abrupt in time and relatively large in magnitude MTTF values are fairly insensitive to the precise failure criteria. A total of 5×10^7 device hours under stress were accumulated during the reliability tests.

The AlGaAs process resulted in unexpectedly low values of median time to failure (MTTF). The MTTF values extrapolated to the maximum operating conditions of $T_j = 150$°C and $J_c = 6 \times 10^4$ A/cm^2 ranged between 3×10^3 and 8×10^4 h with an average activation energy, E_a, of 0.57 eV. To benchmark the AlGaAs process and to establish an existence proof of acceptable AlGaAs HBT reliability, AlGaAs broadband amplifiers and dividers were purchased from other commercial suppliers and tested on HTOL. The

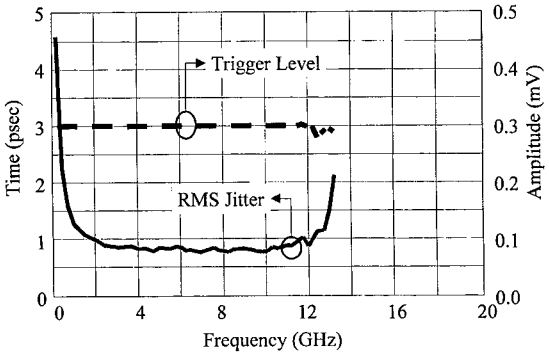


Fig. 5. Trigger jitter versus trigger frequency for an HP83480 option 100 communications analyzer

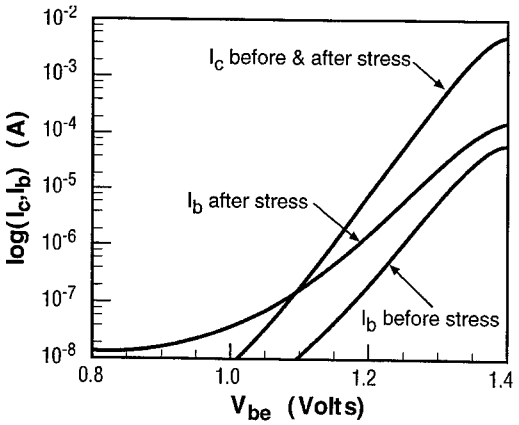


Fig. 6. A representative room temperature Gummel plot for a 3×3 μm^2 InGaP emitter HBT before and after stress at $J_c = 6 \times 10^4$ A/cm^2 and $T_j = 333$°C showing stable I_c and a large increase in I_b after stress which results in a significant decrease in β.

commercial products were stressed at elevated temperatures and biased at their standard operating conditions. MTTFs and activation energies for the commercial AlGaAs HBTs were comparable to the HP process with the amplifier tests resulting in $E_a = 0.5$ eV and MTTF $= 3.5 \times 10^4$ h and the divider tests resulting in $E_a = 0.65$ eV and MTTF $= 2.5 \times 10^4$ h. Arrenhius plots of both the HP and commercially available AlGaAs ICs are compared in Fig. 7.

Similar reliability testing on the InGaP emitter process resulted in dramatically higher values of MTTF. Based on 5×10^6 circuit-hours of HTOL testing, typical values of MTTF ($T_j = 150$°C, $J_c = 6 \times 10^4$ A/cm^2) are 4×10^5 h with activation energies, E_a, between 0.50 and 1.0 eV and sigmas of 0.5. The relatively large values of MTTF and the small value of sigma insure that failure rates will be acceptably low. The projected fraction of the population failing in 10 years is 0.1 %. The reliability results for 4 InGaP wafers (approximately 200 ICs) are plotted in Fig. 8. There are six additional wafers (300 ICs) which have not reached MTTF at $T_j = 363$°C after 3000 h of stress. HTOL tests are continuing on these wafers.

In contrast to previous investigations [7,8] these results indicate that activation energies can be similar for both AlGaAs and InGaP emitter materials. The previously reported InGaP results ($E_a = 2.0$ eV in Ref. [7] and 1.53 eV in Ref. [8]) are plotted together with the results from the present work in Fig. 9. Note that at high stress temperatures MTTF values for the published results all lie below the 0.6 eV line, representing

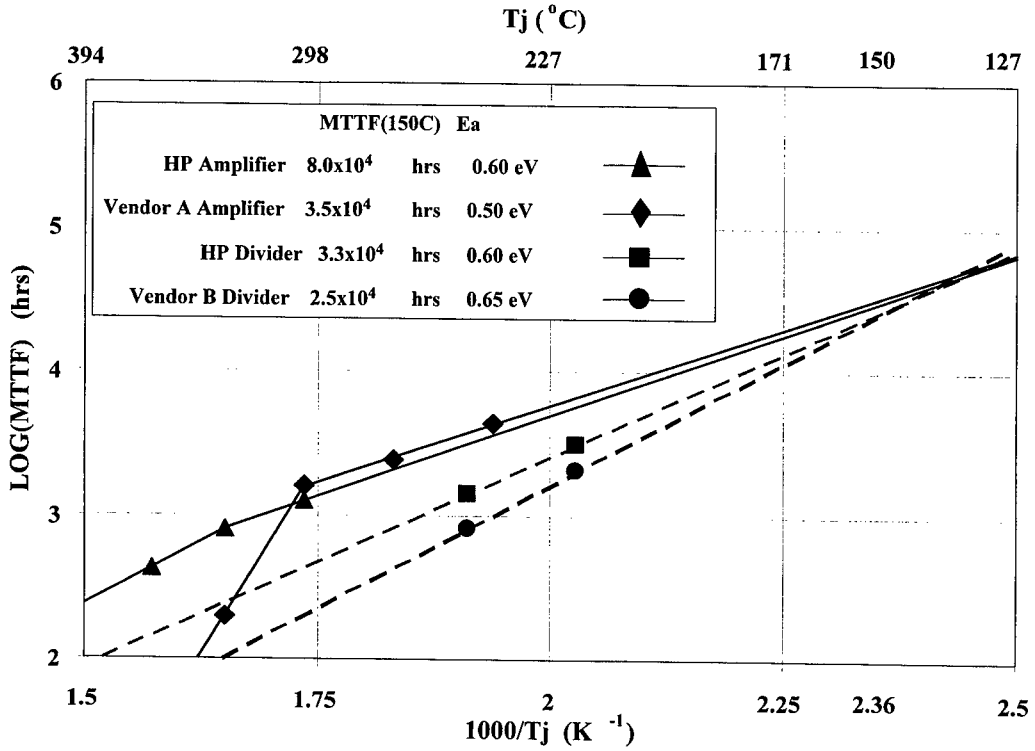


Fig. 7. MTTF versus $1/T_j$ for HP and commercially available AlGaAs emitter amplifiers and dividers. The HP circuits were stressed at $J_c = 6 \times 10^4$ A/cm^2. The commercial circuits were stressed at the standard bias conditions.

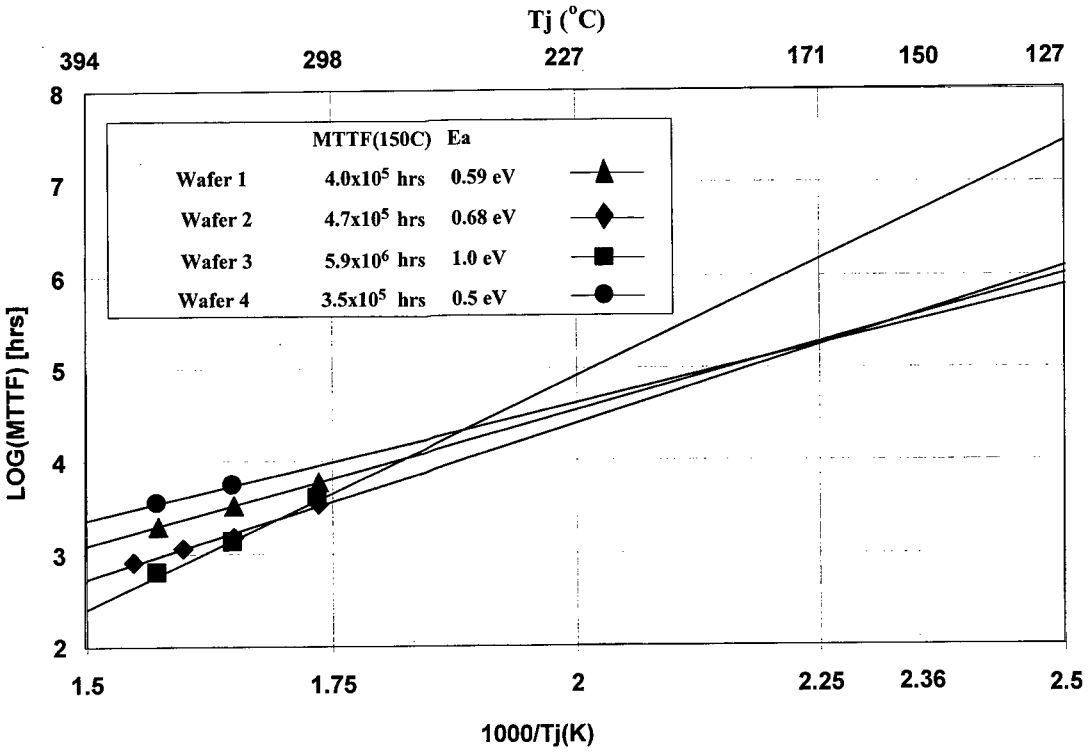


Fig. 8. MTTF versus $1/T_j$ for InGaP emitter Darlington amplifiers from four wafers stressed at $J_c = 6 \times 10^4$ A/cm^2.

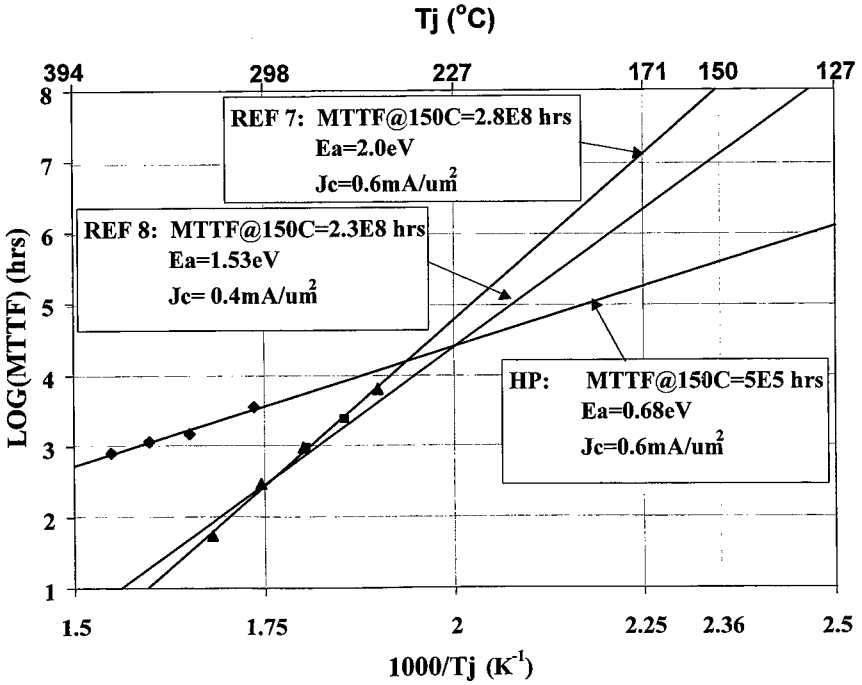


Fig. 9. MTTF versus $1/T_j$ for a representative HP InGaP emitter wafer and data published in Refs. [7,8].

the average InGaP activation energy measured in the present work. This may imply that a low activation energy mechanism is present in the InGaP HBTs in Refs. [7,8] but is masked at the high stress temperatures by an additional higher activation energy mechanism that does not appear to be present in the HP InGaP devices. Reliability test times greater than 10^4 h at temperatures below 230°C would be required to verify the absence of the low E_a mechanism found in both the InGaP and AlGaAs HBTs investigated in this work.

The similar activation energies reported here for both the AlGaAs and InGaP processes suggest that the same failure mechanism dominates in both cases. As has been reported in the literature [7,9] the failure mode is consistent with a rapid increase in emitter-base recombination current potentially initiated by recombination enhanced defect formation. Two-dimensional, physical simulations predict that the recombination rate in the InGaP emitter device should be 1000 times smaller than in the AlGaAs emitter device. The dramatic reduction is primarily due to the larger hole barrier in the InGaP/GaAs heterojunction compared to the AlGaAs/GaAs heterojunction, 0.43 versus 0.25 eV. The smaller recombination current in the InGaP emitter devices offers a possible explanation for the superior reliability of InGaP HBTs. If E_a is interpreted as the formation energy for some recombination center, then the similar E_a values for InGaP and AlGaAs HBTs suggests that β degradation in both device types may be caused by formation of the same recombination center in the p+ base region, which is common to both devices.

5. Conclusion

A high performance InGaP emitter HBT process has been developed for use in RF and microwave instruments. The InGaP HBT is well suited for instrument applications in that it can simultaneously achieve high performance and excellent reliability. The process is based on MOCVD epitaxial material, G-line, stepper aligned lithography and $SiCl_4$ based reactive ion etching. The HBT achieves comparable frequency performance to 0.25 μm PHEMTs with CDs on the order of 1 μm. F_t and F_{max} values are 65 and 75 GHz respectively. Compared to PHEMTs, HBTs offer higher intrinsic gain, lower $1/f$ noise, improved process control and higher levels of integration. Circuit vehicles such as dc to 20 GHz broadband amplifiers and dc to 16 GHz frequency dividers have been designed into numerous instrument applications. The HBT ICs have resulted in improved instrument performance such as extended bandwidth, improved dynamic range and minimized time domain jitter.

An order of magnitude improvement in MTTF has been achieved by replacing the AlGaAs emitter with InGaP. MTTF values exceed 4×10^5 h for the InGaP process. Activation energies for both InGaP and AlGaAs devices are comparable suggesting that a similar failure mechanism is present in both types of devices.

References

[1] Walters P, Lau P, Buehring K, Penney J, Farley C, McDade P, Weller. IEEE GaAs IC Symposium Tech. Digest, 1995. p. 67.
[2] Yoshimasu T, Tanba N, Hara S. IEEE Microwave and Guided Wave Letters 1994;4:65.
[3] Willemsen H. IEEE GaAs IC Symposium Tech. Digest, 1996. p. 10.
[4] Kobayashi, KW, Tran LT, Block T, Cowles J, Oki AK, Streit DC. IEEE GaAs IC Symposium Tech. Digest, 1995. p. 71.
[5] Sugahara H, Nagano J, Ogawa K. IEEE GaAs IC Symposium Tech. Digest, 1993. p. 115.
[6] Pan N, Elliott J, Knowles M, Vu DP, Kishimoto K, Twynam JK, Sato H, Fresina MT, Stillman GE. IEEE Elect Device Letters 1998;19:115.
[7] Takahashi T, Sasa S, Kawano A, Iwai T, Fujii T. IEDM Tech. Digest, 1994. p. 191.
[8] Beaulieu C et al. JEDEC GaAs REL Workshop Proceedings, 1996. p. 11.
[9] Henderson T et al. GaAs, MANTECH, 1997. p. 38.

PERGAMON

Solid-State Electronics 43 (1999) 1445–1450

SOLID-STATE ELECTRONICS

Passivation of InP-based HBTs

R. Driad[a], S.R. Laframboise[a], Z.H. Lu[a,b], S.P. McAlister[a,*], W.R. McKinnon[a]

[a]*Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, Canada, K1A 0R6*
[b]*Department of Materials Science, University of Toronto, Toronto, Canada*

Abstract

The passivation of the exposed semiconductor surfaces in a device is necessary to ensure good device performance. In this study we examine how different surface treatments and geometric device designs can be used to improve the dc performance of InP-based heterostructure bipolar transistors. We show how sulphur and UV-ozone treatment improves the performance of large-area devices and illustrate the effects of the treatment through XPS measurements on the semiconductor surfaces. We also demonstrate that a combination of UV-ozone + HF can lead to improved performance of small, high-frequency devices. This approach is contrasted with results using a layer structure with a thin emitter, where the emitter is left on the device to self-passivate the base surface. In this case the thin emitter layer on the base is depleted and no passivation treatment is needed. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Passivation of the semiconductor surfaces in a device is often necessary to stabilize the surface and prevent the surface (surface states) from adversely affecting the device performance [1]. The passivating layer on a device may not only be used to protect and stabilize the semiconductor interface, but also to provide an isolating dielectric layer for the devices. Thus one wishes to protect against the environment, isolate the contacts from the devices, stabilize the surface and have a stable high resistivity uniform layer covering the devices. Not only are the chemical properties of the semiconductor/dielectric interface important but also the electrical, structural and morphological properties of the passivating/isolating material layers. Along with the choice of passivating material the cleaning and surface preparation of the surface, before the passivating layer is deposited, can also be crucial. All these issues make the passivation and isolation task a demanding and difficult one [2].

Extensive literature has been published on the improvement of III-V semiconductor surface/interface electrical characteristics after sulphur treatment. It has been a common approach since the work of Sandroff et al. [3] on GaAs. Significant reduction of the surface recombination velocity and enhancement of photoluminescence intensity have been observed on III-V compounds and related materials after sulphur treatment [4,5]. It has also been shown by several authors [6,7] that sulphur passivation results in lower interface state densities with improved device performance.

In an alternative approach the combination of ultraviolet (UV) irradiation and ozone has been extensively used in device processing [8]. This process relies on both 253.7 and 184.9 nm wavelengths from the UV source and can be effective in removing organic and

* Corresponding author. Tel.: +1-613-993-2514; fax: +1-613-990-0202.
E-mail address: sean.mcalister@nrc.ca (S.P. McAlister)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00087-8

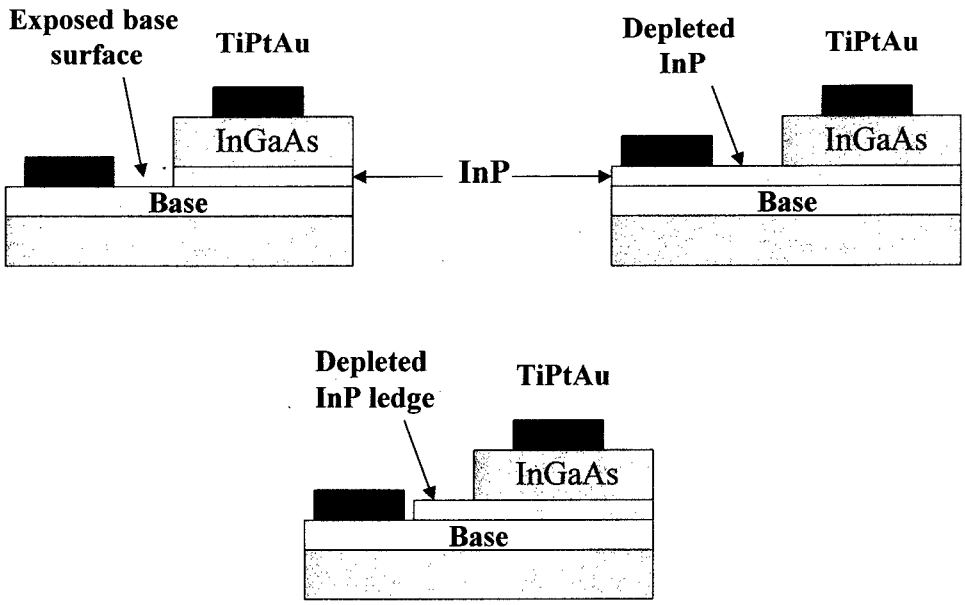


Fig. 1. Schematic cross section of the devices, showing different emitter-base mesa structures. The normal structure has an exposed InGaAs base—other structures either have the base contact through a depleted InP layer (self passivated) or onto the base through a via hole in the InP layer (ledge passivated).

non-organic species from the semiconducting surface. It has been found to produce uniform thin stoichiometric native oxides [2] and excellent passivation properties and good electrical characteristics can be obtained with UV-ozone treatments, for instance, for the fabrication of Schottky MIS diodes [9].

In this paper we report a qualitative study on semiconductor surface passivation of InGaAs/InP based heterostructure bipolar transistors (HBTs) using different techniques. An HBT is a useful test vehicle, because of its sensitivity to surface recombination effects. The III-V semiconductors layers used in these devices can have surfaces that are adversely affected by exposure to the air and capping with dielectric layers such as SiO_2. Most notorious in this respect are many examples in the literature for AlGaAs/GaAs HBTs [10,11]. InP-based HBTs are cited to be better than GaAs-based devices because InGaAs, the narrow bandgap material used in the base, and InP have lower surface recombination velocities than GaAs [12,13]. (InGaAs is the normal lattice-matched composition $In_{0.53}Ga_{0.47}As$.) For InP thermal oxidation is not a viable solution to treat the semiconductor surface since this induces phosphorus vacancies in the bulk and elemental phosphorus at the oxide interface—this giving rise to an electrically poor interface. The passivation problem can be addressed with various processing and materials deposited on the vulnerable semiconductor surface or by geometrical designs of the device to prevent the vulnerable surface(s) being exposed. We examine both approaches in this paper.

2. Experimental details

In a typical emitter-up HBT the surface that is crucially important is the emitter-base region which typically exposes a small area of the base and the edge of the emitter—since the device structure normally involves an emitter-base mesa (Fig. 1). The exposed semi-conductor surfaces of importance in our devices are the InGaAs of the base and the InP emitter edge.

The epitaxial layer structures used to fabricate the devices used in this study are described in Table 1. The major difference in the second layer design is that the InP emitter is 30 nm thick. The HBTs were triple-mesa devices fabricated using wet-chemical etching—no dry etching which could cause surface damage was used. (Any chemical residues after etching and using photoresist can be an avoided if appropriate careful cleaning procedures are followed.) In Fig. 1 we show the different cross-sectional structures of the devices studied. In the normal structure (N) the InP is removed from the InGaAs base. In the other devices the thin InP emitter is left intact on the base and the contact to the base is either through the InP, or through a via hole in the InP layer. The latter two structures are referred to as self-passivated and ledge-passivated devices respect-

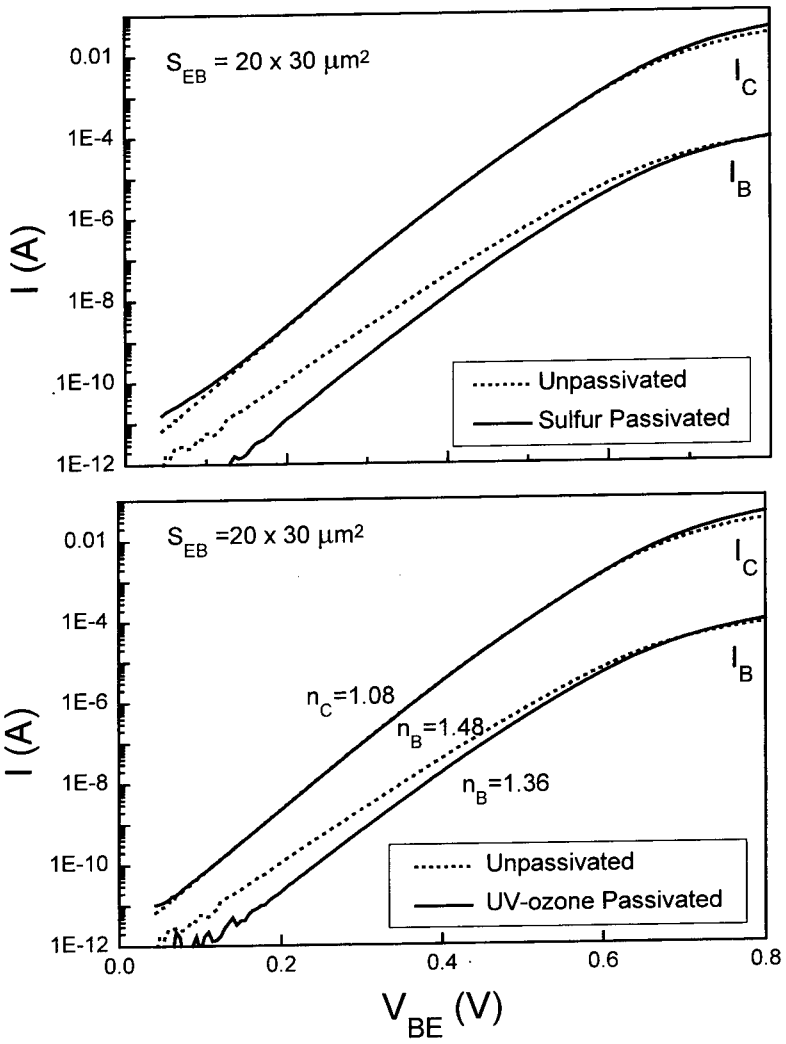


Fig. 2. Gummel plots for normal large area HBTs with and without the surface treatments indicated. There was no oxide coating on these devices.

ively. The use of a thin ledge to protect the base of an HBT has been widely used in AlGaAs/GaAs HBTs [11,13].

For the passivation treatments of the surfaces we used a sulphide treatment and also UV-ozone. The sulphide approach used a $(NH_4)_2S$ aqueous solution (Johnson Matthey yellow sulphide with a S weight concentration of 8.9%). The samples were dipped in the $(NH_4)_2S$ solution for ~1 min at room temperature, then rinsed in acetone solvent, followed by washing in methanol and blown dry in nitrogen. Our UV-ozone exposures were carried out at room temperature (18°C) in a commercial system (model PR-100, UVP Inc.). The samples were placed 10 mm below the UV lamp and the UV ozone generated in air. The samples were exposed to the UV-ozone treatment for times from 5 to 60 min.

The chemical compositions of InGaAs (100) surfaces before and after UV-ozone or sulphur treatment were studied using XPS. The XPS experiments were carried out in a PHI 5500 system equipped with monochromated Al K_α source and a hemispherical electron energy analyzer. As 3 d, In 3d5/2, Ga 2p3/2, and S 2p core levels were recorded and analyzed.

3. Results and discussion

A useful test of the effectiveness of the passivation of an HBT is to examine the low current region of a

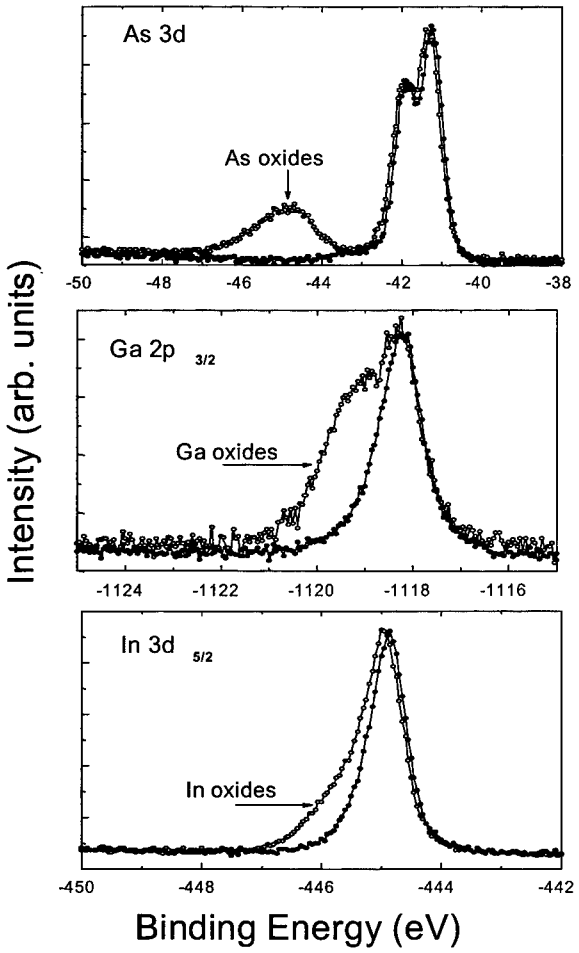


Fig. 3. XPS intensities for InGaAs surfaces: untreated surface data is shown using open data points and the sulfur-treated surface with solid data points. This shows the reduction of the surface oxides using the sulfur passivation.

Gummel plot and the dc gain. Surface recombination causes the collector and base current curves to cross, because the ideality factor of the base current is >1 [10]. (When recombination is a problem the ideality factor is ~ 2 since the recombination centres are most effective if in mid-gap.) Poor Gummel plots can occur even for large-area devices. In our $20 \times 30\ \mu m^2$ devices the perimeter/area for the emitter is $1/6\ \mu m^{-1}$, but since the device is non-self-aligned the exposed base area is at least an order of magnitude larger than the exposed edge of the InP emitter.

In Fig. 2 we show the effect of sulphur passivation on a normal large area device and the effect of a UV-ozone treatment. The XPS data in Fig. 3 clearly show that the air-exposed surface (open circles) is covered with a few nm thick layer of mixed As, Ga and In oxides. These native oxides are known to produce non-radiative recombination defect centers. These surface oxides were completely removed by S treatment, as concluded from the core level spectra recorded from the S-treated surface (solid circles). By comparing S-treated and in situ cleaved surfaces, we concluded in previous studies [14,15] that S forms ordered Ga–S–Ga and In–S–In bridge bonds along [001] on GaAs (100) and InP (100) surfaces, respectively. We suggest that the sulphur treatment results at first in slight etching of the semiconductor surface (this removes the native oxides as well as the surface and near surface defects) and then the passivation of the surface states [16,17]. The overall effect is to reduce the interface state densities and the surface recombination.

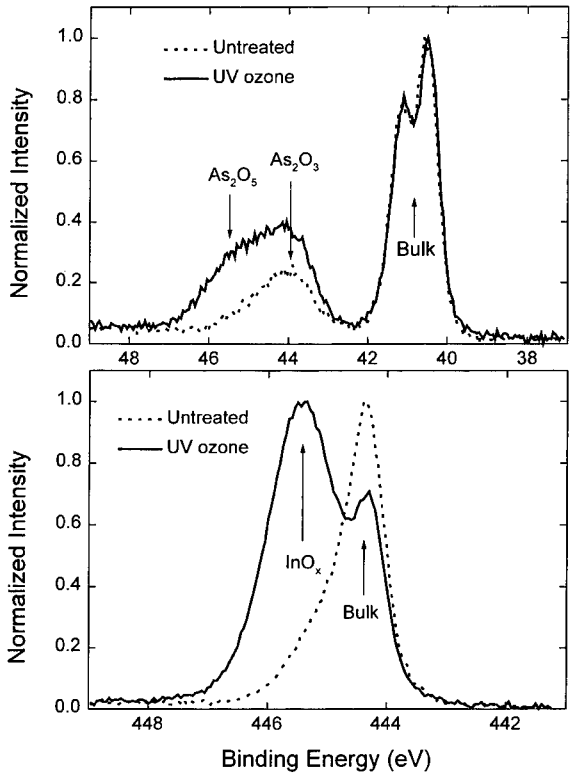


Fig. 4. Normalized XPS intensities for untreated and UV-ozone-treated InGaAs surfaces, showing the growth of As and In oxides.

As shown in Fig. 2, the UV-ozone treatment yields an improvement of the HBT current gain of almost one order of magnitude at low collector currents. The improved characteristics of the UV-ozone treated devices are correlated with an efficient reduction of the excess current originating from surface recombination. This was also confirmed by the decrease of the base ideality factors. Based on XPS studies, we found that the UV-ozone treated InGaAs (100) surface is covered

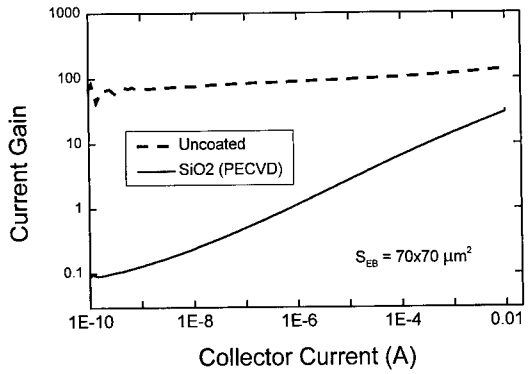


Fig. 5. DC current gain for a large area devices with or without a PECVD oxide coating layer. No passivation treatment was done before the oxide deposition.

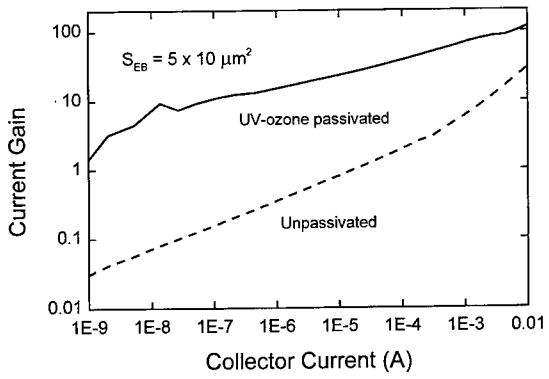


Fig. 6. DC current gain for small-area high frequency devices. Both devices were PECVD-oxide coated, the device with the higher gain had the UV-ozone+HF surface treatment prior to oxide deposition. The layer structure had the 150 nm thick InP emitter, as in Figs. 2 and 5.

with As_2O_3, As_2O_5, Ga_2O_3, and InO_x (Fig. 4). (In contrast for InP Besland et al. [18] claim that after a 10 min UV ozone treatment the interface composition is $InPO_4$.) As reported in our previous study [19], we found that the UV-ozone treatment not only produces a surface oxide film, which removes defective surface layers induced during device processing, but also restores the stoichiometric composition at the semiconductor surface when the UV-ozone exposures are about 10 min.

We suggest that the following occurs during surface passivation by UV-ozone: removal of surface and near surface defects through oxidation (these defects could be interstitial such as elemental As, vacancies, metallic and organic contaminates and could be produced during device fabrication) and restoration of a stoichiometric semiconductor surface, which is achieved by the removal of surface vacancies and interstitials.

The improvements in the electrical characteristics (Fig. 2) are not necessarily maintained after the deposition of an oxide. For unpassivated devices that have reasonable characteristics, deposition of an oxide can adversely alter the characteristics of even large-area devices. This we show in Fig. 5. A newly fabricated unpassivated device is worse after deposition of a PECVD oxide, indicating that the interface between the semiconductor and the deposited oxide is electrically poor. Removing the PECVD oxide and exposing the oxide-semiconductor interface to HF can restore the original performance. Not only does HF remove, selectively, the deposited oxide and also the native oxides, but it also passivates dangling bonds at the semiconductor surface with hydrogen [20,21] or fluorine [22]. However an HF treatment alone on InGaAs does not passivate the surface if an oxide is subsequently deposited on the surface.

It is possible to obtain good results using the normal designed devices in a self-aligned high frequency device, as long as the advantages of the passivation treatment can be preserved after a dielectric layer (such as SiO_2) is deposited. We have accomplished this by using a UV ozone+HF treatment prior to the oxide deposition. The results for the dc gain for a device having this treatment compared with those from an unpassivated device show the effectiveness of this passivation scheme (Fig. 6). We suggest the reasons for the success of this approach include the facts that the 10 min UV ozone treatment grows an oxide which consumes some of the electrically-poor semiconductor surface and returns the semiconductor interface to its original composition, then the HF removes most of the oxides that were created by the UV-ozone treatment—and passivates the semiconductor surface. The surface oxide created by the UV-ozone presumably contained the defects (vacancies, interstitial and other contaminates) trapped at or near the surface.

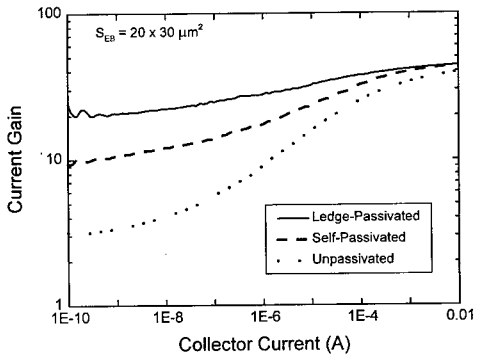


Fig. 7. DC current gain for large area HBTs with the self- or ledge-passivated structures (see Fig. 1). The second wafer design was used for these devices. There was no oxide coating.

A alternative way to deal with the passivation problem directly is not to expose the base region at all, but to leave it covered with a thin layer of the emitter material. (This has been a strategy in GaAs-based devices [13].) This can only work if the layer is depleted (or undoped) or else there will be a simple leakage path from the base contact to the emitter. The contact to the base can be either through a via hole in the InP layer on the base or by a diffused contact through the layer. This we have tried with the two designs (self- and ledge-passivated). The devices employed the second epitaxial layer design with the thin InP emitter. The results are shown in Fig. 7. The ledge passivated device has the highest gain of the three designs. Note there was no oxide or passivation treatment on the normal device. To apply this approach to high frequency devices is more of a challenge. Recently, however, a similar study to ours has been made on InGaP/GaAs HBTs [23]. In this paper a design using a thin emitter that self-passivates the base is shown to be best to improve the gain, and this also helps preserve a good F_{max} as long as the contact through the thin depleted emitter layer is made close to the emitter mesa, in a self-aligned design. Such an approach reduces R_B.

The passivation studies discussed here have not clarified which semiconductor surface in the InP/InGaAs HBTs is most responsible when the performance is poor due to surface passivation problems. Both the InP on the emitter and InGaAs base are involved. For example Kikawa et al. [24] recently reported that use of SiO_2 on InP causes the Fermi level at an InP surface to shift to close to the conduction band edge, resulting in a surface leakage path and excess base current. Because there is a heterojunction between the InP and InGaAs holes cannot get into the emitter especially at low bias. Thus it is most likely that the recombination at the base surface is more important than on the emitter edge. This is supported by the results for the device where InP is left on the base, as in the ledge-passivated design.

4. Conclusions

We have shown how the characteristics of large-area InP/InGaAs HBTs are significantly altered by deposition of oxide on the devices or by common surface treatments. For high frequency devices we have successfully treated the semiconductor surface with UV ozone + HF to passivate the devices and obtain good performance. In an alternative approach we have shown that using a thin InP emitter is a useful approach since it allows a thin ledge to be left on the devices which then do not have the base surface exposed.

References

[1] Nakjima O, Nagata K, Ito H, Ishibashi T, Sugeta T. Jpn J Appl Phys 1985;24:L596.
[2] Ingrey S. J Vac Sci Technol A 1992;10:829.
[3] Sandroff CJ, Nottenburg RN, Bischoff JC, Bhat R. Appl Phys Lett 1987;51:33.
[4] Skromme BJ, Sandroff CJ, Yablonovitch Y, Gmitter T. Appl Phys Lett 1987;51:2022.
[5] Yablonovitch E, Bhat R, Zah CE, Gmitter TJ, Koza MA. Appl Phys Lett 1992;60:371.
[6] Iyer R, Chang RR, Lile DL. Appl Phys Lett 1988;53:134.
[7] Fan JF, Oigawa H, Nannichi Y. Jpn J Appl Phys 1988;27:L–2125.
[8] Vig JR. J Vac Sci Technol 1985;A-3:1027.
[9] Iliadis AA. Elect Lett 1989;25:573.
[10] Tiwari S, Wright SL, Kleinsasser AW. IEEE Trans Electron Devices 1987;34:185.
[11] Liou JJ. Principles and analysis of AlGaAs/GaAs heterojunction bipolar transistors. Norwood: Artech House, 1996.
[12] Jalali B, Pearton SJ, editors. InP HBTs: Growth, processing and applications. Norwood: Artech House, 1995.
[13] Hayama N, Honjo K. IEEE Electron Dev Lett 1990;11:388.
[14] Lu ZH, Graham MJ. Phys Rev B 1993;48:4604.
[15] Lu ZH, Graham MJ, Feng ZH, Yang BX. Appl Phys Lett 1992;60:2773.
[16] Lu ZH, Graham MJ, Feng XH, Yang BX. Appl Phys Lett 1993;62:2932.
[17] Sandroff CJ, Nottenburg RN, Bischoff JC, Bhat R. Appl Phys Lett 1987;50:256.
[18] Besland M-P, Louis P, Robach Y, Joseph J, Hollinger G, Gallet D, Viktorovich P. Appl Surf Sci 1992;56(58):846.
[19] Driad R, Lu ZH, Laframboise SR, Scansen D, McKinnon WR, McAlister SP. International Conference on Indium Phosphide and Related Materials, 1998. p. 459.
[20] Hossain S, Pantano GG, Ruzyllo J. J Electrochem Soc 1990;137:3287.
[21] Yablonovitch E, Alara DL, Chang CC, Gmitter T, Bright TB. Phys Rev Lett 1986;57:249.
[22] Paul JK, Bose DN. J Appl Phys 1991;70:7387.
[23] Sewell J, Bozada C, DeSalvo G, Dettmer R, Gillespie J, Cerny C, Ebel J, Jenkins T, Nakano K, Pettiford C, Quach T, Thompson R, Via D, Welch R, Anholt R, Campman K, Givens M, Tischler M. International Conference on GaAs Manufacturing Technology, 1998. p. 101.
[24] Kikawa T, Takani S, Masuda H, Tanoue T. International Conference on Indium Phosphide and Related Materials, 1998. p. 76.

WIDE BANDGAP DEVICES

PERGAMON

Solid-State Electronics 43 (1999) 1451–1458

SOLID-STATE ELECTRONICS

GaN-based electronic devices

M.S. Shur[a,*], R. Gaska[b], A. Bykhovski[a]

[a]*Electrical, Computer, and Systems Engineering and Center for Integrated Electronics and Electronics Manufacturing, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA*

[b]*APA Optics Inc, 2950 NE 84th Lane, Blaine, MN 55449, USA*

Abstract

We review the parameters and performance of GaN-based pyroelectric and piezoelectric sensors and report on the characteristics of GaN-based High Electron Mobility Transistors (HEMTs) grown on 4H-SiC substrates. The piezoelectric and pyroelectric devices can operate at elevated temperatures, where conventional pyroelectric materials have poor characteristics or even lose pyroelectric properties. The transistors grown on insulating 4H-SiC substrates have a much smaller thermal impedance, exhibit less severe self-heating effects, lower noise, and their microwave performance is less sensitive to a temperature increase than for similar devices on sapphire substrates. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The electron peak velocity in GaN is triple of that for Si, its electron mobility is about twice as large as for Si, and its thermal conductivity is comparable to that of Si (see Fig. 1). With an increase in the doping density, the electron mobility decreases more slowly than, for example, for GaAs. This allows us to obtain record values of the mobility-sheet carrier density products. The AlGaN/GaN material system is capable of supporting the sheet carrier densities of the two-dimensional electron gas up to 1.5×10^{13} cm^{-3} (up to 5×10^{13} cm^{-2} in doped channel structures), which is 5–20 times larger than in the AlGaAs/GaAs materials system [1]. GaN epitaxial layers can be grown on SiC, which allows us to combine superior transport properties of GaN with an exceptional thermal conductivity of SiC [2,3]. Large piezoelectric constants of AlN and GaN can be used in piezoelectric [4–6] and pyroelectric [7,8] sensors and should be accounted in designing conventional electronic devices, such as AlGaN/GaN HEMTs, for enhancing the sheet carrier concentration and reducing leakage current [9]. A recent analysis of hot electrons in GaN quantum wells [10] seems to imply that the breakdown field in such wells will be determined by the AlGaN cladding layers and not by the quantum well material. All of this gives hope that electronic devices based on GaN will reach the same prominence as GaN-based blue/amber, and white light emitters. In this paper, we review recent progress on GaN electronic devices, including piezoelectric and pyroelectric sensors, detectors of microwave radiation [11], and field effect transistors with emphasis on AlGaN/GaN on insulated SiC substrates that can utilize superior physical properties of the GaN-based material system. An example of such a device is a Doped Channel Heterostructure Field Effect Transistor grown on SiC. More recently, GaN-based HEMTs have been demonstrated on insulated 4H-SiC substrates. These devices hold promise for superior power microwave performance and have already yielded record micro-

* Corresponding author. Tel.: +1-518-276-2201; fax: +1-518-276-2990.

E-mail address: shurm@rpi.edu (M.S. Shur)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00088-X

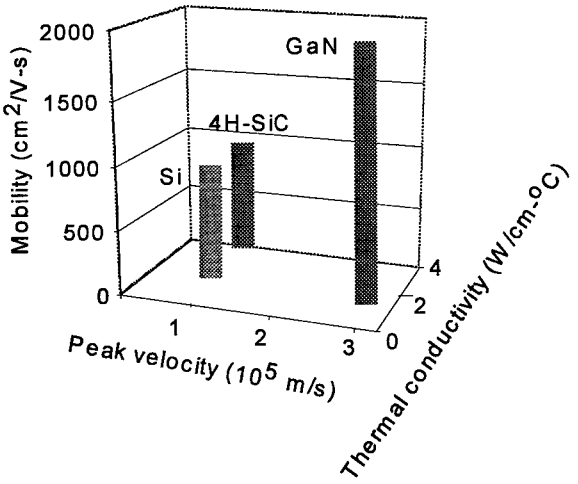


Fig. 1. Materials properties of Si, SiC, and GaN.

wave power [12]. Table 1 gives a brief summary of GaN-based electronic devices.

2. GaN piezoelectric and pyroelectric sensors

GaN-based materials usually have the wurtzite crystal structure and are grown in the (0001) direction. This axis is a polar axis, and these materials have strong lattice polarization effects, which allows us to use them in piezoelectric and pyroelectric sensors. N-type GaN exhibits strong pyroelectric effect even in samples, where the contacts were placed on top of the GaN film grown in the (0001) direction. In this case, the pyroelectric signal is proportional to a small deviation (3–4°) of the direction of the GaN growth from the (0001) axis, see Fig. 2.

The spontaneous polarization in a pyroelectric crystal is usually compensated by ambient positive and negative charges attracted to the charged surface of a pyroelectric crystal from air and/or from the bulk of the material. However, the magnitude of the polarization depends on temperature, and, when temperature changes, the dynamic variations of the

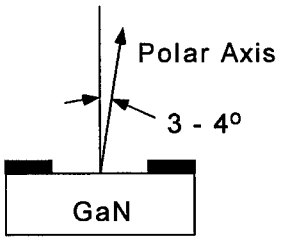


Fig. 2. Schematic of GaN pyroelectric sample [7,8].

Table 1
GaN-based electronic devices

Device	Status	Possible applications
Schottky barrier diode	Demonstrated for variety of metals	Switch, FET building block
p–n junction	Demonstrated on both regular and Lateral Epitaxial Overgrowth material	Photodetector, switch, BJT building block
GaN MESFET	Demonstrated	High temperature digital circuits
GaN MISFET	Demonstrated	High temperature digital circuits, non-volatile memories
AlGaN/GaN HFET	Demonstrated on sapphire and SiC substrates with record power levels	High-temperature, high-power microwave, high-temperature digital circuits
AlGaN/GaN and GaN/SiC HBTs	Demonstrated	High-temperature, high-power microwave, high power switches
GaN-based piezoelectric and piezoelectric sensors	Demonstrated for GaN and AlGaN/GaN devices	Pressure sensors, especially for high temperature applications
GaN pyroelectric sensor	Demonstrated for primary and secondary pyroelectric effect	Temperature sensors, especially for high temperature applications

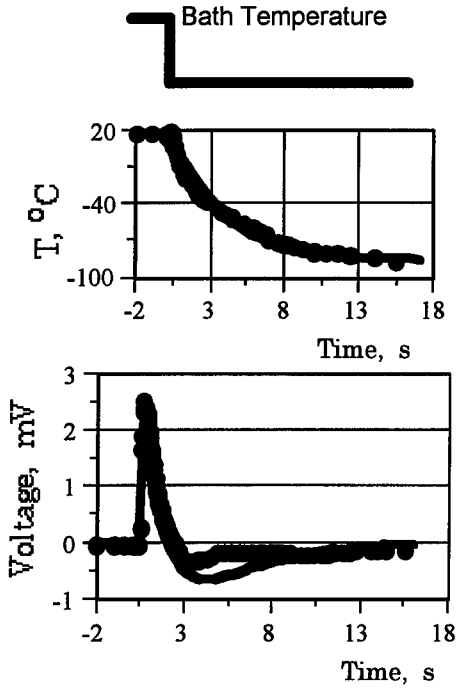


Fig. 3. Bath temperature (a), sample temperature (b), and pyroelectric voltage (c) vs time [7]. Dots are measured data, solid curves are calculated using our model. $\Delta T_0 = 108°C$.

pyroelectric charge may be monitored. These charge variations include the variation of the spontaneous polarization (primary pyroelectric effect) and the variations caused by changes in temperature-related strain (secondary pyroelectric effect). The primary pyroelectric effect is dominant under the condition of a fast heat transfer, which can be implemented, for example, when a sensor is immersed in an ambiance with a small viscosity and /or high flow velocity. In such moving medium (for example, hot gas flow), GaN-based sensors will generate a voltage response which is proportional to the heat (and thus, gas) flow rate. The primary pyroelectric effect has been studied in Ref. [7], and the secondary pyroelectric effect has been investigated in Ref. [8]. The measured pyroelectric voltage coefficient of n-type GaN is on the order of 10^4 V/mK and is comparable to those of the pyroelectric ceramics [13]. The pyroelectric voltage, V, can be described by the simple equation [7,8]:

$$\dot{V} = \alpha_1 \dot{Q} + \alpha_2 \dot{W} + \dot{V}_{\text{piezo}} - \frac{V}{\tau_s} \tag{1}$$

Here dots denote time derivatives,

$$Q = \frac{Sh}{V_0}(T - T_\infty) \tag{2}$$

is a thermal flow per unit volume, V_o and S are the sample volume and surface, h the heat transfer coefficient,

$$W = c_p \rho (T - T_0) \tag{3}$$

is the heat accumulated (dissipated) by the sample, c_p is the sample specific heat, r is the sample density, T is temperature, T_o and T_∞ are the initial sample temperature and bath temperature, respectively, τ_s is the pyroelectric charge relaxation time, α_1 and α_2 are constants. The first term in the right-hand side of Eq. (1) represents the response to the heat flow, the second term accounts for the response to the heat accumu-

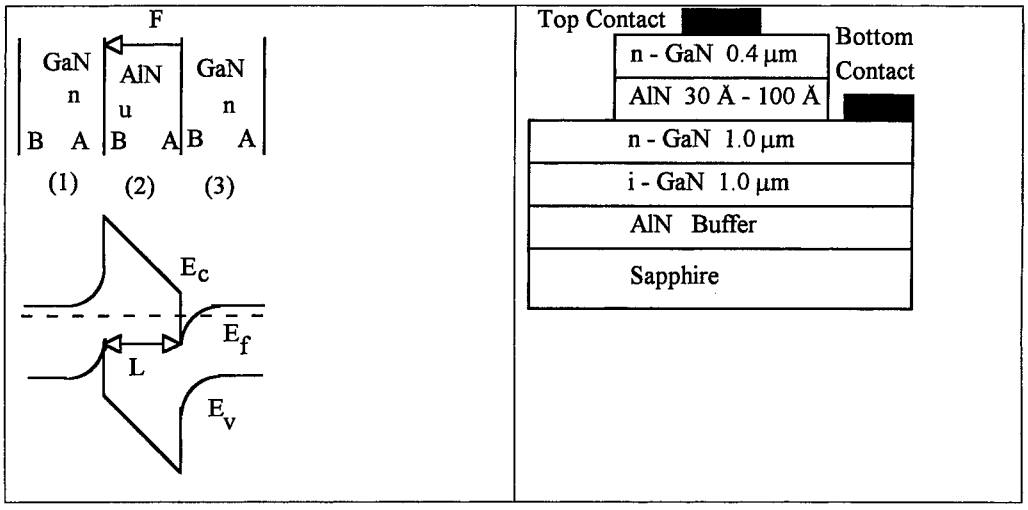


Fig. 4. Band diagram of the symmetrically doped SIS structure (a) and GaN-AlN-GaN epitaxial structure (b) [15].

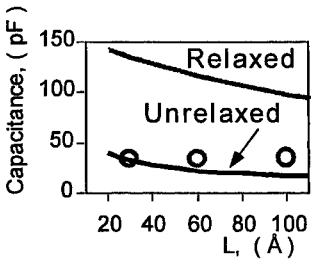


Fig. 5. Capacitance of GaN-AlN-GaN SIS structure at zero bias voltage as a function of AlN layer thickness. Circles are experimental data, solid and dash lines are theory for fully unrelaxed and fully relaxed structures, respectively [15].

lation (dissipation), the third term represents the piezoelectric voltage, V_{piezo}, caused by the strain from non-uniform thermal expansion, i.e. the secondary piezoelectric effect, Refs. [7,8]. Fig. 3 (from [7]) shows that the theory of the pyroelectric effect in GaN samples based on Eqs. (1)–(3) is in good agreement with experimental data. The pyroelectric voltage coefficient extracted from these data is comparable to that of the pyroelectric ceramics such as PZT or $BaTiO_3$, and is proportional to the heat flow density dissipated (accumulated) by the sample. A characteristic value of heat flow conversion into electricity (extracted from our data using our model of the pyroelectric effect in GaN) is $P_f = dF/dQ = 0.0023\ V\ m^2/W$.

3. Piezoelectric sensors

Strongly pronounced piezoelectric properties play a key role in strained GaN-AlGaN and InGaN-GaN multilayer structures. The lattice-mismatch-induced strain generates polarization fields. In GaN-AlN-GaN semiconductor-insulator-semiconductor (SIS) structures with the growth axis along a (0001) crystallographic direction (see Fig. 4), the strain-induced electric fields can shift the flat band voltage by 1.5 V and produce an accumulation region on one side and the depletion region on the other side of the AlN insulator [14]. As a consequence of the asymmetry in the space charge distribution, the capacitance-voltage (C–V) characteristics and the current-voltage (I–V) characteristics of the symmetrically doped SIS structures become asymmetrical. The degree of the asymmetry depends on strain. This allows one to develop C–V and I–V characterization techniques for SIS structures and FETs, extract critical thicknesses and demonstrate the gradual elastic strain relaxation process with an increase in strained layer thickness [15–17]. To determine the elastic strain relaxation, Bykhovski et al. [17] plotted the capacitance at zero bias as a function of AlN thickness, L (see Fig. 5). Donor concentrations were $2 \times 10^{17}\ cm^{-3}$. In the same Figure, we show the theoretical curves for the symmetrically doped structure corresponding to totally relaxed and totally unrelaxed films. The experimental result for 30-Å AlN film is in excellent agreement with the calculation for the totally unrelaxed structure. Thicker AlN films seem to be partially relaxed. Therefore, the starting point for the generation of misfit dislocation corresponds to $L \geq 30$ Å. This conclusion is in agreement with experimental studies of strained-layer GaN/AlN superlattices [18]. In this work the upper limit for the complete relaxation was estimated to be around 100 Å.

Piezoelectric effects strongly affect performance of light emitting devices based on III-N quantum well structures. They can dramatically change the selection rules for the interband transitions in III-Nitride quantum wells and multiple quantum wells (MQWs). Strain-induced electric field causes the spatial separation of electron and hole inside the quantum well [19]. As a result, the optical matrix element for the lowest conduction-band-first heavy hole transition can be dramatically reduced, and transitions forbidden in strain-free structures can exist. Also, strain-induced electric field causes a significant reduction of an apparent band gap, which, in turn, results in a red shift in optical spectra and is reflected in photoemission features [20]. Electroluminescence measurements in InGaN/GaN quantum wells demonstrated the blue shift of the electroluminescence caused by an increase of the driving current, which was partially explained by the screening of the piezoelectric field [21]. A strain-induced modulation of optical transitions in quantum structures can be utilized for development of high-speed UV light modulators and fast switching light emitting devices.

Small static gauge factors (GFs) of n-type GaN can indicate that, because of a small activation energy, the donor states are close to the conduction band, and the donor energy levels shift together with the conduction band under the stress. Consequently, the donor activation energy does not change under the strain, and the static piezoresistive effect is small.

The comparison of the piezoresistance in GaN [6] with the experimental results for SiC [22] and GaAs [23] shows that n-type GaN-based electromechanical sensors are more sensitive than those made from GaAs or SiC. The GaN sensors can be used for time-dependent measurements of vibrations, sudden changes in acceleration, force, pressure, etc. in the wide frequency range, above approximately 1 Hz. As mentioned previously, our data show that a high sensitivity of our samples to the stress variations can be explained by strong piezoelectric effect properties of GaN. In SiC, piezoelectric constants are much smaller than in GaN [24]. That is why GaN has an advantage over SiC for time-dependent applications. These unique properties

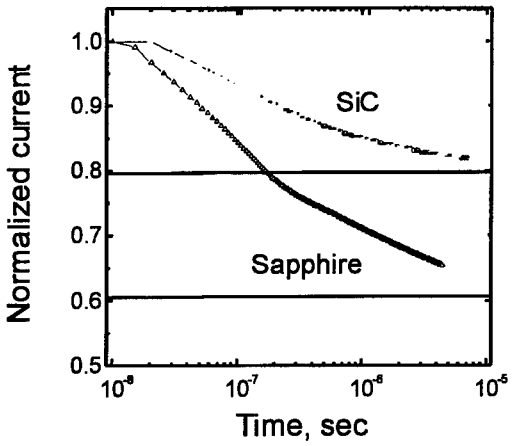


Fig. 6. Normalized source-drain current at zero gate bias for HFETs grown on sapphire and p-type 6H-SiC substrates, measured at source-drain voltages of 7.5 and 10 V, respectively. Horizontal lines show the source-drain current values at time moment $t = 1$ ms [30].

of GaN films can be also useful in high temperature piezoelectric sensors.

4. AlGaN/GaN HEMTs on SiC substrates

GaN/AlGaN HEMTs dissipate a very large power. Therefore, their thermal impedance is very important. Typical AlGaN/GaN FET epilayer structures are deposited over basal plane sapphire substrate using a low pressure MOCVD system. Conventional heat-sinking designs, such as a flip-chip (see, for example [25]) can also be used in order to reduce self-heating. Thibeault et al. [26] obtained excellent results for AlGaN/GaN HFETs using this approach. Another approach involves using SiC substrates. The replacement of sapphire substrates with SiC substrates drastically decreases the thermal impedance of the AlGaN/GaN HFETs and self-heating effects. Such devices have been demonstrated on doped 6H-SiC [1–3,9] and semi-insulating 4H-SiC [27,28]. Recent studies of GaN/AlGaN on SiC substrates showed that these substrates reduce the device thermal impedance by more than an order of magnitude (down to approximately 2°C mm/W, i.e. about 20 times smaller than for typical GaAs power FETs. The dissipated values of the DC power were up to 0.8 MW/cm^2 [3].

Fig. 6 (from [29]) shows the normalized source-drain current at zero gate bias for HFETs grown on sapphire and p-type 6H-SiC substrates, measured at source-drain voltages of 7.5 and 10 V, respectively. This Figure clearly illustrates the advantages of using SiC substrates.

The epitaxial GaN structure grown on SiC also have

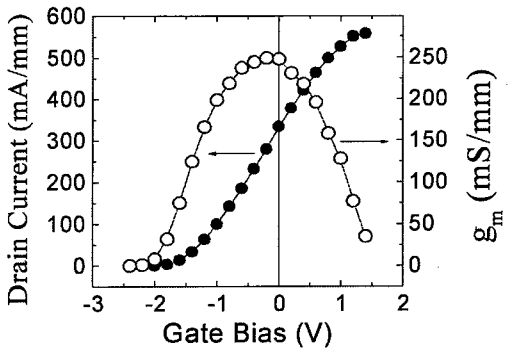


Fig. 7. Transfer characteristics of $Al_{0.25}Ga_{0.75}N$-GaN HFETs on insulating 4H-SiC substrates (after [28]).

an advantage of a much smaller lattice mismatch (only 3–4%). Comparative studies of the Hall mobility, Quantum Hall Effect, and Shubnikov-de-Haas effect on AlGaN/GaN heterostructures grown by MOCVD on sapphire and 6H-SiC substrates using an isolating AlN buffer show that the material quality is much better for the heterostructures grown on 6H-SiC [30]. Semi-insulating SiC substrates yield a much better microwave performance because of sharply reduced parasitics (see, for example [12]).

GaN/AlGaN HFETs on SiC substrates exhibited a strong backgating effect. The application of bias to the SiC substrate also changes the shape of the transconductance dependence on the top gate bias. When the top gate is floating, the SiC substrate acts as the gate with the effective device transconductance on the order of 50 mS/mm, even in the devices with the distance between the substrate and the channel on the order of 0.5 micron [2]. Thinner epitaxial structures should exhibit a larger current enhancement.

Fig. 7 shows typical characteristic of an AlGaN/

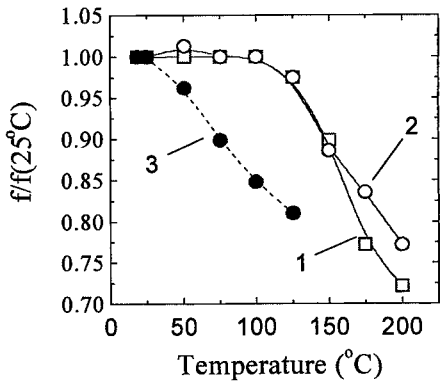


Fig. 8. Normalized f_t (curve 1) and f_{max} (curve 2) of $Al_{0.25}Ga_{0.75}N$-GaN HFET on 4H-SiC. Curve 3 shows $f_t(T)$ for AlGaN-GaN HFET grown on sapphire substrate (after [28]).

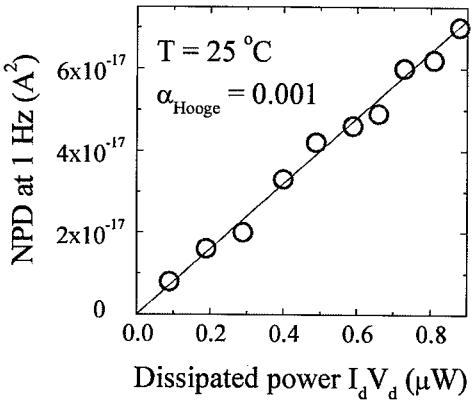


Fig. 9. Spectral noise power density (NPD) of drain current noise at 1 Hz as a function of power dissipated on the transistor (after [28]).

GaN HEMT on an insulating 4H-SiC substrate. This device exhibits a fairly high transconductance. Fig. 8 compares the temperature dependencies of the normalized cutoff frequency, f_t, (curve 1) and the maximum frequency of oscillations, f_{max}, (curve 2) of $Al_{0.25}Ga_{0.75}N$-GaN HFET on 4H-SiC with the $f_t(T)$ dependence for an AlGaN-GaN HFET grown on sapphire substrate (curve 3). As can be seen from the Figure, the device on 4H-SiC substrate is much more tolerant to a temperature increase.

Fig. 9 shows the spectral noise power density (NPD) of the drain current noise at 1 Hz as a function of power dissipated on the transistor [28,31]. The extracted Hooge constant, $\alpha = 0.001$, is about an order of magnitude smaller than that for the device grown on sapphire [32] indicating a much better material quality.

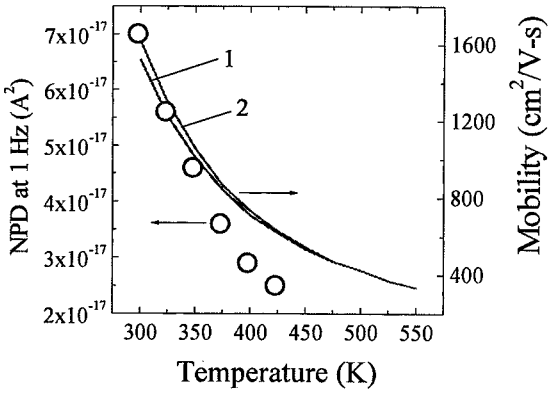


Fig. 10. Noise power density of $Al_{0.25}Ga_{0.75}N$-GaN HFET on 4H-SiC substrate measured at 1 Hz frequency and gate bias $V_G = -1$ V (circles) and calculated 2D electron mobility (solid curves) as a function of temperature. 1–$n_s = 5 \times 10^{12}$ cm^{-2}; 2—10^{13} cm^{-2} (after [28]).

Table 2
Parameters of $Al_{0.25}Ga_{0.75}N$-GaN HFETs on insulating 4H-SiC (after [28])

DC	
Contact resistance	0.3–0.4 Ω mm
Drain current	0.65 mA/mm
Transconductance	250 mS/mm
Threshold voltage	−2 V
Breakdown voltage	100 V
High frequency	
f_t	35 GHz
f_{max}	85 GHz
Thermal	
Thermal impedance	2–2.5°C mm/W
Noise	
Hooge constant	0.001

Fig. 10 compares the temperature dependence of the noise power density of $Al_{0.25}Ga_{0.75}N$-GaN HFET on 4H-SiC substrate measured at 1 Hz frequency and gate bias $V_G = -1$ V (circles) with the calculated temperature dependence of the 2D electron mobility (solid curves) for two different concentration of the 2D-electrons [28]. As can be seen from this Figure, there is a strong correlation between the level of the $1/f$ noise and the low-field mobility. This confirms that the level of the $1/f$ noise is linked to the material quality.

Parameters characterizing the performance of an AlGaN/GaN HEMT on 4H-SiC substrate are summarized in Table 2 [28].

Fig. 11 clearly illustrates the effect of self-heating, which is important even in devices on SiC substrates [28]. The parameter that gives a quantitative measure of self-heating is the negative output conductance in the saturation region. Table 3 compares self-heating parameters of AlGaN/GaN HEMTs grown on differ-

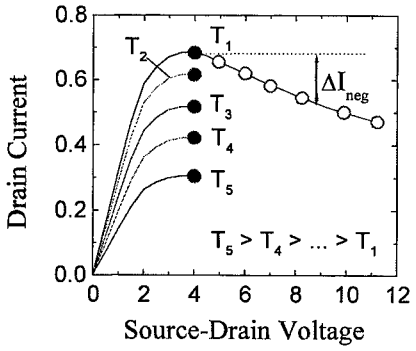


Fig. 11. Effects of self-heating effect and ambient temperature increase (after [28]).

Table 3
Self-heating parameters (after [28])

Sample	V_{ds} for peak current	% decrease in current at $V_{ds}=15$ V	Slope of I_d–V_d in self-heating region
p-SiC	7	3.6	3.25×10^{-3}
n-SiC	8.5	3.0	2.62×10^{-3}
Sapphire	6.5	21.8	2.56×10^{-2}
i-SiC	4.53	3.3	3.07×10^{-3}

ent substrates [28]. This table clearly demonstrates advantages of SiC substrates.

5. GaN/AlGaN HEMT microwave detectors

Large densities of the 2D-electron gas at the AlGaN/GaN heterointerface make these devices attractive for applications in plasma wave electronics devices, such as detectors of microwave and terahertz radiation [33]. Fig. 12 (from [11]) shows measured and calculated frequency dependencies of the detector responsivity for a GaN Heterostructure FET (HFET) with a cutoff frequency of approximately 2 GHz. The periodic variation of the responsivity with frequency is caused by changes in the transmission line impedance with frequency. As can be seen from the Figure, the device operates as broad band non-resonant detector of microwave radiation at frequencies higher than the cutoff frequency and the measured data agree with the theory.

6. Conclusions

Piezoelectric and pyroelectric properties of GaN-based materials allow us to develop a new family of pyroelectric and piezoelectric sensors that might operate at elevated temperatures, where more conventional pyroelectric materials do not work. GaN-based HEMTs dissipate large power and experience self-heating. The self-heating effects are drastically reduced in the GaN-based HEMTs on 4H-SiC semi-insulating substrates. An additional benefit of these substrates is an improved material quality, which is revealed by a lower noise spectral density.

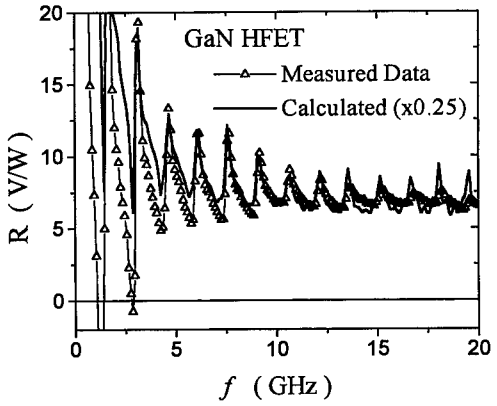


Fig. 12. Measured (symbol) and calculated (solid) frequency dependence of the detector responsivity for a GaN HFET with the gate length $L = 5$ μm, the gate bias $U_G=-1$ V, and the threshold voltage $U_T=-2$ V [11].

Acknowledgements

The Office of Naval Research has supported the work (Project Monitors Dr Colin Wood and Dr John Zolper).

References

[1] Gaska R, Yang JW, Osinsky A, Khan MA, Shur MS IEDM-97 Technical Digest, pp. 565–568, December 1997.
[2] Gaska R, Shur MS, Yang JW, Osinsky A, Orlov AO, Snider GL Proceeding of the International Conference on SiC and Related Compounds, Stockholm, Sweden, September 1997, 1998. 2. p. 1445–1448.
[3] Gaska R, Chen Q, Yang J, Osinsky A, Khan MA, Shur MS. High temperature performance of AlGaN/GaN HFETs on SiC substrates. IEEE Electron Device Letters 1997;18(10):492–4 October.
[4] Bykhovski A, Gelmont B, Shur MS, Spencer M. Piezoelectric effects in SiC/AlN/SiC SIS structures. In: Proceedings of International Semiconductor Device Research Symposium, Charlottesville, VA, 1–3 December, 1993. p. 581–4.
[5] Bykhovski A, Gelmont B, Shur MS, Khan MA. J Appl Phys 1995;77(4):1616.
[6] Gaska R, Yang J, Bykhovski AD, Shur MS, Kaminski VV, Soloviov SM. Piezoresistive effect in GaN-AlN-GaN structures. Appl Phys Lett 1997;71(26):3817–9.
[7] Bykhovski AD, Kaminski VV, Shur MS, Chen QC, Khan MA. Pyroelectricity in gallium nitride thin films. Appl Phys Lett 1996;69:3254.
[8] Bykhovski AD, Kaminski VV, Shur MS, Chen QC, Khan MA Pyroelectric effect in wurtzite n-type GaN, Proceedings of Material Research Society, 423, 75 1996.

[9] Gaska R, Yang JW, Osinsky A, Bykhovski AD, Shur MS. Piezoeffect and gate current in AlGaN/GaN high electron mobility transistors. Appl Phys Lett 1997;71(25):3673–5.
[10] Dyakonov M, Shur MS. Consequences of space dependence of effective mass in heterostructures. J Appl Phys 1998;84(7).
[11] Lu JQ, Shur MS, Weikle R, Dyakonov MI, Khan MA Proceedings of Sixteenth Biennial Conference of Advanced Concepts in High Speed Semiconductor Devices and Circuits, Ithaca, New York, 4–6 August, 1997. p. 211–217.
[12] Sheppard ST, Doverspike K, Pribble WL, Allen ST, Palmour JW, Kehias LT, Jenkins TJ. High power GaN/AlGaN HEMTs on silicon carbide. In: Proceedings of 1998 Device Research Conference, Charlottesville, VA, 1998.
[13] Fraden J. Handbook of Modern Sensors. New York: Springer, 1996.
[14] Bykhovski A, Gelmont B, Shur M. The influence of the strain-induced electric field on the charge distribution in GaN-AlN-GaN structure. J Appl Phys 1993;74(11):6734.
[15] Bykhovski A, Gelmont B, Shur MS, Khan MA. Strain and charge distribution in GaN-AlN-GaN SIS structure. In: Proceedings of International Conference on Silicon Carbide and Related Materials—1993, Washington DC. Institute of Physics Conference Series Number 137: Institute of Physics Publishing, 1994. p. 691–4.
[16] Bykhovski A, Gelmont B, Shur M, Khan MA. Current–voltage characteristics of strained piezoelectric structures. J Appl Phys 1995;77(4):1616–21.
[17] Bykhovski AD, Gaska R, Yang JW, Shur MS. Capacitance–voltage characteristics of strained AlGaN/GaN heterostructures. In: Proceedings of EMC, Charlottesville, 1998.
[18] Sitar Z, Paisley MJ, Yan B, Ruan J, Choyke WJ, Davis RF. J Vac Sci Technol 1990;B8:316.
[19] Wang J, Jeon JB, Sirenko YM, Kim KW Piezoelectric effect on optical properties of pseudomorphically strained wurtzite GaN quantum wells 1997;9(6):728–730.
[20] Martin G, Botchkarev A, Rockett A, Morkoç H. Valence band discontinuities of wurtzite GaN, AlN and InN heterojunctions measured by X-ray photoemission spectroscopy. Appl Phys Lett 1996;68(18):2541–3.
[21] Chichibu S, Azuhata T, Sota T, Nakamura S. Appl Phys Lett 1996;69(27):4188–90.
[22] Shor JS, Bemin L, Kurtz AD. IEEE Trans Electron Dev 1994;41(5):661.
[23] Sagar A. Phys Rev 1958;112:1533.
[24] Gualieri JG, Kosinski JA, Ballato A. IEEE Trans Ultrason, Ferroel, and Freq Control 1994;41:53.
[25] Shur MS. GaAs devices and circuits. New York: Plenum Publishing Corporation, 1987.
[26] Thibeault BJ, Keller BP, Wu Y-F, Fini P, Mishra UK, Nguyen C, Nguyen NX, Le M High performance and large area flip-chip bonded AlGaN/GaN MODFET, IEDM-97 Technical Digest, IEEE, December 1997.
[27] Sriram S, Smith TJ, Rowland LB, Burk AA Jr, Augustine G, Balakrishna V, Hobgood HM, Brandt CD Proceedings of 55th Device Research Conference, 1997. p. 138–139.
[28] Gaska R, Shur MS, Ping A, Adesida I, Dickens A, Kuksenkov V, Temkin H. High-frequency, low-noise performance of AlGaN-GaN HFETs on insulating 4H-SiC at elevated temperatures. In: Nitride Workshop, St Louis, August, 1998.
[29] Gaska RR, Osinsky A, Yang J, Shur MS. Self-heating in high power AlGaN/GaN HFETs. IEEE Electron Device Letters 1998;19(3):89–91.
[30] Gaska R, Yang JW, Osinsky A, Chen Q, Khan MA, Orlov AO, Snider GL, Shur MS. Electron transport in AlGaN-GaN heterostructures grown on 6H-SiC substrates. Appl Phys Lett 1998;72(6):707–9.
[31] Kuksenkov DV, Giudice G, Temkin H, Gaska R, Ping A, Adesida I Low-frequency noise in AlGaN-GaN doped channel heterostructure field effect transistors on insulating SiC substrates, Electronics Letters 1998; 34(23):2274–6.
[32] Kuksenkov DV, Temkin H, Gaska R, Yang JW. Low-frequency noise in AlGaN/GaN heterostructure field effect transistors. IEEE Electron Device Letters 1998;19(7):222–4.
[33] Dyakonov MI, Shur MS. IEEE Transactions on Electron Devices 1996;43(3):380–7.

PERGAMON

Solid-State Electronics 43 (1999) 1459–1464

SOLID-STATE ELECTRONICS

SiC and GaN wide bandgap semiconductor materials and devices

A.A. Burk Jr.[a,*], M.J. O'Loughlin[a], R.R. Siergiej[b], A.K. Agarwal[b], S. Sriram[b], R.C. Clarke[b], M.F. MacMillan[b], V. Balakrishna[b], C.D. Brandt[b]

[a]*Northrop Grumman ESSD Advanced Technology Laboratory, Baltimore, MD 21203, USA*
[b]*Northrop Grumman ESSD Science and Technology Center, Pittsburgh, PA 15235, USA*

Abstract

Wide bandgap semiconducting materials are promising candidates for high-power, high-temperature, microwave and optoelectronic devices because of their superior thermal and electrical properties in comparison to conventional semiconductors. Recent developments in SiC and GaN materials and processing have enabled the demonstration of semiconductor devices with dramatically improved performance. Examples of these developments and new wide bandgap devices are presented. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

Wide bandgap semiconducting materials are promising candidates for high-power [1], high-temperature [2], microwave [3] and optoelectronic [4] devices because of their superior thermal and electrical properties in comparison to conventional semiconductors (Table 1) [5]. SiC, for example, has an order of magnitude greater thermal conductivity and breakdown field strength and a higher saturated electron drift velocity than GaAs. SiC's lower electron mobility is largely offset by closer device geometry afforded by the higher breakdown strength and thermal conductivity. In the case of GaN, in addition to microelectronics applications, the direct bandgap makes highly efficient blue light emitting diodes (LEDs) and laser diodes possible. The lack of high-quality, large-area and lattice matched substrates (in the case of GaN), combined with the difficulties in processing these materials due to their chemical, mechanical and thermal stability, have hampered the development of wide-bandgap devices. In the case of 4H-SiC, up to two-inch diameter, device-quality, substrates are now commercially available. Although work has been done in molecular beam epitaxy [6,7] as well as in ion implantation [8] for device active layer formation, the most successful methods for SiC and GaN device active layer fabrication have been vapor phase techniques. SiC epitaxial layer background doping densities less than 1×10^{14} cm^{-3} and n and p-type intentional doping from 1×10^{15} cm^{-3} to over 1×10^{19} cm^{-3} have been reported [9]. In the case of GaN, the use of AlN buffer layers, SiC substrates and lateral overgrowth [10] have dramatically reduced the $\sim 10^{10}$ cm^{-2} dislocation density initially observed in layers grown on sapphire substrates.

2. Results

2.1. SiC Epitaxial growth

Vapor phase epitaxial (VPE) growth has been a very successful means of SiC active layer formation. Several

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00089-1

Table 1
Electronic and physical properties of Si, GaAs, 4H-SiC and GaN

Property	Si	GaAs	4H-SiC	GaN
Bandgap (eV)	1.12	1.43	3.25	3.4
Breakdown field (MV cm^{-1})	0.25	0.3	~3	~3
Saturated electron velocity (10^7 cm s^{-1})	1	2.0 (peak) 1.2 (sat)	2.0	2.5 (peak) 1.5 (sat)
Electron mobility N_d ~10^{16} cm^{-3} (cm^2 V^{-1} s^{-1})	1200	6500	800	900
Thermal conductivity (W cm^{-1} K^{-1})	1.5	0.5	4.9	1.3 (on sapphire)
Dielectric constant	11.8	12.8	9.7	9
Normalized Johnson figure of merit [28]	1	7	360	560

reactor configurations have been developed for the growth of SiC epitaxial layers, but typically on only one 35 mm diameter wafer at a time [9,11–15]. All reactors are constructed with high temperature materials such as quartz, graphite and SiC. Hydrogen carrier gas along with silane and propane reagents are typically employed at atmospheric and reduced pressure and temperatures ranging 1450–1600°C. Despite these extreme temperatures, device quality SiC epitaxial layers have been grown exhibiting excellent morphology, purity, mobility and doping control.[9] For example, Fig. 1 contains a carrier concentration profile of a MESFET active layer grown in our single wafer reactor with n-type doping layers ranging from ~1×10^{14} cm^{-3} to over 1×10^{19} cm^{-3}.

We are currently developing a SiC multi-wafer horizontal planetary reactor, supplied by Aixtron, AG, to improve uniformity, reproducibility and throughput of the SiC epitaxial growth process. Fig. 2 shows a photograph of the reactor and schematic of the SiC reactor chamber described below. As currently configured, the multi-wafer reactor is capable of growth on seven 2-inch diameter substrates at a time. The planetary reactor concept was originated by Frijlink et al.

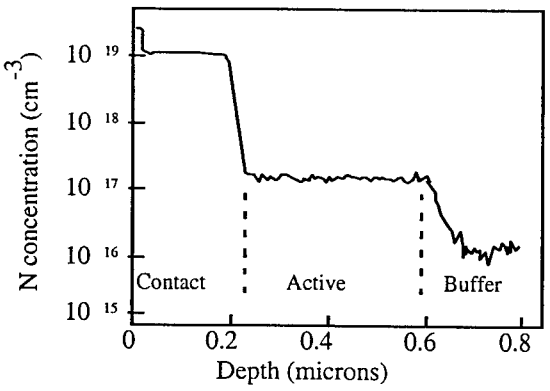


Fig. 1. SIMS profile of nitrogen doping in a SiC MESFET (buffer layer doping of ~1×10^{14} cm^{-3} is less than the SIMS detection limit).

[16] for the growth of highly uniform III–V compound semiconductors. In this design the gases enter from the center of the reactor and flow outward radially. The entire susceptor rotates about its center in order to average any asymmetry in susceptor temperature or reagent flows and the individual wafers also rotate about their centers. Extending this concept to 1600°C has been extremely challenging from a materials and coating perspective[17].

Reasonably low n-type background doping of approximately 1×10^{15} cm^{-3} has been determined by CV measurements on unintentionally doped layers grown in the multi-wafer reactor. Fig. 3 contains a SIMS profile revealing no more than mid-10^{14} atoms cm^{-3} concentrations of aluminum and boron impurities, the primary compensating p-type impurities in SiC. The aluminum and boron detection limits for these measurements were approximately 3×10^{13} cm^{-3} and 6×10^{14} cm^{-3}, respectively. Room temperature mobilities shown in Fig. 4, as high as 1000 cm^2 V^{-1} s^{-1} compare favorably with material grown in our single wafer reactor and elsewhere [18].

The average epitaxial intrawafer thickness and doping uniformities from the multi-wafer reactor are 3.8 and 7.3%, respectively at doping levels of ~1×10^{16} cm^{-3} for 35 mm diameter wafers. As shown in Fig. 5, a superposition of the doping profiles at the center of seven wafers from a single growth run reveals only 7 and 4% standard deviation/mean for thickness and doping respectively. The average wafer-to-wafer thickness and doping uniformities from a series of 10 growth runs are 8.6 and 12.5%, respectively, at ~1×10^{16} cm^{-3}. While still requiring further improvement, this level of uniformity is sufficient for initial pilot line device production.

2.2. Device results

SiC metal semiconductor field effect transistors (MESFETs) and static induction transistors (SITs) exploit SiC's high saturation velocity, breakdown field strength and thermal conductivity for microwave

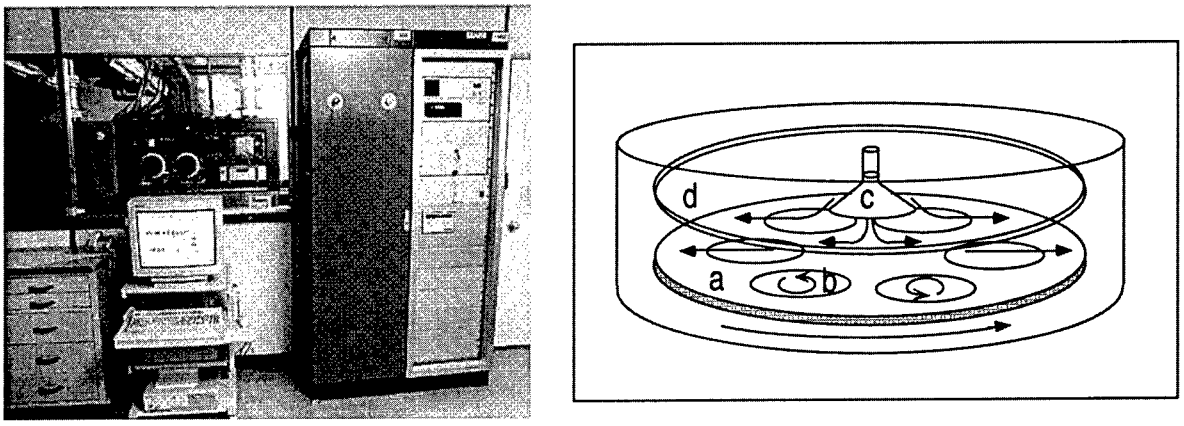


Fig. 2. Planetary SiC VPE Reactor capable of growth on seven 2-inch wafers at a time. Inset contains a schematic of the growth chamber consisting of, (a) susceptor, (b) satellites (wafer holders), (c) injection nozzle and (d) ceiling.

power amplification. Physics based modeling predicts over 12 dB of gain can be achieved at Ku-band (18 GHz) with 0.25 μm-gate 4H-SiC MESFETs [19]. Due to GaN's higher saturation velocity and mobility GaN/AlGaN modulation doped field effect transistors (MODFETs) have demonstrated cut-off frequencies in excess of 97 GHz promising millimeter wave applications [20].

2.2.1. SiC MESFETs

4H-SiC MESFETs with SiC-VPE active layers and semi-insulating SiC substrates have been fabricated as reported previously [21]. In this process, reactive ion etching (RIE) was used for mesa isolation and channel recessing. Ni, sintered using rapid thermal annealing (RTA) is employed as the ohmic source and drain metal. A multi-layer metal system is used for gates with Au as the top layer to reduce gate resistance. The gate length was ~0.5 μm and the gate-source and gate-drain spacings were ~0.5 and ~1 μm, respectively. Air bridges were used to interconnect the source fingers (Fig. 6). Devices up to 6.4 mm total gate width have been fabricated using this process. The 1.92 mm gate periphery SiC MESFET, whose microwave power performance is shown in Fig. 7, exhibited cut-off frequencies of 42 GHz. This device achieved 6.2 W of pulsed power at X-band with over 50% drain efficiency.[22] The power density of 3.3 W mm^{-1} is more than six times higher than that typically obtained with GaAs MESFETs at this frequency. At S-band, a 42 mm periphery 4 H-SiC MESFET recently produced a CW output power of 53 W [23].

2.2.2. SiC SITs

Static induction transistors (SITs) are a class of transistor with a short channel FET structure in which a

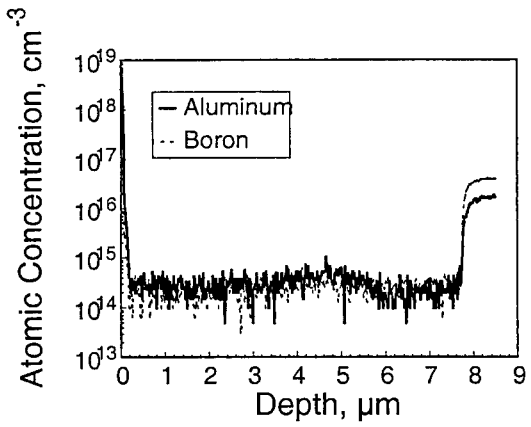


Fig. 3. SIMS measurement showing low aluminum and boron concentrations.

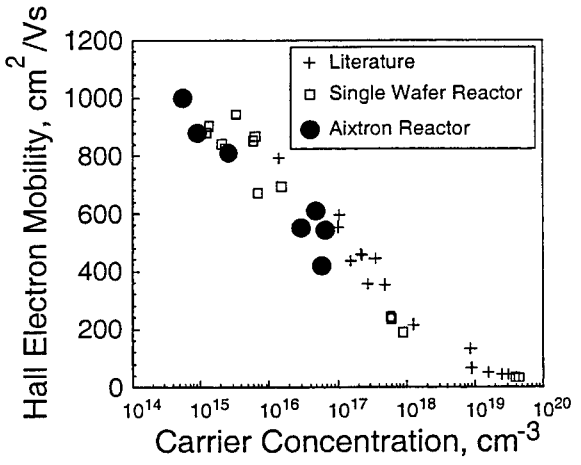


Fig. 4. Hall electron mobility on 4H-SiC layers from the literature [18] (pluses), single wafer reactor (squares) and Aixtron reactor (circles).

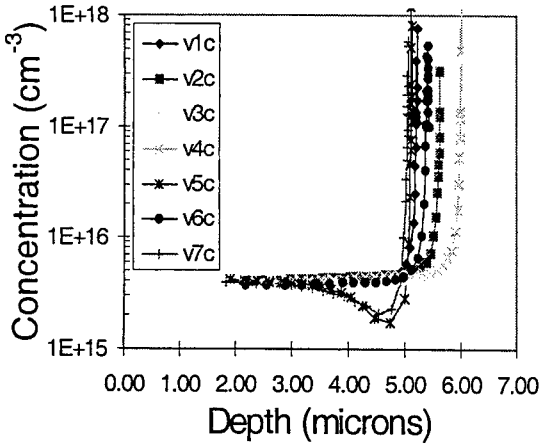


Fig. 5. An overlay of the carrier concentration profiles obtained from the center of seven wafers grown in the same growth run.

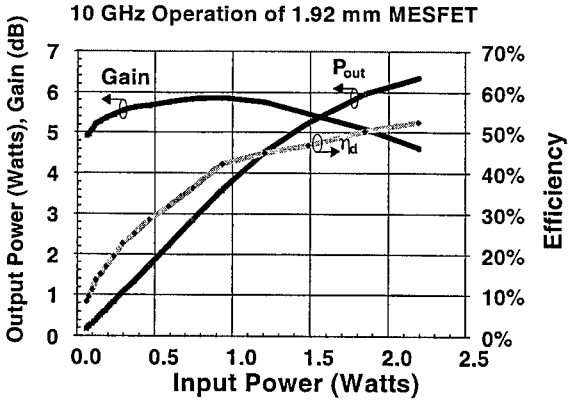


Fig. 7. 4H-SiC, 0.5 μm gate-length MESFET with 6.2 W output power at X-band.

current travelling vertically between source and drain is controlled by the height of an electrostatically induced potential barrier under the source (Fig. 8) [24]. Analogous to a vacuum triode both the gate (grid) and the drain (anode) voltage affect the drain current. Due to the enormous power densities achievable in a SiC SIT, they are ideal for pulsed power applications such as in radar transmitters. SiC SITs with 4 GHz cut-off frequencies have been developed from VHF through S-band, exhibiting over four times the power density of silicon power transistors at S-band [25]. 300 W of power has been obtained from a 32 cm periphery SIT as shown in Fig. 9.

2.2.3. GaN devices

The recent demonstration of room-temperature, CW operation of blue laser diodes [4], only two years after the first demonstration of laser action in this materials system, is representative of the extremely rapid pace of development in GaN. Using active layers provided by Professor Bob Davis of North Carolina State University, we have fabricated 0.25 μm gate-length GaN MODFETs with cut-off frequencies of 80 GHz. The 0.96 mm gate-width MODFET shown in Fig. 10 produces over 2 W of power at X-band. Recently a larger, 2 mm periphery, GaN/AlGaN HEMT is reported to have produced 4 W of CW power at X-band.[26] Smaller, 100 μm, gate width devices exhibit over 3 W mm^{-1} at 18 GHz [27]. Despite these promising results, GaN devices do not yet consistently produce the RF power densities predicted by their DC power triangles. They can also exhibit drain current drift and light sensitivity. These undesirable traits, however, were also exhibited by early GaAs devices. They were successfully eliminated by refinements in active layer formation, passivation and device design. We are currently investigating these approaches in order to improve our GaN devices.

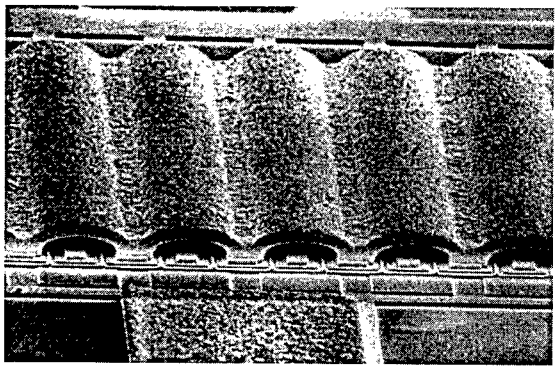


Fig. 6. SEM photgraph of an air-bridged, 0,5μm gate-length, 4H-SiC MESFET.

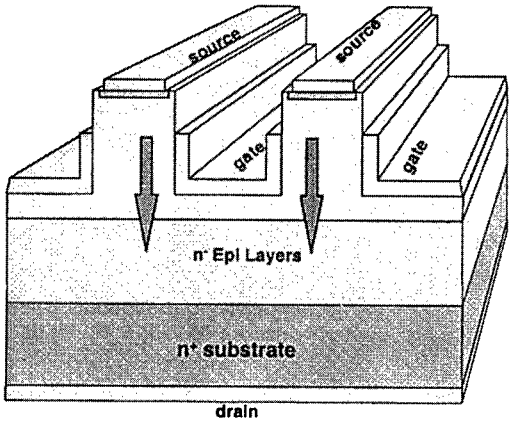


Fig. 8. Schematic of a SiC-SIT.

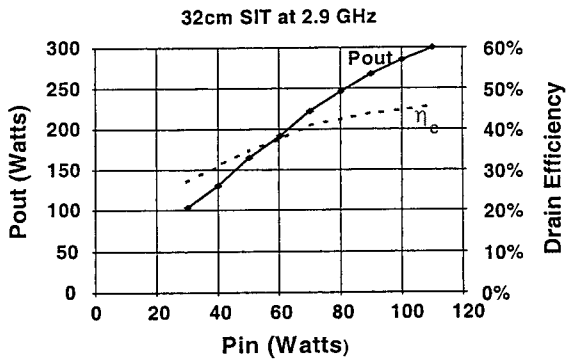


Fig. 9. 4H-SiC 32 cm periphery SIT producing 300 W at S-band with 45% drain efficiency.

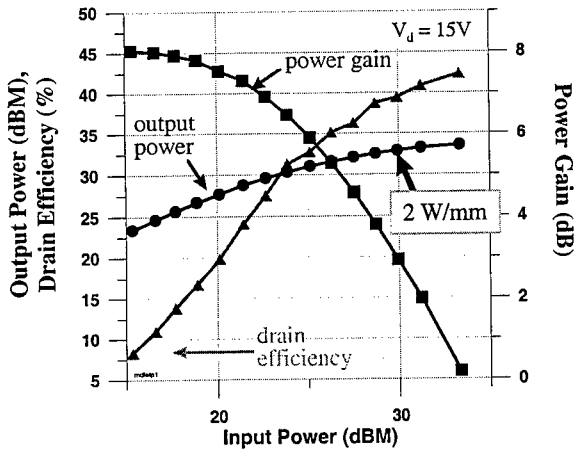


Fig. 10. 0.96 mm AlGaN MODFET shows >2 W output power at 10 GHz.

3. Conclusions

Recent developments in wide bandgap SiC and GaN semiconductor materials (and processing) have resulted in new microelectronic and optoelectronic devices. These devices have already demonstrated previously unobtainable functionality and levels of performance. In particular, high-power microwave devices have exhibited over four times the power density of conventional Si and GaAs semiconductor devices. Further advances in wide bandgap semiconductor materials and processes will result in the insertion of wide bandgap semiconductor devices into advanced microelectronic and optoelectronic products.

Acknowledgements

This work was supported in part by the Department of the Air Force under contract F33615-92-C-5912 and F33615-95-C-5427 (Tom Kensky, contract monitor and Laura Rea, program direction leader).

References

[1] Agarwal AK, Seshadri S, Casady JB, Mani SS, MacMillan MF, Saks N, Burk Jr. AA, Siergiej RR, Augustine G, Balakrishna V, Sanger PA, Brandt CD. Presented at 9th European Conference on Diamond, Diamond-Like Materials, Nitrides and Silicon Carbide (DIAMOND 98), Crete, Greece, Sept. 13–18, 1998. Diamond Rel Mater, in press.

[2] Siergiej RR, et al. In: McCluskey FP, Grzybowski R, Podlesak T, et al., editors. High temperature electronics. CRC Press, 1997. p. 60.

[3] Siergiej RR, Clarke RC, Sriram S, Agarwal AK, Bojko RJ, Morse AW, Balakrishna V, MacMillan MF, Burk Jr. AA, Brandt CD. Presented at 2nd Euro. Conf. On SiC and Rel. Mat., Montpellier, France, Sept. 2–4, 1998. J Mat Sci Eng B, in press.

[4] Nakamura S, Senoh M, Nagahama S, Iwasa N, Yamada T, Matsushita T, Kiyoko H, Sugimoto Y, Kozaki T, Umemoto H, Sano M, Chocho K. Appl Phys Lett 1998;72(2):211.

[5] MorkoÇ H, Strite S, Gao GB, Lin ME, Sverdlov B, Burns M. J Appl Phys 1994;76:1363.

[6] Yoshinobu T, Nakayama M, Shiomi H, Fuyuki T, Matsunami H. J Crystal Growth 1990;99:520.

[7] Rowland LB, Kern RS, Tanaka S, Davis RF. J Mater Res 1993;8:2753.

[8] Ghezzo M, Brown DM, Downey ED, Kretchmer J, Hennessy W, Polla DL, Bakhru H. IEEE Electron Device Lett 1992;10:639.

[9] Burk Jr. AA, Rowland LB. Phys Status Solidi B 1997;202:263 .

[10] Nam O-H, Bremser MD, Zheleva TS, Davis RF. Appl Phys Lett 1997;71(18):2638.

[11] Powell JA, Larkin DJ, Matus LG, Choyke WJ, Bradshaw JL, Henderson L, Yoganathan M, Yang J, Pirouz P. Appl Phys Lett 1990;56:1442.

[12] Karmann S, Suttrop W, Schoner A, Schadt M, Haberstroh C, Engelbrecht F, Helbig R, Pensl G, Stein RA, Leibenzeder S. J Appl Phys 1992;72:5437.

[13] Kong HS, Glass JT, Davis RF. US Pat. 5011549 (30 Apr. 1991).

[14] Kördina O, Hallin C, Glass RC, Janzén E. In: Spencer MG, Devaty RP, Edmond JA, Khan MA, Kaplan R, Rahman M, editors. Silicon carbide and related materials, Institute of Physics Conference Series, No. 137. Bristol: Institute of Physics Publishing, 1994. p. 305.

[15] Rupp R, Lanig P, Völkl J, Stephani D. J Crystal Growth 1995;146:37.

[16] Frijlink PM. J Crystal Growth 1988;93:207.

[17] Burk Jr. AA, O'Loughlin MJ, Mani SS. In: Pensl G, MorkoÇ H, Monemar B, Janzén E, editors. Silicon Carbide, III-Nitrides and Related Materials, Materials Science Forum, No. 264–268. Switzerland: Trans Tech Publications, 1998. p. 83.

[18] Schaffer WJ, Negley GH, Irvine KG, Palmour JW. Mater Res Soc Symp Proc 1994;337:595.

[19] Trew RJ. Case Western Reserve University, private communication.
[20] Khan MA, Chen Q, Shur MS, Dermott BT, Higgins JA, Burm J, Schaff WJ, Eastman LF. IEEE Electron Device Lett 1996;EDL-17:584.
[21] Sriram S, Clarke RC, Burk Jr. AA, Hobgood HM, McMullin PG, Orphanos PA, Siergiej RR, Smith TJ, Brandt CD, Driver MC, Hopkins RH. IEEE Electron Device Lett 1994;EDL-15:458.
[22] Sriram S, Augustine G, Burk Jr. AA, Glass RC, Hobgood HM, Orphanos PA, Rowland LB, Siergiej RR, Smith TJ, Brandt CD, Driver MC, Hopkins RH. IEEE Electron Device Lett 1996;EDL-17:369.
[23] Sadler RA, Allen ST, Alcorn TS, Pribble WL, Sumakeris J, Palmour JW. Proceedings 56th Device Research Conference, Charlottesville, VA, 1998, in press.
[24] Nishizawa J, Terasaki T, Shibata J. IEEE Trans Electron Dev 1975;ED-22:185.
[25] Siergiej RR, Sriram S, Clarke RC, Agarwal AK, Brandt CD, Burk Jr. AA, Smith TJ, Morse A, Orphanos PA. In: Tech Digest Int Conf SiC and Rel Mater '95, 1995. p. 321.
[26] Sheppard ST, Doverspike K, Pribble WL, Allen ST, Palmour JW. Proceedings 56th Device Research Conference, Charlottesville, VA, 1998, in press.
[27] Mishra UK, Wu Y, Keller BP, Keller S, Denbaars SP. Trans Microwave Theory Tech 1998;46(6):756.
[28] Johnson EO. RCA Rev. 1965;163–177.

PERGAMON

Solid-State Electronics 43 (1999) 1465–1471

SOLID-STATE ELECTRONICS

Surface p-channel metal-oxide-semiconductor field effect transistors fabricated on hydrogen terminated (001) surfaces of diamond

A. Hokazono[a], K. Tsugawa[a,b], H. Umezana[a,b], K. Kitatani[a], H. Kawarada[a,b,*]

[a]*School of Science and Engineering, Waseda University, Ohkubo 3-4-1, Shinjuku-ku, Tokyo, 169, Japan*
[b]*Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology, Ohkubo 3-4-1, Shinjuku-ku, Tokyo, 169, Japan*

Abstract

Metal-oxide-semiconductor field effect transistors (MOSFETs) with a surface p-channel have been fabricated on hydrogen-terminated diamond (001) surfaces without doping. The maximum transconductance of the device with the gate length of 6 μm is 16 mS/mm, which is the highest in diamond MOSFETs and comparable to that of silicon n-channel MOSFET with the same gate length. The relatively high transconductance is due to the low density of surface states on hydrogen-terminated diamond. The diamond MOSFETs operate at the temperatures of up to 330°C in air without any passivation of the device surfaces. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

Diamond field effect transistors (FETs) were demonstrated for the first time using oxygen-terminated surfaces and boron doped substrates [1–3]. Although high temperature operation at 500°C was exhibited [4], the current drive capability of the diamond devices was much lower than that of silicon FETs [4]. Using the surface p-type semiconductive layer of hydrogen-terminated diamond, a new type of metal-semiconductor FETs (MESFET) was developed [5] and the performance of the MESFET was improved by device isolation [6]. Recently, the E/D type logic circuit on diamond was demonstrated for the first time [7] and a high power handling capability was also exhibited [8]. The problem in the operation of the MESFETs is a relatively high saturation voltage due to the gate leakage current at negative gate bias (forward bias), because the gate is not electrically isolated. To overcome this difficulty, the metal-oxide-semiconductor (MOS) structure is necessary.

In the MOSFETs fabricated on boron-doped channels with oxygen-terminated surface, the maximum transconductance of 2 μm gate FETs is limited to around 0.1 mS/mm at room temperature and 1 mS/mm at 450°C [4]. It is not only due to the deep acceptor level (0.37 eV) of boron, but also due to the surface states probably originated from the oxygen-terminated surface [9], which disturb the control of the drain current by the gate voltage. On the hydrogen-terminated surface, however, metal-dependent Schottky barrier heights have been measured [10] indicating that the density of the detrimental surface states is very low.

* Corresponding author. Fax: +81-3-3200-2567.
E-mail address: kawarada@mn.waseda.ac.jp (H. Kawarada)

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00090-8

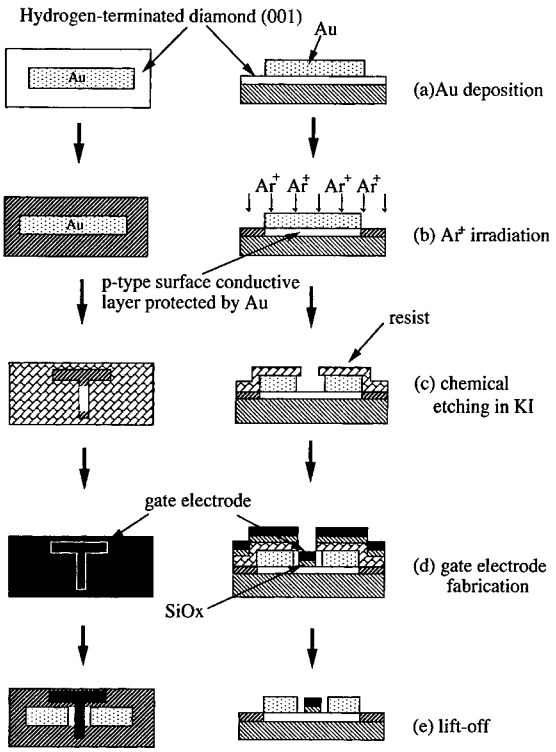


Fig. 1. Schematic cross section of the MOSFET fabrication process. (a) Au deposition on hydrogen terminated (001) surface. (b) Ar ion irradiation to the surface, where the surface conduction was lost and a highly resistive layer (hatched surface) was formed. Under Au, the surface conduction remained. (c) Chemical etching of Au by KI to fabricate a source, a drain, and a channel region. (d) Successive deposition of SiO_x (gate oxide) and gate electrode by vacuum evaporation. (e) MOSFET structure fabrication by lift-off.

Since the surface semiconductive layer associated with hydrogen-termination is very thin, a planar MOS diode structure can not be assumed. A quantitative analysis of the surface state density is difficult as a result. On the other hand, the thin layered structure is suitable for the MOSFET operation. In this work we qualitatively characterize the effect of surface states on the MOSFETs without sacrificing the advantages of hydrogen-terminated surface such as the low density of surface states and the high channel conductance. The temperature stability of the devices is also investigated. The preliminary result has been shortly reported in Ref. [11].

2. Experimental

Homoepitaxial diamond was deposited on high-pressure-synthetic type Ib diamond (001) using microwave plasma-assisted chemical vapor deposition

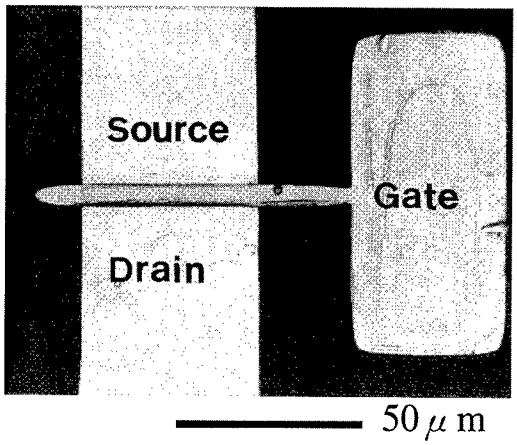


Fig. 2. Optical micrograph of the diamond MOSFET of Pb gate with gate length $L_g = 6$ μm and gate width $W_g = 49$ μm. The insulator is SiO_x with thickness of 20–30 nm.

(MPCVD). The reaction gases were CH4 (2.25%)/O2 (0.75%) diluted with H2. A dopant gas such as a boron-containing gas was not used. The deposition was completed with pure hydrogen plasma for 2–3 min followed by cooling in pure hydrogen ambient. The as-grown homoepitaxial diamond without doping showed p-type surface conduction. Since the valence band is bent upward in the surface region, the p-type layer can be used as a channel in the FET structure.

The basic fabrication steps of the self-aligned insulating gate device are shown in Fig. 1. Au was deposited and patterned in rectangle (Fig. 1(a)). The Au contacts showed ohmic on the hydrogen terminated surfaces with the surface conductive layer and were used as the ohmic contacts for source and drain later in the present fabrication process. The device isolation was then carried out by insulating the surface p-type layer with the exposure of Ar^+ ions, where Au works as a stopping mask of Ar^+ ions. The surface p-type layer was restricted to the region under the Au contact (Fig. 1(b)). The center of the Au rectangle was selectively etched by potassium iodide (KI) to form a source, a drain (two separated Au contacts), and a channel (bared hydrogen-terminated surface) (Fig. 1(c)). On the channel, silicon monoxide (SiO) and metal were successively deposited using a vacuum evaporator (Fig. 1(d)). The gate insulator and gate metal were self-aligned with the source and drain (ohmic contacts) through overhanging photo resist. Finally the self-aligned MOSFET was completed by the lift-off method (Fig. 1(e)). The undercut of the Au beneath the photo resist defines the separation distances between a gate and Au ohmic contacts. The aim of a self-aligned-gate process is to reduce the high parasitic source and drain resistance. Fig. 2 shows an optical micrograph of the fabricated MOSFET. The

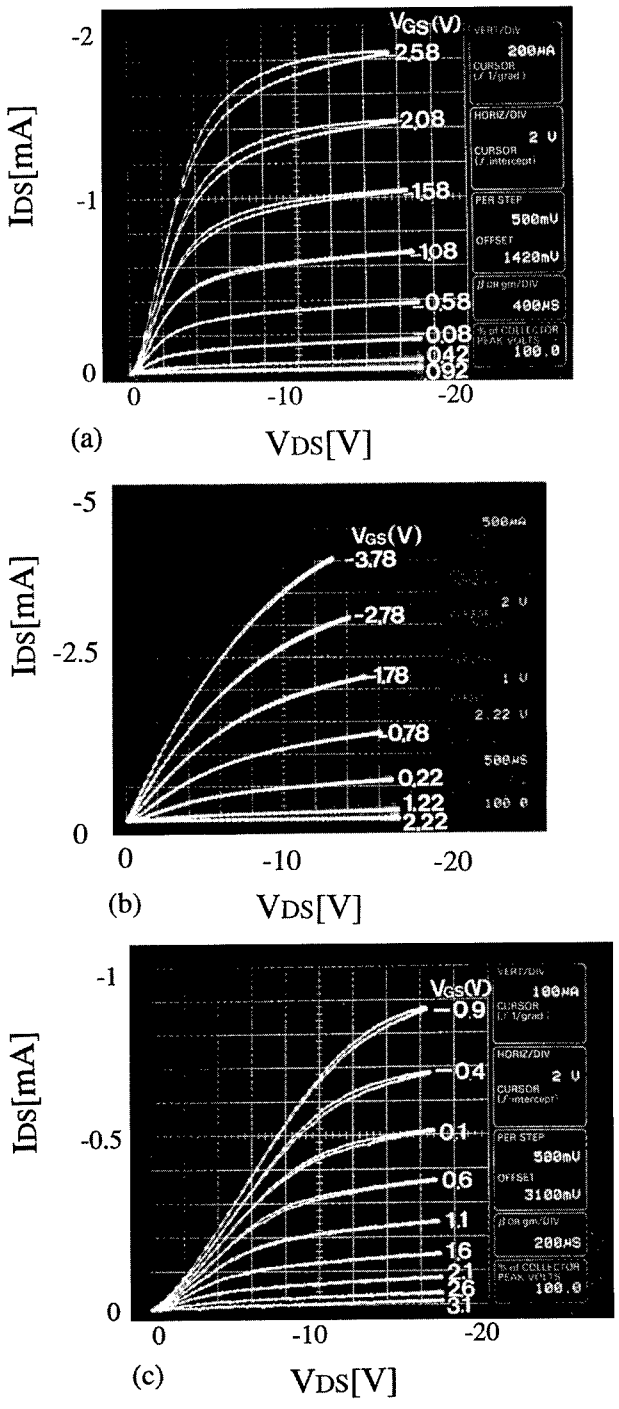


Fig. 3. Room temperature I_{DS}–V_{DS} characteristics of diamond MOSFETs using the SiO_x as insulator and the surface conductive layer as channel. (a) Pb gate with $L_g = 6$ μm and $W_g = 49$ μm on the SiO_x with 20–30 nm in thickness. The transconductance g_m is 16 mS/mm. (b) Al gate with $L_g = 2.5$ μm and $W_g = 60$ μm on the SiO_x with 20–30 nm in thickness. The g_m is 17 mS/mm. (c) Au gate with $L_g = 7$ μm and $W_g = 60$ μm on the SiO_x with 70 nm in thickness. The g_m is 6.5 mS/mm.

layered structure for a MOSFET consists of SiO_x and surface conductive layer under the gate.

The gate insulator was made by thermal evaporation of SiO. Since SiO has a property of sublimation, the deposition of insulator at room temperature and the formation of it without destroying the surface conductive layer (active region) is feasible. Considering the vacuum level of 4×10^{-6} Torr in the chamber during SiO evaporation, the deposited insulator might be SiO_x ($1 < x < 2$). From Rutherford backscattering measurement, x (the ratio of O/Si) is estimated to be more than 1.5, where the insulating property can be expected [12]. The formation of SiO_x does not make any change in the conductivity of active channel. The hydrogen-termination on the diamond surface has been preserved under the SiO_x and the Au contacts.

We have measured the typical drain current versus voltage characteristics for a MOSFET with a curve tracer from room temperature up to 400°C on a probe station with a hot stage.

3. Results and discussion

3.1. Drain I–V characteristics

Drain-to-source current voltage (I_{DS}–V_{DS}) characteristics for diamond MOSFETs measured at room temperature are shown in Fig. 3, where Pb, Al and Au were employed as gate metal, respectively. From all of the I_{DS}-V_{DS} characteristics, it can be seen that very good drain current saturation and pinch-off are achieved as the drain voltage increases. By means of the reduction of the gate leakage current, high current drive was achieved. The Pb and Al gate MOSFETs fabricated using a self-aligned insulated gate process exhibited peak transconductances (g_m) of 16–17 mS/mm (Fig. 3 (a), (b)), which are the highest in diamond MOSFETs. These are comparable values to that of n-channel Si MOSFET with the equivalent gate length. The transconductance of the diamond MOSFETs is as high as that of the MESFET [7,8,13] indicating that the density of surface states at the SiO_x/diamond interface is very low. The effect of hydrogen-termination can be preserved after the oxide deposition.

At $V_{GS} = 0$ V in Fig. 3, all the corresponding I_{DS} is not pinched off indicating normally on (depletion mode) operation. The threshold voltages (V_{ths}) of Pb, Al, and Au gates in Fig. 3 are 0.8 1.4, and 2.8 V, respectively. The V_{th} trend reflects a qualitative change on p-type channel as a function of metal work function, but is not precisely fitted. These V_{th} values are also influenced by the effective oxide thickness and the residual charges in oxide.

The I_{DS}–V_{DS} characteristics of Au gate MOSFET (Fig. 3(c)) certificates that the current control of tran-

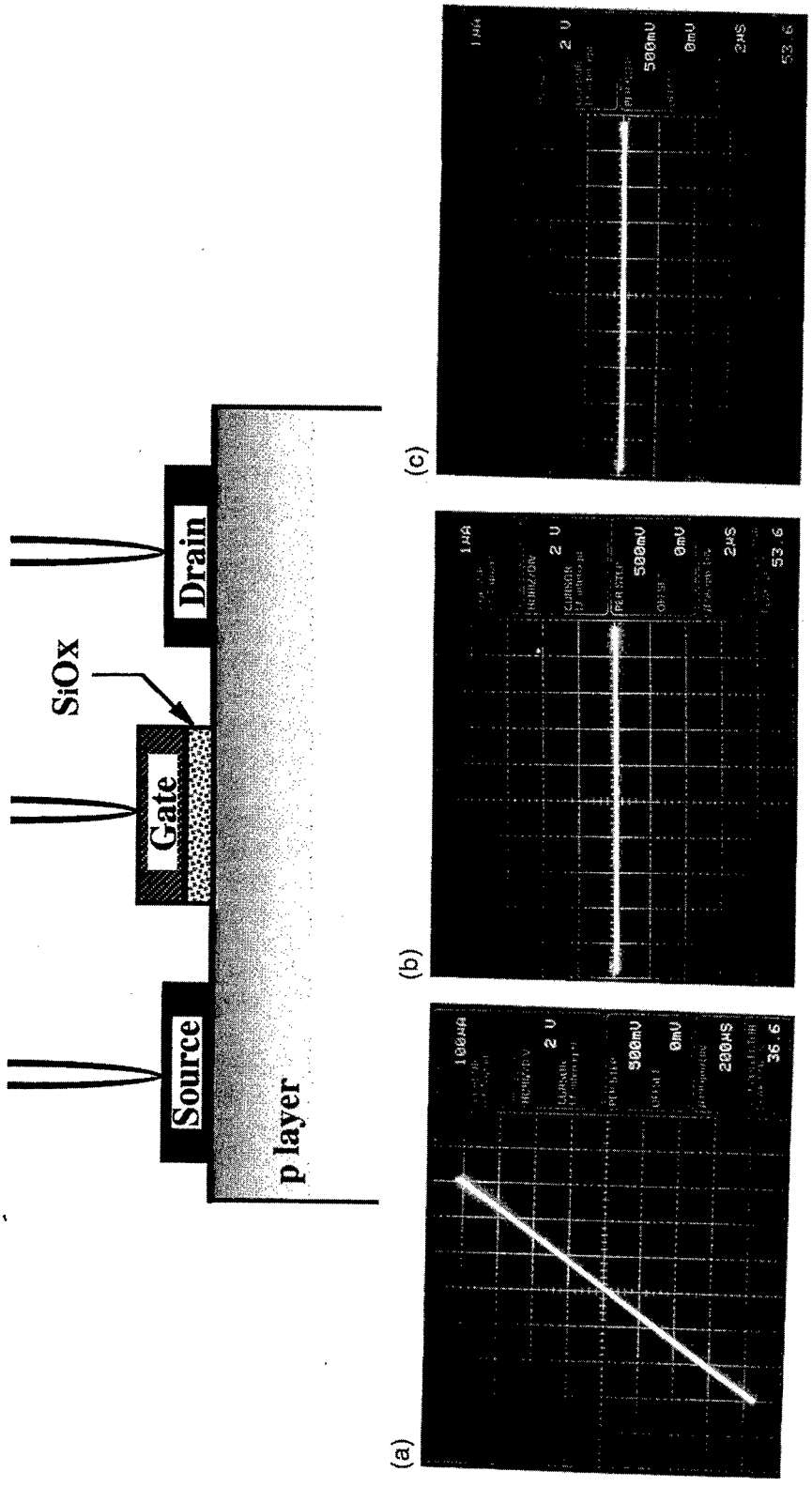


Fig. 4. Current–voltage characteristics between respective electrodes of a Al gate MOSFET shown in Fig. 3(b) at room temperature. (a) Between source and drain. 100 μA/div. in vertical and 2 V/div. in horizontal. (b), (c) Between gate and source and between gate and drain, respectively. In both cases, 1 μA/div. in vertical and 2 V/div. in horizontal.

sistor is not based on metal contact (Schottky) gate through voids of SiO_x films even if present, because Au makes ohmic contacts on the hydrogen-terminated surface [14]. Fig. 4 is the current–voltage characteristics between respective electrodes (gate, source, and drain) of a MOSFET. When SiO_x is inserted between a gate electrode and the surface p-type layer of diamond, the current is completely cut off by the insulator (SiO_x). The blocking of applied voltage by the SiO_x is up to 200 V, indicating the breakdown field strengths of more than 2×10^7 V/cm. It is comparable to the breakdown field strengths of thermally oxidized SiO_2.

From MOSFET device simulation using a drift and diffusion model, acceptor and hole depth profiles hardly identified in experiments can be estimated [11]. Fig. 5 shows the two examples of the in-depth profile of acceptors with equivalent surface concentration (1×10^{13} cm^{-2}) and corresponding hole distribution. One is diffused acceptors expressed in gaussian function with diffusion length of 10 nm and the other is surface acceptors. Higher transconductance and better correspondence of I_{DS}–V_{DS} characteristics of MOSFET to the experimental values as shown in Fig. 2 can be obtained in shallower acceptor distribution less than 1 nm in diffusion length. The surface acceptors, whose hole distribution is shown in Fig. 5, reproduce the I_{DS}–V_{DS} characteristics at best without the effect of surface state trapping. The depth of this carrier (hole) distribution is very shallow and suitable for the most advanced MOSFETs with less than 50 nm in gate length [15] where the junction depth of less than 10 nm is inevitable to sustain uniform electric field between source and drain necessary for FET operation.

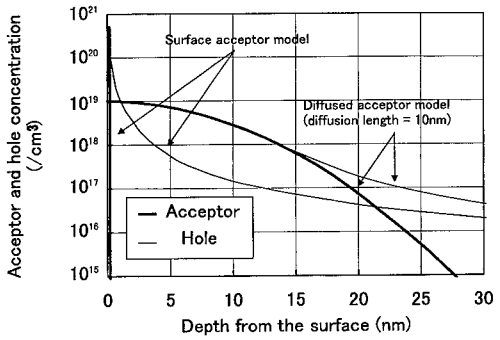


Fig. 5. In-depth profiles of two types of acceptor distribution and corresponding hole distribution used in MOSFET device simulation. One is diffused acceptors expressed in gaussian function with diffusion length of 10 nm and the other is surface acceptors located on the first atomic layer. The surface concentration of the two types of acceptor are equivalent (1×10^{13} cm^{-2}).

3.2. Surface channel properties

In the operation of the FETs, holes near the surface, which are produced by the surface acceptors [16] or the diffused acceptors near the surface [17], are responsible for the p-type surface channel. At the drain side of the channel where channel potential is lower than gate potential, i.e., a negative charge is induced, the acceptors are fully ionized (negative) and the holes (positive) are depleted to form a pinch-off region (Fig. 6).

The channel mobility can be estimated by the following relationship in the slope S of the linear region in the I_{DS}-V_{DS} characteristics,

$$S = \mu_p C_{ox} V_{DS} Z/L \tag{1}$$

where μ_p is the channel hole mobility, C_{ox} is the SiO_x capacitance per unit area, and Z and L are the gate width and length, respectively. From the linear region of $V_{DS} = -1.0$ or -2.0 V in Fig. 3(a), the channel mobility is calculated to be 120 cm^2/V s by the above equation on the condition that the SiO_x thickness is 30 nm and the dielectric constant of SiO_x is equal to that of SiO_2 (3.9). Since the channel mobility was evaluated in the small drain voltage and the large gate voltage, the distribution of holes in the surface channel is deeper than that in the surface conductive layer without gate bias. In the thicker channel, the obtained channel mobility can be estimated to be higher than the Hall mobility (30–40 cm^2/V s) [17] measured on the surface p-type layer of an undoped homoepitaxial film.

On the same surfaces as the depletion mode MOSFETs with Pb or Al gate shown in Fig. 3(a), (b) were fabricated, the MESFETs with gate metals of low electronegativity such as Pb or Al, operated in normally off mode (enhancement mode) [5–8]. The charge transfer from the metals to hydrogen-terminated diamond is large enough to ionize surface acceptors or acceptors near the surface without producing holes. On the other hand, the MESFETs with gate metals of higher electronegativity such as Cu or Ni operated in

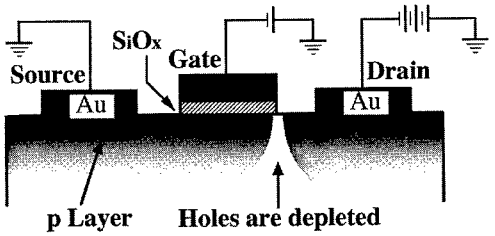


Fig. 6. Schematic cross section of a MOSFET using hydrogen-terminated surface conductive layer, when the channel is pinched off.

depletion mode [7], where the charge transfer is too low to deplete the surface holes completely. Considering that the electronegativity of oxygen is higher than that of carbon (diamond) and that of Si is the same level as those of Cu and Ni, the electron charge transfer from SiO_x to hydrogen-terminated diamond can be hardly expected. The surface holes can not be depleted and the diamond MOSFETs with the SiO_x gate oxide operate in normally on mode (depletion mode). The reduction of carrier concentration of hydrogen terminated diamond surface is needed to obtain the normally off mode (enhancement mode) operation of the diamond MOSFETs.

Generally in GaAs MOSFETs, there a problem in the reduction of transconductance due to the surface state trapping centers which exist in the oxide/GaAs interface. On the other hand, the diamond MOSFETs using hydrogen-terminated surfaces exhibit similar transconductance to those of MESFETs indicating that the scarce density of trapping states is achieved even after the deposition of an oxide layer. This indication is also supported by no hysteresis in the I_{DS}–V_{DS} curve shown in Fig. 3(b). Considering that, in the deep depletion mode GaAs MOSFETs, the degradation of transconductance due to the surface states become prominent above 10^{11} cm^{-2} eV^{-1} [18], the surface states density of hydrogen-terminated diamond is expected to be lower than this value. The low trapping density suggested by the strong dependence of Schottky barrier height (SBH) on the metal electronegativity [14] is also realized under the oxide deposition using SiO evaporation.

3.3. High temperature operations

The temperature dependence of the MOSFET operation has been measured from room temperature to 400°C in air without passivation. The I_{DS}–V_{DS} characteristic of the fabricated Au gate MOSFET operating at an ambient temperature of 300°C in air is shown in Fig. 7. It shows a good characteristic with low leakage and clearly defined saturation and pinch-off. Figure 8 shows the temperature variation of the transconductances of the Au gate MOSFET. At room temperature, the transconductance of the device was 4 mS/mm. The relatively low transconductance was due to the thick gate oxide (70 nm in thickness) and the lower conductivity of the surface p-type layer compared with that of the FET shown in Fig. 2. The transconductance is almost constant until 330°C. The result is consistent with the temperature independence of the carrier concentration obtained by the Hall measurement [17]. Above 350°C, the device operation has been degraded because of the reduction of the surface conduction by oxidation [15]. In order to obtain the operation at higher temperatures, the passivation of the hydrogen-terminated surface by a temperature resistant material is required. Considering that the C—H bondings are stable at least up to 600°C in ultra high vacuum (UHV), the operation temperature is expected to be above 500°C provided that a suitable passivation technology of the active channel is available.

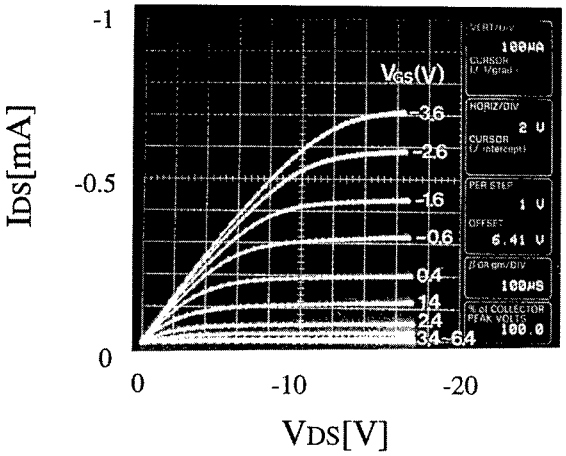


Fig. 7. I_{DS}–V_{DS} characteristics of Au gate MOSFETs at 300°C using the SiO_x as insulator. Au gate with $L_g = 7$ μm and $W_g = 62$ μm on the SiO_x with 70 nm in thickness. The g_m is 2.6 mS/mm.

4. Conclusion

The effect of hydrogen-termination reducing the surface states density is maintained after the oxide deposition using the low-temperature evaporation of SiO. It provides a convenient method for the formation of an insulating film with good interface properties needed for the MOSFET operation in diamond. The transconductance of MOSFET with several μm gate length is 17 mS/mm, which is the highest in diamond

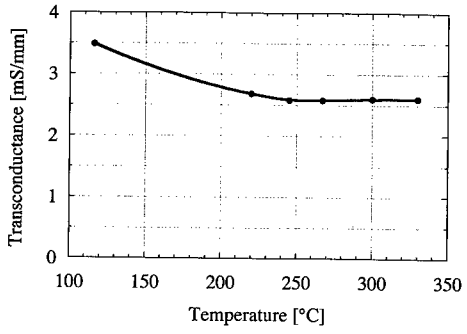


Fig. 8. Temperature dependence of the transconductance g_m of the Au gate MOSFET shown in Fig. 7. The g_m of the FET is 4 mS/mm at room temperature.

MOSFETs and is comparable to the best diamond MESFETs [8,13] using the equivalent surface p-type conductive layer. The I_{DS}–V_{DS} characteristics of the FET were improved by means of the reduction of the gate leakage current in the MOS structure. The stable transistor operation with almost constant transconductance has been observed up to 330°C without the passivation of surfaces. The MOSFETs on the surface p-type conductive layer is already equipped with a very shallow junction structure necessary for very short gate length, e.g., 50 nm. Micro fabrication technology such as small gate length and thin insulator for hydrogen-terminated diamond surfaces will open a new field of nm structure device.

Acknowledgements

The authors are grateful to Prof. I. Ohdomari, Mr S. Morimoto and Mr Y. Shimura of the Micro Technology Laboratory in Waseda University for their cooperation. They also thank Mr T. Shikata and Dr N. Fujimori of Sumitomo Electric Industries, Ltd., and Mr T. Yamashita of Tokyo Gas Co., Ltd. for their collaboration. This work was supported in part by Grant in Aid for Scientific Research (B) (09555103, 10450127) from the Ministry of Education, Science, Sports and Culture of Japan and by the Murata Science Foundation.

References

[1] Shiomi H, Nishibayashi Y, Fujimori N. Jpn J Appl Phys 1989;28:L2153.
[2] Gildenblat G Sh, Grot SA, Badzian AR. Proc IEEE 1991;79:647.
[3] Zeisse CR, Hewett CA, Nguyen R, Zeidler JR, Wilson RG. IEEE Electron Device Lett 1991;EDL-12:602.
[4] Dreifus DL, Tesmer AJ, Holms JS, Plano LS. Proceedings of Second International High Temperature Electronics Conference, vol. 1, 1994. p. 29.
[5] Kawarada H, Aoki M, Itoh M. Appl Phys Lett 1994;65:1563.
[6] Kawarada H, Itoh M, Hokazono A. Jpn J Appl Phys 1996;35:L1165.
[7] Hokazono A, Ishikura T, Nakamura K, Yamashita S, Kawarada H. Diam Relat Mater 1997;6:339.
[8] Gluche P, Aleksov A, Vescan A, Ebert W, Kohn E. IEEE Electron Dev Lett 1997;EDL-18:547.
[9] Yun Y, Maki T, Kobayashi T. J Appl Phys 1997;82:3422.
[10] Kawarada H, Aoki M, Sasaki H, Tsugawa K. Diam Relat Mater 1994;3:961.
[11] Tsugawa K, Hokazono A, Kawarada H. Proceedings of International Conference on Silicon Carbide, III-Nitrides and Related Materials, Stockholm 1997, Material Science Forum, Trans Tech, Winterthur, vols. 266–268, 1998. p. 977.
[12] Bennett AJ, Roth LM. Phys Rev 1971;B4:2686.
[13] Hokazono A, Kawarada H. Jpn J Appl Phys 1997;36:7133.
[14] Kawarada H. Surf Sci Rep 1996;26:205.
[15] Ono M, Saito M, Yoshitomi T, Fiegna C, Ohguro T, Sasaki H, Iwai H. J Vac Sci Technol 1995;B13:1740.
[16] Kawarada H, Sasaki H, Sato A. Phys Rev 1995;B52:11,351.
[17] Hayashi K, Yamanaka S, Ohkushi H, Kajimura K. Appl Phys Lett 1996;68:376.
[18] Lile DL, Collins DA. Thin Solid Films 1979;56:225.

PERGAMON

Solid-State Electronics 43 (1999) 1473–1478

SOLID-STATE ELECTRONICS

GaN MODFET microwave power technology for future generation radar and communications systems

D.E. Grider*, N.X. Nguyen, C. Nguyen

Microelectronics Laboratory HRL Laboratories, LLC, 3011 Malibu Canyon Rd., Malibu, CA 91362, USA

Abstract

In order to gain a better understanding of the role that GaN MODFET technology will play in future generation radar and communications systems, a comparison of the state-of-the-art performance of alternative microwave power technologies will be reviewed. The relative advantages and limitations of each technology will be discussed in relation to system needs. Device results from recent MBE-grown GaN MODFETs will also be presented. In particular, 0.25 μm gate GaN MODFETs grown by MBE have been shown to exhibit less than 5% variation in maximum drain current density (I_{dmax}) from the center to the edge of a 2 inch wafer. This level of uniformity is a substantially higher than that normally found in MOCVD-grown GaN devices (~28% variation). In addition, evidence is also presented to demonstrate the excellent reproducibility of MBE-grown GaN MODFETs. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Comparison of GaN MODFET and other RF power technologies

The recent rapid progress in GaN MODFET technology provides clear convincing evidence that it has the potential to revolutionize the field of microwave power electronics. Consequently, this swiftly evolving technology will likely have a dramatic impact on a wide range of systems such as future generation radar and satellite communications systems. In order to understand the intrinsic advantages of GaN MODFET technology for microwave power electronic applications, it is extremely useful to compare the existing performance of this technology with alternative RF technologies over the frequency band from a few hundred MHz up to 20 GHz.

The most important advantage of GaN MODFET technology is its extremely high RF power densities, as measured in Watts per millimeter of FET gate periphery (i.e. gate width), with that of alternative RF power technologies. Fig. 1 [1–19] shows such a comparison over the frequency band from less than 1 up to 20 GHz. Note that the results for each technology have been grouped as an aid for comparison only. For pHEMT technology, which is currently being used in a number of microwave power applications such as active array radar T/R modules, the typical power density for these devices is approximately 0.5 W/mm over a broad range of frequencies. By stark comparison, the power density of GaN MODFET technology ranges from 4 to 14× larger than pHEMTs. In particular, GaN MODFETs have been demonstrated with RF power densities of 6.8 and 3.3 Watts/mm at 10 and 18 GHz, respectively [16–19]. Both of these values represent the highest power densities ever reported at these frequencies for any FET technology.

Fig. 1 also shows that SiC MESFET technology exhibits CW and pulsed RF power densities from 1.0 W/mm up to approximately 3.3 W/mm. These power

* Corresponding author. Fax: +1-310-317-5450.

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00091-X

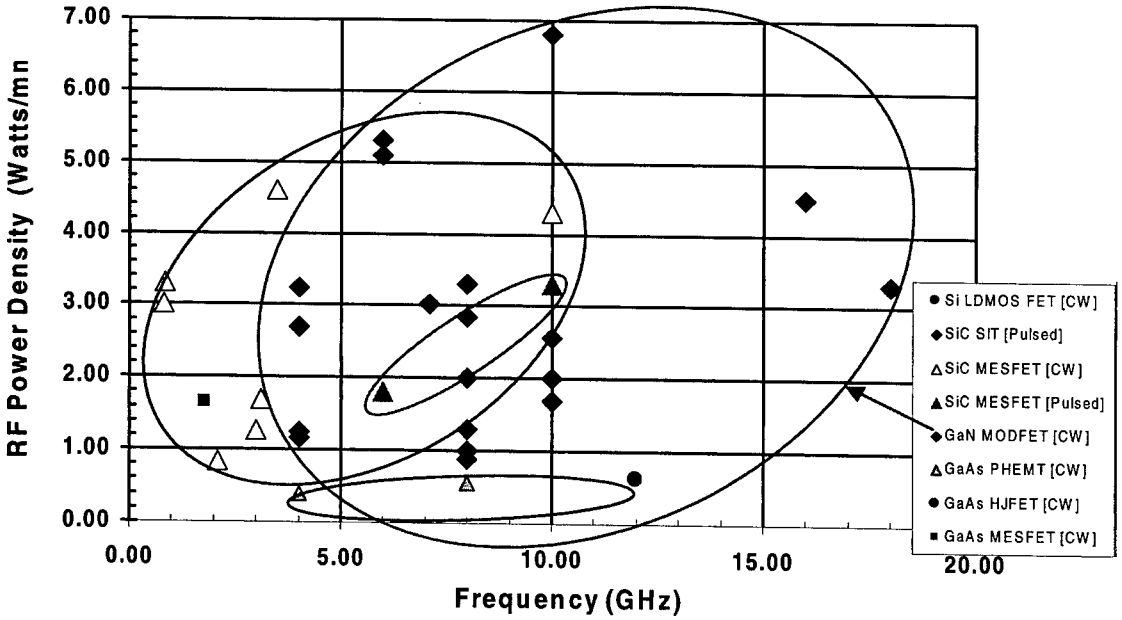


Fig. 1. Comparison of the RF power densities of a GaN MODFETs and a number of alternative power technologies at frequencies up to 20 GHz.

densities are substantially lower than the highest values reported for GaN MODFETs at a given frequency. In addition, these SiC MESFET power densities have only been demonstrated up to frequencies of 1.8 and 10 GHz for CW and pulsed operation, respectively. By comparison, CW operations of GaN MODFETs have been demonstrated with 3.3 W/mm up to 18 GHz. The power density advantage of GaN MODFETs com-

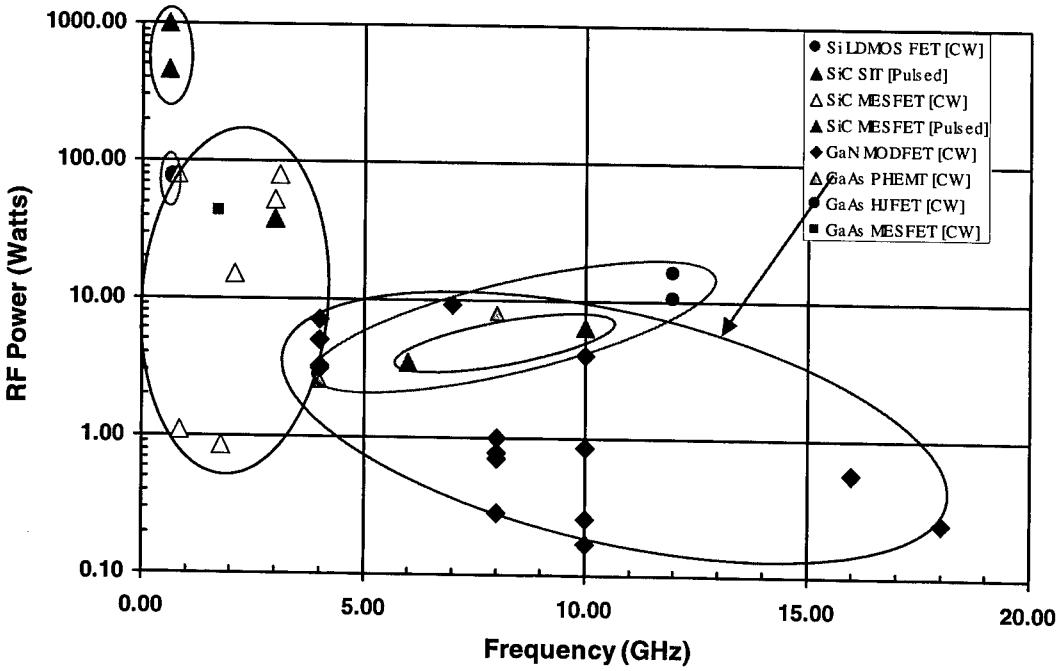


Fig. 2. Comparison of the RF output power for GaN MODFETs and a number of alternative power technologies for frequencies up to 20 GHz.

pared with SiC MESFETs at frequencies above 10 GHz are due in large part to the modulation doping and the resultant improved transport in the GaN devices.

The total RF output power of GaN MODFETs however is still somewhat lower than other alternative RF power technologies as shown in Fig. 2 [1–19]. Again, please note that the results for each technology have been grouped as an aid for comparison only. At low frequency (< 1.0 GHz), SiC SITs and Si-based LDMOS dominate the RF power output performance comparison with values from just under 100 Watts up to 1000 W [7,9]. Up to 3.0 GHz, large gate periphery SiC MESFETs have been demonstrated with RF output powers up to 53 W under CW operation [20]. Under pulsed operation, SiC MESFETs with output powers of over 6 W at 10 GHz. In comparison, the highest RF output powers demonstrated by different groups for GaN MODFET devices is 3.1 and 4.0 W at 4.0 and 10.0 GHz, respectively [14,18,19].

The reason for this apparent disparity is due, in large part, to the relative immaturity of GaN MODFET microwave power technology. Until recently, most of the power measurement reported for this technology have been carried out using devices with gate peripheries of a few tenths of a millimeter. This situation is changing rapidly, however, as GaN MODFETs are beginning to be fabricated with gate peripheries of up to 2 mm and even higher. This scaling up process, however, will also undoubtedly introduce an entire new set of challenges and issues that will need to be addressed.

2. Thermal management

One issue that needs to be addressed, in particular, is the apparent decrease in actual RF power density that is observed as the gate periphery is increased [17,18]. This phenomenon is most likely due to thermal effects or GaN MODFET device/material nonuniformity across a multi-finger device, or some combination of these factors. Thermal management of these high power GaN MODFETs can be addressed by using high thermal conductivity SiC substrates instead of Sapphire substrates for the GaN MODFETs. The thermal conductivity of SiC is 3.8 W/(K cm), which is almost as high as that of copper, in contrast to a value of approximately 0.4 W/(K cm) for sapphire. If the device and circuit topology is such that the self-generated heat is to be removed through the substrate, SiC is the best substrate for high power GaN devices. Furthermore, because of the smaller lattice mismatch with GaN, SiC is also a superior substrate for device heteroepitaxy. The best GaN-based FET performance is achieved when grown on SiC substrates. Most notable is the recent result from Cree Research [19]. In this work, Sheppard and co-workers demonstrated a record power density of 6.8 W/mm with 52.4% power-added-efficiency (PAE) at 10 GHz. The major disadvantages of using SiC as a substrate are the high cost and small wafer size. Since these are current limitations, presumably influenced by economic concerns rather than fundamental difficulties, we believe this approach will be an important one in the development of GaN high power electronics.

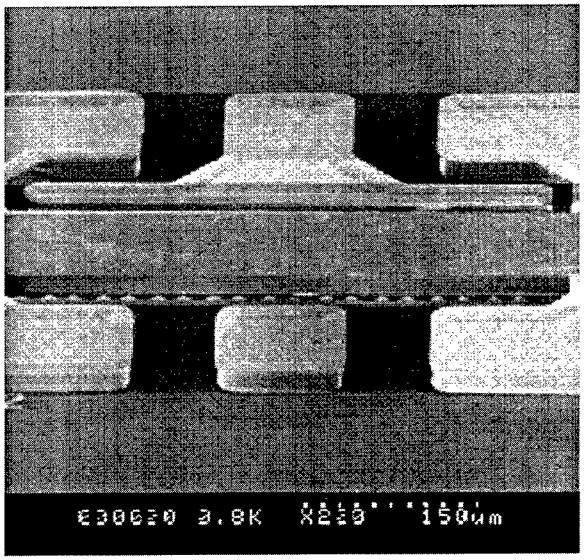


Fig. 3. SEM micrograph of a multi-finger device with 30 μm thick Au heat-spreader.

Alternatively, there are a number of heat-spreader/flip-chip thermal management approaches that effectively take the heat out of the device through the source and/or drain contacts [18,20,21]. The approach at HRL is a combination of (i) lateral heat spreading from the active area of the device to a larger dice area using a thick Au heat spreader and (ii) heat sinking by flip-chip mounting the device on a ceramic AlN substrate. We have developed a fabrication technique for the thick Au heat spreader connecting the source contacts of the individual fingers. Fig. 3 is an SEM micrograph of the contact area and Au heat spreader for a 10-finger device fabricated on a sapphire substrate. The 30 μm thick Au plated air-bridge provided a ground plane as well as a high thermal conductivity spreader to help equalize the temperature distribution among the active fingers.

3. High uniformity MBE-grown GaN MODFETs

The goal of achieving good across-wafer uniformity of GaN MODFET devices is primarily a materials

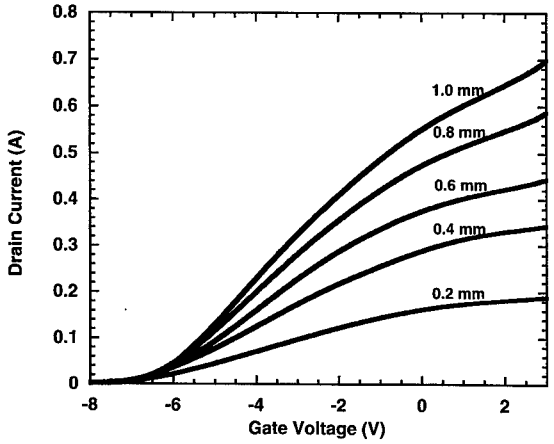


Fig. 4. Transfer characteristics of GaN MODFETs with gate peripheries ranging from 0.2 to 1 mm at $V_{DS}=10$ V.

issue. In this section, recent device results will be presented which indicate that MBE-grown GaN MODFETs exhibit very good across-wafer uniformity and wafer-to-wafer reproducibility in DC and RF characteristics. In particular the maximum DC drain current (I_{dmax}) in MBE-grown GaN MODFETs varies by less than 5% across a 2-inch sapphire wafer [20]. This is far superior to MOCVD-grown GaN MODFETs which have been shown to exhibit I_{dmax} uniformity of 28% and higher across sapphire wafer samples that are $\frac{1}{4}$ of a 2-inch wafer or even smaller [20].

We have recently demonstrated that gas-source molecular beam epitaxy (GS-MBE) can be used to grow device quality GaN MODFET heterostructure materials on sapphire substrates [20,21]. This GS-MBE system makes use of an RF-plasma nitrogen source along with solid source materials for the group III elements and the dopants. The details of the MBE growth process, including the actual MBE-grown GaN MODFET device structure are presented elsewhere [20–21].

The focus of this section will be on the enhanced uniformity and reproducibility of GaN MODFET devices fabricated using GS-MBE. A 0.25 μm gate length baseline fabrication process was utilized to fabricate MBE-grown GaN MODFET devices with gate peripheries of from 0.2 up to 1 mm. For thermal management on the wider devices, this process included a thick 30 μm thick Au heat spreader. Measurements of the DC and RF characteristics of these MBE-grown GaN MODFETs have been carried out.

Fig. 4 shows the transfer characteristics of 0.25 μm MBE GaN MODFETs with gate peripheries from 0.2 up to 1 mm. The drain bias for these measurements was 10 V. The transconductance for these devices peaked at a value of 200 mS/mm, at a gate bias of approximately −4.0 V. The high value of +2.0 V for the forward biased Schottky gate turn-on voltage made

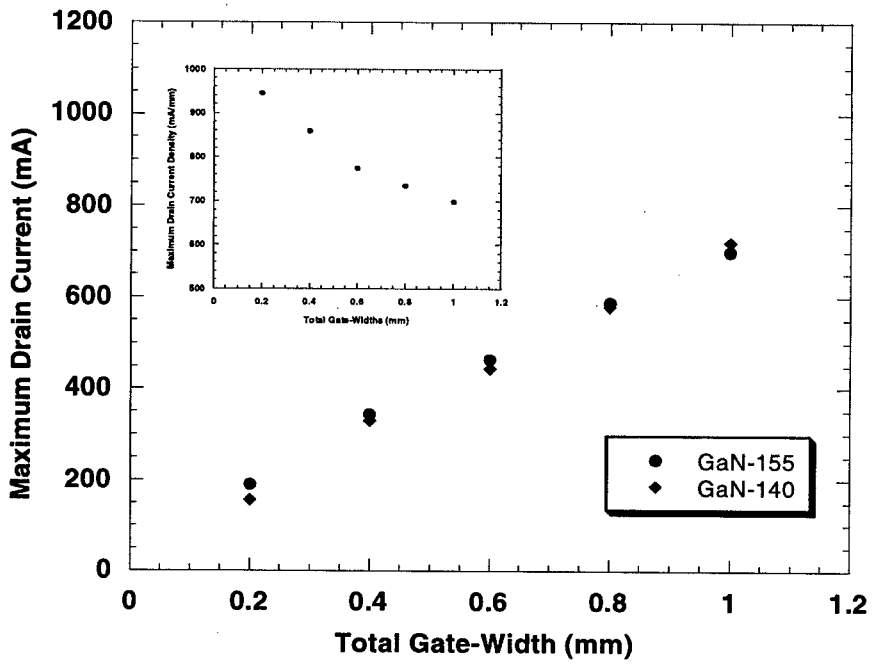


Fig. 5. Dependence of GaN MODFET maximum drain current and current density (see inset) on gate periphery at $V_{DS}=10$ V for two different wafers.

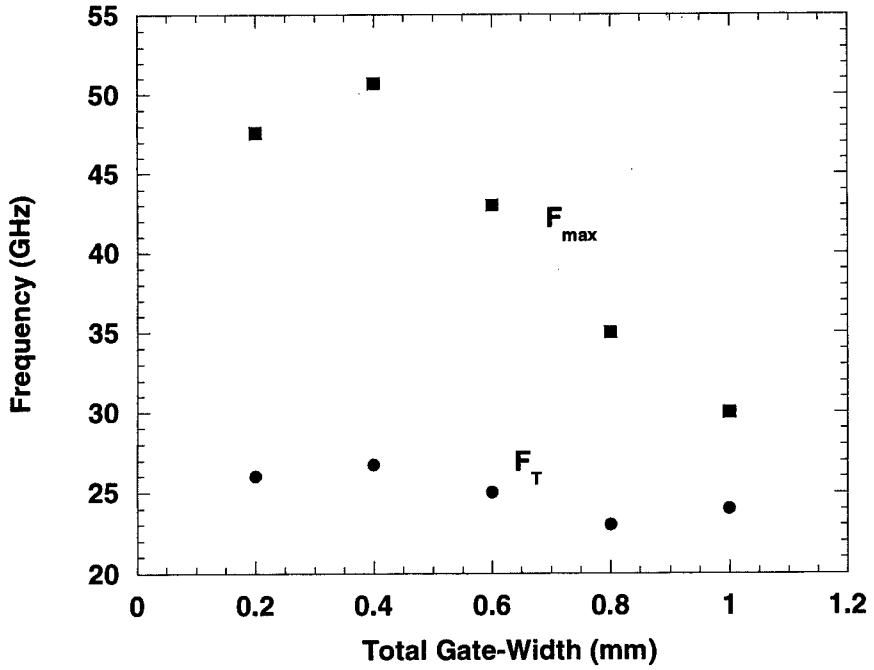


Fig. 6. Cut-off frequencies extrapolated from S-parameter measurements for devices of varying gate peripheries.

possible the relatively large gate voltage swing from pinch-off at −8.0 V up to +3.0 V. These devices also exhibited gate-to-drain breakdown voltages of approximately 40 V.

As can be seen in Fig. 5, the maximum drain current in these GaN MODFET devices scales linearly with gate periphery, up to a maximum drain current of 700 mA for a 1.0 mm wide device. There is, however, some roll-off in the maximum drain current density, from 950 mA/mm for 0.2 mm device down to 700 mA/mm for 1.0 mm device, as shown in the inset in Fig. 5. This roll-off may be attributed to the higher thermal degradation in larger devices, but further studies are needed to quantify this effect. Although these GaN MODFETs include the thick Au heat-spreader to extract heat from the source contacts, for these dc measurements, the devices were not flipped onto a heat-sink (e.g. high thermal conducting substrate AlN).

S-parameter measurements of the RF performance of these GaN devices were also carried out as a function of gate periphery, and the results are shown in Fig. 6. The unity-current-gain-cutoff-frequency (f_T) of these GaN devices remained approximately constant, varying between 23 to 26 GHz as the gate periphery increased from 0.2 up to 1.0 mm. Although these values are lower than the best reported GaN MODFETs grown by MOCVD [15], the independence of f_T on gate widths is another strong indication of the uniformity of MBE-grown GaN. As the growth of GaN materials by MBE mature, f_T of MBE-GaN MODFETs is expected to equal that currently achieved by MOCVD. The extrapolated maximum oscillation frequency (f_{max}), on the other hand, decreased from 50 to 30 GHz with increasing gate periphery from 0.4 to 1.0 mm, respectively. This effect may be associated with increasing parasitic source inductance for the wider GaN MODFETs, and alternative device design is underway to address this issue.

Large signals measurement have been performed on these MBE-grown GaN MODFETs. The power performance is currently limited by the drain current compression at microwave frequencies. A similar phenomenon has been observed in many devices grown by MOCVD [22]. It has been proposed that carrier trapping in either the buffer or the AlGaN barrier layer or both is responsible for this undesirable effect. We are currently investigating this issue.

4. Summary

In summary, a comparison has been presented of the microwave power performance of GaN MODFET technology with a number of other mature and still-developing power technologies including GaAs-based pHEMTs, SiC SITs, SiC MESFETs and others. From a power density perspective, GaN MODFETs represent the obvious technology of choice over a broad range of frequencies from approximately 4.0 up to 18 GHz, and possibly higher. However, the total RF out-

put power performance of GaN MODFETs still lags that of pHEMTs and even SiC MESFETs due to the fact that, until recently, the focus has been on relatively small gate periphery GaN devices. Given the superior power density of GaN MODFET technology, however, the total output power of these devices should scale up with increasing gate periphery in the future.

Recent results from MBE-grown GaN MODFETs with gate peripheries ranging from 0.2 to 1.0 mm have also been presented. These devices exhibited excellent DC and small signal RF characteristics, which appear to scale well with gate periphery. The 1 mm wide GaN MODFET produced a maximum drain current (I_{dmax}) of 700 mA. In addition, the unity-current-gain-cutoff-frequency for MBE-grown GaN MODFETs was 25 GHz and did not vary significantly with increasing gate periphery. These results indicate that MBE GaN MODFETs can be scaled to larger gate peripheries, and thereby higher values for the total RF output power.

Acknowledgements

This work has been supported in part by Dr. Elliot Brown and Dr. Gerald Witt under the MAFET-Thrust III program.

References

[1] Kraemer B, Basset R, Chye P, Day D, Wei J. In: 1994 IEEE MTT-S Digest, 1994. p. 683.
[2] Cardullo M, Page C, Teeter D, Plaztker A. In: 1996 IEEE MTT-S Digest, 1996. p. 145.
[3] LeSage SR, Mendes MJ, Sullivan FJ. In: 1990 GaAs IC Symposium Digest, 1990. p. 353.
[4] Chen YC, Wu CS, Pao CK, Cole M, Bardai Z, Hou LD, Midford TA, Cisco TC. GaAs IC Symposium Digest 1995;1996:281.
[5] Teeter D, Bouthilette S, Aucoin L, Platzker A, Alarfo C, Bradford D. In: 1995 IEEE MTT-S Digest, 1995. p. 323.
[6] Sarkissian R, Basset R, Shingu Z, Ono F, in press.
[7] Weitzel CE, private communication.
[8] Morse AW, Esker PM, Clarke RC, Brandt CD, Siergiej RR, Agarwal AK. In: 1996 IEEE MTT-S, 1996. p. 677.
[9] Morse AW, Esker PM, Sriram S, Hawkins JJ, Chen LS, Ostop JA, Smith TJ, Davis CD, Barron RR, Clarke RC, Siergiej RR, Brandt CD. In: 1997 IEEE MTT-S Digest, 1997. p. 53.
[10] Allen ST, Sadler RA, Alcorn TS, Palmour JW, Carter Jr. CH. In: 1997 IEEE MTT-S Digest, 1997. p. 57.
[11] Allen ST, Palmour JW, Carter CH, Weitzel CE, Moore KE, Nordquist KJ, Pond LL. In: 1996 IEEE MTT-S Digest, 1996. p. 681.
[12] Sadler RA, Allen ST, Alcorn TS, Pribble WL, Sumakeris J, Palmour JW, Kehias LT. SiC MESFET with Output Power of 50 W CW at S-Band. In: 56th Annual Device Research Conf., Charlottesville, VA, 1998.
[13] Wu YF, Keller BP, Keller S, Nguyen NX, Le M, Nguyen C, Jenkins TJ, Kehias LT, Denbaars SP. IEEE Electron Device Lett 1997;18:438.
[14] Wu YF, Keller BP, Keller S, Fini P, Mishra U, Le M, Nguyen NX, Nguyen C, Widman D. Electron Lett 1997;33:1742.
[15] Nguyen NX, Keller BP, Keller S, Wu YF, Le M, Nguyen C, Denbaars SP, Mishra UK, Grider D. Electron Lett 1997;33:334.
[16] Wu YF, Keller S, Kozodoy P, Keller BP, Parikh P, Kapolnek D, Denbaars SP, Mishra UK. IEEE Electron Device Lett 1997;18:290.
[17] Wu YF, Keller BP, Keller S, Nguyen NX, Le M, Nguyen C, Jenkins TJ, Kehias LT, Denbaars SP, Mishra UK. High Speed and High Power AlGaN/GaN MODFETs. In: 55th Annual Device Research conf, Ft Collins, CO, 1997.
[18] Thiebault BJ, Keller BP, Wu YF, Ffini P, Mishra UK, Nguyen C, Nguyen NX, Le M. Performance and Large Area Flip-Chip Bonded AlGaN/GaN MODFETs. In: International Electron Devices Meeting (IEDM-97) Digest, 1997.
[19] Sheppard ST. High Power Microwave GaN/AlGaN HEMTs on Silicon Carbide. In: 56th Annual Device Research Conf., Charlottesville, VA, 1998.
[20] Nguyen C, Nguyen NX, Le M, Grider DE. Electronics Lett 1998;34:309.
[21] Nguyen NX, Nguyen C, Grider DE. Electronics Lett 1998;34:811.
[22] Khan MA, Shur MS, Chen QC, Kuznia JN. Electronics Lett 1994;30:2175.

PERGAMON

Solid-State Electronics 43 (1999) 1479–1482

SOLID-STATE ELECTRONICS

Progress towards ultra-wideband AlGaN/GaN MMICs

J.C. Zolper*

Office of Naval Research, Electronics Division, Code 312, 800 North Quincy Street, Arlington, VA 22217, USA

Abstract

The AlGaN/GaN material system possesses fundamental properties that make it an ideal candidate for high power microwave devices. This has been known for over 30 years, however, significant device progress has only come recently. This article gives a summary of AlGaN/GaN electronic device progress, outlines future device directions, and discusses the use of these new devices in power amplifier circuits. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Gallium nitride; AlGaN/GaN; Microwave power devices

1. Introduction

The potential of the wide bandgap, nitride-based, semiconductors (GaN, AlN, and AlGaN) for the realization of high power, high frequency transistors has been well documented [1–3]. This potential is due to the advantageous material properties, summarized in Table 1, along with the existence of the AlGaN/GaN heterostructure [4,5]. The later allows modulation doping to form a high mobility two dimensional electron gas (2DEG) and, more importantly, the formation of piezoelectronically induced sheet carriers. The piezoelectric effect is at least three times stronger in these materials than in GaAs (see values for e_{31} and e_{33} in Table 1) and contributes to the realization of high sheet electron densities (up to 5×10^{13} cm^{-2} is predicted for an AlN/GaN interface) in AlGaN/GaN HEMT [6,7]. Coupling this high sheet carrier density with the high breakdown field of GaN yields predictions of microwave power densities greater then 10 W/mm at 10 GHz. This is approximately 10 times higher than that achievable with a AlGaAs/GaAs power HEMT. The status of AlGaN/GaN HEMTs is reviewed in section 2.

2. AlGaN/GaN microwave transistors

The first reports of GaN based transistors were by Khan et al., with the demonstration of a GaN MESFET and an AlGaN/GaN HEMT [8,9]. Both transistors had gate lengths of 4 μm with the MESFET having a g_m of 23 mS/mm and I_{DS}(max) of ~180 mA/mm (@V_{GS}=0 V, V_{DS}=20 V). The AlGaN/GaN HEMT achieved a g_m of 28 mS/mm at 300 K (46 mS/mm at 77 K) and I_{DS}(max) of ~50 mA/mm (@V_{GS}=0.5 V, V_{DS}=25 V). The HEMT structure had a 2 DEG mobility of 563 cm^2/V s at 300 K and 1517 cm^2/V s at 77 K. The first microwave results were published by Binari et al., for a GaN MESFET with a demonstrated f_t of 8 GHz and a f_{max} of 17 GHz for a gate length of 0.7 μm [10].

Since these early results, significant improvements have been made in material quality and device processing. AlGaN/GaN 2DEG mobilties up to 2019 cm^2/V s have been reported for growth on 6H-SiC substrates

* Fax: +1-703-696-2611.
E-mail address: zolperj@onr.navy.mil (J.C. Zolper)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00092-1

Table 1
Summary of key material parameters for AlGaAs/GaAs, 4H SiC, and AlGaN/GaN transistors (after Ref. [4–6])

Metric	AlGaAs/GaAs	4H SiC	AlGaN/GaN
Maximum sheet electron density (cm^{-2})	$2–3 \times 10^{12}$	N/A	$1–5 \times 10^{13}$
Breakdown field (V/cm) ($\times 10^5$)	4	20	33
2D electron mobility (cm^2/V s)	8500	N/A	2000
Saturated electron velocity ($\times 10^7$ cm/s)	1.0	2.0	2.2
Thermal conductivity (W/cm K)	0.53	4.9[a]	1.3[b]
Piezoelectric coefficient (C/m^2)			
e_{31}	0.093		−0.36
e_{33}	−0.185	0.2	1.0

[a] This is for undoped SiC. Doped or SI SiC has a thermal conductivity of 3.3 W/cm K.
[b] This is for highly dislocated GaN. Theory predicts 1.7 W/cm K for low dislocation material [25].

and ~1600 cm^2/V s for growth on sapphire substrates [11]. Small gate width transistors have achieved a f_t of up to 75 GHz and a f_{max} of up to 140 GHz [12]. While these small signal results are impressive, the true test of these devices is their output power at microwave frequencies. Fig. 1 shows the evolution of total output power and power density of AlGaN/GaN HEMTs versus time. The best power density reported is for HEMTs grown on a semi-insulating 4H SiC substrate with 6.8 W/mm demonstrated at 10 GHz and 4.1 W/mm at 16 GHz [13]. Total power results have also been pushed up with 7.6 W achieved at 4 GHz for HEMTs grown on sapphire and flip-chip mounted on AlN carriers [14]. It is notable that both the total power and power density have seen marked increases in the last year as more research is shifting from photonic (LEDs and lasers) to electronic applications. The improvements have partially resulted from an increased understanding and application of the piezoelectric effect at the AlGaN/GaN interface that induces large sheet carrier concentrations [7]. This has resulted in current densities up to 1.6 A/mm for a AlGaN/GaN HEMT with a gate length of 0.25 μm [15]. A comparable AlGaAs/GaAs HEMT would have a current density of ~0.5 A/cm^2.

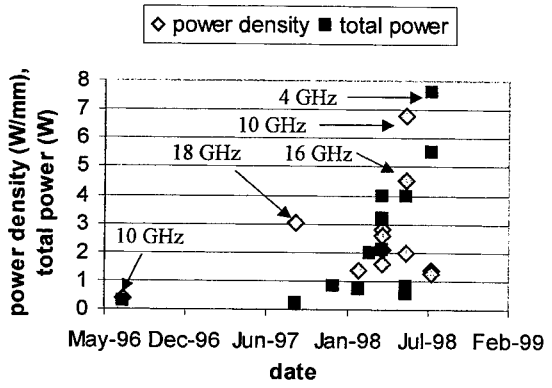


Fig. 1. Total power (W) and power density (W/mm) for AlGaN/GaN HEMTs versus time of publication. The frequency of measurement is noted for key results.

The impressive results discussed above can be considered 'hero' results. Continued material and device improvements are needed to make this material system a reproducible and manufacturability product. Foremost, trapping effects which cause current and power compression must be eliminated [16,17]. The source of the traps is under investigation; however, the traps are likely due to one of (or a combination of) four sources: (1) surface states, (2) deep levels in the AlGaN barrier, (3) states in the highly defective low temperature buffer, and (4) states associated with dislocations.

If the traps are due to surface states then some of the same techniques applied to passivating surfaces in GaAs and InP based devices may be employed. These include dielectric passivation, undoped surface layers, or depleted p-type surface layers. If the traps are in the AlGaN barrier, this may be minimized by improving the growth conditions (eliminating undesirable background impurities such as oxygen or DX-like centers with Si). If the traps are due to the low temperature buffer, this may be reduced by proper device design (e.g. increasing the high temperature buffer thickness). To reduce detrimental effects of dislocations, lateral epitaxial overgrowth (LEO) may be employed [18–20]. LEO is a technique whereby dislocation 'filtering' is accomplished by selective area nucleation and growth followed by lateral growth over a mask. The laterally overgrown region can have many orders-of-magnitude lower density of threading dislocations than the window region. Already p/n diodes fabricated in reduced dislocation density LEO GaN have shown three to four orders-of-magnitude lower reverse leakage current than highly dislocated 'conventional' material [21]. Work is ongoing in all the above areas.

3. Future device directions

The progress in field effect AlGaN/GaN transistors has been impressive but these may not meet the linearity requirements of future military electromagnetic systems. Therefore, III-nitride based heterostructure bipolar transistors (HBTs) are now being studied since, in GaAs and InP, HBTs have demonstrated improved linearity over the FET counterpart. The first AlGaN/GaN HBT has been reported, however, the current gain was only 3 [22]. The initial result was limited by a high base resistance [22]. Key issues remain to be resolved for group III-nitride HBTs. First among these is how to overcome the high ionization energy of p(Mg)-GaN that causes a high based resistance. A potential solution includes invoking the piezoelectric field to induce free holes and thereby reduce the base resistance. A second concern is achieving sufficiently long minority carrier lifetimes in the base to realize good current gain.

4. Amplifier circuit impacts

The material and device parameters discussed above enable new approaches to wide bandwidth power amplifiers. First, since the characteristic impedance of the AlGaN/GaN transistors is roughly three times that of a similar size AlGaAs/GaAs device, impedance matching and power combining is simpler. Second, the factor of three increase in thermal conductivity and the factor of eight increase in dielectric breakdown strength for AlGaN/GaN compared to AlGaAs/GaAs enables high-efficiency, class B, push/pull amplifier architectures. Class B push/pull high power amplifier designs with AlGaAs/GaAs transistors require a compromise in operating power or frequency to accommodate the two times higher breakdown needed in class B over class A circuits since the thermal and dielectric limits are comparable in this material. This is not the case for AlGaN/GaN transistors which should operate at full power and frequency in a class B push/pull design [23]. Third, the higher power density capability of AlGaN/GaN allows narrower transistor fingers (smaller W_g) for the same power level thereby reducing C_{GS}, since C_{GS} is proportional to W_g. Reducing C_{GS} allows a larger bandwidth since the upper frequency limit (f_{high}) is given by:

$$f_{high} = \frac{1}{2\pi Z_o C_{GS}} \quad (1)$$

where Z_o is the characteristic load impedance.

The improved transistor performance described above will significantly improve the bandwidth of amplifiers incorporating them, however, to realize still larger bandwidths, new circuit approaches are also required. While traditional matching networks are effective at single frequencies, distributed approaches or traveling wave designs will be needed to achieve multi-octave bandwidths. In these approaches, however, care must be taken to not to loose efficiency via backward wave propagation. Furthermore, the application of analog circuit design approaches to microwave circuits will enable conversion of f_t-limited lumped element circuits to f_{max}-limited distributed designs [24]. This area is the focus of a Multidisciplinary University Research Initiative (MURI) managed by the Office of Naval Research[1].

5. Conclusions

Rapid progress in group III-nitride material has enabled impressive demonstrations of high power and high frequency field effect transistors. While further work is needed to improve repeatability and scalability, the results show the clear potential of these materials for high power amplifiers. Future work will involve the study of HBTs in this system to explore the potential for improved linearity. Moreover, the true utility will be seen when these devices are used to enable wide bandwidth solid state power amplifiers.

Acknowledgements

This article was prepared by a US Government employee. Pursuant to Title 17, US Code Section 05, it is not subject to copyright. Approved for public release, distribution unlimited.

References

[1] Baliga BJ. IEEE Electron Device Lett 1989;10:455.
[2] Chow TP, Tyagi R. IEEE Trans Elec Dev 1994;41:1481.
[3] Trew RJ, Shin MW, Gatto V. Solid State Electronics 1997;41:1561.
[4] Yoder MN. IEEE Trans Elec Dev 1996;43:1633.
[5] Chow TP, Ramungul N, Ghezzo M. Materials Research Society, Symposium E, Boston MA, Fall 1997. p. 89.
[6] Bykhovski AD, Gelmont BL, Shur MS. J Appl Phys 1997;81:6332.

[1] MURI teams are led by U. Mishra at University of California at Santa Barbara and L. Eastman at Cornell University. This is funded by the Director of Defense Research and Engineering (DDR&E), Dr Robert Trew and managed by J. C. Zolper at ONR.

[7] Asbeck PM, Yu ET, Lau SS, Sullivan GJ, Van Hove J, Redwing J. Electronic Letts 1997;33:1230.
[8] Khan MA, Kuznia JN, Bhattarai AR, Olsen DT. Appl Phys Lett 1993;62:1786.
[9] Khan MA, Bhattarai AR, Kuznia JN, Olsen DT. Appl Phys Lett 1993;63:1214.
[10] Binari SC, Rowland LB, Kruppa W, Kelner G, Doverspike K, Gaskill DK. Electronics Lett 1994;30:1248.
[11] Gaska R, Yang JW, Osinsky A, Chen Q, Khan MA, Orlov AO, Snider GL, Shur MS. Appl Phys Lett 1998;72:707.
[12] Eastman LF. In: Materials Research Society, San Francisco, CA, Symposium F, 13–15 April, 1998.
[13] Sheppard ST, Doverspike K, Pribble WL, Allen ST, Palmour JW. In: 56th Device Research Conference, Charlottesville, VA, 22–24 June, 1998, paper V.B-5, 1998.
[14] Wu Y-F, Thibeault BJ, Keller BP, Keller S, Denbaars SP, Mishra UK. In: Topic Workshop on Heterostructure Microelectronics, Kanagawa, Japan, 31 August–2 September, 1998 paper S4-5, 1998.
[15] Thibeault BJ, Keller BP, Wu Y-F, Fina P, Mishra UK, Nguyen C, Nyuyen NX, Le M. IEEE IEDM Washington, DC, 7–10 December, 1997. p. 569.
[16] Kruppa W, Binari SC, Doverspike K. Elec Letts 1995;31:1951.
[17] Thibeault BJ, Wu Y-F, Keller S, Keller BP, Denbaars SP, Mishra UK. In: Nitride Workshop, 4–6 August, 1998, St Louis, MO, 1998.
[18] Usai A, Sunakawa H, Sakai A, Yamaguchi A. Jap J Appl Phys, part 2 1997;36:L899.
[19] Kapolnek D, Keller S, Vetury R, Underwood RD, Kozodoy P, Denbaars SP, Mishra UK. Appl Phys Lett 1997;71:1204.
[20] Zheleva TS, Nam O-H, Bremser MD, Davis RF. Appl Phys Lett 1997;71:2472.
[21] Kozodoy P, Ibbetson JP, Marchand H, Fini PT, Keller S, Speck JS, DenBaars SP, Mishra UK. Appl Phys Lett 1998;73,975.
[22] McCarthy LS, Kozodoy P, Denbaars SP, Rodwell M, Mishra UK. In: 25th International Symposium on Compound Semiconductors, 12–16 October, 1998, Nara, Japan, paper We2A-1, 1998.
[23] Yoder MN. In: Proceedings of IEEE/Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits, Ithaca, New York, 4–6 August, 1997. p. 3.
[24] Rodwell M. UCSB (private communication).
[25] Slack GA. J Phys Chem Solids 1973;34:321.



PERGAMON

Solid-State Electronics 43 (1999) 1483–1488

SOLID-STATE ELECTRONICS

Formation processes and properties of Schottky and ohmic contacts on *n*-type GaN for field effect transistor applications

Yuji Koyama, Tamotsu Hashizume*, Hideki Hasegawa

Research Center for Interface Quantum Electronics (RCIQE) and Graduate School of Electronics and Information Engineering, Hokkaido University, Sapporo, 060-8628, Japan

Abstract

In view of applications to FETs, an attempt was made to characterize and optimize vacuum deposition-based formation processes of Schottky and ohmic contacts on *n*-type GaN. Detailed I–V and XPS measurements were made, changing surface treatments and metal species. Pre-deposition surface treatments were found to be crucial to obtain good electrical characteristics. Au Schottky contacts formed on warm NH_4OH treated surfaces showed nearly ideal thermionic-emission I–V characteristics with a highest Schottky barrier height of 1.03 eV. For ohmic contacts, Ti/Al contacts were investigated, using the circular transmission line model (TLM) method. Excellent ohmic characteristics with a minimum contact resistance of $5–8 \times 10^{-5}$ Ω cm^{-2} were obtained by annealing at 600°C for 1 min in N_2 atmosphere. Finally, narrow ring Au Schottky gates with Ti/Al ohmic electrodes were successfully fabricated, demonstrating applicability of the present contact formation processes for FET applications. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The GaN-based material system is of great interest, not only for recently reported high-efficiency blue/green light emitting diodes (LEDs) and laser diodes (LDs), but also for realization of devices for high temperature electronics and high power microwave electronics [1]. Significant progress has recently been made in growing high-quality GaN epitaxial layers. However, substantial efforts are required in the area of processing techniques. In particular, developments of processing technologies for formation of stable Schottky gate contacts with high Schottky barrier heights (SBHs) and low leakage currents as well as for formation of ohmic contacts with low contact resistances are crucial for realization of the GaN-based electronic devices such as metal semiconductor field effect transistors (MESFETs), high electron mobility transistors (HEMTs), etc. Research on Schottky [2–7] and ohmic contacts [8,9] of GaN are of current interest. However, their interface properties have not been well understood, nor have their formation processes been amply optimized yet.

In this paper, an attempt was made to characterize and optimize the formation processes of Schottky and ohmic contacts for *n*-type GaN in view of applications to field effect transistors (FETs). Both Schottky and ohmic contacts were prepared by standard vacuum deposition processes. The effects of pre-deposition surface

* Corresponding author. Tel.: +81-11-706-6873; fax: +81-11-716-6004.

E-mail address: hashi@rciqe.hokudai.ac.jp (T. Hashizume)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00093-3

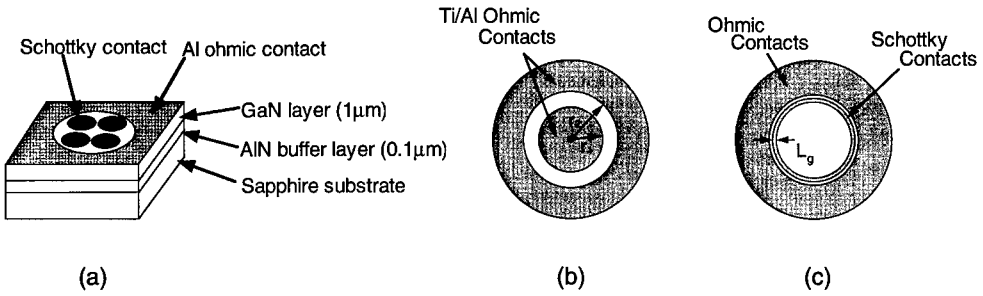


Fig. 1. (a) Sample structure for characterization of Schottky interfaces, (b) ohmic contact and (c) narrow ring Au Schottky diodes.

treatments and barrier metals on the Schottky characteristics were systematically investigated by current-voltage (I–V) and X-ray photoelectron spectroscopy (XPS) methods. As a result, Au contacts formed on NH_4OH treated surfaces showed nearly ideal thermionic-emission I–V characteristics with a highest Schottky barrier height of 1.03 eV. For ohmic contacts, Ti/Al contacts were investigated, using the circular transmission line model (TLM) method [10–12]. Excellent ohmic characteristics with a minimum contact resistance of $5–8 \times 10^{-5}\ \Omega\ cm^{-2}$ were obtained by annealing at 600°C for 1 min in N_2 atmosphere. Finally, narrow ring Au Schottky gates with Ti/Al ohmic electrodes were successfully fabricated, demonstrating applicability of the present contact formation processes for FET applications.

2. Experimental

Si-doped *n*-type GaN (0001) layers ($n = 5–7 \times 10^{16}\ cm^{-3}$) grown on sapphire substrates by MBE were used. The sample structures used in this study for characterization of Schottky contacts, ohmic contacts and their combination are shown in Fig. 1(a)–(c). Since GaN is grown on an insulator, the conventional geometry using back contacts could not be used, and all the electrode patterns were defined on the surface, using the standard photolithography with a positive photoresist (Shipley Co Ltd, Microposit S2400). For the study of Schottky contacts, large circular Schottky dots shown in Fig. 1(a), having a large diameter of 400 μm, was used together with a surrounding large ohmic contact made of an evaporated Al film. Evaporated Al films without annealing were reported to be a good ohmic contact material to *n*-type GaN [4], and our preparatory experiments also confirmed that Al electrodes indeed gave ohmic characteristics with contact resistance low enough at least for such a large dot geometry. For Schottky contacts, a series of high and low workfunction metals (Pt, Au and Ag) was deposited onto *n*-GaN surfaces by a conventional vacuum deposition process using an Electron-beam evaporation source or a resistance heater. For study of ohmic small electrodes for use in FET applications, the contact resistance obtainable by the above mentioned simple Al contact without annealing was found to be too high and inadequate. Thus Ti/Al alloy contacts [9] were evaluated using the circular geometry shown in Fig. 1(b), and applying the circular TLM method. In order further to confirm the applicability of the optimal processes for FET fabrication, narrow ring Au Schottky diodes with ohmic electrodes shown in Fig. 1(c) were fabricated and tested. Since mesa-etching of GaN with standard wet chemical agents is known to be difficult, such ring shaped structures seem to provide useful alternatives for FET feasibility study.

Due to the difficulty of wet chemical etching of GaN, special attention was paid in this study to the surface treatments prior to metal deposition. Three surface treatments, (1) cleaning in organic solvents only (hereafter referred to as the organic treatment); (2) dipping into an $HF:HCl:H_2O = 1:5:5$ solution for 1 min after cleaning in organic solvents (hereafter referred to as the HF/HCl treatment); and (3) dipping into a warm NH_4OH solution [13] hold at 50°C for 15 min after cleaning in organic solvents (hereafter referred to as NH_4OH treatment), were applied just prior to putting the sample into the evaporation chamber.

The electrical properties of the fabricated Schottky and ohmic contacts were investigated by current-voltage (I–V) characteristic measurements, using a HP 4156A parameter analyzer. Chemical status of GaN surfaces and metal/GaN interfaces were characterized by X-ray photoelectron spectroscopy (XPS) system (Perkin Elmer Co Ltd, ESCA5100) with a spherical capacitor analyzer (SCA) and a monochlomated Al k α radiation (h $\nu = 1486.6$ eV). The binding energies of the spectra were carefully calibrated by the separate measurements of $Au4f_{7/2}$ peak position.

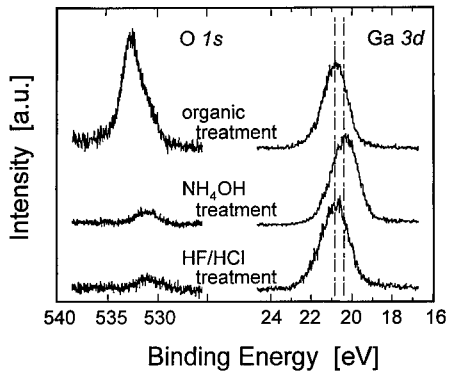


Fig. 2. O1 s and N1 s XPS spectra of the GaN surfaces after three kinds of surfaces treatment.

3. Results and discussion

3.1. XPS study of surfaces after pre-deposition surface treatments and interface after metal contact formation

Fig. 2 shows the Ga3 d, N1 s and O1 s XPS spectra obtained from the GaN surfaces after the three kinds of treatments. The values of XPS integrated intensity ratios of O1 s to N1 s and those of Ga3 d to N1 s were summarized in Fig. 3.

As seen in Fig. 3, the samples treated by the organic treatment exhibited non-stoichiometric (Ga-rich) surfaces including large amounts of oxide and/or oxynitride components. On the other hand, after both of the HF/HCl and NH_4OH treatments, significant decrease of the oxide components was observed as seen in Fig. 3. However, the detailed behavior was very different. Namely, after the NH_4OH treatment, a peak shift of Ga3 d spectrum toward a lower energy and the reduction of the Ga3 d/N1 s intensity ratio took place. This can be interpreted in terms of removal of native gallium oxide layer, making the surface oxide-free and near stoichiometric.

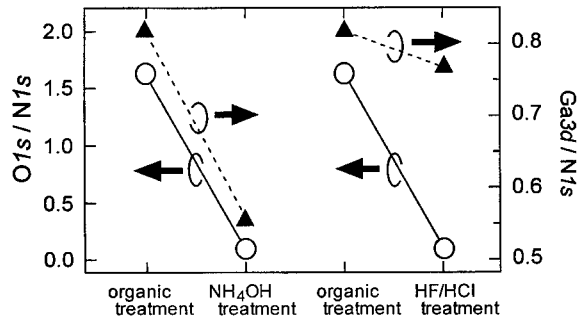


Fig. 3. Change in XPS integrated intensity ratios of O1 s to N1 s and Ga3 d to N1 s of the GaN after two kinds of surfaces treatment.

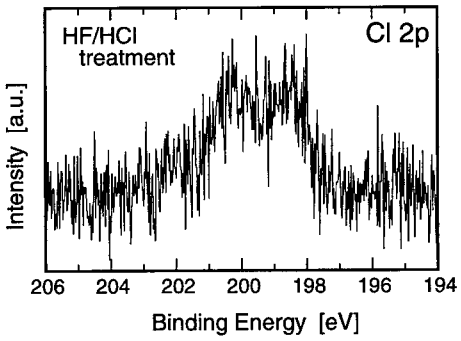


Fig. 4. Cl2p spectrum of the GaN surface treated by HF/HCl solution.

In contrast to this, neither of these two changes took place after the HF/HCl acid treatment. Furthermore, Cl-related XPS peaks appeared on the HF/HCl treated surface as shown in Fig. 4. This indicates that gallium chloride component was formed at the surface simultaneously with oxide removal, giving a Ga-rich surface.

XPS analysis was also made for the characterization of metal/GaN interfaces. For this measurement, the thickness of the evaporate metal film was kept to about 1 nm so that photoelectrons could escape from the interface. Fig. 5 summarizes the XPS integrated intensity ratios of Ga3 d to N1 s obtained from the Au/GaN and Ag/GaN interfaces formed after the HF/HCl and NH_4OH treatments. For comparison, the values of the ratio just after the surface treatments are also shown. For the diodes treated in the NH_4OH solution, the intensity ratio was almost the same before and after metal deposition. This means that the nearly stoichiometric surface was maintained after metal deposition in the case of the NH_4OH treatment. On the other hand, the intensity ratio was drastically changed after metal deposition for the HF/HCl treated surfaces. This strongly indicates that the Ga-rich surface after the HF/HCl treatment triggers interfacial reactions

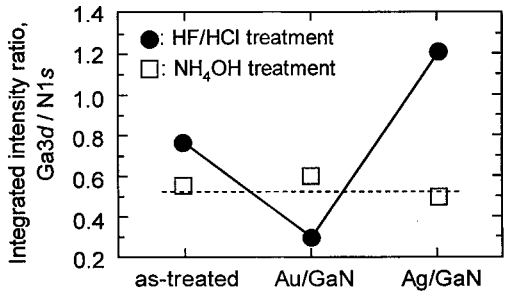


Fig. 5. XPS integrated intensity ratios of Ga3 d to N1 s obtained from the as-treated surfaces and the gold/GaN and silver/GaN interfaces.

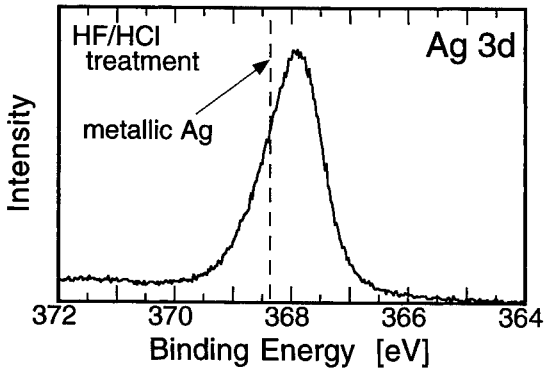


Fig. 6. Ag3d spectrum of the GaN surface treated by HF/HCl solution.

during metal deposition, leading to formation of interfacial layers. As a direct evidence of such an interfacial reaction, the Ag3 d spectrum of the Ag/GaN interface was shifted toward a position lower than that of the metallic Ag peak (Fig. 6).

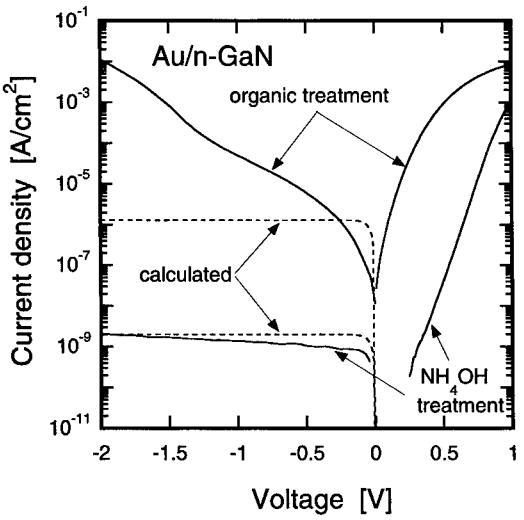


Fig. 8. I–V characteristics of the Au-Schottky contacts formed by the evaporation process after organic treatment and NH_4OH treatment.

3.2. Electrical properties of Schottky contacts

Figs. 7(a)–(c) show the measured forward I–V characteristics of the Ag, Au and Pt Schottky contacts formed by a vacuum deposition process after three different surface treatments. The values of the Schottky barrier height (SBH) and the ideality factor n were obtained by analyzing the measured I–V curves with the thermionic emission theory, and they are attached to the curves in Fig. 7(a) and (c), where meaningful values could be obtained. For the analysis, an effective electron mass of $m*_n$ of 0.22 m_0 was used, and this gave the value of the Richardson constant of 26.4 A cm^{-2} K^{-2} [14].

For the diodes treated only in organic solvents, high n-values and poor linearity of I–V curve were obtained. No clear dependence of SBH on metal work-function was observed. For the diodes formed after the HF/HCl treatment, almost all of the diodes exhibited ohmic-like characteristics, as shown by the linear scale I–V curves given in the inset of Fig. 7(b).

On the other hand, good thermionic emission characteristics with the value of n close to unity were realized on the diodes formed after the NH_4OH treatment. In particular, Au Schottky diode gave a highest

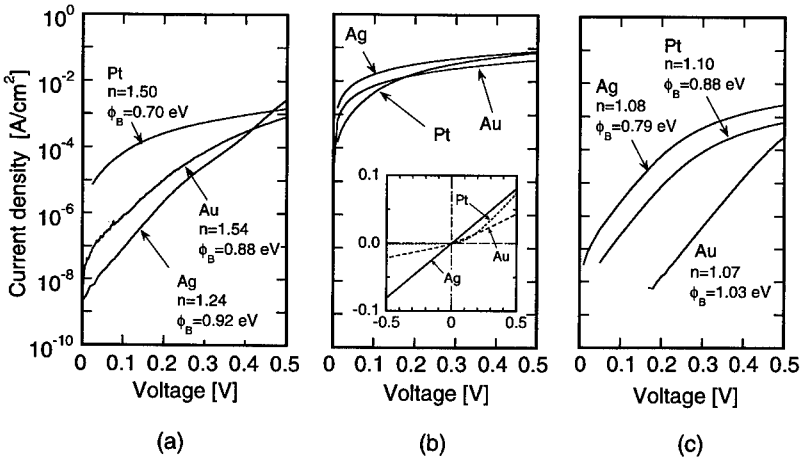


Fig. 7. Typical forward I–V characteristics of the Ag, Au and Pt Schottky contacts to n-GaN formed by a conventional vacuum deposition process: (a) organic treatment, (b) HF/HCl treatment and (c) NH_4OH treatment.

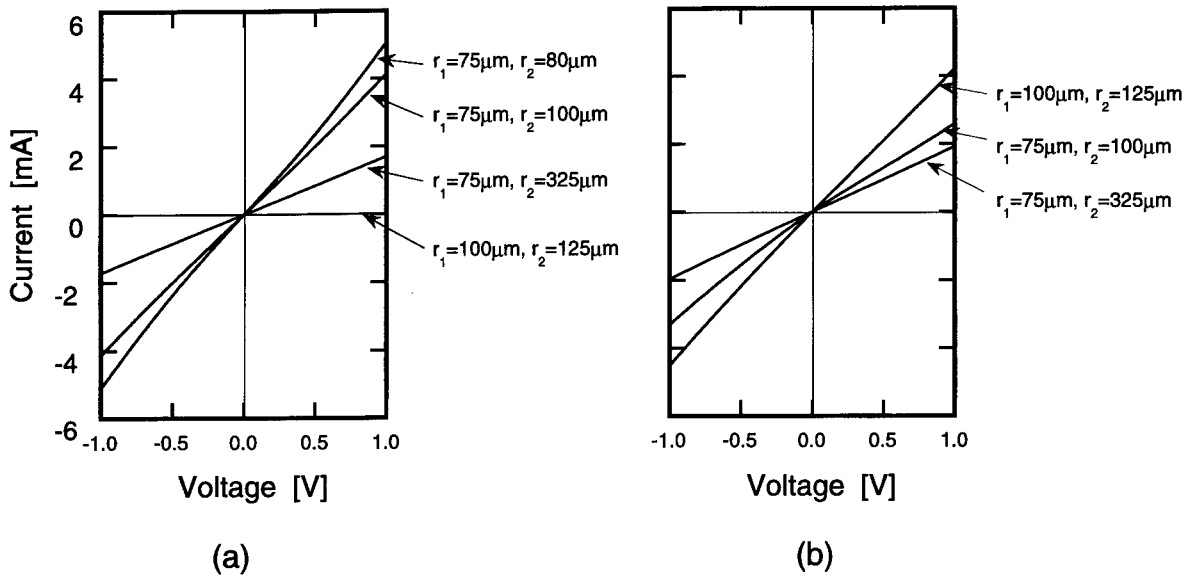


Fig. 9. I–V characteristics of (a) the as-deposited and (b) annealed Ti/Al contacts.

SBH of 1.03 eV with $n = 1.07$. No systematic dependence of SBH on the metal workfunction was found.

From the viewpoint of FET device application, reverse I–V characteristics are important, because they directly affect the FET gate controllability of the depletion layer width. Reverse I–V characteristics of the Au/*n*-GaN diodes formed after the organic and NH_4OH treatments are shown in Fig. 8, together with the calculated curves. In the calculation, the values of the SBH and the ideality factor were used, taking account of the image force barrier lowering effect. The diodes formed after the NH_4OH treatment showed a reverse I–V curve with a low leakage current which agree well with the calculated curve. On the other hand, anomalous reverse I–V behavior was observed for the samples formed after organic treatment. High reverse currents flow even at low reverse bias region, and the currents suddenly increased at a bias around −1.5 V. This appears to be the case that thin insulator-transition layer exists at the interface, and/or inhomogeneous metal/semiconductor contacts are formed [15].

Thus, it was found that there exists a strong correlation between the chemical and electronic properties of metal/GaN interfaces formed by vacuum deposition. Namely, the oxide-free and near stoichiometric surface obtained after the NH_4OH treatment remained stable during metal contact formation, and gave rise to excellent Schottky characteristics. On the other hand, oxide containing and non-stoichiometric surfaces obtained after the organic and HF/HCl treatments were unstable during contact formation, and exhibited poor Schottky characteristics, being dominated by the interfacial layers.

3.3. Electrical properties of ohmic contacts

For preparation of Ti/Al ohmic contacts, a 40 nm-thick Ti film and a 100 nm-thick Al film were deposited by a standard electron-beam deposition and vacuum deposition, respectively, after the NH_4OH treatment. Examples of the I–V characteristics of the as-deposited and annealed Ti/Al contacts are shown in Fig. 9(a) and (b). The behavior of the as-deposited contacts ranged from ohmic to rectifying. However, after high temperature annealing all contacts exhibited ohmic characteristics. The best result was obtained

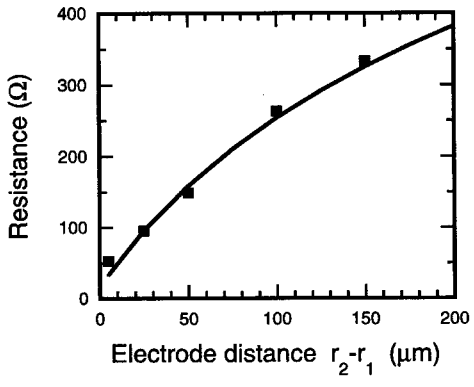


Fig. 10. An example of a curve fitting to experimental data obtained from the circular TLM method.

after annealing at 600°C for 1 min in N_2 atmosphere. In order to estimate the specific contact resistance, the measured data for various values of inner and outer radii, r_1 and r_2, of the circular electrode shown in Fig. 1(b) were fitted into the following equation [10].

$$R = \frac{R_{sh}}{2\pi}\left[\ln\frac{V_2}{r_1} + L_T\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\right] \quad (1)$$

where R_{sh} is the sheet resistance of GaN, and L_T is the transfer length. The transfer length L_T is related to the contact resistance r_c by

$$r_c = R_{sh}L_T^2 \quad (2)$$

An example of such a fitting vs the electrode distance (r_2-r_1) is shown in Fig. 10. The values of the contact resistance obtained after the best annealing condition mentioned above were in the range of $5–8 \times 10^{-5}$ Ω cm^{-2}. These values are somewhat larger than those reported recently by Luther et al. [9], however this may be due to the fact that the carrier concentration of GaN used in this study was 5 times lower than that in ref. [9].

Finally, in order to confirm applicability of the present Schottky and ohmic contact formation processes for FET applications, attempts were made to realize narrow ring Au Schottky gates with Ti/Al ohmic electrodes on the same chip, using the NH_4OH surface treatment and the standard photolithography process. For the fabricated Au Schottky gate with $L_g = 2$ μm and the spacing of 2 μm, good I–V behavior was also obtained. The reverse leakage current was very small and was of the order of 10^{-10} A at a reverse bias voltage of −10 V. Furthermore, electrical breakdown was not observed up to −40 V. Thus, the present contact formation processes seem to be promising for FET fabrication.

4. Conclusion

In this paper, an attempt was made to characterize and optimize the formation processes of Schottky and ohmic contacts on *n*-GaN by vacuum deposition in view of applications to field effect transistors (FETs).

1. From the viewpoints of oxide removal and maintenance of stoichiometry, it is concluded that etching in a warm NH_4OH (50°C) solution gives the best result.
2. Au, Ag and Pt Schottky contacts formed on NH_4OH treated surfaces showed nearly ideal thermionic-emission I–V characteristics with Au contacts giving a maximum Schottky barrier heights of 1.03 eV.
3. Schottky diodes treated only in organic solvents or those treated in an HF/HCl solution after organic cleaning showed poor I–V curve with high *n*-values and large leakage currents. Interface transition layers formed at the metal/GaN interfaces seem to be responsible for poor behavior.
4. For ohmic contacts, Ti/Au contacts after annealing at 600°C for 1 min in N_2 atmosphere, showed excellent ohmic characteristics with a minimum contact resistance of $5–8 \times 10^{-5}$ Ω cm^{-2}.
5. Narrow ring Au Schottky gates with Au/Ti ohmic electrodes were successfully fabricated, demonstrating applicability of the present contact formation processes for FET applications.

Acknowledgements

This work was partly supported by a grant-in-aid for Scientific Research on Priority Areas **A** (##10131201) and Scientific Research **(B)** (##09450118) from Ministry of Science, Education, Sports and Culture, Japan.

References

[1] Trew BJ, Shin MW, Gatto V. TWHM'96, Solid-State Electron 1997;41:1561.
[2] Hacke P, Detchprohm T, Hiramatsu K, Sawaki N. Appl Phys Lett 1993;63:2676.
[3] Binari SC, Dietrich HB, Kelner G, Rowland LB, Doverspike K, Gaskill DK. Electronics Lett 1994;30:909.
[4] Wang L, Nathan MI. Appl Phys Lett 1996;68:1267.
[5] Guo JD, Feng MS, Guo RJ, Pan FM, Chang CY. Appl Phys Lett 1995;67:2657.
[6] Suzue K, Mohammad SN, Fan ZF, Kim W. J Appl Phys 1996;80:4467.
[7] Schmitz AC, Ping AT, Khan MA, Adesida I. Mat Res Soc Symp Proc 1996;395:831.
[8] Fan Z, Mohammad SN, Kim W, Aktas Ö, Botchkarev AE, Morkoç H, Duxstad K, Haller EE. J Elec Mat 1996;25:1703.
[9] Luther BP, Mohney SE, Jackson TN, Kahn AA, Chen Q, Yang JW. Appl Phys Lett 1997;70:57.
[10] King DJ, Zhang L, Ramer JC, Hersee SD, Lester LF. Mat Res Soc Symp Proc 1997;468:421.
[11] Marlow GS, Das MS. Solid-State Electron 1982;25:91.
[12] Reeves GK. Solid-State Electronics 1980;23:487.
[13] Prabhakaran K, Andersson TG, Nozawa K. Appl Phys Lett 1996;69:3212.
[14] Strite S, Morkoc H. J Vac Sci Technol 1992;B10:12.
[15] Sullivan JP, Tung RT, Pinto MR, Graham WR. J Appl Phys 1991;70:7403.

CIRCUIT APPLICATIONS, DEVICE FABRICATION AND CHARACTERIZATION

PERGAMON

Solid-State Electronics 43 (1999) 1489–1495

SOLID-STATE ELECTRONICS

Transferred-substrate HBT integrated circuits

Mark Rodwell[a,*], Q. Lee[a], D. Mensa[a], J. Guthrie[a], S.C. Martin[b], R.P.Smith[b],R.Pullela[a,1],B.Agarwal[a,2],S.Jaganathan[a],T.Mathew[a],S.Long[a]S.Long[a] S. Long[a]

[a]*Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA*
[b]*Jet Propulsion Labs., Caltech, Pasadena, CA, USA*

Abstract

Using substrate transfer processes, we have fabricated heterojunction bipolar transistors with submicron emitter-base and collector-base junctions, minimizing *RC* parasitics and increasing f_{max} to 500 GHz. The process also provides a microstrip wiring environment on a low-ϵ_r dielectric substrate. Demonstrated small-scale ICs in the process include lumped and distribute amplifiers with bandwidths to 85 Hz, 48 GHz static frequency dividers, and 50GHz AGC/limiting amplifiers. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Heterojunction bipolar transistor; HBT; Transferred-substrate

1. Introduction

Applications—present and potential—for heterojunction bipolar transistors (HBTs) include RF/microwave analog-digital conversion, microwave direct digital frequency synthesis, and fiber-optic transmission in the range of 40–120 Gigabits/second.

The motivation for high circuit bandwidth varies with the application. In fiber-optic transmission, fast ICs will directly enable increased time-division-multiplexed transmission rates, complementing increases in channel capacity provided by wavelength-division-multiplexing. For oversampled ($\Delta-\Sigma$) analog-digital converters, in-band quantization noise power decreases in proportion to the 5th-power of the oversampling ratio (2nd-order modulators). Very high speed IC technologies offer the potential of $\Delta-\Sigma$ ADCs with clock frequencies in the 10 s of GHz, providing high dynamic range and large instantaneous bandwidth over radio-frequencies and lower microwave frequencies. In direct digital frequency synthesis (DDS), increases in logic IC clock rates and DAC bandwidths will result in increased synthesizer bandwidth. Used in radar transmitters, modulation bandwidths can then be increased and frequency agility improved.

A 100 GHz clock rate IC technology would permit significant advances in the performance of such signal conversion ICs. To permit clock rates exceeding 100 GHz, significant issues in transistor design, interconnects, packaging, and thermal design must be addressed. The transistor current gain (f_τ) and power gain (f_{max}) cutoff frequencies must be several hundred GHz. Wiring parasitics must be minimized. The interconnects must have small inductance and capacitance per unit length, and the wire lengths, hence transistor spacings, must be small. Given that fast HBTs operate

* Corresponding author.
[1] Present address: Lucent Technologies, Murray Hill, NJ, USA
[2] Present address: Rockwell Semiconductor Systems, Newport Beach, CA, USA

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00094-5

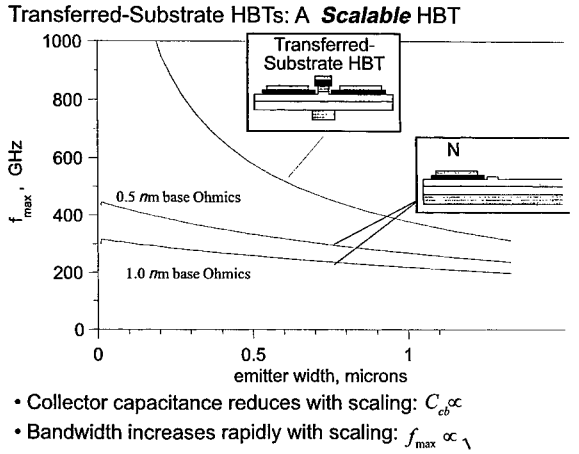


Fig. 1. Scaling of transferred-substrate and triple-mesa HBTs.

at $\sim 10^5$ A/cm^2 current density, efficient heat sinking is then essential. To provide predictable performance, interconnects of more than a few ps length must have a controlled characteristic impedance. To prevent circuit–circuit interaction through ground-circuit common-lead inductance ("ground loops"), the IC technology must provide an integral ground plane for ground-return connection. Similarly, to prevent circuit–circuit interaction between the IC's input and output lines, common-lead inductance between the IC and package ground systems must be made vanishingly small.

2. Transferred-substrate HBTs

Transferred-substrate HBT ICs [3] offer high transistor bandwidth, low wiring capacitance, and low wiring inductance. In this process a substrate transfer step permits two-side processing of the device epitaxial layers, allowing definition of a transistor with narrow and aligned emitter-base and collector-base junctions.

In HBTs, f_τ is primarily determined by the base transit time τ_b, the collector transit time τ_c, and the emitter charging time $C_{je}(kT/qI_e)$. Increases in f_τ are obtained by thinning the collector, by thinning and grading the base, and by increasing the emitter current density. Unfortunately, thinning the base and collector epitaxial layers increases the collector capacitance C_{cb} and the base resistance R_{bb}, decreasing $f_{max} = (f_\tau/8\pi R_{bb}C_{cbi})^{1/2}$. The base-collector junction is a distributed network and $R_{bb}C_{cbi}$ represents an effective, weighted time constant [4]. *For a fixed emitter stripe length*, decreasing the width W_e of the emitter-base junction decreases the base spreading resistance, but does not decrease the base contact resistance. Scaling W_e reduces R_{bb} towards a minimum set by the contacts.

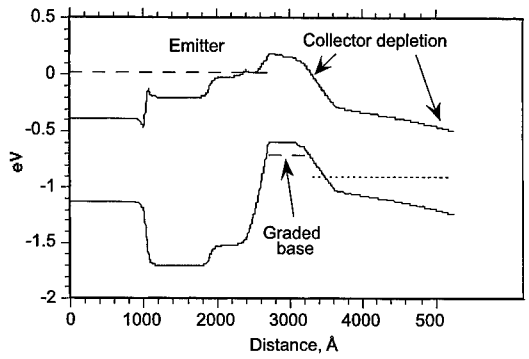


Fig. 2. Band diagram, under bias, of a typical device.

Decreasing the width of the base-collector junction W_c decreases C_{cb}.

In normal double-mesa HBTs (Fig. 1), the collector-base junction and base Ohmic contact are defined in a single process step. The Ohmic contacts must be at least one contact transfer length, setting a minimum collector junction width, and a minimum collector capacitance. $R_{bb}C_{cb}$ has a minimum value, independent of lithographic limits, and f_{max} does not improve with scaling. In the transferred-substrate device, the collector width need only be larger than the emitter width. Reducing W_c and W_e progressively reduces $R_{bb}C_{cb}$, and f_{max} increases rapidly with scaling (Fig. 1). With lateral scaling alone, f_{max} should approach 1 THz as dimensions are scaled to $\sim$0.1 μm. Subsequently thinning the base and collector layers increases f_τ at the expense of f_{max}. Simultaneous high values for both f_τ and f_{max} are thus obtained.

3. Growth and fabrication

Fig. 2 shows the band diagram associated with a typical transferred-substrate HBT layer structure. The InGaAs base is nominally 400–500 Å thick, has $kT - 2kT$ bandgap grading, and is Be-doped at $5{\cdot}10^{19}$ cm^3. The InGaAs collector is 2500 Å thickness. A collector N^+ pulse-doped layer placed 400 Å from the base delays the onset of base pushout at high collector current densities. Devices typically use Schottky collector contacts [6], although HBTs with N+ subcollector layers (Ohmic-collector devices) have also been fabricated. While Ohmic-collector devices have non-zero collector series resistance, hence lower f_{max} [4], the 0.2 V barrier present in the Schottky-collector device increases the V_{ce} required to suppress base pushout at high current densities. Ohmic-collector devices thus show higher f_{max} under the low-V_{ce} conditions associated with current-mode-logic (CML). Schottky-collector devices are used for emitter-coupled-logic (ECL), where the operating V_{ce} is higher.

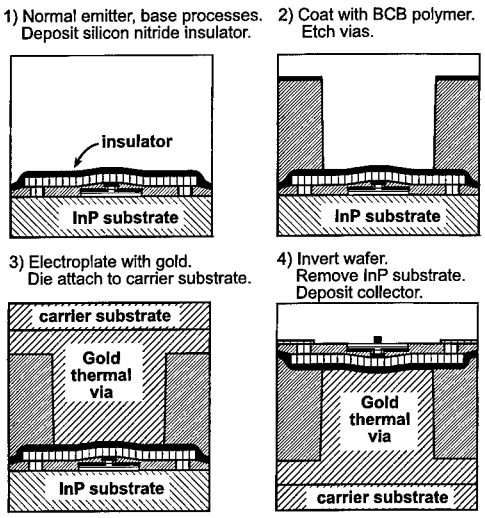


Fig. 3. Transferred-substrate HBT process flow.

Fig. 3 shows the process flow. Standard fabrication processes [2] define the emitter-base junction, the base mesa, polyimide planarization, and the emitter contacts. The substrate transfer process commences with deposition of the PECVD Si_3N_4 insulator layer and the Benzocyclobutene (BCB) transmission-line dielectric (5 μm thickness). Thermal and electrical vias are etched in the BCB. The wafer is electroplated to metalize the vias and to form the ground plane. Subsequently, the wafer is Indium-bonded to a GaAs carrier substrate, and the InP substrate removed in HCl. Schottky collectors are then deposited, completing the process.

The transferred-substrate process permits high device bandwidths f_{max} through simultaneous scaling of emitter and collector junctions. For the emitter-base junction, deep submicron scaling requires tight control of lateral undercutting during the base contact recess etch. To form the emitter, reactive-ion etching in $CH_4/H_2/Ar$, monitored with a HeNe laser, first removes the N^+ GaInAs emitter contact layer. A HCl/HBr/Acetic selective wet etch then removes the AlInAs emitter, stopping on the AlInAs/GaInAs emitter-base grade. By etching at 10°C, the etch rate is slowed, and a controlled emitter undercut is formed. The undercut both narrows the emitter and serves (as normal) to define the liftoff edge in the self-aligned base contact deposition. A timed 30 s nonselective wet Citric/H_3PO_4/H_2O_2 etch then removes the base-emitter grade. Etch selectivity in both the RIE and HCl/HBr/Acetic etches aids in etch-depth control, and we are able to reproducibly etch ~100 Å into the base without use of surface contact resistance probing as a process monitor. Fig. 4 shows the cross-section of a 0.15 μm emitter-base junction.

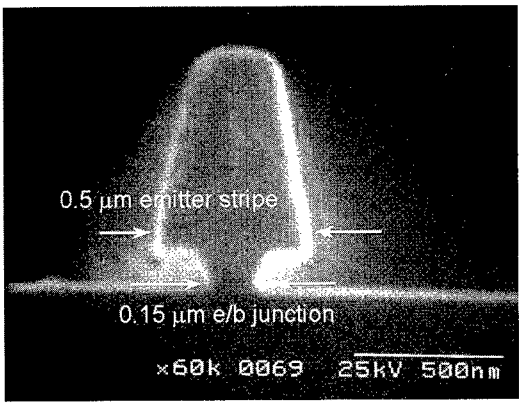


Fig. 4. Cross-section of emitter-base junction. The 0.5 μm emitter metal was defined with a projection lithography system.

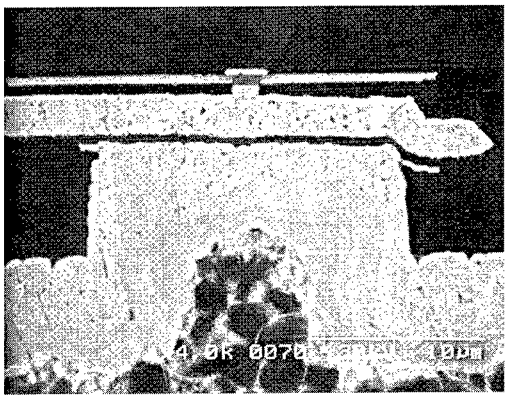


Fig. 5. Cross-section of a transferred-substrate HBT showing the thermal via.

In defining submicron collector-base junctions, use of the Schottky-collector contact eliminates the need for an etch of similar precision through an N^+ collector Ohmic contact layer. The collector junction is defined by the stripe width of the deposited metal. Subsequent to collector deposition, a self-aligned wet etch of ~1000 Å depth removes the collector junction sidewalls (eliminating fringing fields) and reduces the collector junction width by ~2000 Å. The step, intended to reduce C_{cb}, generally provides a greater increase f_{max} than would be expected from the observed reduction in collector junction width.

In addition to HBTs with narrow emitter and collector stripes Fig. 5, the process provides thermal vias for HBT heatsinking, NiCr resistors, Si_3N_4 MIM capacitors, and microstrip wiring on a $\epsilon_r = 2.7$ dielectric with vias, ground plane, and 3 levels of interconnects. To tolerate high power densities, the NiCr resistors must have thermal vias, which results in significant parasitic capacitance. Pull-up resistors in ECL do not require the thermal via.

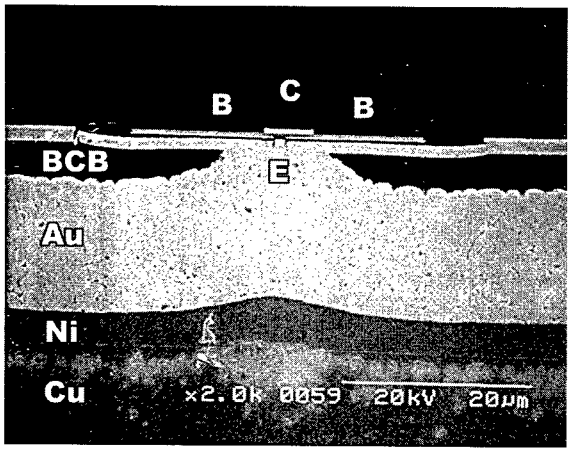


Fig. 6. Cross-section of a transferred-substrate HBT with a electroplated copper substrate.

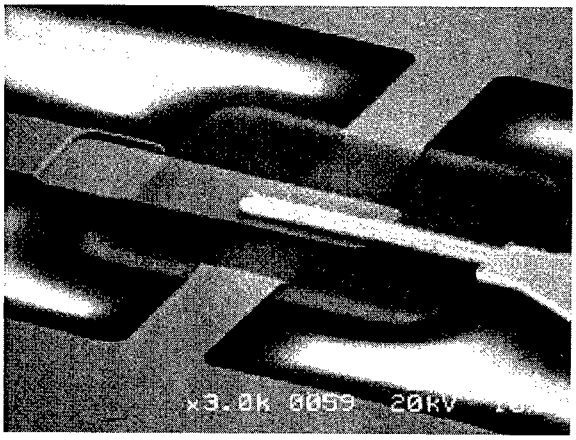


Fig. 7. Transferred-substrate HBT defined by contact lithography.

A significant missing feature is in packaging. In the present process the microstrip ground plane is isolated from the wafer back surface by the GaAs transfer substrate, increasing thermal resistance and preventing low-inductance connections between the IC and package ground. A modified process with a fully metallic electroplated copper substrate is in development [5]. This will provide highly effective heatsinking (395 W/M·K for Cu vs 74 for InP) and, with a package-IC ground interface over the full IC back surface, very low package-chip ground inductance. Presently, devices have been fabricated in the copper-substrate process (Fig. 6) but difficulties with residual stress limit IC yield.

4. Device performance

Devices have been fabricated using contact lithography at 1–2 μm resolution, using a 0.5 μm stepper, and using electron-beam lithography. Fig. 7 shows a device with a 0.6 by 8 μm emitter and a 1.6 by 12 μm collector. The device exhibits 215 GHz peak f_τ, and peak f_{max} above 400 GHz. Fig. 8 shows HBT emitter-base and collector-base junctions defined by electron-beam lithography. Submicron devices fabricated by E-beam lithography (Fig. 9) exhibit 500 GHz. f_{max} [7]. Neither contact lithography nor electron-beam lithography is suitable for fabrication of large ICs. We have recently fabricated HBT ICs using a 0.5 μm projection lithogra-

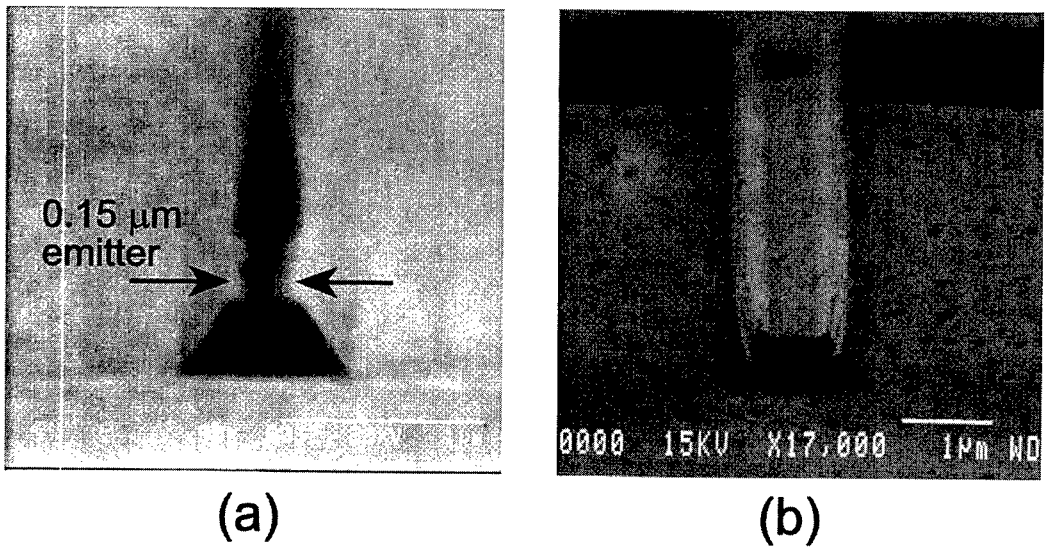


Fig. 8. E-beam HBT: test structure with 0.15 μm emitter-base junction (a), and 0.4 μm Schottky collector stripe (b).

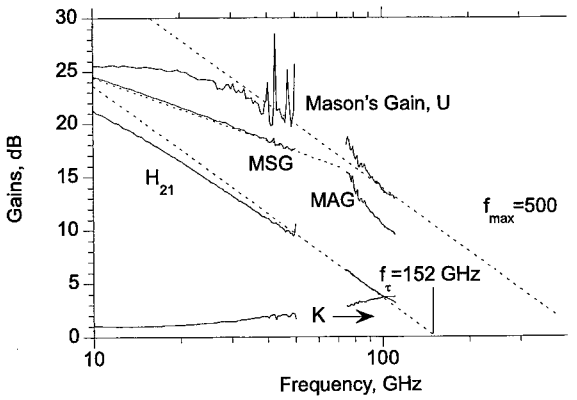


Fig. 9. W-band gains of device with a 0.4 × 25 μm emitter and a 1.0 × 29 μm collector.

phy system, and have obtained microwave gains (and f_{max}) somewhat higher than that shown in Fig. 9.

With the exception of reactively-tuned circuits, for which f_{max} is the sole determinant of circuit bandwidth, circuit design generally requires high values for both f_τ and f_{max}. Examining significant terms in $1/2\pi f_\tau$, with base bandgap grading, $\tau_b + \tau_c$ is small at 0.4–0.5 ps, while the Schottky collector eliminates $R_c C_{cb}$. At peak f_τ bias, $(C_{je} + C_{cb})/g_m \simeq 0.1$ ps. Presently $R_{ex} C_{cb} \simeq 0.1$ ps, and has significant impact upon f_τ. To obtain >300 GHz f_τ, base bandgap grading must be increased, the collector thinned, and InAs emitter Ohmic contact layers employed.

Device scaling also reduces D.C. current gain and breakdown voltage. Base current in narrow-emitter InAlAs/InGaAs HBTs is predominantly due to conduction on the exposed InGaAs base surface between the emitter mesa and the base Ohmic contact. β decreases with emitter width, but increases as the base is thinned, as base bandgap grading is increased, and (at the expense of f_{max}) as the emitter-base spacing is increased. $\beta > 50$ has been obtained with 0.2 μm emitters. The 2500 Å InGaAs collectors have very low breakdown, 1.5 V BV_{ceo}, 2 V BV_{cbo} *at* 1–2 $\cdot 10^5$ A/cm^2 *current density*. Clearly, InP collectors should be used for highly scaled devices.

5. Integrated circuits

As first digital ICs demonstrations, we fabricated ECL and CML master-slave D flip-flops, configured as 2:1 static frequency dividers [8]. Circuits were fabricated using contact lithography, producing devices with 0.6 × 8 μm emitters and 1.6 × 12 μm collectors. The devices operate at 1.25 mA/μm^2. The differential logic swing is 600 mV. The collector pull-up resistors are 50 Ω, hence the divider outputs directly drive 50 Ω output lines without buffering. Circuit design is entirely

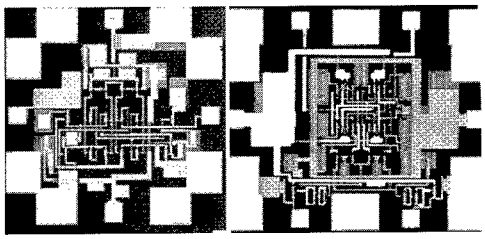


Fig. 10. CML (a) and ECL (b) master-slave D-flip-flops.

standard. The CML divider uses series-gated master and slave latches. Emitter-follower buffers are added to the CML clock and data ports to form the ECL divider. The ICs are shown in Fig. 10.

The CML and ECL static frequency dividers operated at maximum clock frequencies of 47 and 48 GHz. The circuits dissipate 380 mW (ECL) or 75 mW (CML) from a -5 V supply, and have pad-limited 0.6 × 0.6 mm die areas. The peak ECL clock speed of 70 GHz predicted from SPICE simulation (in which efforts were made to model all significant device and interconnect parasitics) does not correlate with the measured 48 GHz. We suspect this discrepancy arises from strong emitter-follower gain peaking at 50 GHz. We have observed very strong ~50 GHz gain peaking inconsistent with SPICE simulation in resistive feedback amplifiers [9,10], and have suppressed it with shunt resistive loading at the emitter-follower output [11]. Design studies now in progress suggest the feasibility of 100 GHz clock rates with the present HBTs, but the device models—and the circuits—remain to be validated.

Various analog ICs have been demonstrated in the process, including 85 GHz distributed amplifiers (Fig. 11), 50 GHz broadband differential amplifiers for optical fiber receivers (Fig. 12), broadband Darlington feedback amplifiers (Fig. 13), and very broadband f_τ-doubler feedback amplifiers. (Fig. 14).

6. Conclusions

100 GHz digital ICs require a high-bandwidth transistor, low-parasitic interconnects, and effective heat-

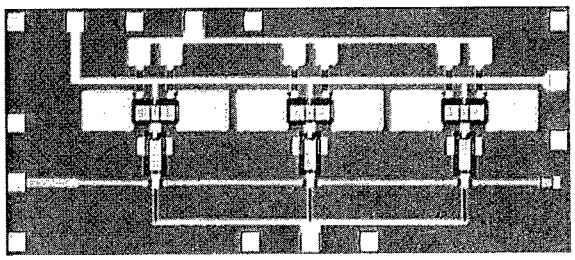


Fig. 11. Distributed amplifier in the transferred-substrate process. The amplifier exhibits 7 dB gain and 85 GHz bandwidth.

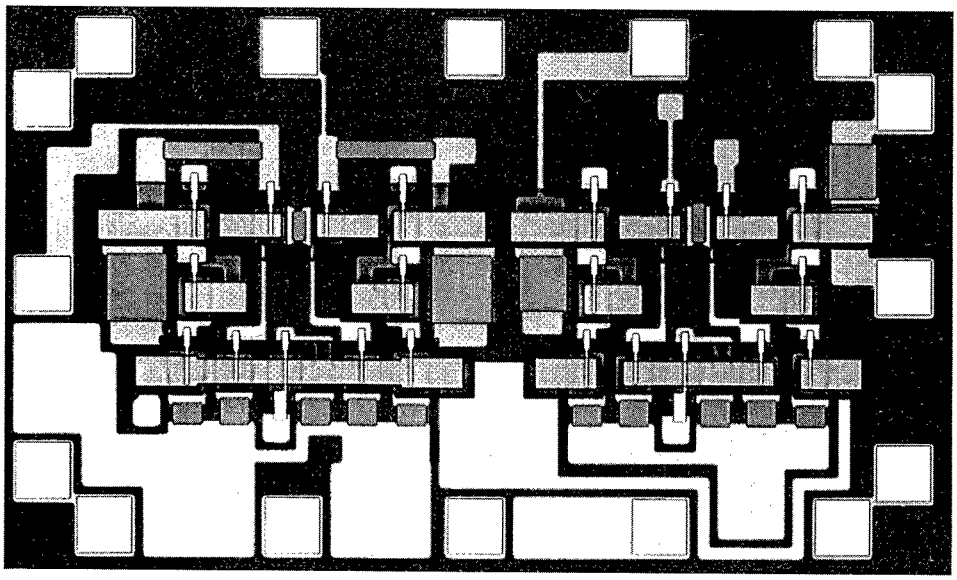


Fig. 12. D.C.-50 GHz, 11 dB broadband differential amplifier.

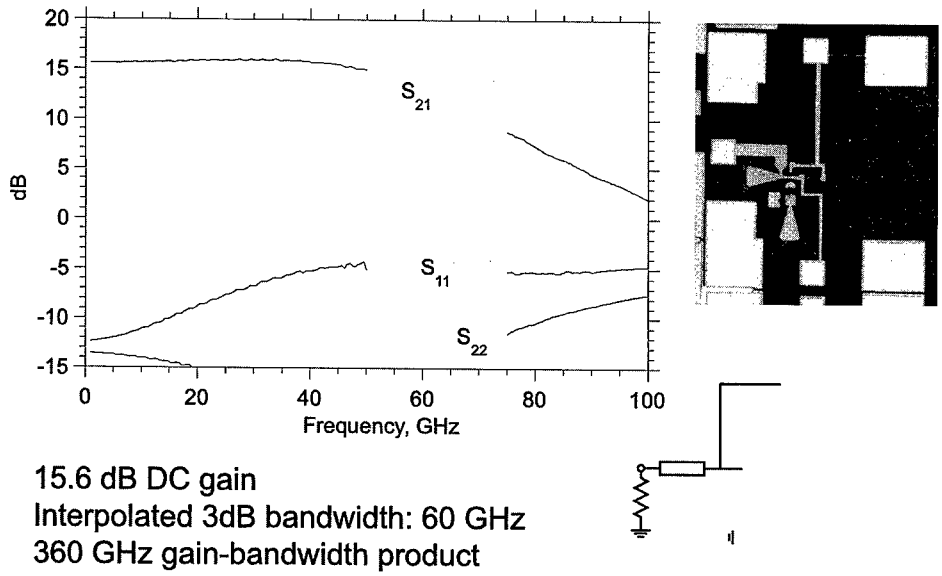


Fig. 13. Broadband Darlington feedback amplifier.

sinking. Using substrate transfer processes, HBTs can be fabricated with highly scaled lithographic and epitaxial dimensions, giving both high f_τ and high f_{max}. With further scaling and improved circuit design, 100 GHz digital ICs will be feasible.

Acknowledgements

Work at UCSB supported by the ONR under grants N00014-95-1-0688 and N00014-98-1-0068 (J. Zolper, M. Yoder), and by the AFOSR under grant F4962096-

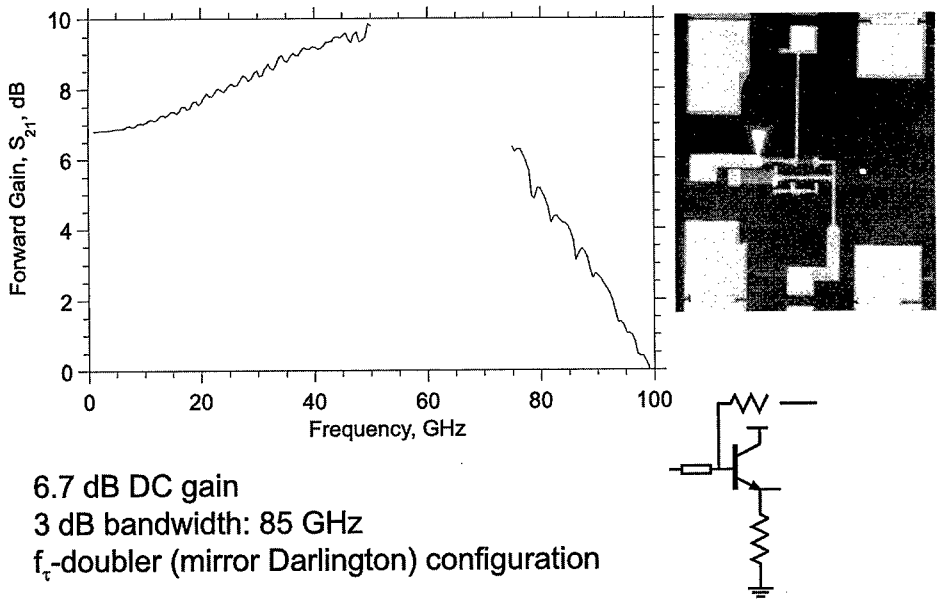


Fig. 14. Broadband feedback amplifier. The amplifier consists of a mirror f_τ doubler and emitter-follower input buffer.

1-0019 (H. Schlossberg). JPL work was performed at the Center for Space Microelectronics Technology, JPL, Caltech, and sponsored by the NASA office of Space Science. The author would like to acknowledge the use of references [1] and [12] in this study

References

[1] Matsuoka Y, Yamahata S, Kurishima K, Ito H. Ultrahigh-speed InP/InGaAs double-heterostructure bipolar transistors and analyses of their operation. J of Appl Phys 1996;35:5646–54.

[2] Stanchina WE, Jensen JF, Walden RH, Hafizi M, Sun H-C, Liu T, Raghavan G, Elliot KE, Kardos M, Schmitz AE, Brown YK, Montes ME, Yung M. An InP-based HBT fab for high-speed digital, analog, mixed-signal, and optoelectronic ICs. In: GaAs IC Symp. Tech. Dig, 1995. p. 31–4.

[3] Lee Q, Agarwal B, Mensa D, Pullela R, Guthrie J, Samoska L, Rodwell MJW. A >400 GHz f_{max} transferred-substrate heterojunction bipolar transistor IC technology. IEEE Electron Dev Lett 1998;19:77–9.

[4] Vaidyanathan M, Pulfrey DL, Extrapolated f_{max} of heterojunction bipolar transistors, IEEE Transactions on Electron Devices (submitted).

[5] Guthrie J, Mensa D, Agarwal B, Lee Q, Pullela R, Rodwell M. HBT IC process with a copper substrate. IEE Electronics Letters 1998;34(5):467–8.

[6] Smith RP, Allen ST, Reddy M, Martin SC, Liu J, Muller RE, Rodwell MJW. 0.1 μm Schottky-Collector AlAs/GaAs resonant tunneling diodes. IEEE Electron Device Letters 1994;15(8).

[7] Lee Q, Martin SC, Mensa D, Pullela R, Smith RP, Agarwal B, Guthrie J, Rodwell M. Deep submicron transferred-substrate heterojunction bipolar transistors. In: 1998 Device Research Conference, June Charlottesville, VA, 1998.

[8] Pullela R, Mensa D, Agarwal B, Guthrie J, Rodwell M. 47 GHz static frequency divider in ultrafast transferred-substrate heterojunction bipolar transistor technology. In: 1998 Conference on InP and Related Materials, May, Tsukuba, Japan, 1998.

[9] Agarwal B, Mensa D, Lee Q, Pullela R, Guthrie J, Samoska L, Rodwell MJW. A 13 dB, 50 GHz feedback amplifier with AlInAs/GaInAs transferred-substrate HBT. In: 1997 IEEE International Electron Device Meeting, December, Washington, DC, 1997.

[10] Agarwal B, Lee Q, Pullela R, Mensa D, Guthrie J, Rodwell MJW, A transferred-substrate HBT wideband differential amplifier to 50 GHz, IEEE Microwave and Guided Wave Letters, June 1998.

[11] Mensa D, et al. Baseband amplifiers in the transferred-substrate HBT technology. In: 1998 GaAs IC symposium. Nov. 1–4, Atlanta, GA, 1998.

[12] Agarwal B, Pullela R, Lee Q, Mensa D, Guthrie J, Rodwell MJW. 80 GHz distributed amplifiers with transferred-substrate heterojunction bipolar transistors. In: 1998 IEEE MTT Microwave Symposium, June, Baltimore MD, 1998.

PERGAMON

Solid-State Electronics 43 (1999) 1497–1506

SOLID-STATE ELECTRONICS

SiGe heterostructure CMOS circuits and applications

E.H.C. Parker*, T.E. Whall

Department of Physics, University of Warwick, Coventry, CV4 7AL, UK

Abstract

A review is given of the 300 K electron and hole mobilities in Si/SeGe heterostructures and the potential applications of these materials in CMOS technology. Prospects for further enhancements in carrier mobility and CMOS process design options are discussed for Si/SiGe strained layers on Si and on relaxed SiGe 'virtual substrates'. Recent work on heterointerface quality, limited area growth of virtual substrates, carrier mobility and velocity-field characteristics is also reported. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction: the limits of CMOS and the role for SiGe

Silicon CMOS has a unique range of features which makes it widely applicable, facilitating the design of both analogue and digital circuits. Low power consumption, high input impedance, excellent noise immunity, high integration levels, and proven reliability are amongst its attributes. With each new generation there are improvements in speed, current drive and noise performance along with reductions in supply voltage. As the technology approaches 0.1 μm feature size further gains in performance are to be had, with promises of digital and analogue operation in the GHz regime, ultimately replacing the bipolar circuits currently used in high speed, high gain, low noise applications. Even so, such deep submicron technologies set very ambitious process targets.

The technology is not without its problems. With shrinking geometries carriers are increasingly squeezed into a narrow region adjacent to the Si/SiO_2 interface and their mobilities become miserably low—in the range 50 to 200 $cm^2\ V^{-1}\ s^{-1}$, the holes being especially sluggish—at least partly due to their large effective masses. Such low mobilities reduce device speed and levels of integration, and make for increased power consumption. With huge new markets being created in mobile communications and other wireless systems operating at frequencies in the range 900 MHz–2.2 GHz, demanding highly linear, low noise, high speed devices with low power supply voltages, Si-CMOS is starting to show its age. Even with clever circuit design it cannot by itself meet the specification levels required for the market-leading products, and III–V or bipolar technologies have to be latched on to or integrated with CMOS—solutions which lack elegance and are costly. Many of the problem areas in the quest for an MOS one-chip solution to cellular phones are being researched with some success, but the design bottleneck, preventing further integration, is the RF section, hampered by inadequate transistor performance.

However, with SiGe there are excellent prospects of rejuvenating the technology. If SiGe can be utilized to create the active channels in Si-CMOS it brings to it very significant increases in mobilities—extending the reach of CMOS comfortably into RF circuitry. For example, even at this very early stage a crude SiGe p-channel FET has shown a high frequency performance superior to any other p-channel transistor [1]. Carrier mobility is a key performance parameter and improvements give, crucially, much better linearity, higher cur-

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00095-7

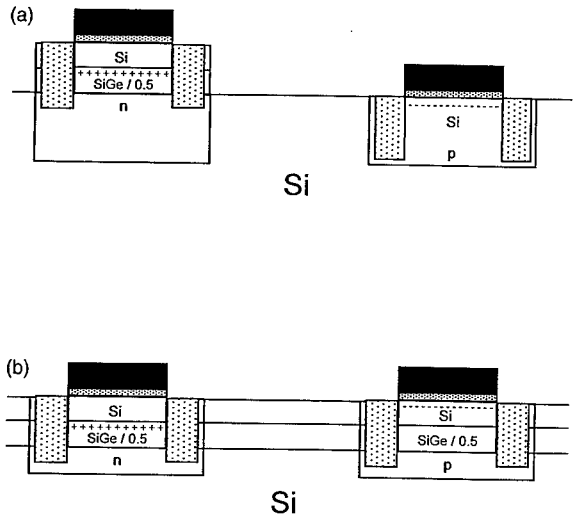


Fig. 1. Schematics of two possible process options incorporating pseudomorphic SiGe layers for enhanced p-channel CMOS involving: (a) processing the n-channel device first, requiring relatively high thermal budgets prior to selective growth of SiGe for the p-channel device; (b) processing from a blanket coverage epilayer, with restricted thermal budget to maintain the strain integrity of the SiGe.

rent drive, better noise performance and reductions in the supply voltage. Such developments would impact on both digital and analogue circuitry and in the case of the LNAs, local oscillators, analogue-to-digital converters and power amplifiers used in mobile communication, systems there would be particular benefit from using SiGe transistors. Moreover, the possibility of a matched n- and p-channel performance in CMOS considerably facilitates the design of amplifiers, mixers and filters. The enhanced carrier mobilities seen particularly at low vertical fields auger well for high speed low voltage micropower circuits involving MOSFETS biased in the weak inversion regime. This would enable micropower circuits to assess the technologically important 1–10 MHz baseband region and be used for video signal processing. Furthermore, the huge mobile communications and LAN market and the seemingly endless applications requiring high speed logic provide enormous leverage for such a mobility-enhanced CMOS technology.

The increased mobilities in Si/SiGe heterostructures are largely due to strain-induced reductions in effective mass and concomitant conduction and valence subband splittings. High germanium content in the case of holes produces further significant reductions in effective mass. These improved band structures can lead to much reduced phonon, interface charge, interface roughness and alloy scattering [1].

In this review we discuss Si/SiGe heterostructure MOSFET devices, with particular emphasis on p-channels where, because of the low hole mobility in bulk Si pMOS, the performance gains are expected to be largest. We look at potential CMOS configurations involving SiGe and consider the critical issues which need to be addressed if SiGe is to be incorporated into the CMOS process. The paper includes work carried out by The University of Warwick on effective mobility, velocity-field characteristics, heterointerface quality and the development of strain-tuning relaxed SiGe alloys ('virtual substrates') of low defect density.

2. Enhanced p-channel HCMOS

Mobility improvements in the p-channel device in CMOS, with the ultimate aim of producing symmetric n- and p-channel characteristics, will yield very significant performance (power-delay product) enhancements, (up to around 100%), in inverter-based circuitry, and enable considerably improved integration levels. Such a scenario can be envisaged through the addition of a thin defect-free (pseudomorphic) strained Si_{1-x} Ge_x layer in the p-channel device in CMOS and two process design options are shown in Fig. 1. The SiGe lowers the channel resistance and may also allow significant reductions in the spreading resistance intrinsic to B-implanted p-channel contacts. To date, thermal oxides formed directly on the SiGe are not of device quality and the gate oxide is provided by oxidis-

Table 1
Calculated 300 K mobilities in the $Si/Si_{0.6}Ge_{0.4}/Si$ two-dimensional hole gas, assuming effective mass $m^* = 0.27\ m_0$, acoustic phonon scattering potential $E_1 = 5.8$ eV, alloy thickness = 5 nm, carrier density = $1–2 \times 10^{12}$ cm^{-2}. Other parameters are taken from Ref. [3]. Lower row of figures indicates the mobility attainable if interface scattering can be reduced to a negligible contribution

Component mobilities				Total mobility
Interface roughness scattering $cm^2\ V^{-1}\ s^{-1}$	Alloy scattering $cm^2\ V^{-1}\ s^{-1}$	Acoustic phonon scattering $cm^2\ V^{-1}\ s^{-1}$	Non-polar optical phonon scattering	$\mu = 1/\Sigma_i^{[FSI]}\ 1/\mu_i$ $cm^2\ V^{-1}\ s^{-1}$
951	2428	4277	2835	486
00	2428	4277	2835	1000

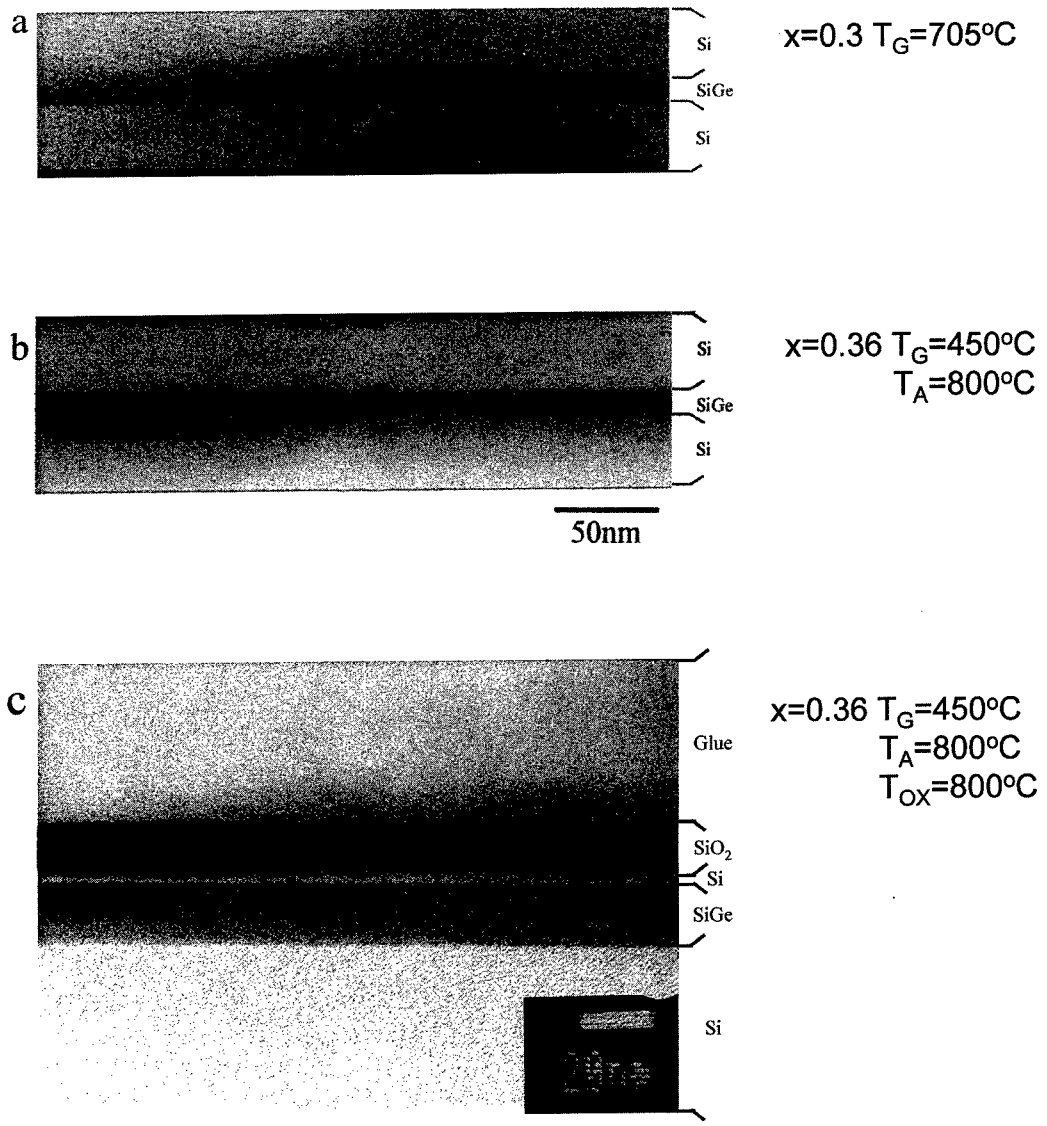


Fig. 2. XTEM micrographs of unoxidised (a) and (b) and oxidised (c) Si/SiGe/Si structures grown by solid source-MBE: (a) Si/$Si_{0.7}Ge_{0.3}$/Si grown at 705°C (T_G) showing strain-induced long range (~100 nm) roughness on the top SiGe/Si interface; (b) Si/$Si_{0.64}Ge_{0.36}$/Si grown at low temperature (450°C) and annealed in situ (800°C (T_A) for 30 min) showing no discernible roughness at the top interface; (c) thermal oxidation (800°C for 120 min) of (b) to leave a 3 nm thick Si cap layer.

ing an overlaying Si layer and leaving a very thin Si cap layer between SiGe and the Si/SiO_2 interface. The cap layer however, provides a conduit for parasitic hole conduction at high gate voltages which detracts from the enhanced SiGe p-channel performance. An n^+ poly Si gate with a B-doping layer below the channel to adjust the threshold voltage is one possible solution to the problem of parallel conduction [2].

Considering 0.1 μm technology the value of x in the active p-channel $Si_{1-x}Ge_x$-layer must yield a low field 300 K mobility in the range of 300–500 $cm^2\ V^{-1}\ s^{-1}$ at an effective (vertical) field of 1 MV cm^{-1} to be of consequence, and the SiGe layer must be able to withstand CMOS process thermal budgets. We anticipate such mobility enhancements for $x = 0.4$ as material quality improves. Our optimism is based on the results of calculations shown in Table 1. The interface roughness, alloy and acoustic phonon scattering mobilities have been obtained for a strained, $x = 0.4$ alloy of thickness 5 nm corresponding to the Matthews–Blakeslee equilibrium critical thickness, by scaling of the calculations of Laikhtman and Kiehl [3], using what we believe are

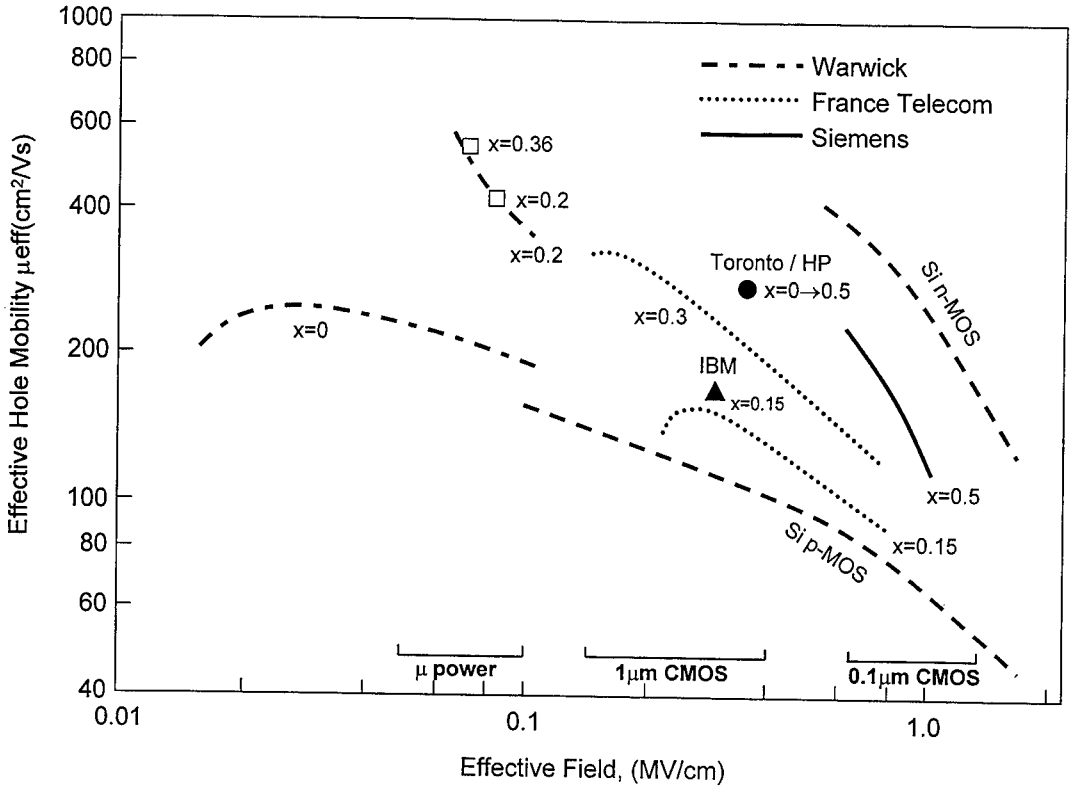


Fig. 3. Available experimental data on 300 K hole effective mobilities obtained in pseudomorphic $Si/Si_{1-x}Ge_x/Si$ structures plotted against effective field (E_{eff}) (see text). All data refer to buried SiGe channels except for the IBM sample where the gate dielectric was produced by plasma oxidation. The alloy composition in the Toronto/HP sample was graded. The squares refer to modulation-doped structures. The bars indicate the range of E_{eff} values present in micropower, 1 and 0.1 μm CMOS technologies. E_{eff} is defined in Refs. [19,20].

more realistic values of the acoustic phonon scattering potential and effective mass [4]. The non-polar optical phonon scattering is a worst case calculation, again based on the LK paper, but with complete relaxation of the selection rule which would otherwise forbid intra-subband NPO phonon scattering. These figures suggest that interface micro-roughness typical of that currently seen in SiGe/Si structures is the major factor limiting carrier mobility in samples with $x \geq 0.4$. In fact the roughness-limited mobility is almost certainly an underestimate being based on roughness associated fluctuations in energy eigenvalues in a rectangular well, whilst neglecting comparable contributions due to roughness-induced fluctuations in strain and coulomb potential [5] and also the influence of roughness at the nearby SiO_2/Si interface. Improvements in heterointerface quality, grading of Ge concentration and increased alloy thickness values could make for substantial gains in mobility. There has been some controversy over the magnitude of alloy scattering in SiGe but work by The University of Warwick suggests it is not limiting either the observed low temperature or 300 K mobilities in these structures [1]. Erroneous interpretation of experimental data on alloy scattering and mobilities in SiGe has occurred because a Hall scattering factor (r) of unity has been assumed, whereas McGregor et al. [6], for example, have recently shown that $r \approx 0.3$ at 300 K in an $x = 0.2$ alloy. A further problem concerns the long range (100 nm) roughness which develops on the surface of the compressively strained SiGe layer during epitaxial growth and which is not compatible with a planar technology, (see Fig. 2a). We have adopted a new approach in which a strained SiGe layer ($x = 0.36$) is grown epitaxially and capped with a 10 nm silicon layer at a low growth temperature (450°C) [7]. This yields a planar interface on the 100 nm scale and, following a high temperature (800°C) anneal, a high mobility hole channel in the SiGe whose planarity and strain are preserved (see Fig. 2b). Thermal oxidation of the Si layer can leave a very thin highly planar (3 nm) Si cap between the SiO_2 and the SiGe as shown in Fig. 2c.

Fig. 3 shows all the available data [8–12] on effective

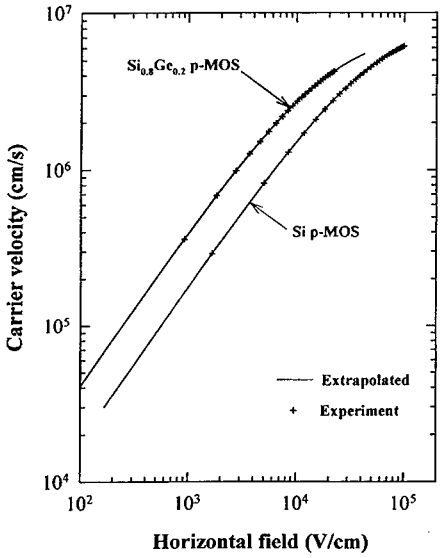


Fig. 4. Experimentally deduced velocity-field characteristic of an $x = 0.2$ $Si/Si_{1-x}Ge_x/Si$ p-channel MOSFET device, compared to a Si pMOS control. Vertical field is 0.1 MV cm^{-1}.

hole mobilities plotted against effective (vertical) field, E_{eff}, for fully pseudomorphic structures of nominally uniform alloy composition (except the Toronto/HP sample) produced by epitaxial growth, compared to mobilities observed at the unstrained Si/SiO_2 interface. Clearly, enhancements are now being obtained through the use of SiGe but much of this is recent work and improvements in material quality and heterostructure design should see further substantial increases. Most significant is the factor of 2 improvement seen by Siemens [8] at high E_{eff} values for $x = 0.5$. Some of the data presented in Fig. 3 (including that of Siemens) are symptomatic of a contribution from parasitic conduction in the Si cap layer at high E_{eff} values. Commensurate enhancements in the hole-velocity field characteristics have also been observed in an $x = 0.2$ MOS structure (see Fig. 4) [13]. These experimental results suggest a saturation velocity in $Si_{0.8}Ge_{0.2}$ pMOS equal to or in excess of that in Si pMOS. (see also, the implications for device performance, discussed in Section 3.)

3. SiGe HCMOS on virtual substrates

A more ambitious SiGe HCMOS technology involves driving up the n-channel performance using tensile-strained silicon, with a comparable p-channel

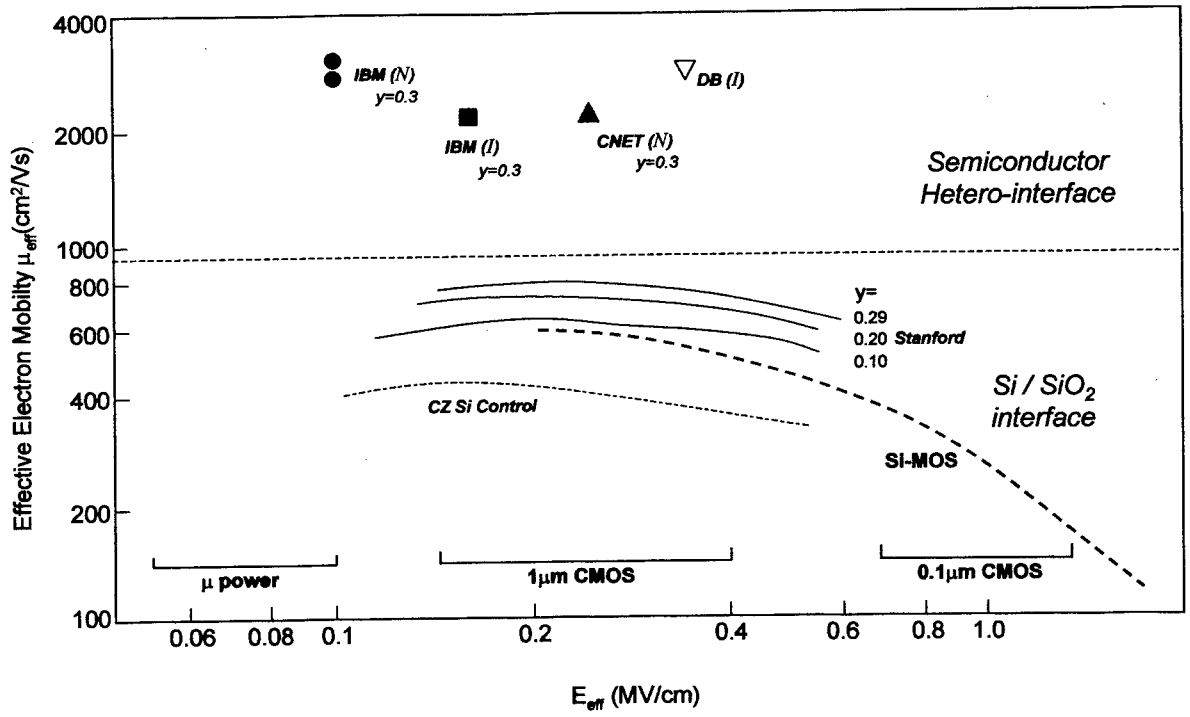


Fig. 5. Available experimental data on 300 K electron effective mobilities in strained Si grown on virtual substrates with terminating composition $Si_{1-y}Ge_y$. The upper section shows mobilities at remote doped hetero-interfaces and the lower section refers to oxide-gated/(tensile strained) Si interfaces. I denotes 'inverted' modulation-doped structure (doping supply layer below strained silicon) and N denotes 'normal' case (doping above silicon). The oxide-gated structures are operated in the inversion mode.

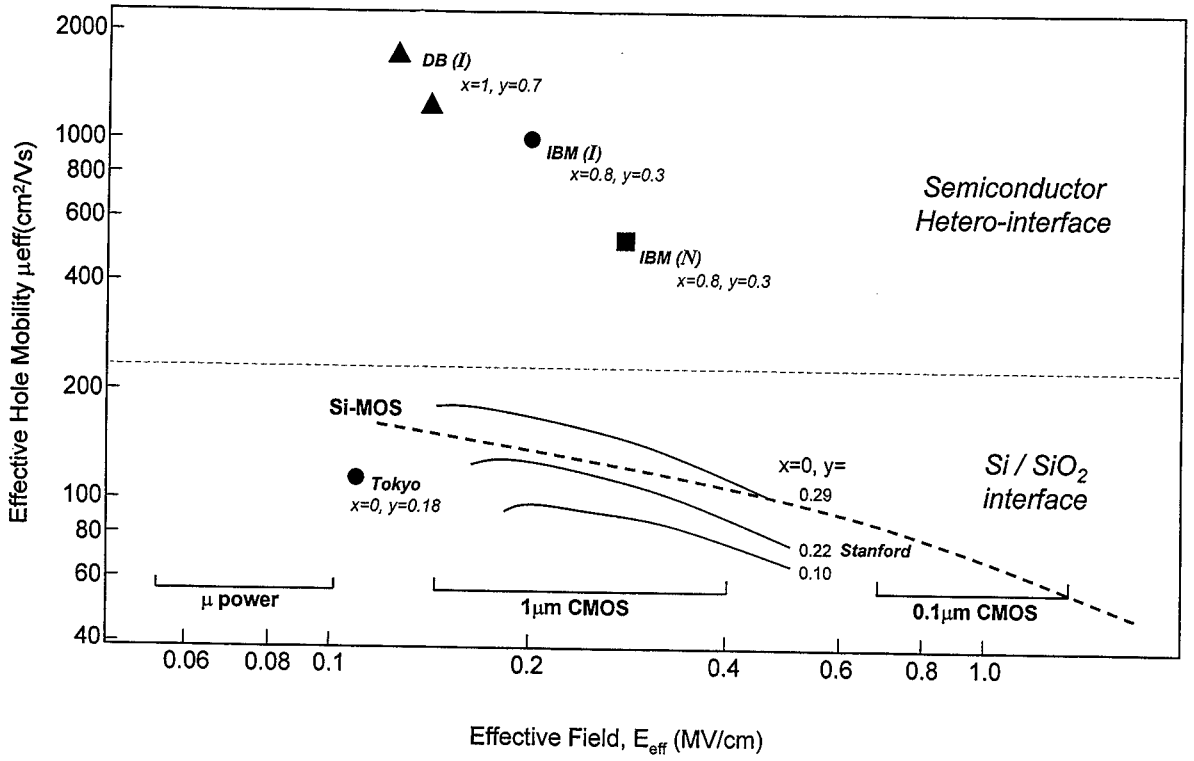


Fig. 6. Available experimental data on 300 K hole effective mobilities in compressively strained $Si_{1-x}Ge_x$ and tensile strained Si grown on virtual substrates with terminating composition $Si_{1-y}Ge_y$. The upper section shows mobilities for remote-doped hetero-interfaces and the lower section for oxide-gated/(tensile strained) Si interfaces. I denotes inverted and N denotes normal interface. The oxide-gated data refer to inversion layers.

performance obtained possibly using the same Si-channel or a strained channel of high Ge content (not excluding pure germanium). Here, a relaxed strain-tuning SiGe 'virtual substrate' (VS) needs to be formed on a silicon wafer prior to formation of the active channel. The VS is formed by grading the germanium composition (y) from $y = 0$ at the silicon substrate to a terminating composition typically of $y = 0.3$, but up to $y = 0.7$ for a pure germanium strained-channel. In order for such a process to be considered by MOS applications engineers the VS should have no detrimental influence on device performance or reliability, be able to withstand subsequent processing and be economic to produce.

Calculations of the electron mobility in biaxial tensile strained silicon have been carried out for $y = 0.3$ [14]. Electrons have $m^* = 0.2\ m_0$ and the degeneracy splitting of the conduction band suppresses f-phonon (intervalley) scattering, leaving only g-phonon scattering between two valleys, leading to predicted 300 K mobilities at $E_{\text{eff}} = 0.5$ MV cm^{-1} of 4000 cm^2 V^{-1} s^{-1}—around six times that in ordinary Si-n MOS. The calculation assumes no interface scattering—a reasonable approximation when tensile strained silicon is capped with a $Si_{0.7}$ $Ge_{0.3}$ layer. On the other hand, interface roughness will be important when capped with SiO_2. Calculations of the hole mobilities in tensile Si have only been made for the bulk 3D case for which a strong enhancement factor is indicated [15]. Preliminary experimental work suggests that mobilities up to 2000 cm^2 V^{-1} s^{-1} at $E_{\text{eff}} = 0.5$ MV/cm are likely in high Ge content layers [16]. Here again kinetically or stress-induced roughness could be a dominant influence especially for transport at the normal heterointerface.

All the available data on electron and hole mobilities in VS structures are presented in Figs. 5 and 6, again compared to mobilities at the conventional Si/SiO_2 interface [1,12,15–20]. The majority of the measurements have been made on modulation doped (MD) structures where the carriers are produced from a remote doping layer incorporated during epitaxial growth and reside at the Si/SiGe heterointerface. The E_{eff} values for MD structures have been calculated assuming reasonable values for background doping (10^{16} cm^{-3}). It is evident the MD structures values show the largest enhancements, albeit at low E_{eff} values, with up to a factor of five for electron mobili-

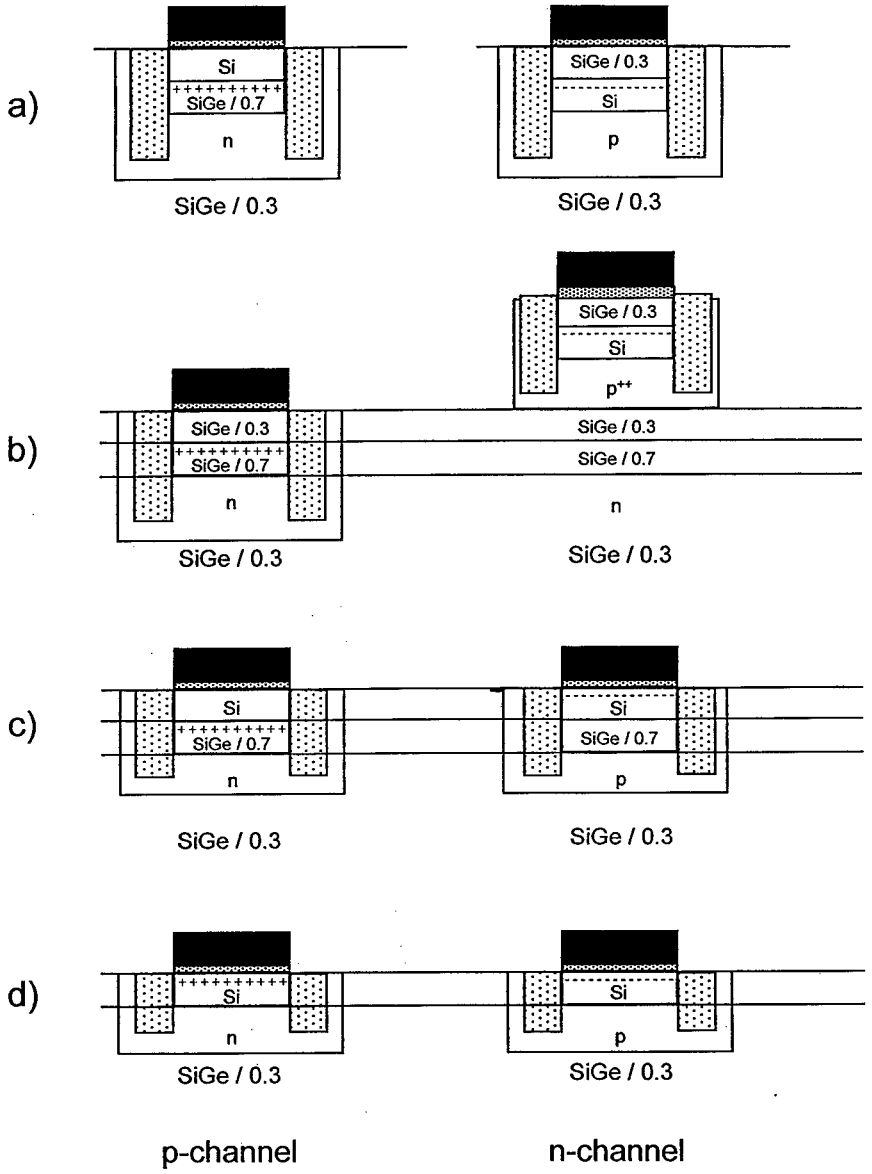


Fig. 7. Schematic of various process options for full HCMOS on a blanket coverage virtual substrate with terminating composition $y = 0.3$: (a) involves separate selective epitaxial depositions for the n- and p-channel devices after formation of deep n- and p-wells; (b) involves growth of a stacked p-channel/n-channel structure, followed by selective etching to remove the n-channel layers, hence revealing the p-channel device. The p^{++} layer can be used for isolation or, (c) the n-=channel is formed at the (tensile strained) Si/SiO_2 interface; (d) similar to (c) but both n- and p-channels are formed at the (tensile strained) Si/SiO_2 interface.

ties and ten for holes. Only a few studies have been made on oxide-gated structures and these were mostly for conduction at the Si/SiO_2 interface. The comparatively low electron and hole mobilities can possibly be attributed to the presence of oxide interface charge, Si/SiO_2 interface roughness, strain reduction during processing and parasitic conduction. Also, we would not expect the hole effective masses in tensile strained silicon to be as low as in the compressively strained SiGe channels referred to in Section 2 [15]. A further possibility is that whereas most of the silicon-layer may be strained, there could be some reduction at the Si/SiO_2 interface. Nevertheless, it is clear from the figures that mobility enhancements of more than a factor two at high E_{eff} values should be possible for both electrons and holes if the processing challenges can be met.

There have been several proposed designs for a VS-based HCMOS [20] and undoubtedly the optimum

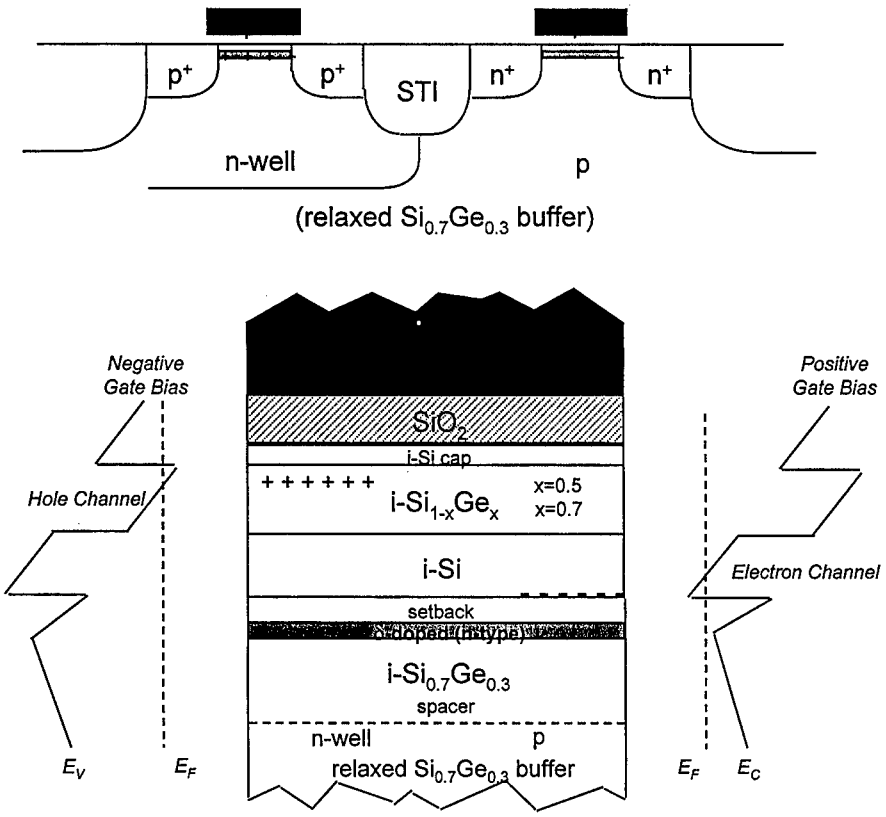


Fig. 8. Schematic of a stacked Si/SiGe HCMOS technology showing the n- and p-channel device layout (top) and the n- and p-channels formed in the layer sequence by applying symmetric gate biases just above threshold (bottom).

process would be to use the tensile strained Si/oxide interface for both the n- and p-channel devices (see Fig. 7 d) since this has the greatest simplicity and maximum gate capacitance. However, oxidation-induced interfacial roughness must not be allowed to compromise the carrier transport (possibly by using a deposited oxide) and there is uncertainty as to the maximum hole mobility achievable in tensile strained Si. Use of oxide-gated buried Si/SiGe hetero-interfaces for both the n and p channels may be beneficial in reducing interface scattering but gate capacitance will be reduced. In this case optimum performance might be obtained when the p- and n-channel devices are separately processed as shown in Fig. 7a.

Anticipating the day when carrier mobilities at high vertical fields might match those in MD structures Armstrong and co-workers [21] have proposed and modelled the stacked CMOS heterostructure for digital applications shown in Fig. 8—and their results are revealing. The structure involves a planar design and uses an n-type delta-doped layer to bend the energy band so as to avoid inversion of the low mobility surface channel. A p^+ polysilicon gate is used for each device and shallow trench isolation (STI) is used to avoid prolonged high-temperature processing which could give rise to strain relaxation and diffusion. Both devices operate in the enhancement mode, and have channel lengths of 0.2 μm and gate oxide thicknesses of 5 nm. The authors assume low-field mobilities of 2500 $cm^2\ V^{-1}\ s^{-1}$ and 800 $cm^2\ V^{-1}\ s^{-1}$ for SiGe HnMOS and HpMOS respectively, and that the saturation velocities are equal to 10^7 cm s^{-1}, the value for bulk (unstrained) Si. Factors of 1.23 and 2.25 increases are predicted for nMOS and pMOS current drive respectively, with proportional increases in extrinsic transconductances. In addition the drain currents in the Si/SeGe devices saturate at low drain biases, 0.4 V for HnMOS and 0.8 V for HpMOS. This cuts down the active power consumption by up to a staggering factor of three to four. The high carrier mobilities in the HCMOS structure result in a 6.4 times improvement in power-delay product at a stage delay of 28 ps for the unloaded case and a 4.6 times improvement at a delay of 55 ps for the loaded case compared to bulk Si. The minimum delay is 22 ps for unloaded Si/SiGe CMOS operating at 1.5 V. The performance advantage derives largely from the increased current drive of the Si/SiGe pMOSFET. CMOS delay depends on drive

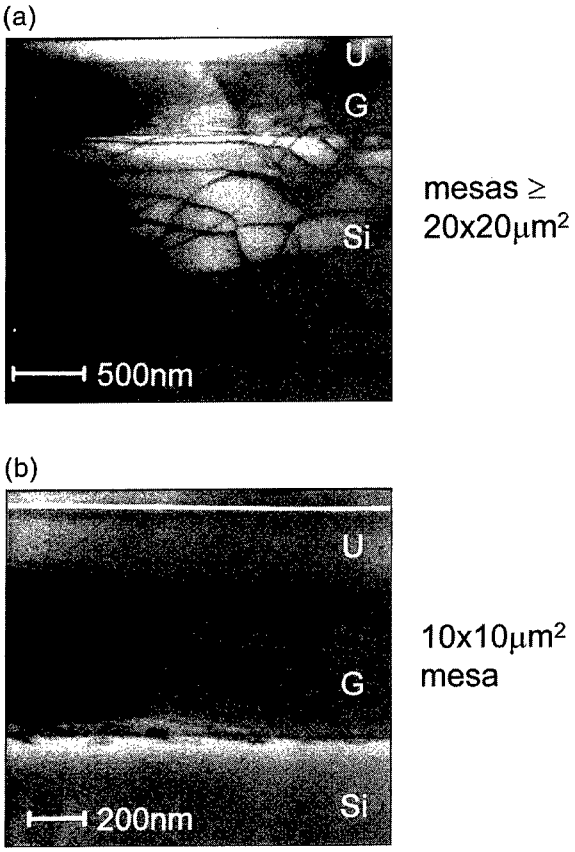


Fig. 9. XTEM of a graded $Si_{1-y}Ge_y$ virtual substrate grown at 550°C by solid source MBE with x varying linearly from 0.05 to 0.23 over 500 nm (G), followed by a uniform (U) 200 nm $Si_{0.77}Ge_{0.23}$ layer, then a 10 nm (tensile strained) Si layer and topped with a 50 nm uniform $Si_{0.77}Ge_{0.23}$ cap layer: (a) growth on Si mesas $>20 \times 20$ μm^2 (including blanket coverage) showing a number of dislocation pileups and a dislocation network penetrating deep into the Si mesa. The top surface has a crosshatched topography; (b) growth on a 10×10 μm^2 mesa—showing a threading dislocation network confined to the graded SiGe region and no evidence of threading components penetrating either below or above the graded layer. The silicon layer is highly planar and the topmost surface is free of crosshatch.

current, so it is encouraging to note that the mobility gains to be derived from SiGe are not offset by increases in threshold voltage. Crucially the off-state current and punchthrough performance should not be worse than bulk silicon, and the predicted subthreshold swing and drain induced barrier lowering (DIBL) figures are, therefore, similarly encouraging. Simulations of how mobility improvements impact on analogue performance have also been made indicating that increases of between 50% (n-channel) and 100% (p-channel) in f_T can be expected for VS-based devices [21]. To date, however, there has been very little published on either the designs, or performances at circuit level of systems involving SiGe transistors.

Clearly, major technological challenges remain, not least of these being the one of producing a commercially-viable low-defect strain-tuning virtual substrate. There is considerable activity world-wide in this regard, and several recent developments indicate encouraging progress and real grounds for optimism. Hammond et al. [22], for example have shown that by restricting growth to small areas (i.e. 10×10 μm^2) defined by mesas, relaxed defect-free, linearly graded, SiGe can be obtained with smooth surface topography (see Fig. 9).

4. Conclusions and outlook

At this stage we can be confident that very significant performance gains in Si CMOS can be accessed by incorporating high mobility Si/SeGe channels as part of the vertical architectures of the p- and n-type devices. Which interface is optimum for CMOS—the buried channel heterointerface or the semiconductor/oxide interface remains an issue and further investigations of oxide formation on SiGe heterostructures are urgently required. Recent studies show that a new enhanced performance CMOS technology can be produced in current and future generations simply by incorporating a thin SiGe layer in the p-channel device and this seems to be the first entry point for SiGe CMOS. With the inevitable advances in materials research, virtual-substrate-based MOS technology promises stunning performance improvements impacting on all applications involving CMOS and allowing Si-MOS technologies to comfortably access huge new markets requiring devices and circuits operating well into the GHz regime.

Acknowledgements

The authors have had very useful discussions with A. Asenov, H. Fischer, T. Grasby, M. J. Kearney, A. O'Neill, C. Papavassiliou, L. Risch and T. Thornton and we thank them for that. We are particularly grateful to Lothar Risch, Hermann Fischer, T. Grasby and S. Kiatgamolchai for allowing us to use their device data prior to publication.

References

[1] Whall TE, Parker EHC. J Phys D Appl Phys 1998;31:1397 and references therein.

[2] Verdonckt-Vandebroek S. IEEE Trans Elec Dev 1994;41:90.
[3] Laikhtman B, Kiehl RA. Phys Rev B 1993;47:10,515.
[4] Whall TE. Thin Solid Films 1997;294:160.
[5] Feenstra RM, Lutz MA. J App Phys 1995;78:6391.
[6] McGregor B, Parry C, Phillips PJ, Parker EHC, Whall TE. App Phys Lett (in press).
[7] Grasby T, Parry C, Phillips PJ, Whall TE, Parker EHC. Appl Phys Lett 1999;74:10,1848.
[8] Fischer H, Risch L, Siemens AG, private comm.
[9] Boullion P, Skotnicki T, Kaladis C, Gwoziecki R, Dollfus P, Regolini JL, Sagnes I, Bodnar S. IEDM 96, Tech. Digest, 959.
[10] Voinigescu SP, Salama CAT, Noel JP, Kamins TL. IEDM 94, Tech. Digest 369.
[11] Li PW, Yang ES, Fang YF, Chu JO, Meyerson BS. IEEE El Dev Lett 1994;15:402.
[12] Taur Y, Nowak EJ. IEDM 97-215.
[13] Ansaripour G, Asenov A, Braithwaite G, Kaya S, Parker EHC, Whall TE. (in preparation).
[14] Miyata H, Yamada T, Ferry DK. App Phys Lett 1993;62:2661.
[15] Nayak DK, Goto K, Tutani A, Murota J, Shiraki Y. IEEE Trans Elec Dev 1996;43:1709 and references therein.
[16] Höck G, Glück M, Hackbarth T, Herzog HJ, Kohn E. EMRS Spring Meeting, June 16–19 1998, Strasbourg (France).
[17] Garchery L, Warren P, Sagnes I, Badoz PA. MRS Symp Proc 1995;379:321.
[18] Welser J, Hoyt JL, Takagi S, Gibbons F. IEDM Tech Digest 1994;373.
[19] Rim K, Welser J, Hoyt JL, Takagi S, Gibbons JF. IEDM Tech Digest 1995;517.
[20] O'Neill AG, Antoniadis DA. IEEE Trans Elec Dev ED 1997;44:80.
[21] Armstrong MA, Antoniadis DA, Sadek A, Ismail K, Stern F. IEDM Tech Digest 1995;761.
[22] Hammond R, Phillips PJ, Whall TE, Parker EHC, Graf T, von Känel H, Shields AJ. Appl Phys Lett 1997;77:2517.

PERGAMON

Solid-State Electronics 43 (1999) 1507–1512

SOLID-STATE ELECTRONICS

InP-based complementary HBT amplifiers for use in communication systems

Donald Sawdai*, Dimitris Pavlidis

Solid-State Electronics Laboratory, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI, 48109, USA, (http://www.eecs.umich.edu/dp-group/)

Abstract

This paper addresses a method to improve the linearity characteristics of power amplifiers by developing a PNP HBT technology and combining the PNP HBTs with NPN HBTs in a push–pull amplifier. InP-based PNP HBTs were fabricated with $h_{fe} > 30$, $BV_{ECO} = 5.6$ V, and f_T and f_{max} of 11 and 31 GHz, respectively, which is the best reported for InAlAs/InGaAs PNP HBTs.

Common-collector push–pull amplifiers were simulated using these HBTs, demonstrating an improvement of 14 dB in second harmonic content under Class B operation. A common-emitter push–pull amplifier fabricated with the same HBTs demonstrated the best IM3 (by ~7 dBc) and smaller second harmonic content (by ~9 dBc) compared with NPN HBTs. In addition, the circuit produced 1.32 dBm more output power than the NPN HBT alone at 1 dB of gain compression. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Wireless communication systems require power amplifiers with high linearity, high power-added efficiency (PAE), and high power-handling capability. HBTs offer this capability, and AlGaAs/GaAs devices have already been introduced in wireless systems. Another promising device approach is InP-based HBTs, which have demonstrated very good high-frequency performance and have been implemented in various integrated circuits for electronic and optoelectronic applications. Their power capability is promising, with power levels up to 1.4 mW/μm^2 reported by the authors using NPN single HBT designs [1]. Further enhancement is possible by means of double heterojunction designs, and a power density of 3.6 mW/μm^2 with PAE of 54% at 9 GHz has been reported with the latter approach from Hughes Research Labs [2]. InP/InGaAs HBTs are consequently of interest for power amplification in communication systems.

This paper addresses another method to improve power performance by developing a PNP HBT technology and combining the PNP HBTs with NPN HBTs in a push–pull amplifier. While not as impressive as their NPN counterparts, the current state-of-the-art performance for PNP HBTs is sufficient for applications up to X-band, and further enhancement in capability can be achieved. The best-published InAlAs/InGaAs PNP HBTs demonstrated $\beta = 30$, $f_T = 13$ GHz, and $f_{max} = 35$ GHz as reported by the authors [3], while the best-published InP/InGaAs PNP HBTs were reported by AT&T with $\beta = 20$, $f_T = 11$ GHz, and $f_{max} = 25$ GHz [4].

Complementary push–pull amplifiers have several advantages over single-transistor power amplifiers, and they are much simpler to design than NPN-only push–

* Corresponding author.
E-mail address: dsawdai@engin.umich.edu (D. Sawdai)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00096-9

pull amplifiers. Since the output voltage swing is generated across two transistors in the push–pull amplifier, approximately twice the output voltage is possible. In addition, push–pull amplifiers can produce linear output in Class AB or Class B, which can have efficiencies as high as 78%. To achieve the same linearity, single transistor amplifiers must operate in Class A, with efficiency limited to 50%. InP-based HBTs offer an additional advantage for push–pull operation: the turn-on voltage V_{BE} is approximately 0.75 V, versus 1.4 V for GaAs-based HBTs. This smaller turn-on voltage reduces both crossover distortion and overall power consumption.

In this paper, we present issues and results on complementary push–pull circuits fabricated in the InP material system. A first report on such circuits was recently presented by the authors [5]. This paper provides further details on their design and performance. First, device results are presented for both the NPN and the PNP HBTs used in this study. Then, simulations of complementary amplifiers are presented to demonstrate the advantages of push–pull amplifiers. Finally, the results from a hybrid common-emitter push–pull amplifier are presented and discussed.

2. Integration of NPN and PNP HBTs

Several technologies are available for integrating NPN and PNP HBTs in order to fabricate circuits. In the InAlAs/InGaAs material system, first PNP and then NPN HBT layers have been grown uninterrupted on a mesa-patterned substrate, which resulted in a planar wafer with NPN and PNP HBTs after device fabrication [6]. In addition, several NPN/PNP integration technologies have been demonstrated in the AlGaAs/GaAs material system. Hybrid common-emitter push–pull amplifiers using $6f \times 2 \times 20$ μm^2 HBTs have demonstrated 6 dB of gain, 0.5 W of output power, and 42% PAE at 10 GHz [7]. Selective low-pressure OMVPE has also been used to create monolithic common-emitter push–pull amplifiers that demonstrate power combination at 10 GHz with power cancellation in the second harmonic [8]. Finally, selective MBE has been used to monolithically create common-collector push–pull amplifiers that produce 7.2 dB gain and 7.5 dBm output power at 2.5 GHz using $4f \times 3 \times 10$ μm^2 HBTs [9].

InP-based HBTs offer several advantages over GaAs-based HBTs in satisfying the requirements of wireless communications. In general, the improved frequency performance of InP-based HBTs produces higher gain at high frequencies. The lower contact and sheet resistances of the emitter cap and subcollector layers, along with the smaller offset voltage (0.2 V vs 0.4 V), reduce the saturation "knee" voltage of InP-based HBTs and allow the use of low-voltage batteries. Since InP has ten times smaller surface recombination velocity than GaAs, the current gain is more uniform with bias, which should produce better amplifier linearity. Less surface recombination also increases the gain of large power HBTs with many emitter fingers (large total periphery), and it also allows for very small HBTs for digital applications. Finally, the higher thermal conductance of the InP substrate (0.7 W/cm-K, vs 0.5 for GaAs) allows for more power dissipation in any given HBT design. InP-based single HBTs do suffer from low breakdown voltages, typically around 2 to 7 V. This limitation can be overcome in double HBTs with InP collectors. An additional benefit of double HBTs is that the offset voltage is reduced to almost 0 V.

3. InP-based NPN and PNP HBT performance

Both the NPN HBTs and the PNP HBTs used in this study were grown and fabricated on semi-insulating InP wafers at The University of Michigan. The NPN InP/InGaAs epilayers were grown using an EMCORE low-pressure Metalorganic Chemical Vapor Deposition system, while the PNP InAlAs/InGaAs epilayers were grown by solid-source Molecular Beam Epitaxy. The HBTs were fabricated using our standard self-aligned processes, which produce a 0.2-μm base contact-to-emitter separation. The base-collector capacitance was minimized by using the base contact as a self-aligned etch mask for the base mesa. Electroplated gold airbridges were used to connect the HBTs to coplanar interconnect pads for off-wafer bonding or for on-wafer probing, and micro-trenches were wet-etched to isolate the semiconductor junctions under the airbridge pads from the intrinsic HBTs. Detailed HBT results from these NPN [1] and PNP [10] layers can be found elsewhere.

The HBTs used in this study had 5×10 μm^2 emitters. The NPN HBTs had a DC small-signal gain $h_{fe} > 70$. Their maximum collector current density was about $J_C = 1.4 \times 10^5$ A/cm^2, and the breakdown for large NPN HBTs was as high as $BV_{CE0} = 7.2$ V. Their optimal f_T and f_{max} were 97 and 51 GHz, respectively, at $V_{CE} = 2.0$ V and $I_C = 46$ mA. Similarly, the PNP HBTs had a DC small-signal gain $h_{fe} > 30$. Although the maximum collector current density was about $J_C = 9 \times 10^4$ A/cm^2, the current gain compressed rapidly for $J_C > 5 \times 10^4$ A/cm^2. The breakdown for these HBTs was $BV_{EC0} = 5.6$ V at 10 A/cm^2. Microwave measurements resulted in optimal f_T and f_{max} of 11 GHz and 31 GHz, respectively, at $V_{EC} = 4.0$ V and $I_C = 11.7$ mA.

On-wafer power characterization was performed at 10 GHz using a system developed in-house with

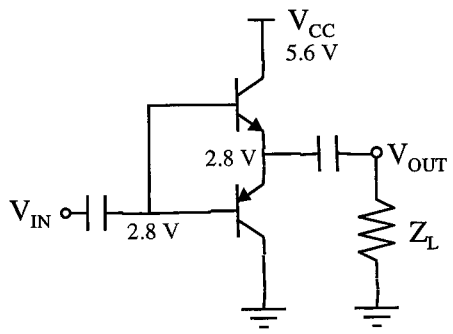


Fig. 1. Circuit diagram of Class B push–pull amplifier with single bias supply used in simulations.

FOCUS electromechanical tuners on both the source and the load. The PNP HBTs were characterized using load pull techniques and demonstrated a small-signal gain of 10 dB, peak power-added-efficiency of 24%, and maximum output power density of 0.49 mW/μm^2 [10]. These characteristics are very similar to the NPN HBTs, which had slightly higher gain (+1dB) and efficiency (+5%) but produced less output power than the PNP HBTs (−3dBm). Also note that these results are very similar to power performance published for PNP AlGaAs/GaAs HBTs [11]: $P_{out}=0.57$ mW/μm^2 and PAE = 33% at 10 GHz.

Further studies of the PNP HBTs indicated that output power scaled linearly with the number of emitter fingers up to 10 fingers, which was the largest HBT measured. The microwave gain was constant up to 4 emitter fingers, and the gain degraded by 3 dB when the number of emitter fingers was increased to 10. Finally, common-base PNP HBTs provided 3–6 dB more gain than similarly biased common-emitter PNP HBTs, which resulted in 5% higher efficiency. These results confirmed the suitability of the PNP HBTs for high-frequency power applications.

4. Push–pull simulations and issues

In order to investigate the benefits of using PNP HBTs together with NPN HBTs in various integrated circuits, the large-signal characteristics of several NPN-only and complementary amplifiers were simulated. For these simulations, non-linear large-signal models were extracted for the HBTs from the previous section. These extractions accurately reproduced most HBT characteristics that dominate the power, efficiency, and linearity performance, such as parasitic resistances and capacitances, turn-on voltage, saturation voltage, Early effect, gain non-uniformity with I_C, and base-collector breakdown. These large-signal HBT models were then used in all of the circuit simulations.

The first circuit investigated was a Class B common-collector push–pull output stage (Fig. 1) consisting of one NPN and one PNP HBT. DC blocking capacitors were used to allow for a single 5.6-V DC power supply to the circuit. The circuit was simulated using the non-linear HBT models in the time domain over a wide range of input power at 1.9 GHz. The results were then compared to similar simulations of single NPN and PNP common-collector power amplifiers with the base unbiased for Class B operation. For the individual NPN and PNP amplifiers, the collector was biased through an RF choke to a single 1.8 V or −3.8 V supply, respectively, which gave optimal performance without HBT breakdown.

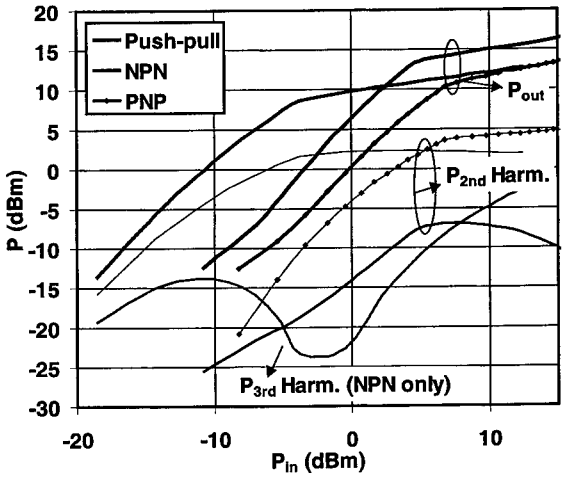


Fig. 2. Simulated power output at fundamental, second, and third harmonics for Class B push–pull, NPN-only, and PNP-only amplifiers at 1.9 GHz.

The input/output power characteristics of the Class B push–pull amplifier are shown together with those of the individual HBTs in Fig. 2. Since the DC gain of the HBTs is low at very small bias, the microwave gain for the Class B amplifiers is low for small P_{in}. However, as P_{in} increases, self-biasing causes the DC I_C to increase, which in turn increases the microwave gain of the amplifier. The peak gain for the NPN and PNP amplifiers were 12.4 and 3.5 dB, respectively, with the push–pull gain at an intermediate 8.6 dB. For larger input power, the waveform peaks pushed the HBTs into saturation, limiting further output power. Note, however, that the maximum output power from the push–pull circuit was roughly 3 dB more than from the individual HBTs, since at that point that output power from both saturated HBTs in the circuit are combined. Although its higher gain gave the NPN amplifier higher PAE at low P_{in}, the push–pull amplifier's higher saturated P_{out} allowed it to achieve the same peak PAE as the NPN amplifier of 40%. Note that these Class B circuits are off and consume no DC power when there is no input signal.

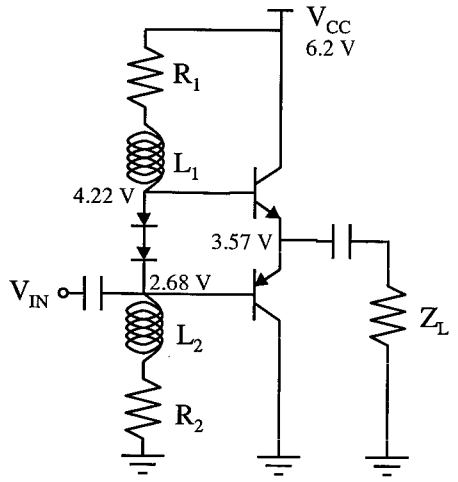


Fig. 3. Circuit diagram of Class AB push–pull amplifier used in simulations.

The advantage of the push–pull amplifier can be seen in the harmonic content of the signal from Fig. 2. At the beginning of P_{out} saturation, the second harmonic content of the output of the NPN and PNP amplifiers are both −7 dBc. The theoretical even-order harmonic cancellation in the push–pull amplifier reduces its second harmonic content to −21 dBc. However, since the odd-order harmonics do not inherently cancel in push–pull amplifiers, its third harmonic shows little variation from the NPN and PNP amplifiers at similar levels of self-biasing.

The second circuit investigated is a Class AB version of previous circuit, with two diode-connected HBTs between the bases of the amplifying NPN and PNP HBTs for biasing (see Fig. 3). Two resistors and two inductors serve to bias these diodes. This Class AB push–pull amplifier was compared to Class AB single NPN and Class AB single PNP amplifiers, both which had the base biased by a similar two-resistor two-inductor network.

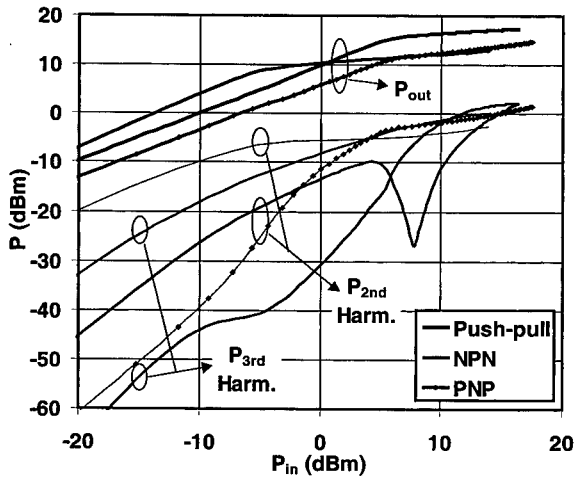


Fig. 4. Simulated power output at fundamental, second, and third harmonics for Class AB push–pull, NPN-only, and PNP-only amplifiers at 1.9 GHz.

The input/output power characteristics of the Class AB push–pull amplifier are shown together with those of the individual HBTs in Fig. 4. Since the Class AB amplifiers keep the HBTs slightly on when no input signal is present, the low-gain region at low I_C is avoided, so that the microwave gain is almost flat for all P_{in} below output power saturation. This microwave gain, 10.3/13.0/6.8 dB for the push–pull/NPN/PNP amplifiers, is somewhat higher than for the Class B case. Otherwise, the gain of the Class AB amplifiers behave similar to their Class B counterparts. Even with no input signal present, the Class AB amplifiers draw significant DC power—for example, the push–pull circuit draws 13.0 mW through the bias network and 12.0 mW through the HBT pair. Therefore, the Class AB peak PAE, 36%, is slightly lower than that of the Class B amplifier. However, the PAE for $P_{in} < +3$ dBm is higher for the Class AB amplifier than for Class B due to its much higher gain.

The Class AB push–pull amplifier demonstrates second harmonic cancellation: −25 dBc at P_{out} saturation, versus −16 dBc for the NPN and PNP amplifiers. The third harmonic is also reduced from −20 dBc for the NPN and PNP amplifiers to −30 dBc for the push–pull amplifier at the same power levels. For the push–pull amplifier, the second harmonic content is slightly reduced (−4 dB) when moving from Class B to Class AB, while the third harmonic content is greatly reduced (−14 dB) due to the elimination of the cross-over distortion.

Similarly, the Class AB circuit was also simulated using large-signal models derived from published characteristics of typical NPN and PNP AlGaAs/GaAs HBTs [12–14]. When using the same $V_{CC}=6.2$ V, the peak gain of the GaAs-based circuit (7.7 dB) was 2.6 dB less than the InP-based circuit, primarily due to the better frequency performance of the InP NPN HBT. Both the larger turn-on voltage and the larger knee voltage of the GaAs HBTs limited the output power at 1 dB of gain compression to 12.7 dBm, which is 3.1 dB lower than for the InP push–pull circuit. This reduced P_{out}, together with increased power consumption due to the larger turn-on voltage, limited the GaAs-based PAE to 25%, as compared to the InP-based PAE of 36%. Although these results do not necessarily reflect the full potential of each technology, since they are based on typical rather than optimized characteristics, they do indicate the power advantages

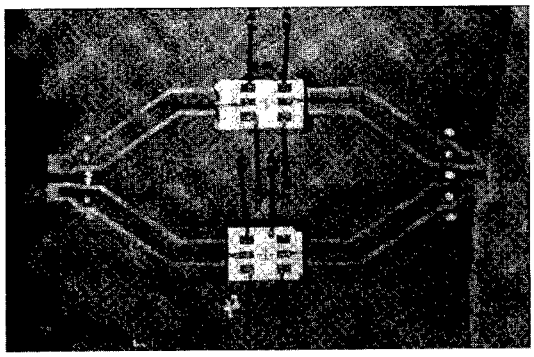


Fig. 5. Micrograph of fabricated NPN/PNP push–pull common-emitter amplifier.

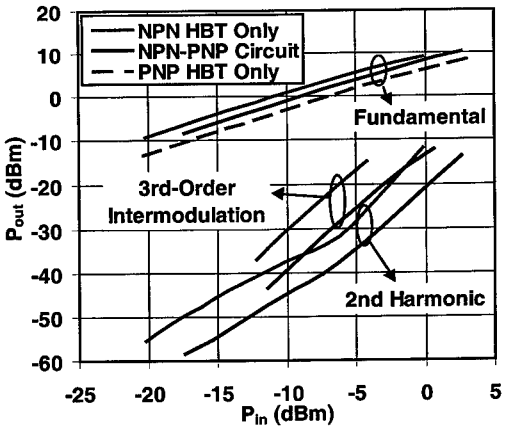


Fig. 6. Measured power output at fundamental, second harmonic, and third-order intermodulation for NPN HBT, PNP HBT, and push–pull amplifier at 8 GHz.

of InP-based HBTs over GaAs-based HBTs in push–pull amplifiers.

The InP-based simulations demonstrate several advantages of the push–pull amplifiers over single-HBT amplifiers. The main advantage is the reduced distortion due to harmonic cancellation, which ideally would show higher improvements if the NPN and PNP HBTs were better matched in terms of gain. Another advantage is the 3-dB increase in output power, which is primarily achievable due to the 5.6-V supply used in the push–pull amplifier. These NPN HBTs have $BV_{CE0}=7.2$ V, but they degrade rapidly when conducting more than a few milliamps for $V_{CE}>3.0$ V. While the push–pull creates up to 4 V of V_{CE} across the NPN HBT when the NPN is off and the PNP is on, the V_{CE} across the NPN is less than 2.5 V when the NPN is on. The same argument also holds for the PNP, allowing the push–pull circuit a greater supply voltage without breakdown and hence greater output power than single-HBT amplifiers.

Both Class B and Class AB push–pull amplifiers had only moderate PAE when compared to the theoretical maximum of 78% for Class B. The primary limitation on PAE was the large knee voltage for the HBTs: 0.5 V and 2.1 V at $I_C=30$ mA for the NPN and PNP, respectively, which together reduced the available output voltage swing from the 5.6-V DC supply to 3.0 V. Reducing the collector and emitter parasitic resistances will reduce the knee voltage, especially for the PNP. Also, using a graded emitter-base junction or using a large bandgap collector can reduce the offset voltage, which would also reduce the knee voltage. The other limitation is the low gain at low I_C, which reduces the gain and linearity of the Class B amplifier. Enhanced emitter-base junction designs and fabrication technologies can increase the gain at low I_C, which would enable the Class B amplifier to improve its linearity and gain while maintaining its high PAE.

5. Push–pull experiment

In order to verify the benefits of complementary technologies, a complementary common-emitter push–pull amplifier was designed and fabricated using the NPN and PNP HBTs described above. Although the common-collector amplifier with single bias supply described in the previous section would be used more often as a power amplifier in a real communication system, a common-emitter amplifier with four bias supplies was fabricated in this experiment. The common-emitter design is simpler for a first-run test, and it directly demonstrates the advantages of complementary amplifiers.

A passive coplanar circuit was designed and fabricated on 10-mil alumina substrates to permit feeding of NPN and PNP common-emitter HBTs from a common input signal probe (see Fig. 5). The HBTs were thinned to 200 μm, cleaved, and then mounted in the circuits on the alumina substrates. Gold bond wires connected the HBT chips to the electroplated gold interconnects on the circuit. An X-band test bench was designed for a variety of on-wafer measurements on the push–pull amplifier.

The circuit characteristics were measured at 8 GHz with both NPN and PNP in Class A ($I_{C0}=20.3$ mA and 11.6 mA, respectively). The input versus output power curves are shown in Fig. 6, which also shows the power output at the second harmonic for single-tone excitation and the third-order intermodulation (IM3) power output for two-tone excitation at 500 kHz separation. Note that for the circuit, this graph shows the *total* power input to the circuit and *total* power output from the circuit; therefore, each HBT in the circuit is receiving half of the input power. The data for the individual HBTs were measured independently for

each HBT. At 1 dB of gain compression, the NPN and PNP HBTs produced 7.7 and 7.5 dBm of output power, respectively, while the circuit produced 9.0 dBm. This ~1.4 dB increase in output power is limited by the gain mismatch between the NPN and the PNP HBTs; when the circuit is at 1−dB compression ($P_{in} = +0.9$ dBm), each HBT receives −2.1 dBm input power, causing the high-gain NPN to be over-saturated and the low-gain PNP to be under-saturated. If the HBTs were perfectly matched, both HBTs would produce the same gain and saturate at the same P_{in}, which would result in 3 dB increase in output power for the push–pull circuit. Note that the simulations do demonstrate the full 3 dB increase in output power, mainly due to the different bias scheme used there.

Fig. 6 also demonstrates the improvements in linearity for the circuit when compared to the NPN HBT alone. At the highest power levels measured, the circuit demonstrated less IM3 (by ~7 dBc) and smaller second harmonic content (by ~9 dBc) when compared to the NPN HBT. These numbers indicate an advantage for the push–pull circuit even under Class A operation.

6. Conclusions

In this paper, we present the characteristics of PNP InP-based HBTs, simulated results on complementary push–pull HBT amplifiers, and first results on the implementation of similar complementary HBT amplifiers. The PNP HBTs demonstrated 10 dB of gain and output power that scaled well with HBT area at 10 GHz. Both the simulations and experiments demonstrate an improvement to the linearity characteristics of NPN HBT amplifiers by including a PNP HBT in a push–pull configuration. Simulations also indicate that reducing the parasitic emitter and collector resistances can substantially improve the amplifier efficiency, and more advanced emitter-base junction designs can increase the gain and linearity of Class B amplifiers.

Such complementary amplifiers could lead to improved wireless communication systems. The high efficiency and Class B operation increases battery life, and the decreased harmonic distortion allow the use of popular modulation techniques with reduced signal bandwidth, such as QPSK. In addition, integration of NPN and PNP HBTs offers other benefits, such as higher gain and less power consumed in small-signal amplifier stages through the use of active loads.

Acknowledgements

The authors would like to thank Dr. K. Hong for the growth of the NPN HBT wafers and Dr. X. Zhang and Professor P. Bhattacharya for the growth of the PNP HBT wafers. This work was supported by ARO MURI under Contract DAAH04-96-1-0001.

References

[1] Sawdai D, Plouchart J-O, Pavlidis D, Samelis A, Hong K. Power performance of InGaAs/InP single HBTs. In: 8th Int Conf InP and Related Materials, April, 1996. p. 133–6.

[2] Nguyen C, Liu T, Chen M, Virk R. Bandgap engineered InP-based power double heterojunction bipolar transistors. In: 9th Int Conf InP and Related Materials, 1997. p. 15–9.

[3] Sawdai D, Zhang X, Pavlidis D, Bhattacharya P. Performance optimization of PNP InAlAs/InGaAs HBTs. In: Proc IEEE/Cornell Conf on Advanced Concepts in High Speed Semiconductor Devices and Circuits, 1997. p. 269–77.

[4] Lunardi LM, Chandrasekhar S, Hamm RA. High-speed, high-current-gain P-n-p InP/InGaAs heterojunction bipolar transistors. IEEE Electron Dev Lett 1993;14:19–21.

[5] Sawdai D, Pavlidis D. Power amplification using NPN and PNP InP HBTs and application to push-pull circuits. In: 22nd Workshop on Compound Semiconductor Devices and Integrated Circuits (WOCSDICE), 1998. p. 45–6.

[6] Stanchina W, Metzger R, Pierce M, Jensen J, McCray L, Wong-Quen R, Williams F. Monolithic fabrication of NPN and PNP AlInAs/GaInAs HBTs. In: 5th Int Conf InP and Related Materials, 1993. p. 569–71.

[7] Tserng H, Hill D, Kim T. A 0.5-W complementary AlGaAs-GaAs HBT push-pull amplifier at 10 GHz. IEEE MW & Guided Wave Lett 1993;3(2):45–7.

[8] Slater Jr D, Enquist P, Hutchby J. Harmonic canellation in monolithic AlGaAs/GaAs Npn/Pnp HBT push-pull pairs. In: Proc IEEE/Cornell Conf on Advanced Concepts in High Speed Semiconductor Devices and Circuits, 1991. p. 305–14.

[9] Kobayashi K, Umemoto D, Velebir J, Oki A, Streit D. Integrated complementary HBT microwave push-pull and Darlington amplifiers with p-n-p active loads. IEEE J Solid-State Circuits 1993;28(10):1011–7.

[10] Sawdai D, Zhang X, Pavlidis D, Bhattacharya P. Power performance of PNP InAlAs/InGaAs HBTs. In: 10th Int Conf InP and Related Materials, 1998. p. 72–5.

[11] Hill D, Kim TS, Tserng HQ. X-band power AlGaAs/InGaAs P-n-p HBT's. IEEE Electron Dev Lett 1993;14:185–7.

[12] Slater D, Enquist P, Hutchby J, Morris A, Trew R. Pnp HBT with 66 GHz f_{max}. IEEE Electron Dev Lett 1994;15:91–3.

[13] Umemoto D, Velebir J, Kobayashi K, Oki A, Streit D, Integrated npn/pnp GaAs/AlGaAs HBTs grown by selective MBE. Electronics Lett 1991;27:1517–1518.

[14] Hill D, Kill T, Tserng H. AlGaAs/GaAs Pnp HBT with 54 GHz F_{max} and application to high-performance complementary technology. In: 51st Ann Device Research Conf, 1993 Paper IIIA-6.

PERGAMON

Solid-State Electronics 43 (1999) 1513–1518

SOLID-STATE ELECTRONICS

A 77 K analog monolithic HEMT amplifier for high-speed Josephson-semiconductor interface circuit

Naoki Harada*, Yuji Awano, Kohki Hikosaka, Naoki Yokoyama

Fujitsu Limited, 10-1 Morinosato-Wakamiya, Atsugi, 243-0197, Japan

Abstract

We developed 0.5 μm-gate 77 K analog monolithic HEMT amplifier that is to be used in an interface from a Josephson IC to a semiconductor IC. The HEMT was built from InGaP/InGaAs/GaAs materials that provide a stable system at 77 K. The amplifier includes a differential amplifier as its first stage to cancel out ground level fluctuations in the Josephson IC, and high gain source-grounded amplifier. An output of 0.7 $V_{p\text{-}p}$ was obtained from a complementary input signal of 30 $mV_{p\text{-}p}$, 3 Gbit/s, that was in RZ format. We successfully used this HEMT amplifier for transferring a voltage signal from 10-stack Josephson high-voltage drivers to a room temperature 50 Ω system with an amplitude of 0.7 $V_{p\text{-}p}$ and a clock frequency of 300 MHz. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Josephson junction devices offer such desirable features as fast switching speed and extremely low-power dissipation. However, it is difficult to construct a large-scale system that would be based on this technology because of difficulties in large scale integration. Creating hybrid system combining Josephson and semiconductor devices is a realistic approach to constructing a Josephson system for practical use. A hybrid system would make it possible to exploit both the high-speed processing capabilities of Josephson circuits and the large-scale processing capability of semiconductor circuits concurrently. For example, a group at UC Berkley proposed a CMOS memory with high-speed superconducting current sensor [1] and CMOS calibrated Josephson A/D converter [2]. Tahara proposed and tested a high-speed superconductive interconnection network circuit which works as a data switching system between room-temperature silicon processor elements [3]. However, the output voltage of Nb-based Josephson device is only 2.8 mV, while room-temperature semiconductor operation requires an operational voltage from at least several hundreds of millivolts up to 1 volt. Therefore, it would be necessary to increase the output voltage from the Josephson device by more than two orders of magnitude. Some Josephson-semiconductor interface circuits have been proposed and developed [4,5] to date. Suzuki built a first interface circuit using a Josephson high-voltage driver and a commercial GaAs comparator [4]. However, his group needed 104 Josephson junctions to drive the comparator, which caused an increase of power consumption in liquid-helium, and the clock frequency of this device was limited to value below 1 GHz due to a deterioration of the input sensitivity of the comparator. The purpose of our work was to develop high-performance output interface circuits that would be able to transfer the voltage signal from Josephson circuits to semiconductor circuits at the speeds that current GHz clocks require.

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00097-0

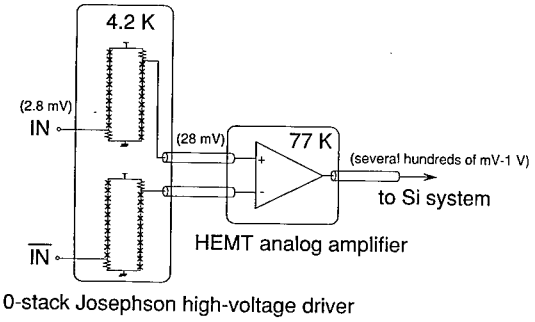


Fig. 1. Concept of the output interface circuit we propose, which provides an interface between the output of a Josephson IC to a semiconductor IC.

2. Interface configuration

It is difficult to amplify the 2.8 mV output voltage signal from a Josephson IC when using only Josephson devices or semiconductor transistors that run at a clock rate comparable to that for Josephson circuits (1–3 GHz [6]). For this reason, we propose the output interface circuit illustrated in Fig. 1. The circuit combines 4.2 K Josephson high-voltage drivers [4] with a 77 K analog HEMT monolithic amplifier. A Josephson driver with 10 stacks boosts the signal voltage from 2.8 to 28 mV, and the HEMT amplifier converts the voltage from 28 mV to nearly 1 V, which is sufficient for driving semiconductor devices at room temperature. This way, we gain both the features of high voltage-gain and of high frequency operation. HEMT device technology was employed for its superior high-speed performance. We choose the operational temperature of the interface circuits of 77 K (i)to suppress power-consumption in liquid-helium, (ii)to prevent direct thermal flow from room-temperature to the liquid-helium environment through signal cables, (iii)to enhance HEMT performance, and (iv)in anticipation of future 77 K superconducting electronics.

We used a system of InGaP/InGaAs/GaAs pseudomorphic materials for HEMT since such systems are free of DX centers and show no instabilities at 77 K.

To construct 77 K analog HEMT amplifiers, we employed the basic wide-band circuit configuration shown in Fig. 2. It includes a differential amplifier with complementary input ports and 50 Ω termination resistors as a first stage, 4-stage high-gain single-ended amplifiers, and an output buffer. A differential amplifier with complementary inputs was employed to cancel out the fluctuations in the output signal voltage that would be caused by ground-level voltage fluctuations in AC driven Josephson circuits.

3. Circuit fabrication

Fig. 3 shows a cross section of the InGaP/InGaAs/GaAs HEMT amplifier we propose. The InGaP/InGaAs/GaAs epitaxial layers are grown by OMVPE. The heterostructures consist of an undoped GaAs buffer, an undoped $In_{0.2}Ga_{0.8}As$ channel, a Si-doped *n*-$In_XGa_{1-X}P$ layers and a Si-doped *n*-GaAs cap layer. The cap layer contains a layer that acts as a stopper for the recess etching process so as to obtain a threshold voltage difference for depletion-mode and enhancement-mode devices. The doping densities in the *n*-InGaP layer and in *n*-GaAs are both 2×10^{18} cm^{-3}. The device fabrication process is based on the selective dry recess etching technique developed at our laboratories [7]. It consists of device isolation by oxygen ion implantation, AuGe/Ni/Au source and drain metalization, 0.5 μm Al gate metalization, WSiN resistor layer formation, SiON insulation layer formation, and Ti/Au interconnection layer metalization. The Al-gate metal layer was used as the first interconnection layer as well. An *n*-GaAs cap layer in the gate region is removed using selective dry etching before gate metal

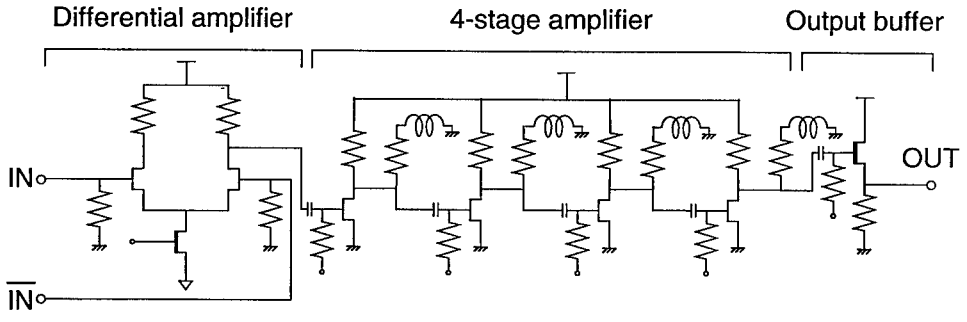


Fig. 2. Schematic view of the HEMT analog amplifier that is to be used for the Josephson-semiconductor interface.

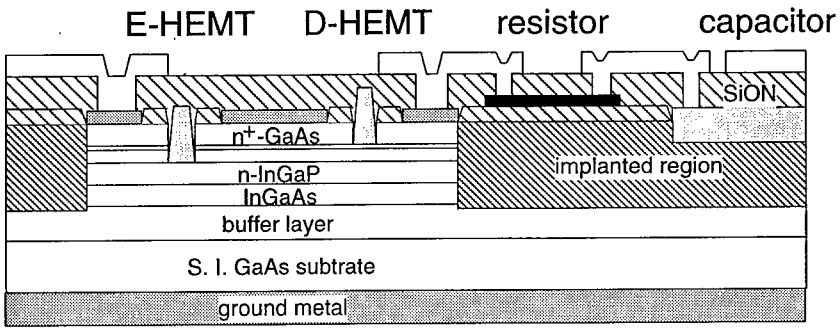


Fig. 3. Cross section of the HEMT amplifier.

formation is initiated. Capacitors are made by sandwiching the SiON layer as dielectric material between the first interconnection layer of Al and the second layer of Ti/Au.

A photograph of the chip is shown in Fig. 4. Total chip size is 1.75 × 6.375 mm.

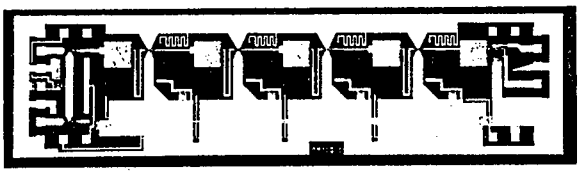


Fig. 4. Photograph of the chip for the monolithic HEMT analog amplifier.

4. Device and circuit performance

We measured the DC and RF performance for 100-μm wide HEMTs. Fig. 5 shows drain I–V characteristics of an E-mode HEMT at room-temperature and at 77 K. The sample has excellent I–V properties without showing any unstable behavior such as the collapse of current which often occurs in AlGaAs/GaAs HEMTs. The *S*-parameters of the E-mode HEMT are measured in the range from 0.2 to 20.2 GHz with a cryogenic wafer prober. Fig. 6 shows how the current gain cutoff frequency (f_T) depends on the drain current for a drain voltage of 1 V. Maximum f_T was 41 GHz, which is about 20% higher than at room-temperature. Cutoff frequencies higher than 40 GHz were maintained over the wide drain current range of 12–24 mA.

Table 1 summarizes the performance of the HEMT. InGaP/InGaAs HEMTs usually have a larger transconductance than AlGaAs/InGaAs HEMTs, since their *n*-InGaP supply layer can be thinner while maintaining the same threshold voltage as compared to an *n*-AlGaAs HEMT. This results in a high voltage gain

Table 1
Device performance of a 0.5 μm-gate InGap/InGaAs E-mode HEMT

	77 K	300 K
Transconductance	534 mS/mm	432 mS/mm
Transconductance factor	1.6 S/Vmm	0.67 S/Vmm
Threshold voltage	0.06 V	−0.10 V
Threshold voltage uniformity σ		10 mV

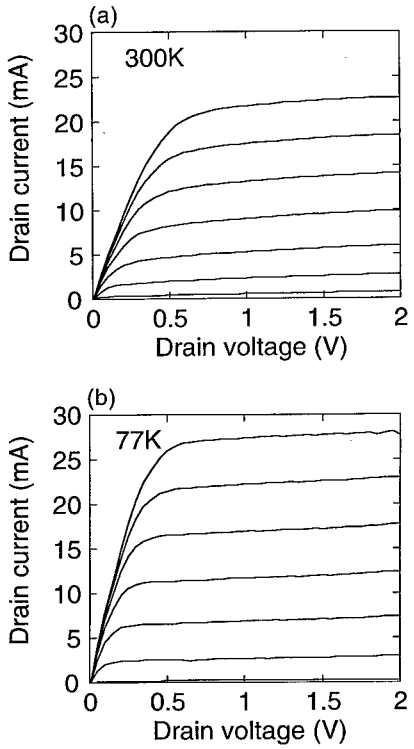


Fig. 5. Drain I–V characteristics of an 100 μm wide E-mode InGaP/InGaAs HEMT. The gate-to-source voltage is increased in 0.1 V steps up to a maximum of 0.7 V. (a) at room-temperature (b) at 77 K.

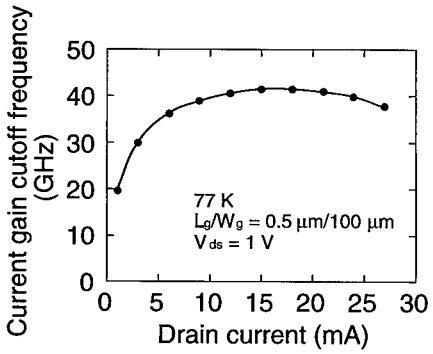


Fig. 6. Dependency of the current gain cutoff frequency on the drain current for a 100 μm wide E-mode HEMT at 77 K.

of InGaP/InGaAs HEMT amplifiers. The standard deviation of the room-temperature threshold voltage on a 3″ wafer is as low as 10 mV, which demonstrates the excellent structural uniformity provided by our crystal growth technique and the dry recess etching process.

The amplifier was tested at 77 K using cryogenic wafer prober with microwave probes. The probes have a pair of signal pins that supported frequencies up to 40 GHz, and some DC bias pins. The bias pins were grounded at the base of the needles through 2200 pF capacitors and 50 Ω resistors in order to prevent a fluctuation of DC bias voltage. The amplifier responds in a linear fashion over an input voltage range from 0.4 to 30 $mV_{p\text{-}p}$, and output saturated at about 0.7 $V_{p\text{-}p}$. The voltage gain of the amplifier at 3 GHz, as calculated in the linear region, was 23 dB. The output waveform response for high-frequency operation was also tested. We added complementary input signals which simulate outputs of 10-stack Josephson high-voltage drivers to the two input ports. Input was Return-to-Zero (RZ) format, with 3 Gbit/s and 2^{23}–1 PRBS, and its amplitude was 30 $mV_{p\text{-}p}$. As shown in Fig. 7, the output amplitude reaches 0.7 $V_{p\text{-}p}$, which makes it possible to drive room-temperature GaAs DCFL gates and/or future scaled CMOS gates.

5. Examination of an interface circuit

The operation of the overall interface circuit was examined by connecting the 77 K HEMT amplifier to liquid-helium cooled 10-stack Josephson high-voltage drivers. The driver was fabricated using Nb/AlOx/Nb junction technology. The experimental setup is illustrated in Fig. 8. Complementary input signals generated by a room-temperature pattern generator were applied to two Josephson chips. The output signal of the drivers were sent to two input terminals of the HEMT amplifier, which was kept at 77 K, via coaxial cables which were at room temperature. The repetition frequency was 300 MHz, which was limited by cross-

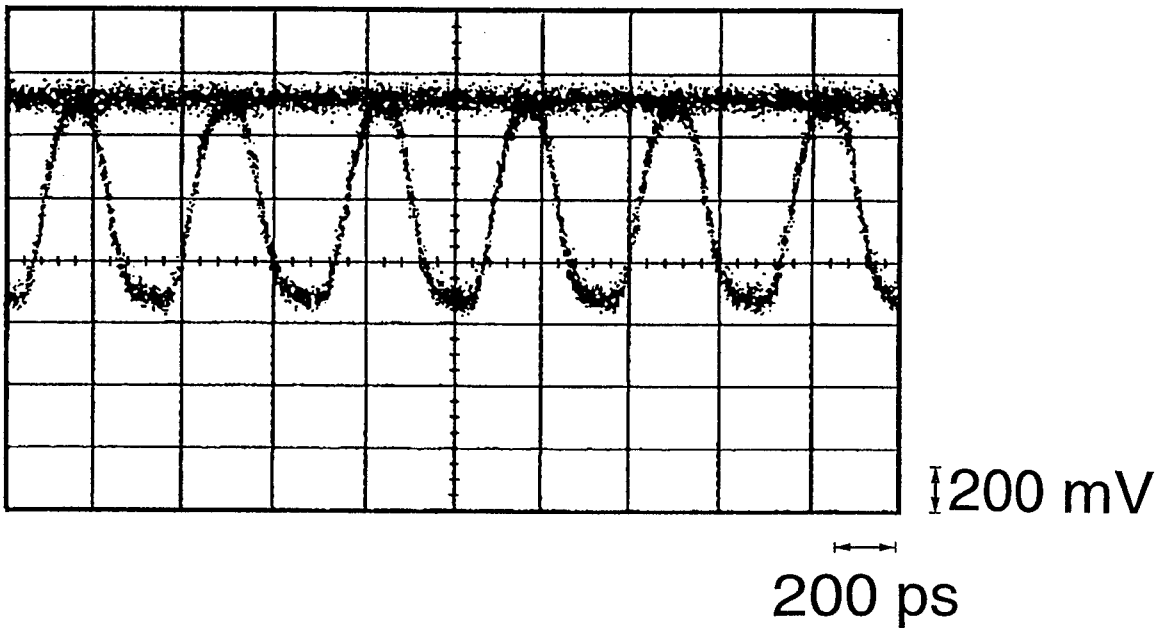


Fig. 7. Output waveform of the HEMT amplifier at 3 Gbit/s. The input signal of 30 $mV_{p\text{-}p}$ is applied to a single port in differential inputs.

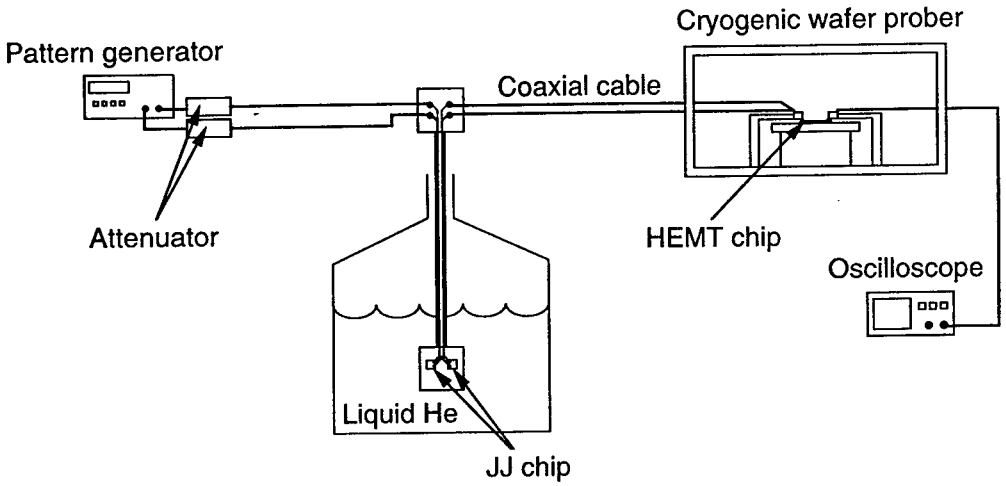


Fig. 8. Experimental setup for testing the operation of the entire interface.

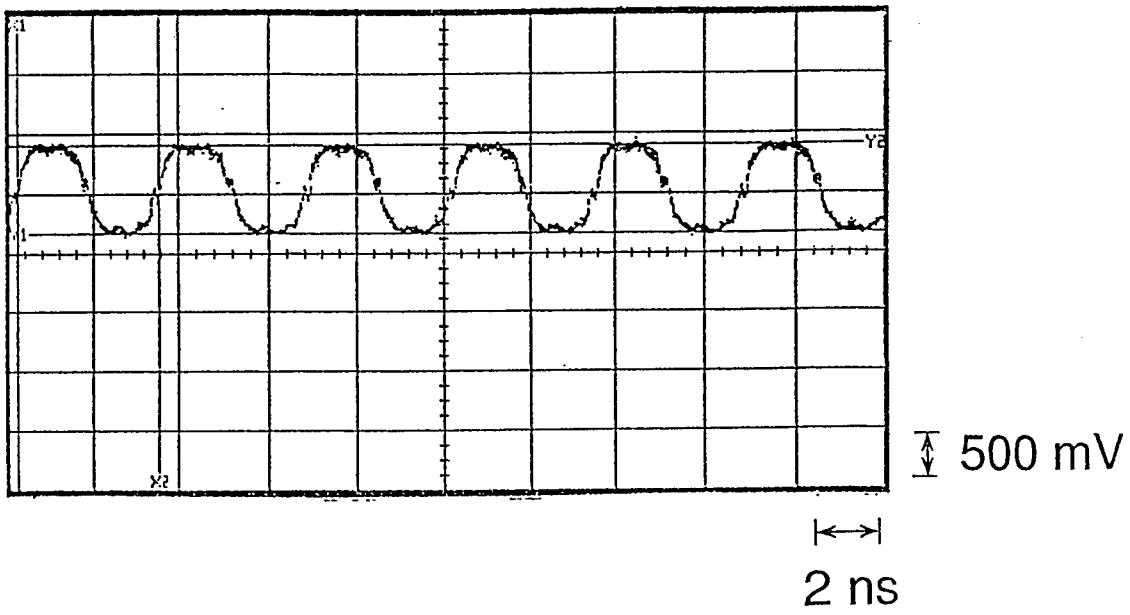


Fig. 9. Output waveform of the interface system illustrated in Fig. 8 at room temperature.

talk between the signal lines in a chip holder which was kept in liquid He. Fig. 9 shows the waveform of the output. The amplitude of the output signal reached 0.7 $V_{p\text{-}p}$. Since we confirmed by circuit simulations that Josephson high-voltage drivers operate at several gigaherz, we conclude that our interface circuit can indeed transfer a high-speed voltage signal from Josephson circuits to a device at room temperature with an amplitude of 0.7 $V_{p\text{-}p}$.

6. Conclusion

We were the first to develop high-speed interface circuit that is a combination of 4.2 K Josephson high-voltage drivers and a 77 K analog HEMT amplifier. We employed a basic wide-band circuit configuration for the 77 K analog HEMT amplifier, which consists of a differential amplifier with complementary input and high-gain single-ended amplifiers. We successfully

demonstrated that this circuit can transfer the voltage signal from a 10-stack Josephson driver to a room-temperature device at an amplitude of 0.7 $V_{p\text{-}p}$.

Acknowledgements

The authors would like to thank Drs. H. Nishi, M. Takikawa, S. Kuroda, and H. Tanaka for their encouragement and support. They are especially grateful to Dr H. Suzuki, Dr K. Imamura and Mr A. Watanabe for their valuable suggestions. They are also grateful to Dr H. Suehiro, Mr Y. Yamaguchi, Mr T. Kikkawa, Mr Y. Matsukura, Mr M. Nihei, and Mr N. Ohnuki for epitaxial wafer growth and supports of device fabrication.

This work was performed under the management of FED as a part of the MITI R&D program (Josephson Device Hybrid System Technologies Project) supported by NEDO.

References

[1] Feldman AR, van Duzer T. IEEE Trans Appl Supercond 1995;5:2648.

[2] Kishore SV, Goshal U, Huynh L, van Duzer T. IEEE Trans Appl Supercond 1995;5:2644.

[3] Tahara S, Yorozu S, Hashimoto Y, Numata H, Nagasawa S. Extended Abstracts of the 5th International Workshop on High-Temperature Superconducting Electron Devices, 1997;29.

[4] Suzuki H, Imamura T, Hasuo S. Cryogenics 1990;30:1005.

[5] Ghoshal U, Kishore SV, Feldman AR, Huynh L, Van Duzer T. IEEE Trans Appl Supercond 1995;5:2640.

[6] Hioe W, Hosoya M, Kominami S, Nagaishi H, Nishino T. IEEE Trans Appl Supercond 1997;7:2484.

[7] Kuroda S, Mimura T, Suzuki M, Kobayashi N, Nishiuchi K, Shibatomi A, Abe M Proceedings of The IEEE Gallium Arsenide Integrated Circuit Symposium, 1984. p. 125.



PERGAMON

Solid-State Electronics 43 (1999) 1519–1526

SOLID-STATE ELECTRONICS

Improvement of 0.1 μm-gate InGaAs/AlGaAs HEMT performance by suppression of electro-chemical etching in deionized water

Tomoyuki Ohshima*, Ryoji Shigemasa, Masakatsu Sato, Masanori Tsunotani, Tamotsu Kimura

Semiconductor Technology Laboratory, Oki Electric Industry Co Ltd, 550-5 Higashiasakawa, Hachioji, Tokyo, 193-8550, Japan

Abstract

The recess gate region of 0.1 μm-gate InGaAs/AlGaAs HEMT is anomalously etched during the deionized water rinsing. When the Au-based ohmic electrodes and GaAs on recess gate region are simultaneously exposed to deionized water, the electro-chemical etching occurs resulting in a non-flat recessed gate profile. Also, the U-shaped etched groove is formed besides the Al/Ti gate after the lift-off process due to the electro-chemical etching effect between the Al/Ti and GaAs. Such an electro-chemical etching effect is strongly related to dissolved oxygen in deionized water, which causes OH^- ion by the electro-chemical reaction with H_2O. The reduction of dissolved oxygen in deionized water suppresses the electro-chemical etching phenomenon. Using the deionized water with reduced dissolved oxygen, device performance of the 0.1 μm-gate InGaAs/AlGaAs HEMT was improved and the current gain cut-off frequency as high as 104 GHz has been obtained. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Recently, the demand on high speed devices operating at mm-wave frequencies has increased because of the application of such devices in high capacity optical fiber communication systems and the mm-wave wireless communication systems. Pseudomorphic InGaAs/AlGaAs high electron mobility transistors (HEMTs) are the most promising devices for these applications because of their excellent high frequency performance [1].

In the fabrication of the monolithic microwave integrated circuits (MMICs) based on such HEMTs, one of the most important fabrication steps is the gate recess etching process. In order to obtain the designed device characteristics with a good uniformity, the recessed depth and shape must be accurately controlled [2,3].

Recently, the electro-chemical etching effect during the recess etching process has been reported to produce significant non-uniformity in the recessed gate profile [4,5]. In the recess etching process of HEMTs, the electro-chemical potential is established in the etching solution between the n-type GaAs on the gate region and other materials such a Au on the ohmic electrode.

It has been also reported that the recessed shape of InAlAs/InGaAs HEMT is strongly affected by the ohmic metal structure exposed to the etching solution because the electro-chemical potential depends on the chemical nature of the material [6]. Therefore, in order to suppress the electro-chemical etching effect during

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00098-2

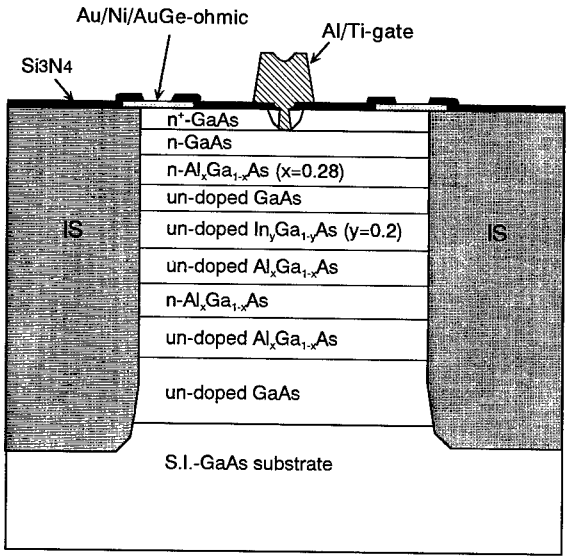


Fig. 1. Cross-section of the pseudomorphic InGaAs/AlGaAs HEMT.

the recess etching process, only the gate region to be recessed should be exposed to the etching solution. Other regions, including the ohmic electrodes, must be covered.

However, we found that the electro-chemical etching effect also occurs during the deionized water rinse process, since the deionized water acts as an electrolytic solution. As a result the gate recess region is unintentionally etched even before the recess etching [7,8]. This electro-chemical etching in deionized water results in the non-uniform recessed gate depth and shape.

This paper describes the mechanism of the electro-chemical etching in deionized water, which seriously affects the device performance.

2. Fabrication process

Fig. 1 shows the cross-section of the 0.1 μm-gate pseudomorphic InGaAs/AlGaAs HEMT used in this study. The HEMT structure consists of a 100 nm GaAs/100 nm AlGaAs buffers, 10 nm AlGaAs doped layer (3×10^{18} cm^{-3}), 3 nm AlGaAs spacer, 10 nm InGaAs channel, 5 nm GaAs spacer, 26 nm AlGaAs doped layer (3×10^{18} cm^{-3}), 50 nm GaAs Schottky layer (3×10^{17} cm^{-3}) and 50 nm GaAs cap layer (4×10^{18} cm^{-3}). All the layers were grown by molecular beam epitaxy (MBE) on a semi-insulating GaAs substrate.

Fig. 2 shows the fabrication process of the HEMT. After the growth of the heterostructure by MBE, device region was isolated by the oxygen ion implantation at the acceleration energies of 80 and 160 keV. Then, Au/Ni/AuGe ohmic metals for the source and the drain electrodes were deposited and alloyed at 420°C. Next, the silicon nitride (Si_3N_4) film of 0.1 μm-thick was deposited by plasma enhanced chemical vapor deposition (P-CVD) and the 0.1 μm resist spacing for the gate pattern was formed using electron beam lithography. Si_3N_4 was etched by reactive ion etching (RIE) using SF_6 as an etching gas to form a 0.1 μm gate opening and contact pads on the ohmic electrodes. Next, the resist pattern for the top portion of the T-shaped gate was formed using i-line lithography.

The recess etching was performed by phosphoric acid (H_3PO_4) and hydrogen peroxide (H_2O_2) based etchant in order to control the threshold voltage of the device using the Si_3N_4 opening as a mask. During the recess etching, only the gate recess region was exposed to the etchant and the other regions including the ohmic electrodes were covered by the photo-resist. After the recess etching, Al/Ti gate electrodes were evaporated and lifted off using organic solvents.

In order to observe the spatial recessed gate profiles, atomic force microscopy (AFM) was used. After the gate formation, transmission electron microscopy (TEM) was used to investigate the recessed shape hidden by the T-shaped gate electrode.

3. Electro-chemical etching phenomenon in deionized water

After the recess etching process, we found a large difference in the recessed gate profiles measured by AFM between the devices with and without the Si_3N_4 opening on ohmic electrodes. As shown in Fig. 3, a flat profile was obtained for the device without the Si_3N_4 opening on the ohmic electrode. On the other hand, the boundary between the channel region and the isolated region was anomalously etched and a non-flat profile was obtained for the device with the Si_3N_4 opening on the ohmic electrode.

Because the ohmic electrodes were covered by photo-resist during the recess etching process, there was no difference in the pattern geometry between those devices. This fact suggests that the non-flat profile of the recessed gate region for the device with the Si_3N_4 opening on the ohmic electrode was formed before the recess etching process.

The only wet process performed before the recess etching is water rinsing for 5 min after the resist removal process using organic solvents, which was followed by RIE of Si_3N_4 film to form the gate opening and the contact pads on ohmic electrodes. The water rinse is required to effectively remove the contamination caused by RIE such as a sulfur, and to wash away the etching residue such as a polymer. It is quite

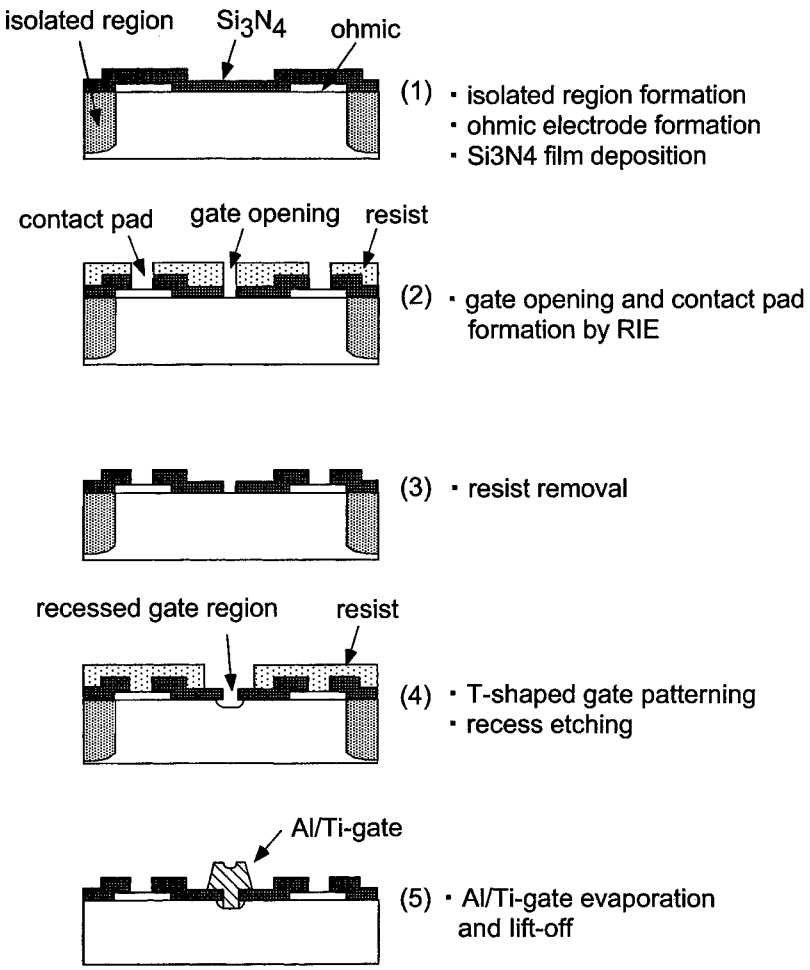


Fig. 2. Fabrication process of the 0.1 μm-gate pseudomorphic InGaAs/AlGaAs HEMT.

important to obtain a clean GaAs surface by water rinsing in order to perform the uniform recess etching. The deionized water used had the resistivity of 18 MΩ cm and the dissolved oxygen concentration, incorporated from air, of 8 ppm.

Fig. 4 shows the profile of the gate region after the water rinsing for 5 min followed by RIE. It was found that the non-flat recessed profile was already formed even before the recess etching. The etching rate by the H_3PO_4 based etchant was uniform for all the regions, and the anomalous profile itself did not change before and after the recess etching, as shown in Fig. 4. We have also confirmed that the non-flat profile was not formed only by the organic solvent treatment used for the resist removal.

Therefore, it is clear that GaAs on the gate region was anomalously etched during the deionized water rinsing. Since the non-flat recessed profile was observed only for the device with the Si_3N_4 opening on the ohmic electrode, this phenomenon can be understood as the electro-chemical effect in deionized water for the ohmic electrode and GaAs [9]. During the deionized water rinsing after the RIE process, the ohmic electrode and GaAs are simultaneously exposed to deionized water.

This electro-chemical etching phenomenon is thought to be related to the electrical properties of GaAs because the isolated region was more deeply etched compared with the channel region. Then, the dependence of the etching depth of the isolated region after the deionized water rinse on the condition of the oxygen ion implantation was investigated. Fig. 5 shows the dependence of the etching depth of the isolated region and the leakage current measured between two ohmic electrodes of 100 μm wide and 2 μm gap on the oxygen ion implantation dose. The deionized water rinse was performed for 5 min.

It was found that the leakage current was minimum

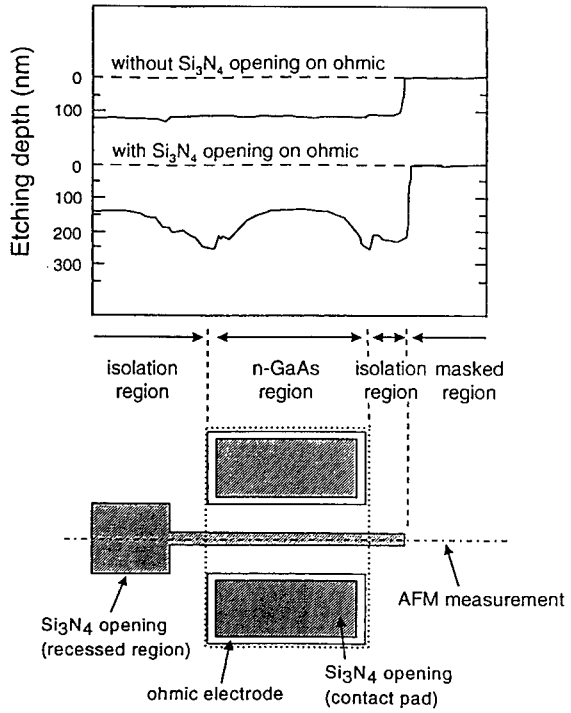


Fig. 3. Recessed gate profiles measured by AFM for the HEMTs with and without Si_3N_4 opening on ohmic electrodes. The water rinse after RIE of Si_3N_4 was performed for 5 min.

at the dose of $1–2 \times 10^{12}$ cm^{-2}, which corresponds to the maximum of the etching depth of the isolated region. The isolated GaAs by the oxygen ion implantation behaves like p-type. Therefore, this result indicates that the electro-chemical etching in deionized water is enhanced by holes in GaAs.

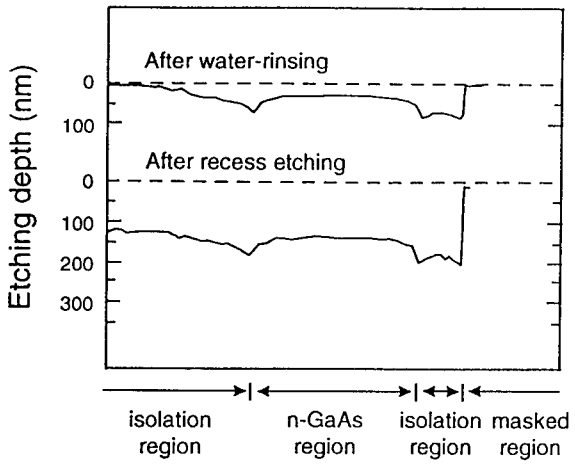


Fig. 4. Profiles of gate recess region after deionized water rinsing and recess etching. The water rinse time is 5 min.

Based on the results mentioned above, the etching mechanism of GaAs in deionized water is graphically shown in Fig. 6. Due to the cathode reaction on the ohmic electrode, OH^- ions are generated from the reaction between the deionized water (H_2O) and the dissolved oxygen (O_2). The OH^- ions generated on the ohmic electrode quickly move to GaAs surface regardless of the high resistivity of deionized water, because OH^- ions in water can move by the exchange of an electron through the hydrogen bonds between H_2O molecules. Therefore, at the same time with the cathode reaction, the anode reaction on GaAs surface with OH^- ions, oxidation and etching, occurs. Because holes in GaAs enhance the anode reaction, the isolated region, which behaves like a p-type region, is deeply etched. As shown in Fig. 6, the etching current decreases as the distance from the ohmic electrode becomes large because of the high resistivity of the isolated GaAs. Therefore the etching occurs near the boundary between the channel region and the isolated region, resulting in the non-flat profile.

The electro-chemical etching effect in deionized water can be understood in terms of the existence of OH^- ions based on the model presented in Fig. 6. In order to establish the influence of OH^- ions on the electro-chemical etching of the isolated region, the etching current measurement was performed for the samples exposed to the diluted solution of ammonia (NH_3) and hydrochloric acid (HCl) with various values of pH. The measurement configuration is shown in Fig. 7. The sample was the oxygen ion implanted n-type GaAs. The ohmic electrode was formed on the back side of the sample. The isolated GaAs surface was exposed to the pH controlled diluted solution. When the electro-chemical etching occurs, the electro-chemical potential is established between the counter electrode made of carbon and the GaAs, and the etching current flows through the external circuit. Fig. 8 shows the dependence of the etching current on pH of the diluted solution. As shown in Fig. 8, the etching rate was negligibly small for $pH < 7$, but becomes larger as pH increases. This result clearly indicates that OH^- ions are strongly related to the electro-chemical etching effect of GaAs. It also supports the model shown in Fig. 6.

Such an electro-chemical etching effect also occurs during the water rinsing in the gate electrode formation process. Fig. 9(a) shows the cross-sectional TEM image of the recessed region around the Al/Ti gate after the lift-off. During the Al/Ti lift-off process using organic solvents, deionized water rinse was performed for 5 min. From Fig. 9(a), it was found that GaAs of the recessed region beside the Al/Ti gate was anomalously etched and the U-shaped groove was formed. This anomalous etching can be also understood as the electro-chemical etching phenomenon

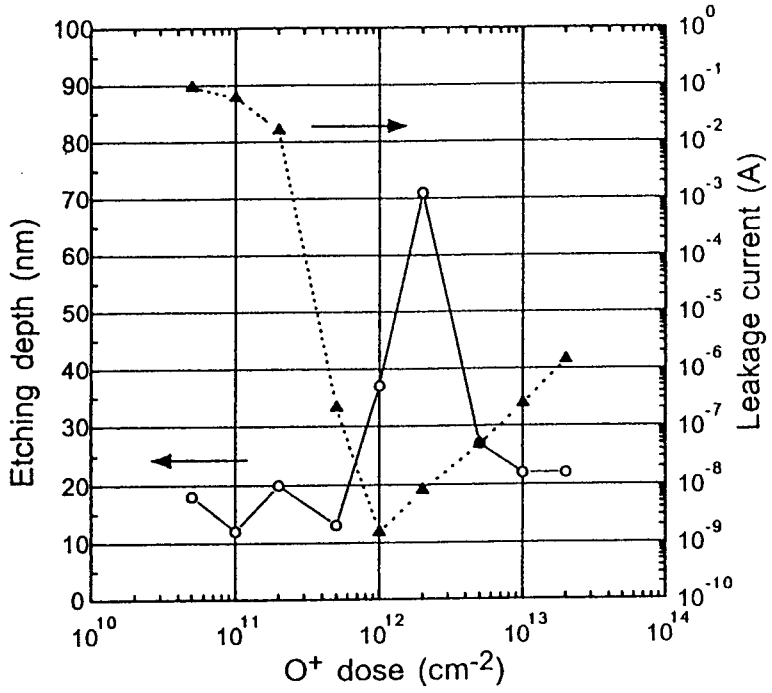


Fig. 5. Dependence of etching depth and leakage current on oxygen ion implantation dose. The water rinse was performed for 5 min.

between the Al/Ti gate and GaAs because both the materials were simultaneously exposed to deionized water during the lift-off process.

As mentioned previously, the electro-chemical etching phenomenon in deionized water strongly affects the recessed shape of the InGaAs/AlGaAs pseudomorphic HEMT. In order to obtain the designed device charac-

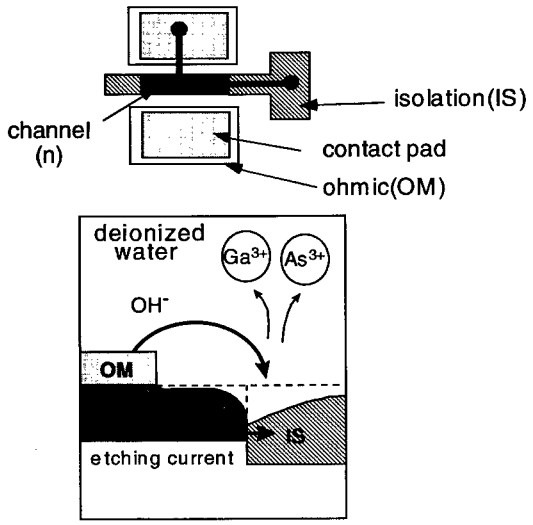
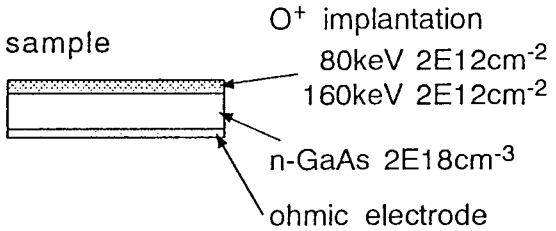


Fig. 6. Model for the electro-chemical etching effect.



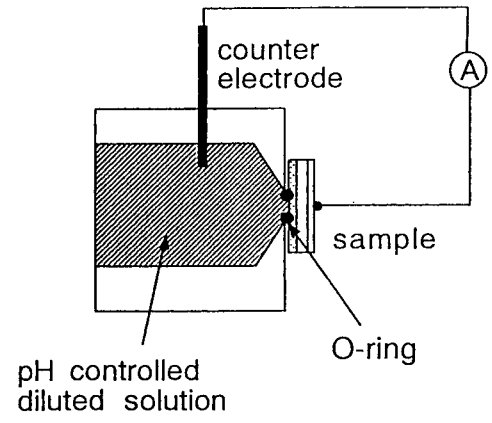


Fig. 7. Measurement configuration of etching current.

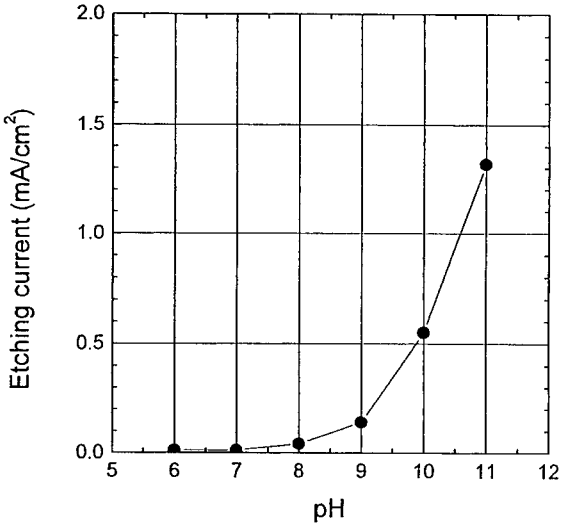


Fig. 8. Dependence of etching current on pH.

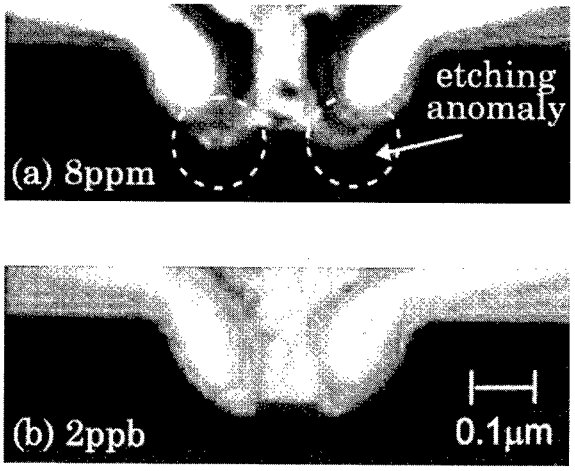


Fig. 9. TEM images of recessed region around the Al/Ti gate after deionized water rinse with dissolved oxygen of (a) 8 ppm and (b) 2 ppb. The water rinse time is 5 min.

teristics, such an electro-chemical etching phenomenon must be suppressed.

4. Suppression of electro-chemical etching effect

Based on the model shown in Fig. 6, the electro-chemical etching effect can be suppressed by reducing the dissolved oxygen in deionized water which causes OH^- ion. To confirm the effect of dissolved oxygen in deionized water on the electro-chemical etching effect, the deionized water with reduced dissolved oxygen was used for the water rinse processes.

The membrane degassing unit was used in order to reduce the dissolved oxygen in deionized water. The dissolved gases including oxygen in deionized water were exhausted by a vacuum pump through the membrane. The concentration of dissolved oxygen in deionized water was reduced and controlled to less than 1 ppb using the unit. Then, the rinse process was performed in N_2 ambiance to prevent incorporation of oxygen from air.

Fig. 10 shows the dependence of the etching depth at the isolated region on the concentration of the dissolved oxygen in deionized water after the RIE of Si_3N_4 film. The water rinse was performed for 5 min. As shown in Fig. 10, the etching depth decreases as the concentration of the dissolved oxygen decreases. The dependence in Fig. 10 is quite similar to that of the etching current on pH in Fig. 8. This result clearly shows that the generation of OH^- ions in deionized water can be suppressed by the reduction of dissolved oxygen.

Fig. 11 shows the profile of the gate recess region after the rinse for 5 min in two different deionized water with the amount of dissolved oxygen of 8 ppm and 2 ppb. As shown in Fig. 11, the electro-chemical etching was successfully suppressed and a flat profile was obtained. The anomalous etching beside the Al/Ti gate after lift-off process shown in Fig. 9(a) can be also suppressed by the deionized water with dissolved oxygen of 2 ppb as shown in Fig. 9(b), resulting in an identical recessed profile.

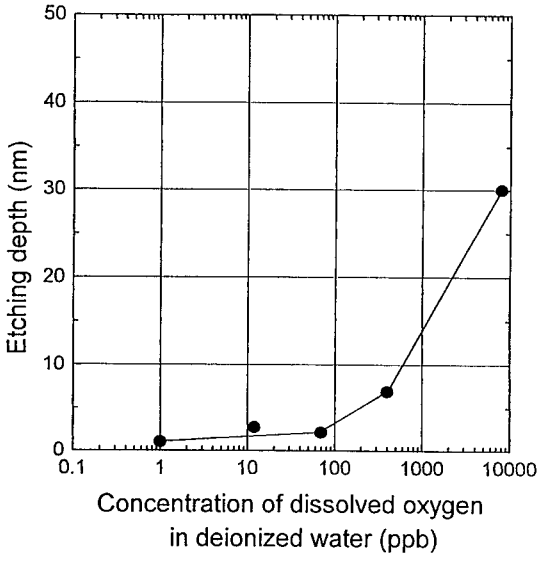


Fig. 10. Dependence of etching depth on the concentration of dissolved oxygen in deionized water. The water rinse was performed for 5 min.

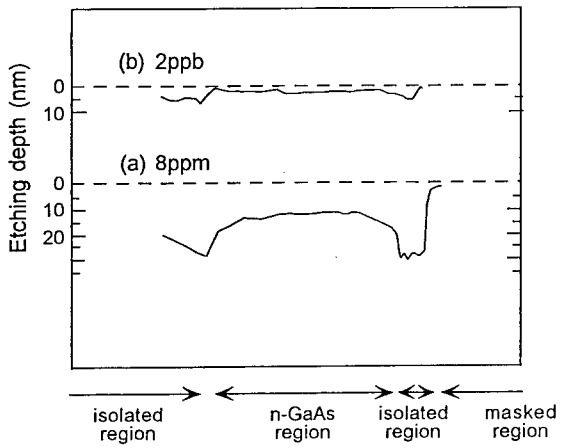


Fig. 11. Profiles of gate recess region after deionized water rinse with dissolved oxygen of (a) 8 ppm and (b) 2 ppb. The water rinse time is 5 min.

5. Device performance

We applied the process using the deionized water with reduced dissolved oxygen for the fabrication of the 0.1 μm-gate pseudomorphic HEMT.

Fig. 12 shows the current gain cut-off frequencies (f_T) extracted from S-parameter measurements for the HEMTs with 100 μm-wide fabricated by the process using deionized water with dissolved oxygen of 2 ppb and 8 ppm. The threshold voltages of those devices were almost the same (−800 mV). The measurement conditions were the drain voltage of 2 V and the drain current of 40 mA. As shown in Fig. 12, by reducing the dissolved oxygen in deionized water, f_T of as high as 104 GHz was obtained, which was an increase of about 20% over that of the degraded device due to the electro-chemical etching. From the small signal equivalent circuit extraction by fitting the S-parameters, it was found that the improvement resulted from the reduction of the source-to-gate capacitance (C_{gs}), and it was 91 fF for the improved device, while 122 fF for the degraded device. The anomalous recessed gate profile caused by the electro-chemical etching is believed to be related to the increase of C_{gs} for the degraded device.

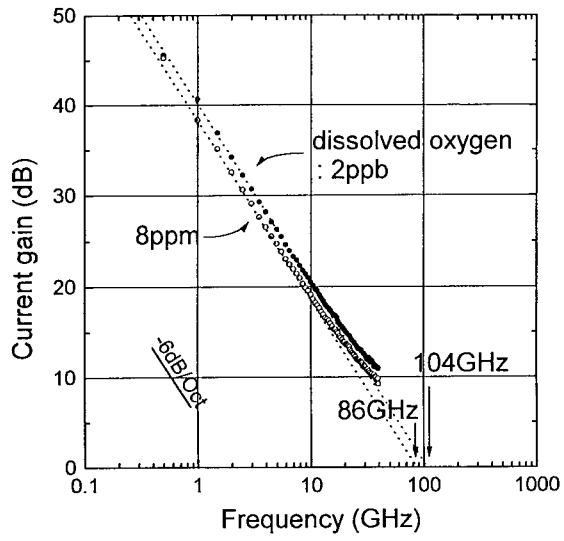


Fig. 12. Frequency characteristics for the HEMTs rinsed in deionized water with dissolved oxygen of 8 ppm and 2 ppb.

6. Conclusion

It was found that the recess gate region of 0.1 μm-gate InGaAs/AlGaAs HEMT was anomalously etched during the deionized water rinse process. When the Au-based ohmic electrodes and GaAs on recess gate region are simultaneously exposed to deionized water, the electro-chemical etching occurs. Due to the cathode reaction on the ohmic electrodes, OH^- ions are generated from the reaction of deionized water (H_2O) and dissolved oxygen (O_2). At the same time with the cathode reaction, the anode reaction on GaAs surface, oxidation and etching occurs, resulting in a non-flat recessed gate profile. Such an electro-chemical etching also occurred during the deionized water rinsing after the gate electrode formation, and the U-shaped etched groove was formed besides the Al/Ti gate.

By reducing the dissolved oxygen in deionized water during the rinse process, the electro-chemical etching phenomenon has been successfully suppressed and the controlled recessed gate profile was obtained.

Using the deionized water with reduced dissolved oxygen, the 0.1 μm-gate InGaAs/AlGaAs pseudomorphic HEMT was fabricated. The current gain cut-off frequency of as high as 104 GHz was obtained, which was an increase of about 20% over the degraded device due to the electro-chemical etching.

References

[1] Arai Y, Sato M, Yamada HT, Hamada T, Nagai K, Fujishiro HI. IEEE Trans Microwave Theory Tech 1997;45:2261.

[2] Tam M, King M, Wu CS, Sanna M, Edwards H, McGlothin R GaAs MANTECH Conference Digest 1997 p. 18.

[3] Ponti ES, Hu M, Brown JJ, Nguyen LD, Vajo JJ, Le MV, Wetzel MD GaAs MANTECH Conference Digest, 1997. p. 14.

[4] Metze GM, McPhilmy S, Laux P. IEEE Electron Dev Lett 1995;16:23.
[5] Kim CT, Lee KW, Chung KW GaAs MANTECH Conference Digest, 1997. p. 10.
[6] Xu D, Enoki T, Suemitsu T, Ishii Y. Proceedings of 10th International Conference on Indium Phosphide and Related Materials, 1998. p. 797.
[7] Nitta Y, Ohshima T, Shigemasa R, Nishi S, Kimura T Technical Digest IEDM, 1996. p. 47.
[8] Ohshima T, Shigemasa R, Nitta Y, Tsunotani M, Kimura T. Proceedngs of the 16th Electronic Materials Symposium, Japan, 1997. p. 241.
[9] Hagio M. J Electrochem Soc 1993;140:2402.



PERGAMON

Solid-State Electronics 43 (1999) 1527–1533

SOLID-STATE ELECTRONICS

Short gate-length InAlAs/InGaAs MODFETs with asymmetry gate-recess grooves: electrochemical fabrication and performance

D. Xu*, T. Suemitsu, H. Yokoyama, Y. Umeda, Y. Yamane, T. Enoki, Y. Ishii

System Electronics Laboratories, Nippon Telegraph and Telephone Corporation (NTT), 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa, 243-0198, Japan

Abstract

We report the realization of asymmetric recess etching in the fabrication of gate grooves for MODFETs based on InAlAs/InGaAs heterojunctions. The novel technology employs asymmetric electrochemical etching when the surface metal of Pt and Ni are deposited on the source and drain pads, respectively. Devices fabricated with this technology, though still immature, show similar DC performance to those fabricated by conventional technology for symmetric devices. Furthermore, due to the formation of a 'step' on the InP etch-stopper as a result of the asymmetric etching, the effective gate length of the devices has been made shorter than its nominal value. This, combined with the asymmetric groove profiles, produces higher f_T and f_{max} values than the usual symmetric devices. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Electrochemical process; Gate recess; Etching; InAlAs; InGaAs; Hetero-junction

1. Introduction

InAlAs/InGaAs heterojunction materials grown on InP substrates have made high performance possible for modulation-doped field-effect transistors (MODFETs) in high-speed and high-frequency applications [1,2]. Device performance can be further improved by employing asymmetrically recessed gates, whose side etching is small on the source side and large on the drain side. The benefits include, without obvious simultaneous deterioration of the parasitic resistance R_{sd}, the reduction of the excessively high output conductance g_o [3], which is typical for MODFETs based on InAlAs/InGaAs heterostructures lattice-matched to InP substrates, and the reduction of feedback capacitance [4]. All these will in turn lead to an enhancement of the gain of devices. If the side etching on the drain side can be made sufficiently large, the gate-drain breakdown characteristics will also be improved [5], which are important for the power application of the MODFETs. With asymmetrically recessed gates, we can take full advantage of the superior transport properties and the resulting excellent performance of MODFETs without having to make a trade-off concerning the width of side etching.

* Corresponding author. University of Notre Dame, Department of Electrical Engineering, 200 Fitzpatrick Hall of Engineering, Notre Dame, IN 46556-5637, USA. Tel.: +1-219-631-8245; fax: +1-219-631-4393.

E-mail address: dxu@nd.edu (D. Xu)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00099-4

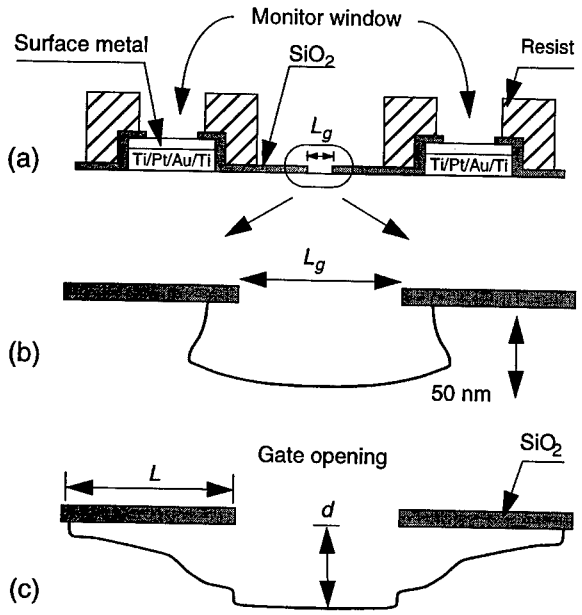


Fig. 1. Schematics of electrochemical etching in the wet-chemical gate recess for MODFET fabrication: device ready for recess etching (a), the gate-groove profile formed by 25-s etching when Ni surface metal is used (b), and the gate-groove profile formed by 160-s etching when Pt surface metal is used (c).

The standard approach to fabricating asymmetric gate grooves is to form asymmetrical resist profiles. Usually, this can be achieved either by the so-called 'double-recess' process [6] or by electron-beam lithography based on a four-layer stack of resist [7,8]. The former methodology needs two-step lithography, in which the first recess defines the wide groove, and the second defines the gate with offset towards the source. The multilayer–resist technology, on the other hand, requires the control of various parameters in electron-beam lithography, which can be difficult if applied for fabricating devices with gate lengths below 0.1 μm. In contrast, the new approach we will describe here is based on asymmetrical etching induced by electrochemical effects in the course of gate-recess etching for MODFET fabrication [9]. With only one-step electron-beam lithography based on single-layer resist, asymmetric gate grooves can be formed via a self-controlled process. Besides the simplification of processing procedures, this single-layer technology is feasible to fabricate devices that have gate lengths shorter than 0.1 μm with good reproducibility. Furthermore, as has already been demonstrated previously, the electrochemical effects are more pronounced in the recess etching when the dimensions of the gate openings are reduced and, as a result, this technology should be more suitable for short gate-length devices. Although the technology needs further improvement to attain a higher asymmetry ratio, it should be considered as a possible alternative to current technologies.

2. Asymmetrical etching in gate recess

2.1. Basics of asymmetric etching

Fig. 1 illustrates the basic points to induce an asymmetric etching in gate openings. One should notice from Fig. 1(a) that, in addition to the gate opening, the metal surfaces of the source and drain pads are also exposed during recess etching. This is the key to the asymmetric etching. Without the exposure of the ohmic pads, the employed citric-acid-based etchant gives an etching rate of 0.2 nm/s for both InAlAs and InGaAs. However, the electrode exposure will significantly modify the etching rate, because an electrochemical potential will be established between the electrodes and the semiconductors in the gate region. Figs. 1(b) and (c) show the gate grooves that are formed when the surface metals of the ohmic pads are Ni and Pt, respectively. The groove for the Ni sample extends at greatly enhanced rates of 10 and 4 nm/s vertically and laterally during the etching. The higher electrode potential of Pt surface metal, on the other hand, will lead to excessive oxidation and a consequent slower semiconductor etching than that when Ni is used.

The basic consideration in inducing asymmetric etching in the gate opening is to respectively deposit Pt and Ni on the source and drain electrodes. This is because these two metals have very different electrode potentials: 1.2 V for Pt and −0.25 V for Ni at 25°C [10]. The oxidation of semiconductors in the gate region is stronger on the source side because the electrode potential of Pt is higher than that of Ni. As a result of the subsequent wet-chemical etching, the gate-recess etching should be slower on the source side than the drain side in terms of both the vertical etching and side etching. In this way, the asymmetric gate groove can be formed with one-step etching.

2.2. Fabrication of asymmetric grooves

The MODFET epitaxial layers for this study were grown by metalorganic chemical vapor deposition (MOCVD) on 3″ wafers. They consist of a 200-nm *i*-InAlAs buffer, a 15-nm *i*-InGaAs channel, a 3 nm *i*-InAlAs spacer, a Si δ-doping, a 5-nm *i*-InAlAs layer, a 3-nm *i*-InP etching stopper, and a 3 nm *i*-InAlAs barrier. The cap layers, an n^+-InAlAs layer and an n^+-InGaAs layer, are for the realization of non-alloyed ohmic contact. All the layers are lattice-matched to InP substrates, and the doping densities of Si are

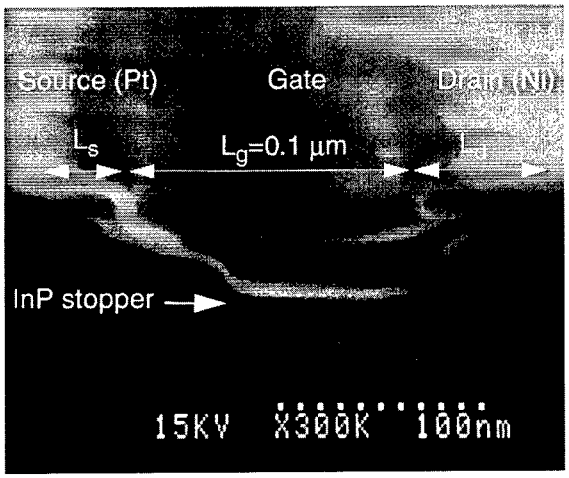


Fig. 2. SEM micrograph of an asymmetrical gate groove for a 0.1-μm MODFET that was fabricated using the principle described in the text. L_s and L_d denote the side etching on the source side and the drain side, respectively. The flat groove bottom results from using InP etch-stopper.

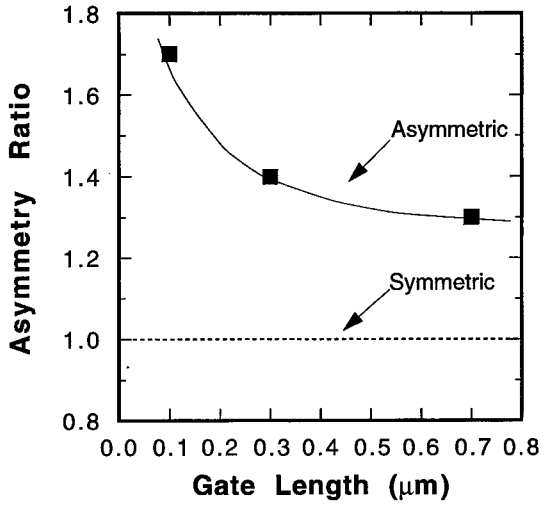


Fig. 3. The asymmetry ratio as a function of the gate length. The dashed line is for the standard conventional devices with symmetric gate grooves.

2×10^{19} and $6 \times 10^{18}/cm^3$ for the n^+-InGaAs and n^+-InAlAs layers, respectively. TLM measurements show that the sheet resistance of all layers and the ohmic contact resistance were around 90 Ω/square and 0.07 Ωmm, respectively.

Each chip (4.2 × 4.2 mm in size) has about 20 electron-beam-defined active devices, including devices for monitoring the drain current during wet-chemical recess etching. The devices for RF characterization have a T-finger layout with the unit gate-finger width of 50 μm. The mesa was fabricated with a non-selective citric-acid-based etchant. The InP etching stopper was etched with a mixture of acetic acid, phosphorus acid, salt acid, hydrogen peroxide, and DI water [11]. We employed Ti/Pt/Au/Ti, from bottom to top, as the base of the non-alloyed ohmic electrodes, on which Ni was firstly deposited. Afterwards, the photo resist was patterned so that only the source ohmic electrodes were opened, and Pt was deposited. Dielectric films were deposited across the wafer and the openings in the films were fabricated by reactive ion etching (RIE). After the definition of gate openings by electron-beam lithography and RIE, photo resist was coated to define the top of the T-gates, and, simultaneously, resist openings on ohmic electrodes were made in order to induce the electrochemical formation of asymmetric gate grooves and to facilitate the measurement of the drain current I_{ds} during the gate-groove fabrication. The recess etching was performed with a mixture of citric acid and hydrogen peroxide.

A SEM (scanning electron microscope) photo of the fabricated gate-groove asymmetry is shown in Fig. 2 [12]. The width of the side etching is defined as the width of the groove where the highest side etching occurs, i.e., in n^+-InGaAs. Accordingly, the widths of the side etching are 30 ($=L_s$) and 50 nm ($=L_d$) on the source and drain sides, producing an asymmetry ratio, defined as L_d/L_s, of around 1.7 for the 0.1-μm gate opening. The asymmetry ratio will decrease with a longer gate as shown in Fig. 3, and therefore this process is suitable for fabricating asymmetric gate grooves for short-gate devices. However, the asymmetry ratio achieved with this technology at the current stage is still not sufficiently high for improving gate-drain breakdown characteristics. This ratio should be increased by enlarging the electrode-potential difference between the semi-conductor and the surface metals of the source and drain electrodes. Possible methods include the use of other combinations of surface metals on the source and drain and the employment of other etchants for recess.

MODFETs were fabricated with the above-described technology. As a reference, another wafer was processed using the standard fabrication procedures for symmetric devices, in which the exposed surface metal is Ni on both the source and drain pads [2].

3. Device performance

3.1. DC performance

The I–V characteristics of 0.1-μm MODFETs with symmetric and asymmetric gate grooves are shown in Figs. 4(a) and (b), respectively. The R_{sd} for the asymmetric devices is found to be around 0.8 Ωmm, in com-

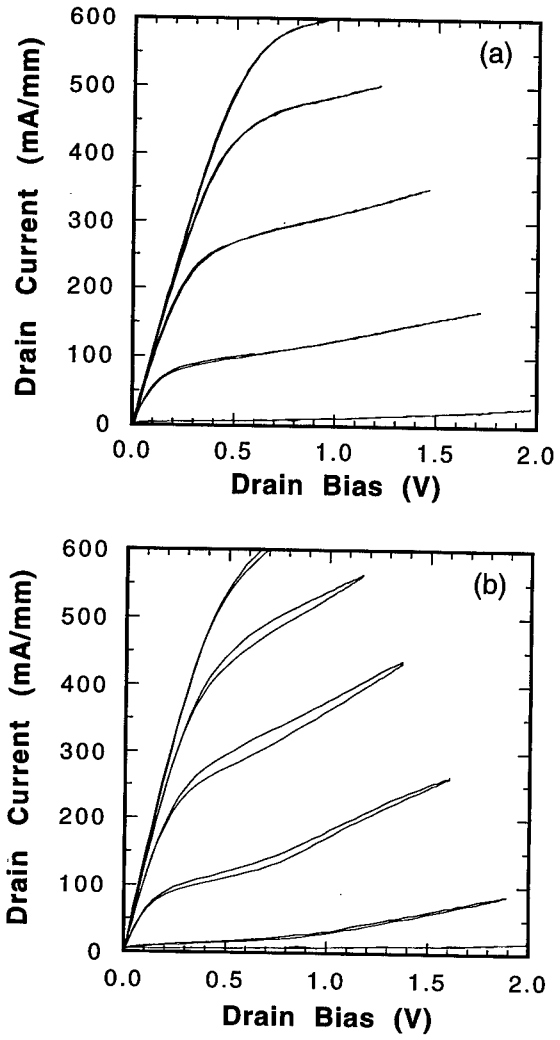


Fig. 4. I–V characteristics for 0.1-μm MODFETs with the conventional symmetric gate groove profile (a) and with the asymmetric gate groove profile shown in Fig. 2(b). Top curves: $V_{ds}=0.4$ V, $\Delta V_{gs}=-0.2$ V.

parison with 0.9 Ωmm for the symmetric devices. This R_{sd} difference is partly responsible for the 10% higher maximum current I_{dss}, defined as the I_{ds} at gate bias $V_{gs}=0.4$ V and drain bias $V_{ds}=1$ V, for the asymmetric devices. As shown in Fig. 5, the asymmetric devices also exhibit about a 6% higher maximum transconductance g_m than their symmetric counterparts for gate lengths ranging from 0.07 to 0.1 μm at $V_{ds}=1$ V; the difference in the parasitic resistance is considered to the main reason for this g_m difference as the gate-to-channel distance should be identical due to the use of the InP etch-stopper in the fabrication.

One may notice the hysteresis in Fig 4(b). We think this should be related to the use of Pt during the formation of the asymmetric grooves: the Pt-induced oxidation in the semiconductors of the gate region is so strong (due to its high electrode potential as mentioned in section 2.1) that a thin layer of oxide may still remain even after recess etching. An indirect evidence for this argument is Fig. 4(a), which doesn't show hysteresis in I–V characteristics due to the only use of Ni as ohmic surface metal.

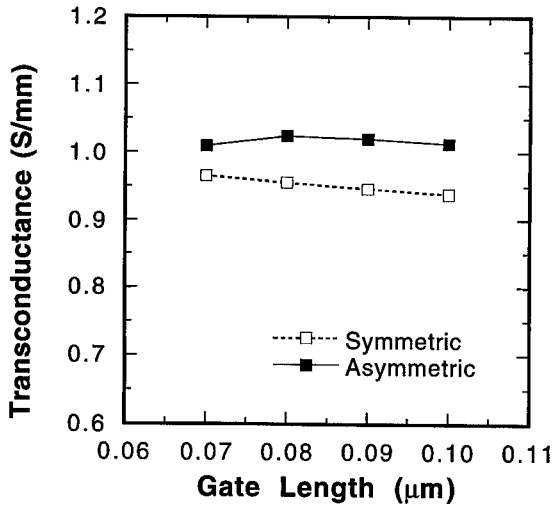


Fig. 5. Maximum transconductance as a function of nominal gate length for MODFETs with conventional symmetric (dashed line) and asymmetric grooves (solid line). The measurements were carried out at $V_{ds}=1$ V.

Another remarkable feature in Fig. 4 is the output

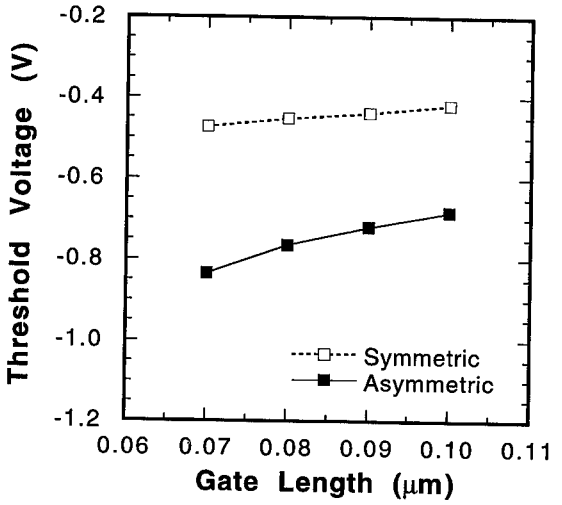


Fig. 6. Threshold voltage as a function of nominal gate length for MODFETs with conventional symmetric (dashed line) and asymmetric gate-groove profiles (solid line). The measurements were carried out at $V_{ds}=1$ V.

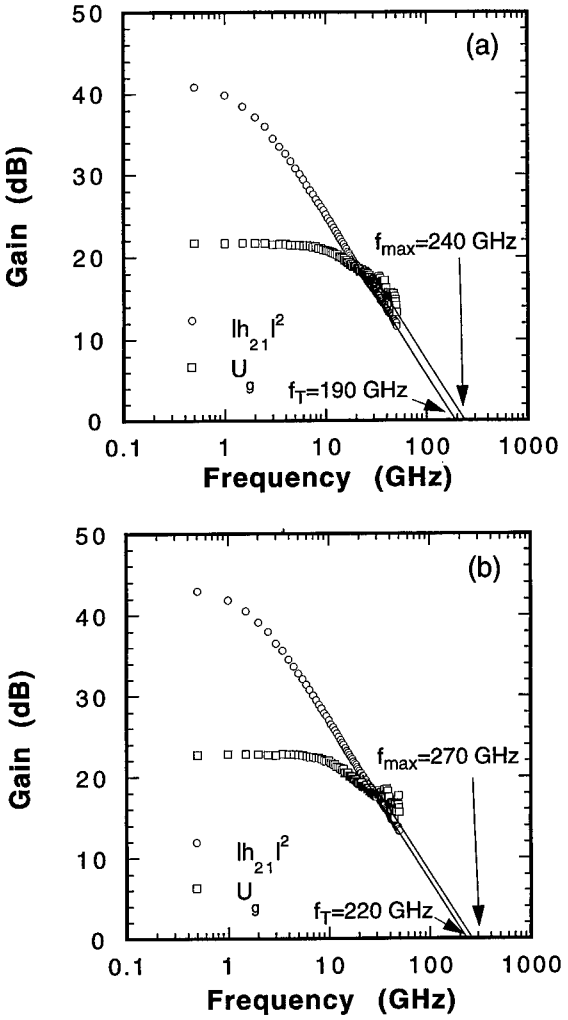


Fig. 7. Current gain $|h_{21}|^2$ and Mason's unilateral power gain U_g as a function of frequency for devices with conventional symmetric (a) and asymmetric gate grooves (b). The devices have a nominal gate length of 0.07 μm and a T-finger layout with a total gate width of 100 μm. The measurements ere carried out at $V_{ds}=1$ and the V_{gs} for peak transconductance.

conductance of 160 mS/mm for the 0.1 μm-asymmetric MODFETs, which is in sharp contrast to the 80 mS/mm for the symmetric devices with the same nominal gate length. The higher output conductance is due to a 'step' that is formed after the recess etching, which can be seen in Fig. 2. The 'step' results in a smaller effective gate length than that defined by the electron-beam lithography. The smaller effective gate length for the asymmetric devices could be another reason for the higher I_{dss} seen in Fig. 4(b). The reduced effective gate lengths for asymmetric MODFETs is further evidenced by their more negative values of threshold-voltage V_{th}, which are due to short channel effects, compared with

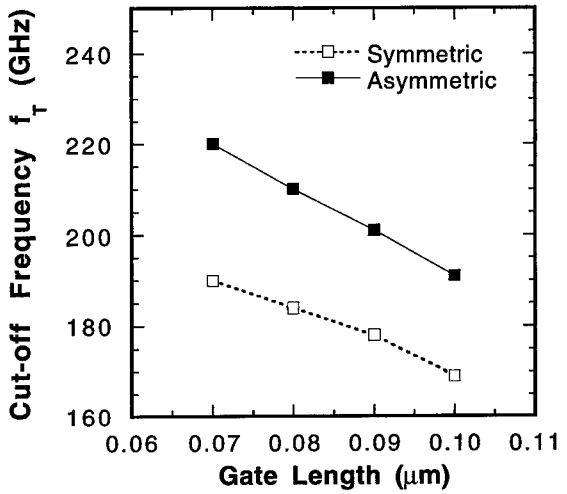


Fig. 8. Current gain cut-off frequency as a function of nominal gate length for MODFETs with conventional symmetric (dashed lined) and asymmetric gate grooves (solid line). The devices have a T-finger layout with a total gate width of 100 μm. The measurements were carried out at $V_{ds}=1$ V and the V_{gs} for peak transconductance.

those with the identical nominal gate lengths but symmetric gate grooves. In Fig. 6, one can also see that the asymmetric devices shown a V_{th} shift of 150 mV when the nominal gate length is reduced from 0.1 to 0.07 μm, while the symmetric devices demonstrate a smaller shift of 50 mV when the gate length is reduced in the same way.

3.2. RF performance

The device S-parameter characteristics were measured using a Cascade on-wafer probe station. Fig. 7 shows current gain $|h_{21}|^2$ and Mason's unilateral power gain U_g as a function of frequency at $V_{ds}=1$ and with the V_{gs} for peak transconductance for the nominal 0.07-μm devices. The f_T and f_{max} obtained by extrapolating the $|h_{21}|^2$ and U_g with −20 dB/decade are 190 and 240 GHz for symmetric devices and 220 and 270 GHz for asymmetric devices, respectively. As mentioned before, the higher f_T for the asymmetric devices is due to the reduced effective gate length. This is strongly supported by the f_T dependence on the nominal gate length for these devices shown in Fig. 8; the asymmetric device always show a higher f_T when the nominal gate length is the same. The higher f_{max}, on the other hand, is attributed to the higher f_T and the lowered feedback capacitance that results from the asymmetric gate groove, although a marked improvement in gate-drain breakdown characteristics has not yet been observed, which is due to the asymmetry ratio

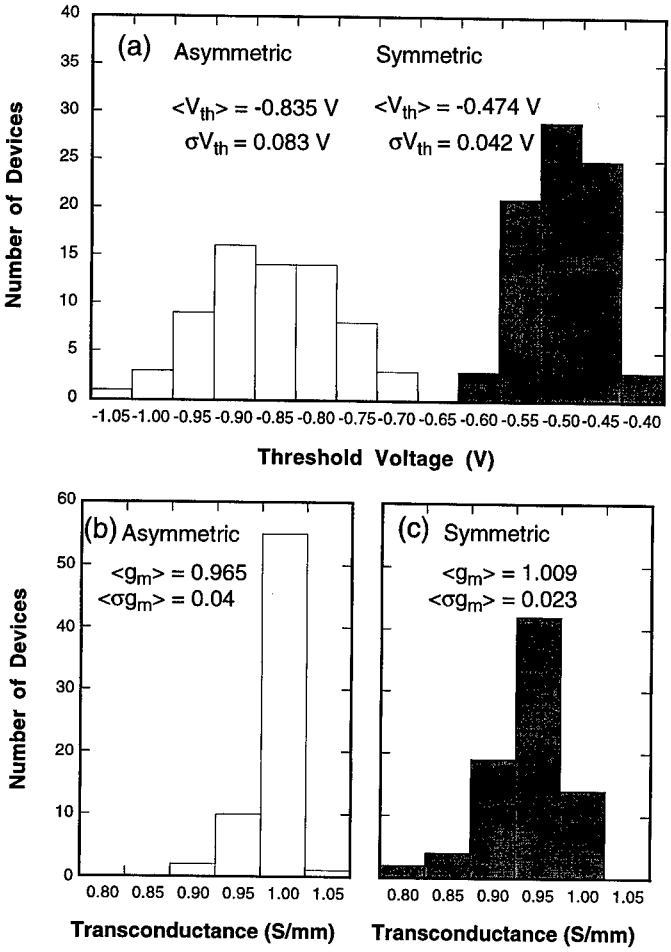


Fig. 9. Histogram of the threshold voltage for 0.07-μm MODFETS with conventional symmetric (grey bars) and asymmetric gate grooves (white bars) (a), and histograms of the transconductance for 0.07-μm MODFETs with asymmetric (b) and conventional symmetric gate grooves (c). The measurements were carried out at $V_{ds} = 1$ V. In (b) and (c) the mean value of the transconductance and the standard deviation of transconductance are given in S/mm.

achieved so far not being sufficiently large.

3.3. Statistics of device performance

The reproducibility of the technology was examined by calculating the standard deviations of V_{th} and g_m across the 3″ wafers. As shown in Fig. 9(a), the 0.07-μm symmetric devices have a V_{th} deviation of around 42 mV, which is a good value and can be understood by the use of the etch-stopper. On the contrary, for the asymmetric devices the V_{th} scattering is as large as 83 mV for the devices with nominal gate length of 0.07 μm. However, it should be noted that the g_m scattering for both wafers are very similar, around 30 mS/mm as shown in Figs. 9(b) and (c). Since g_m is relatively insensitive to gate length as shown in Fig. 5, this small g_m deviation can be taken as an indication of the good uniformity of barrier thickness. Therefore, the large scattering of V_{th} for the asymmetric devices should be attributed to the variations in the effective gate length due to the fluctuations in the wet-chemical etching.

Another concern in device fabrication is the yield. The open squares in Fig. 10 show that the fabrication yield is still higher than 95% when the gate length is reduced to 0.07 μm, if the standard process (for fabricating symmetric devices) is used. In contrast, the asymmetric devices show a remarkably lower yield for gate lengths ranging from 0.07 to 0.1 μm. The exact reason for this remains unclear right now, but this

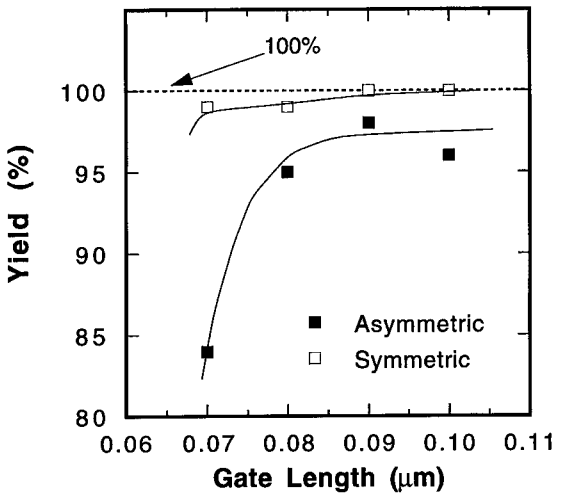


Fig. 10. Fabrication yield as a function of the nominal gate length for MODFETs with conventional symmetric (open squares) and asymmetric gate grooves (solid squares).

lower yield should be related to the asymmetric etching that is employed to fabricate the asymmetric grooves.

4. Conclusions

Asymmetric recess etching can be achieved in the fabrication of gate grooves for MODFETs based on InAlAs/InGaAs heterostructures. The electrochemical etching with the surface metals of Pt and Ni respectively deposited on the source and drain pads leads to an aspect ratio of 1.7 for 0.1-μm devices. Devices fabricated with this technology display a higher f_T value, due to the reduced effective gate length, and a higher f_{max} due to the asymmetric gate groove. Further study is needed in order to increase the asymmetry ratio, uniformity and the yield of devices.

Acknowledgements

The authors would like to thank T. Maruyama, Y. Kondo, T. Makimura, T. Kusumoto, and H. Kitabayashi for their technical assistance in the MODFET fabrication. They are also grateful to T. Ishikawa for his efforts in MOCVD growth.

References

[1] Nguyen LD, Larson LE, Mishra UK. Proc IEEE 1992;80:494.
[2] Enoki T, Tomizawa M, Umeda Y, Ishii Y. Jpn J Appl Phys 1994;33:798.
[3] Lester LF, Smith PM, Ho P, Chao PC, Tibirio P, Duh KG, Wolf ED. IEDM Tech Dig, 1988. p. 172.
[4] Ho P, Kao MY, Chao PC, Duh KHG, Ballingall JM, Allen ST, Tessmer AJ, Smith PM. Electron Lett 1991;27:325.
[5] Gaquiere C, Bonte B, Theron D, Crosnier Y, Arsene-Henri P, Pacou T. IEEE Trans Electron Devices 1995;42:209.
[6] Hur KY, McTaggart RA, Miller AB, Hoke WE, Lemonias PJ, Aucoin LM. Electron Lett 1995;31:135.
[7] Ballegeer DG, Adesida I, Caneau C, Bhat R. Proceedings International Conference on InP and Related Materials (Santa Barbara, California, USA, 1994). 1994. p. 331.
[8] Grundbacher R, Ballegeer D, Ketterson AA, Kao YC, Adesida I. IEEE Trans Electron Devices 1997;44:2136.
[9] Xu D, Enoki T, Ishii Y. IEEE Electron Device Lett 1998;19:10.
[10] Latimer WM. Oxidation potential, 2nd ed, 1952.
[11] He Y, Liang BW, Tien NC, Tu CW. J Electrochem Soc 1992;139:246.
[12] Xu D, Enoki T, Suemitsu T, Umeda Y, Yokoyama H, Ishii Y. J Electron Mater 1998;27:L51–L53.

PERGAMON

Solid-State Electronics 43 (1999) 1535–1540

SOLID-STATE ELECTRONICS

First demonstration of low temperature grown InP-channel HFET transferred onto GaAs substrate

M. Kunze[a,*], L.H. Lee[a], H.Y. Chung[b], E. Kohn[a]

[a]*Department of Electron Devices and Circuits, University of Ulm, D-89069 Ulm, Germany*
[b]*Department of Optoelectronics, University of Ulm, D-89069 Ulm, Germany*

Abstract

Here, we will demonstrate a new approach to merge high performance III/V-material properties with optical display technology. In a first step a low temperature grown and processed InP-FET structure not surpassing a temperature of 300°C is transferred to GaAs with the potential of future transfer to Si and other substrate materials. Future large area display panels require electronic overlay circuitry for optical arrays with high performance. FETs with high current densities grown on various substrate materials optimized to drive simultaneously a large number of light sources will be required. Furthermore each process step, including materials growth, has to be performed at extremely low temperatures to prevent degradation of the low-cost substrates, mainly glass and polymer. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction and new concept

The current state-of-the-art in display device engineering is divided into two parts: the III/V-technology mainly used for GaAs VCSELs with FET devices of high performance, expensive materials and small area dimensions. The other one using large area substrate materials like glass or polymers, with low performance Si-TFT devices, aiming at large areas for consumer applications. Up to now no technology was available to merge the high performance III/V-FETs with the large area display approach. The disadvantage of Si-TFTs, namely low carrier mobilities, leads to low current densities and large switching voltages. Moderate carrier mobilities, usually still below 200 cm^2/V s, are obtained after crystallization at elevated temperatures, which in turn may cause degradation of the amorphous substrate properties. III/V-materials, which would promise higher mobilities, could not be used because of two reasons:

1. a technology to realize FETs on low grade highly dislocated material has not been evaluated.
2. at a growth temperature below 450°C III/V layers contain a high intrinsic antisite defect density. In GaAs this native defect is located within the band-gap and therefore leads to insulating characteristics. Doping activation is inhibited.

In contrast, growth of InP at temperatures below 300°C produces P_{In} antisite defects with a first excited state located 120 meV above the conduction band edge [1]. This leads to auto-doped n-type layers and makes Low Temperature Grown InP (LTG-InP) a promising candidate for active FET layers. The concentration of

* Corresponding author. Tel.: +49-731-502-6181; fax: +49-731-502-6155.
E-mail address: kunze@ebs.e-technik.uni-ulm.de (M. Kunze)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00100-8

this P_{In} antisite defect is connected to growth temperature and V/III flux ratio. Therefore, a strong influence of growth temperature on sheet carrier concentration and saturation current is found. First devices grown on InP substrates at 300°C and below were presented earlier [2,3]. They have demonstrated the feasibility of fully low temperature grown and processed devices on InP substrates not exceeding 300°C. High dc output current densities above 100 mA/mm were reached. Despite the various defects found in LTG-devices such as intersticials, vacancies and the antisites, RF measurement results revealed similar potential for high speed applications as compared to conventionally grown InP channel FETs. For FETs with a $l_g = 0.4$ μm gate length, a device current cutoff frequency of $f_t = 18$ GHz and a maximum oscillation frequency $f_{max} = 40$ GHz were obtained [3]. The materials growth studies of the entire heterostructure including LTG-InP and LTG-AlInAs barrier and buffer layer showed that control of P-accumulation at the LTG-InP channel interface and the control of electronically active defects at the LTG-AlInAs Schottky barrier contact need specific optimization within a narrow field of parameters [4]. However, the study had also indicated that the material needs not to be of high crystalline perfection and that dislocations will not prevent FET channel modulation despite a reduction of channel mobility. This had encouraged us to transfer the LTG-FET-structure onto GaAs within a materials regime of deliberately highly dislocated thin layers on the GaAs substrate (approx. 0.1 μm). Already such a FET device would be unique, since LTG-GaAs is insulating and no LTG-GaAs FET can be produced. In case, FET-characteristics are obtained on highly dislocated material the structure may be transferred to Si using an semi-insulating LTG-GaAs buffer [5], which is also grown at a temperature at or below 300°C as previously used for GaAs-on-Si devices [6].

A key role will play the behaviour of grain boundaries in the LTG-InP channel. In GaAs such grain boundaries represent Schottky barriers with a pinned surface potential and thus no noticeable current can be driven across such a junction. On the other hand along the boundary current flow will also be negligible [7]. In the contrary, the InP-surface is fairly unpinned near the conduction band edge so that n-channel MOS-inversion layers can be formed [8]. This behaviour may indicate that band bending at InP grain boundaries may be negligible as in the case of Si.

Such a study may then also indicate whether FET characteristics can be obtained on polycrystalline LTG-InP material, which may then be grown on recrystallized Si-on-glass. Ultimately, this may allow to assess, whether amorphous LTG-InP TFTs are possible.

2. Key problems of active LTG-InP devices

2.1. Device elements

The development of LTG-FET devices consists of various effects, which need to be combined in the electrical FET characteristics. The main issues are (1) the active channel leading to the desired output current, (2) the gate barrier for channel current modulation and low leakage pinch-off and (3) insulating buffer material for low subthreshold current. These are discussed shortly in the following.

- Active LTG-InP channels grown on InP substrate revealed carrier concentrations above 10^{16} cm^{-3} at 350°C growth temperature up to a maximum of 4×10^{18} cm^{-3} below 260°C[2]. Therefore, a channel thickness of approx. 60 nm would lead to a sheet carrier concentration in the range of 10^{12} cm^{-2} applicable for FET channel operation. However, displacement of phosphorous towards the channel surface and accumulation at the interface leads to problems with channel modulation due to a high interface state density.
- Second, for channel current modulation of the FETs suitable Schottky gate barrier materials which also have to be grown at a low temperature have to be found.
- Negligible buffer leakage current will be needed for TFT applications with low standby power. This also imposes tight limits on the LTG-buffer layer material with its high defect density.

Temperature stability of LTG-InP layers The temperature stability of LTG-InP layers was assessed on InP substrate only HenlePHD, but it is expected that the stability on GaAs is governed by the same mechanisms. Fortunately, up to 400°C no degradation effect could be detected and therefore the application of LTG-InP layers in a display technology seems to be possible. Increased annealing temperatures lead to reduced carrier concentrations measured by Hall effect. This effect is accompanied by an increasing lattice mismatch (Fig. 1). In the limit a maximum mismatch (1.33×10^{-3}) is extracted. To the first order an interplay of two effects, that of the antisites and that of the interstitials, was used for explanation. The antisites can be considered as superimposed second lattice onto the InP lattice and lead to a reduced lattice constant, whereas interstitials cause a lattice widening. Therefore, increasing annealing temperatures from 400 to 520°C reduces the antisites concentration and leads to an increased lattice mismatch caused by the increased influence of interstitials. The reduced carrier concentration is a consequence of the loss of antisites. This decrease of carrier concentration is connected with an activation energy of approx. 1 eV [9]. At

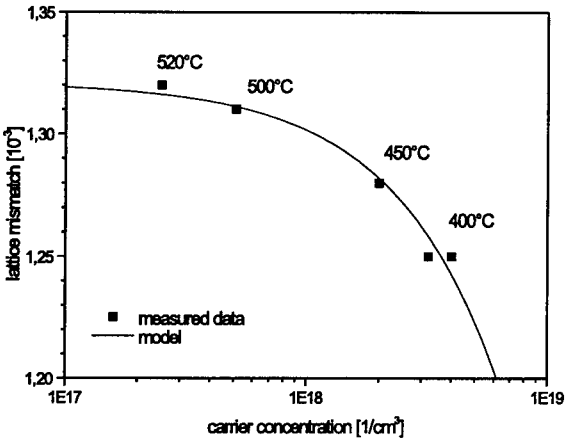


Fig. 1. Lattice mismatch of LTG-InP FET heterostructure on InP substrate annealed at different temperatures versus carrier concentration. The solid line shows a first order approximation model (details in text).

further increased temperatures above 560°C phosphorous starts to segregate into clusters leading to a disappearing lattice mismatch [10]. Due to the wide range of change in carrier concentration with post growth annealing temperature annealing offers a possibility of after-growth carrier concentration redesign.

3. Experimental

3.1. Material growth and processing

The heterostructure devices were grown on 100-oriented semi insulating GaAs substrates by gas source MBE (GSMBE) using PH_3 as P-source. Growth was performed at 280 to 300°C. For sufficient reproducibility of low growth temperatures a set of Mo-wafer-carriers was calibrated by heating a substrate from 400°C, the lower limit of the used pyrometer setup, up to 600°C leading to a specific dependency of the pyrometer temperatures versus thermocouple temperatures. Growth temperatures below 400°C were then realized by an extrapolation of the thermocouple calibration curves.

As buffer layer previously reported insulating LTG-AlInAs buffer layers were used [11]. At this growth temperature the InP free carrier concentration is between 4×10^{17} and 5×10^{16} cm^{-3}. The device structure (Fig. 2) consists of three layers:

1. 60 nm LTG-AlInAs buffer layer on GaAs substrate,
2. 55 nm LTG-InP active channel layer,
3. 20 nm LTG-AlInAs gate contact layer.

Usually, pretreatment of substrate materials to eliminate oxides is performed for a short time at high temperatures, but using low temperatures equivalent surface pretreatment can be achieved deoxidizing the substrates for a long time [4]. The devices are isolated by wet chemical mesa etching. For optimized ohmic contact series resistance windows are etched into the AlInAs top layer to deposit the contact metal directly on the LTG-InP channel surface. The ohmic contacts of Ge/Ni/Au metal are alloyed between 280 and 340°C in N_2-atmosphere. The higher temperature is needed, if the contact needs to be alloyed through a remaining AlInAs layer. The 1.5 μm long gates were defined by optical lithography. Gate metallization was Ti/Pt/Au.

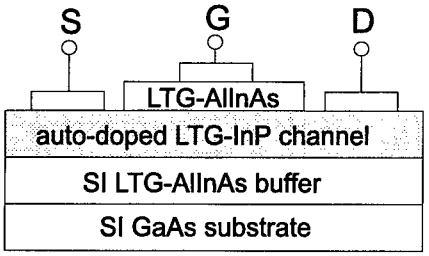


Fig. 2. Schematic device structure of LTG-InP FET on GaAs substrate.

3.2. Results

At first the individual elements of the structure shall be discussed.

Fig. 3 shows a HRXRD rocking curve of an LTG-InP heterostructure on GaAs substrate. The lattice mismatch is 4.41×10^{-2}. This leads to high defect concentrations as can be seen from the SEM micrograph in Fig. 4. Despite the dislocations and defects the measured mobility still remains between 500 and 600 cm^2/V s.

In Fig. 5 the IV-characteristics of an LTG-InP chan-

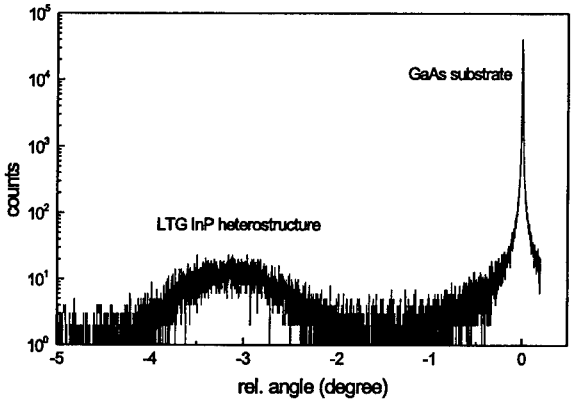


Fig. 3. HRXRD Rocking curve of LTG-InP heterostructure on GaAs substrate.

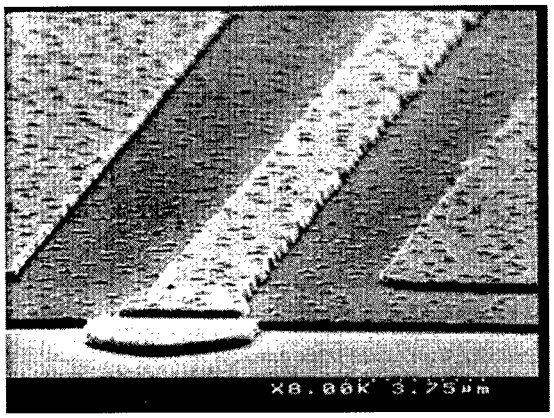


Fig. 4. SEM picture of LTG-InP HFET on GaAs substrate showing high defect density.

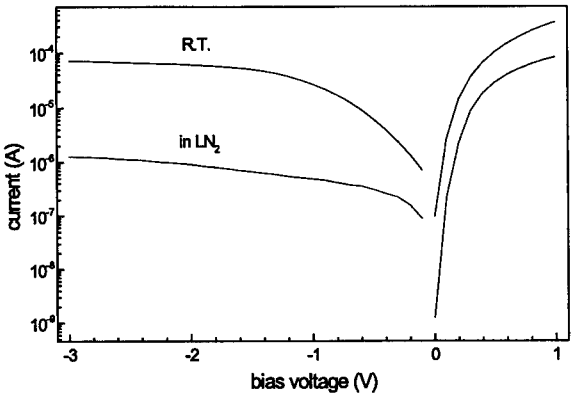


Fig. 6. Ti/Pt/Au-LTG-AlInAs diode on LTG-InP channel at RT and in LN_2 leading to improved breakdown characteristics.

nel with a floating gate positioned in the center is shown. The channel parameters were $W_g = 25$ μm, $l_g = 1.5$ μm and $t = 55$ nm. It was grown on a 60 nm thick LTG-AlInAs buffer on SI-GaAs substrate. The surface was covered by a 20 nm LTG-AlInAs layer, grown using identical parameters to those of the buffer layer. The surface reveals a high defect density similar to that as shown in Fig. 4. At low bias a nonlinearity is observed due to the slightly different contact technology for the samples on GaAs substrate, where the metallization was alloyed through the LTG-AlInAs cap layer (at 340°C for 60 s in N_2-atmosphere). Current saturation is observed supported by the floating gate field distribution. This indicates that the current is indeed confined to the channel and that drift velocity saturation is obtained. In combination with a sheet carrier concentration of $n_S = 0.7 \times 10^{12}$ cm^{-2} from Hall measurement this results in a saturated velocity of 1.3×10^7 cm/s, if carrier depletion effects are neglected. Thus, it seems that the high defect density does not lead to a noticeable deterioration of the saturated velocity. Unfortunately, the channel current could not be modulated by the gate, most likely primarily due to P-accumulation at the channel/cap-layer interface.

Previously for LT-growth on InP it was seen that the Schottky barrier characteristics on LTG-AlInAs show high leakage current levels and early breakdown; this was found to be the critical element in LTG-InP FET devices on InP substrate [2]. Thus, on highly dislocated buffer layer and channel material on GaAs this was also expected to be critical. Indeed, Fig. 6 shows high gate Schottky contact leakage and soft breakdown of Ti/Pt/Au-AlInAs diode with 20 nm LTG-AlInAs barrier layer on an LTG-InP channel on GaAs. However, cooling to liquid nitrogen yields in a two orders of magnitude reduction in leakage current and improved breakdown characteristics (Fig. 6). This shows that the Schottky barrier contains a high number of low barrier injection paths, but not a dominating tunneling component. Thus, passivation of the defects (for example through an interfacial layer) may improve the room temperature characteristics.

Due to the limitations discussed above first FET structures were grown and fabricated on rather low doped channel material yielding low channel current densities, but improved Schottky leakage characteristics ($n_s \approx 10^{11}$ cm^{-2} at 300°C growth temperature). Furthermore, all devices were grown using an optimized channel configuration with reduced excess P towards the channel surface. Therefore, the maximum channel current reaches only a maximum open channel current density of 3 mA/mm and a maximum transconductance of 5 mS/mm.

The dc output characteristics of a device with a gate width of $W_g = 100$ μm and $l_g = 1.5$ μm are shown in

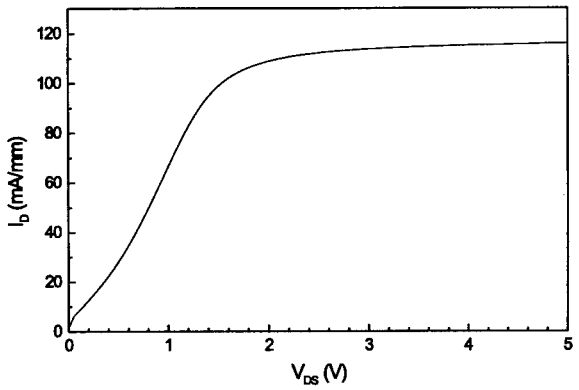


Fig. 5. LTG-InP channel with floating gate (dimensions see text).

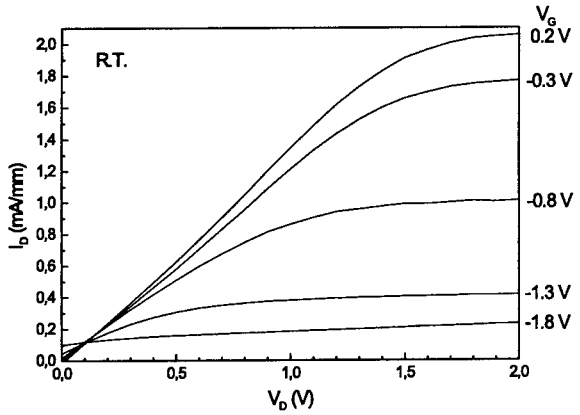


Fig. 7. Output characteristics of LTG-InP FET grown on GaAs ($W_g = 100$ μm and $l_g = 1.5$ μm, $T_{Growth} = 300$°C).

Figs. 7 and 8 at two operation temperatures (RT and LN_2). Due to the above discussed difficulties with the Schottky layer, pinch-off is only partially possible at room temperature. However, in saturation near pinch-off no significant output conductance is observed. Therefore, parasitic conduction paths caused by an LTG-AlInAs buffer layer conductivity and gate to drain leakage can be neglected. Thus, the limiting factor is source to gate leakage. Using LN_2, gate leakage can be reduced and pinch-off was improved. In the best device the improvement in gate leakage was a factor of 40 and a subthreshold behaviour as shown in Fig. 9 was obtained.

RF measurements on these first structures resulted in a $f_t = 0.29$ GHz and $f_{max} = 0.61$ GHz. These cutoff frequencies are still parasitic dominated; they may only confirm free electron transport in the channel. It is expected that the next redesign will essentially improve the present characteristics further.

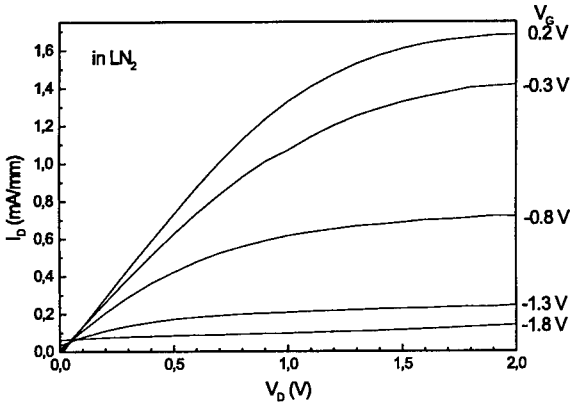


Fig. 8. Output characteristics of LTG-InP FET grown on GaAs substrate in LN_2 ($W_g = 100$ μm and $l_g = 1.5$ μm).

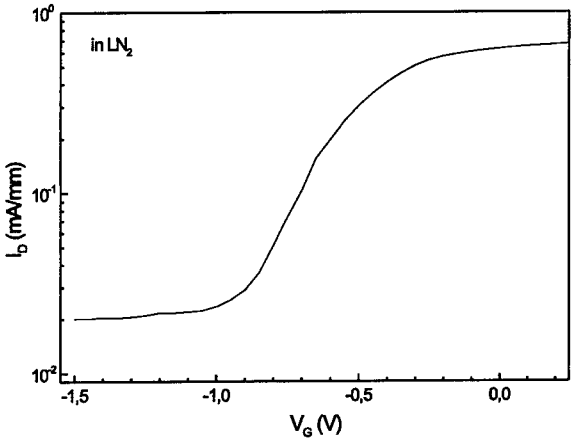


Fig. 9. Semilog transfer characteristics of LTG-InP FET on GaAs substrate in LN_2 showing subthreshold behaviour.

4. Conclusion and outlook

In conclusion, this proof of concept may be considered the starting point for further improvements. Here, the fundamental possibility of transferring the growth of LTG-InP channel HFETs to GaAs substrates was shown. Even transfer onto Si substrate seems possible due to the existing LTG-GaAs buffer technology. Further improved Schottky layer characteristics using a different barrier layer concept may lead to LTG-InP HFETs with the high output current and improved on–off ratio. The results may again act as a starting point for growth on other non-InP substrates.

References

[1] Dreszer P, Chen WM, Seendripu K, Wolk JA, Walukiewicz W, Liang BW, Tu CW, Weber ER. Phosphorous antisite defects in low temperature InP. Physics Review 1993;47:4111–4.

[2] Henle B, Kohn E. InP-based FET structure grown and processed at extremely low temperatures of 280°C. In: Proceedings IPRM 97, Hyannis, MA, 1997. p. 165–8.

[3] Lee LH, Kunze M, Henle B, Kohn E. Characteristics of InP-based HFETs grown at extremely low temperatures of 300°C and below. In: Proceedings IPRM 98, Ibaraki, Japan, 1998. p. 513–6.

[4] Henle B. Wachtsum nichtstöchiometrischer Epitaxieschichten auf InP-Basis für FET-Anwendungen. Ph.D. thesis University of Ulm, Germany, 1997.

[5] Smith FW, Calawa AR, Chen C-L, Manfra MJ, Mahony LJ. New MBE buffer used to eliminate backgating in GaAs MESFET's. IEEE Electron Device Letters 1988;9:77–80.

[6] Frankel MY, Tadayon B, Carruthers TF. Integration of low temperature GaAs on Si substrates. Applied Physics Letters 1993;62:255–7.

[7] Yang JJJ, Dapkus PD, Dupuis RD, Yingling RD. Electrical properties of polycrystalline GaAs films. Journal of Applied Physics 1980;51:3794–800.
[8] Lile DL, Collins DA, Meiners LG, Messick L. n-Channel inversion mode InP MISFET. Electronics Letters 1978;14:657–9.
[9] Henle B, Reitemann G, Kohn E. Low temperature grown InP-based heterostructures and their device application. In: Int. Symp. Compound Semiconductors, St. Petersburg, Russia, 1996. p. 527–30.
[10] Claverie A, et al. Formation of P precipetates during annealing of InP grown by gas source molecular epitaxy at low temperature. Applied Physics Letters 1993;62:1638–40.
[11] Brown AS, Mishra UK, Chou CS, Hopper CE, Melendes MA, Thompson M, Larson LE, Rosenbaum SE, Delaney MJ. AlInAs-GaInAs HEMTs utilizing low temperature AlInAs buffers grown by MBE. IEEE Electron Device Letters 1989;10:565–7.

PERGAMON

Solid-State Electronics 43 (1999) 1541–1546

SOLID-STATE ELECTRONICS

Growth of device quality InGaP/GaAs heterostructures by gas source molecular beam epitaxy using tertiarybutylphosphine

Hironobu Sai, Hajime Fujikura*, Atsushi Hirama, Hideki Hasegawa

Research Center for Interface Quantum Electronics and Graduate School of Electronics and Information Engineering, Hokkaido University, Sapporo 060-8628, Japan

Abstract

The feasibility of preparing device quality InGaP/GaAs heterostructures by the gas source molecular beam epitaxial (GSMBE) growth using tertiarybutylphosphine (TBP) was investigated. Undoped and Si doped InGaP layers with GaAs buffer layers as well as InGaP/GaAs quantum well (QW) samples were grown and characterized by atomic force microscopy (AFM), photoluminescence (PL) and Hall measurements. All of the undoped and doped InGaP layers showed narrow PL peaks and electron mobilities comparable to those reported for InGaP layers grown by other methods. Undoped samples showed a residual carrier concentration of 1×10^{15} cm^{-3} and a mobility of 3300 cm^2/V s at room temperature. In the Si-doped samples, the highest electron density of 2×10^{19} cm^{-3} was achieved without carrier saturation. On the other hand, InGaP/GaAs QW samples showed intense and narrow PL emission lines, indicating formation of sharp heterointerfaces. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

$In_{0.48}Ga_{0.52}P$/GaAs heterojunction bipolar transistors (HBTs) and high electron mobility transistors (HEMTs) have been attracting significant attention due to their advantages over conventional AlGaAs/GaAs-based devices, such as a lower defect density and higher reliability of the InGaP/GaAs system than the AlGaAs/GaAs system as well as absence of DX centers [1]. As compared with standard metal organic vapor phase epitaxy (MOVPE) and gas source molecular beam epitaxy (GSMBE) approaches using highly toxic phosphine (PH_3) to grow InGaP, GSMBE using tertiarybutylphosphine (TBP), which combines the advantages of MBE with the low toxic nature of TBP, seems to be attractive for high volume production of heterostructure wafers. However, crystal qualities of InGaP layers grown by GSMBE using TBP reported so far have been inferior to those by the standard methods and no systematic optimization of the growth has been made to achieve high quality InGaP/GaAs heterostructures by this method [2–4].

From such a viewpoint, we have recently attempted to optimize GSMBE growth of InGaP layers on (001) GaAs substrates using reflection high energy electron diffraction (RHEED) oscillation, double crystal X-ray diffraction (XRD), photoluminescence (PL) and Hall effect measurements [5,6]. The result has indicated that high quality undoped InGaP layers can be grown even by the GSMBE using TBP under carefully optimized growth conditions.

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00101-X

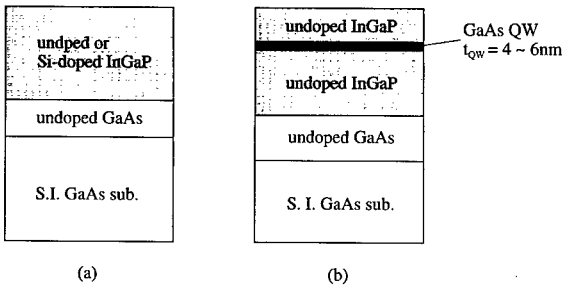


Fig. 1. Sample structures. (a) Undoped or Si-doped InGaP layer. (b) Undoped InGaP/GaAs quantum well.

The purpose of this paper is to investigate the feasibility of the growth of device quality InGaP/GaAs heterostructures by the present GSMBE growth method using TBP. Undoped and Si doped InGaP layers with GaAs buffer layers as well as undoped InGaP/GaAs quantum well (QW) samples were grown and characterized from the viewpoint of surface morphology, crystal quality, carrier concentration, mobility and band lineup. Atomic force microscopy (AFM), PL and Hall measurements were used for this purpose.

(A) Undoped samples showed a residual carrier concentration of 1×10^{15} cm^{-3} and a mobility of 3300 cm^2/Vs at room temperature.

In the Si-doped samples, the highest electron density of 2×10^{19} cm^{-3} was achieved without carrier saturation. InGaP/GaAs QW samples showed intense and narrow PL emission lines, indicating formation of sharp heterointerfaces.

2. Experimental

All growth was carried out using a EIKO EVC500 GSMBE system with a 2800 l/s diffusion pump. Growth was monitored by observing RHEED patterns and their oscillations taken at an electron acceleration energy of 15 kV. The substrate temperatures, Tg, were measured by an infrared pyrometer which was calibrated with the melting point of InSb (525°C).

As the sources of In, Ga and As, conventional metallic In, Ga and As were used. Beam flux intensities of In and Ga sources were calibrated for various temperatures of In and Ga K-cells by measuring the growth rates of various binary compounds such as GaAs, InAs, InP and GaP from the RHEED oscillation period. On the other hand, TBP was used as the phosphorus source and was decomposed to obtain a P_2 supply by a newly designed high efficiency cracker cell [5,6] at cracking temperatures, Tcr, of 800 or 900°C. As for the n-type dopant, elemental Si was used.

In this study, two different kinds of samples were grown on (001)-oriented semi-insulating GaAs substrates as shown in Fig. 1. One is undoped or Si-doped single InGaP layers, as shown in Fig. 1(a) and the other is an undoped InGaP/GaAs/InGaP QW samples as shown in Fig. 1(b). Thickness of the InGaP layers in the single layer sample shown in Fig. 1(a) were about 1 μm, unless otherwise noted. On the other hand, the thickness of the GaAs buffer layer for all the single layer sample was 100 nm.

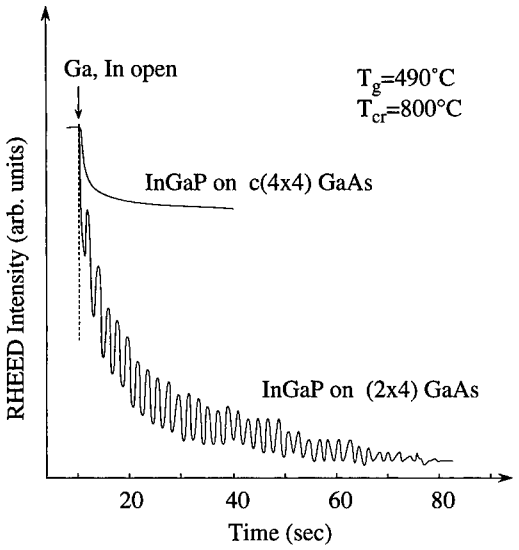


Fig. 2. RHEED intensities taken during the growth of InGaP layers on As-rich c(4 × 4) and As-stabilized (2 × 4) GaAs surfaces.

As for the QW samples, three QW samples having different thickness of GaAs QW layers, t_{QW}, of 4, 5 and 6 nm were grown in this study. The thickness of the undoped GaAs buffer layer, that of the bottom InGaP barrier layer and that of the top InGaP barrier layer were 500, 100 and 150 nm, respectively, in all the QW samples.

In this study, a fixed TBP flow rate of 4 sccm was used for all the growth of InGaP layers. The beam flux intensities of In and Ga were chosen so as to grow $In_{0.48}Ga_{0.52}P$ layers lattice-matched to GaAs with a growth rate of 0.63 ML/s. The growth rate for GaAs layers was 0.33 ML/s, which was achieved without changing the Ga K-cell temperature for InGaP growth.

All the InGaP layers on the GaAs buffer layers were grown on the As-stabilized (2 × 4) GaAs surfaces, which led to the stable layer-by-layer growth of InGaP layers, showing clear and persistent RHEED oscillation as shown in Fig. 2 [6]. On the other hand, growth of the InGaP layer on an As-rich c(4 × 4) GaAs surface was found to result in the growth without RHEED oscillations as also shown in Fig. 2. Therefore, in the growth of InGaP layers below 510°C,

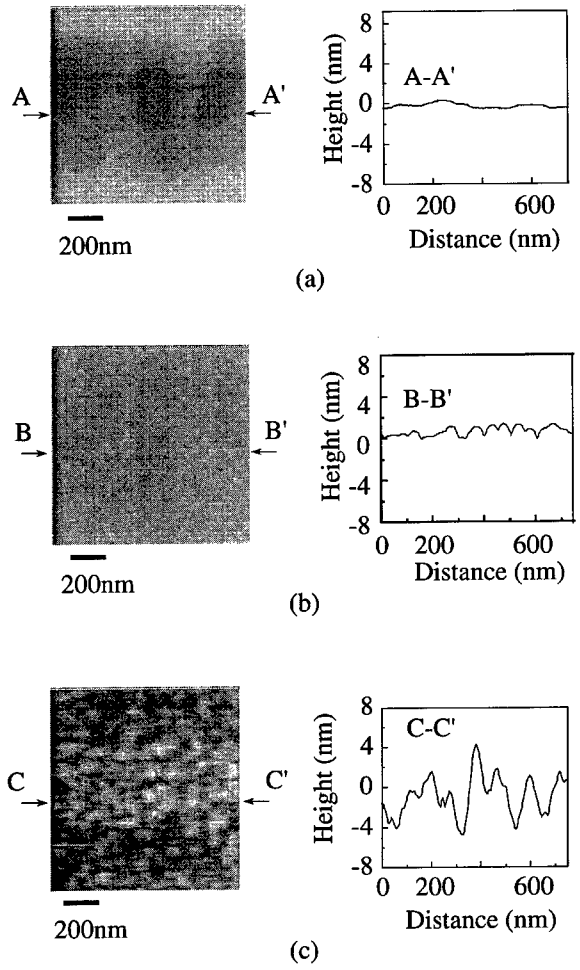


Fig. 3. AFM images of (a) GaAs buffer layer and those of InGaP layers grown at (b) 490°C and (c) 510°C. AFM line-profiles along lines A–A′, B–B′ and C–C′ are also shown.

As_4 flux was shut off at 510°C and left the surface under the residual As pressure in the chamber in order to prevent change of the GaAs surface reconstruction from As-stabilized (2 × 4) to As-rich c(4 × 4).

As is well known for the conventional MBE growth of GaAs layer, a persistent RHEED oscillation is an indication of realization of a stable layer-by-layer growth mode. According to our previous study, the number of observable RHEED oscillation cycles becomes a maximum at 480–500°C. Therefore, in this study, growth of InGaP layers was carried out in this temperature range. The GaAs buffer layers were grown at 580°C, whereas the GaAs QW layers were grown at the same temperature range for InGaP growth.

Optical and electric properties of the undoped and Si doped single InGaP layers shown in Fig. 1(a) were characterized by PL and Hall measurements at 300 K.

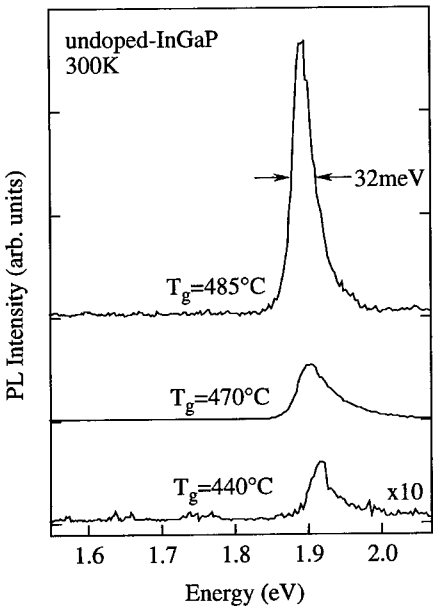


Fig. 4. Room-temperature PL spectra of the InGaP layers grown at various temperatures.

Smoothness of the InGaP layers were characterized by AFM observations and their line-profile analysis. Characteristics of the InGaP/GaAs heterointerfaces were investigated by performing low-temperature PL measurements for InGaP/GaAs QW samples. The PL measurements were made using an optical multi-channel analyzing system (Acton Research Corporation SpectraPro-750) under irradiation of 514.5 nm emission line of the Ar^+ laser.

3. Results and discussion

3.1. Morphology and optical properties of $In_{0.48}Ga_{0.52}P$ layers

Fig. 3 shows AFM images of InGaP layers grown at 490 and 510°C together with that of the GaAs buffer layer. In this case, the thickness of the InGaP layers were 28 nm. AFM line-profiles along lines A–A′, B–B′ and C–C′ are also shown in Fig. 3. It is found from Fig. 3 that a relatively smooth surface morphology of InGaP was obtained at T_g=490°C, whereas growth at a higher temperature of 510°C resulted in a rough surface with 3-dimensional island like structures. Moreover, all the growth of InGaP layers above T_g=510°C resulted in quite rough surfaces due to the In-droplet formation on the surface caused by thermal desorption of phosphorous during the growth [5].

Band edge PL emission is known to provide information on the quality of the crystal. Fig. 4 shows the

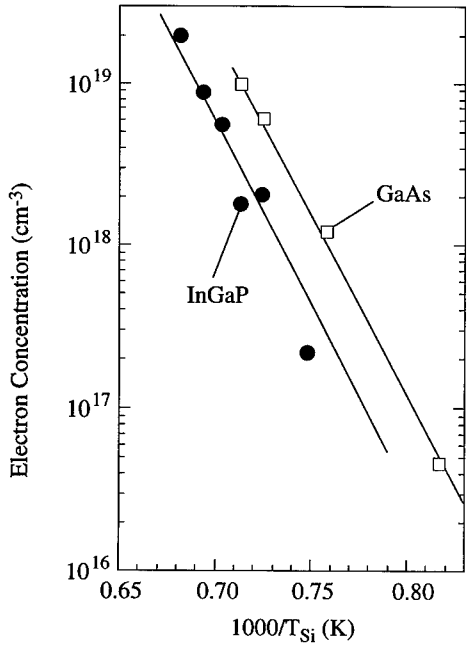


Fig. 5. Dependences of the electron concentrations of Si-doped InGaP layers and GaAs layers on the Si K-cell temperatures.

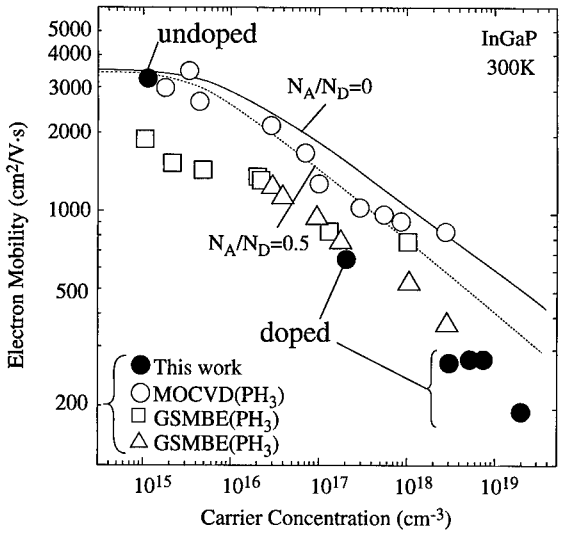


Fig. 6. Relationship between the carrier concentrations and Hall mobilities of the undoped and Si doped InGaP layers measured at 300 K. The best values reported so far for various growth methods are also plotted. Theoretically expected μ–n relationship for two compensation ratios of $N_A/N_D=0$ and 0.5 are also shown [7].

room-temperature PL spectra of the InGaP layers grown at various temperatures. In all the PL spectra, the dominant peaks around 1.9 eV were assigned as band edge emission of the InGaP layers. These peak positions almost agree with the data obtained for the $In_{0.48}Ga_{0.52}P$ layer with no natural ordering grown by GSMBE using TBP [2], indicating that substantial ordering does not exist in the present InGaP layers. It is also seen in Fig. 4 that the peak intensity increases with increasing growth temperature up to 485°C, indicating improvement of the crystal quality of the InGaP layer. Furthermore, the narrowest PL line width of 32 meV was achieved at $T_g=485$°C. On the other hand, the InGaP layers grown above 510°C showed no PL emission at room temperature.

According to the above results, growth within a narrow temperature range around 490°C gives InGaP layers with a good crystal quality and surface morphology.

3.2. Electrical properties of $In_{0.48}Ga_{0.52}P$ layers

From the Hall measurements, the undoped InGaP layer grown under the optimum growth condition was found to have n-type conduction with a low carrier concentration value of 1×10^{15} cm^{-3}. As for the Si-doped InGaP layers, the electron concentrations were plotted versus the reciprocal Si K-cell temperatures in Fig. 5 together with the data for GaAs. As seen in Fig. 5, an exponential doping behavior in the plot of electron concentration versus the Si K-cell temperature was obtained for n-type doping of InGaP layers. It is noted that there is no indication of carrier density saturation up to 2×10^{19} cm^{-3}. The capability of a wide range of electron concentrations from 1×10^{15} to 2×10^{19} cm^{-3}, as well as a simple exponential doping behavior versus the reciprocal Si K-cell temperature in the present InGaP growth, are obviously favorable for the growth of wafers for actual devices.

The slopes of both the doping characteristic lines for the InGaP and the GaAs were almost identical. However, the electron concentration values for InGaP are about 3.5 times smaller than those for GaAs at the same Si K-cell temperatures, indicating the existence of a certain compensation mechanism. Such a discrepancy between the doping characteristics for InGaP grown by GSMBE using TBP and GaAs was also reported by Beam III et al. [2].

The relationship between the carrier concentrations and Hall mobilities of the undoped and Si doped InGaP layers measured at 300 K are plotted in Fig. 6 together with the best values reported so far for various growth methods. The theoretically expected μ–n relationships for two compensation ratios of $N_A/N_D=0$ and 0.5 are also plotted in Fig. 6 [7].

As seen in Fig. 6, the undoped InGaP layer grown in this study locates the left-upper side of Fig. 6 and lies on the theoretical curve, indicating realization of high quality InGaP layer. Moreover, its electron mobi-

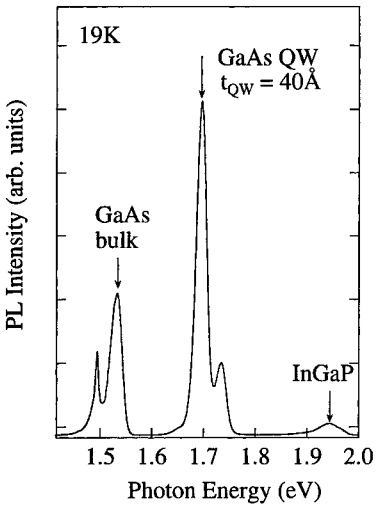


Fig. 7. PL spectrum of InGaP/GaAs quantum well having a well thickness of $t_{QW}=4$ nm taken at 20 K.

lity value of 3300 cm^2/V s is not only the best data for $In_{0.48}Ga_{0.52}P$ layers grown by TBP-based MBE but also among the best data obtained by various growth methods [2–4,8,9]. To arrive at this conclusion, the possibility of conduction contribution from the underlying undoped GaAs buffer layer was carefully checked and ruled out. Namely, the conduction type of the MBE grown undoped GaAs was a weak p-type with a carrier concentration of 10^{14}–10^{15} cm^{-3} and a hole mobility value of a few 100 cm^2/V s. This provides a high parallel sheet resistance separated by a p–n barrier and, thus, should be negligible. On the other hand, in the unlikely case where the electron conduction in the GaAs buffer layer dominates over the conduction within the 1μm thick InGaP layer by some mechanism, the observed electron mobility should be 7000–8000 cm^2/V s, as routinely observed in high quality pure GaAs sample. The observed mobility value of 3300 cm^2/V s is much lower than this.

The Hall mobilities of the Si doped InGaP layers were smaller than those of the theoretical μ–n curve assuming $N_A/N_D=0.5$ as seen in Fig. 6. This is probably due to an increase of compensation ratio which took place with Si doping. As already mentioned in explaining Fig. 5, electron concentrations in InGaP were smaller than those in GaAs for the same Si cell temperature. This may be due to incorporation of acceptors during growth, thereby increasing the compensation ratios, whether Si themselves act partially as acceptors or some other acceptors are produced is not yet clear. However, although these mobility values of the Si doped InGaP samples are smaller than those by MOVPE, they were almost the same as those by GSMBE using PH_3 [9,10].

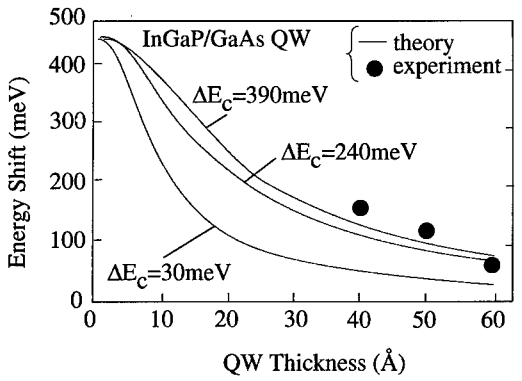


Fig. 8. Dependence of the position of the PL peak due to the InGaP/GaAs QWs on the QW thickness. Theoretically expected relationship between the PL peak position and the QW thickness assuming various values of ΔE_C are also shown.

3.3. Characteristics of $In_{0.48}Ga_{0.52}P$/GaAs heterointerface and band lineup

The PL spectrum of the InGaP/GaAs quantum well having a well thickness of $t_{QW}=4$ nm taken at 20 K is shown in Fig. 7. In Fig. 7, PL peaks around 1.52 eV and that at 1.92 eV can be assigned as emissions from the bulk GaAs and the InGaP barrier layers, respectively. Therefore, peaks observed around 1.7 eV seem to come from the GaAs QW layer itself. The sharp spectra indicate formation of a sharp heterointerface. Existence of the double peaks may be due to one mono-layer thickness variation in the GaAs QW. The PL spectra of the QW samples having a QW layer thickness t_{QW} of 5 and 6 nm also showed the peaks due to the bulk GaAs and to the InGaP barrier layers around 1.52 and 1.92 eV, respectively, whereas the position of the PL peak due to the GaAs QWs shift systematically to a lower energy side with increasing QW thickness as shown in Fig. 8. In Fig. 8, the theoretically expected relationships between the PL peak position and the QW thickness assuming various

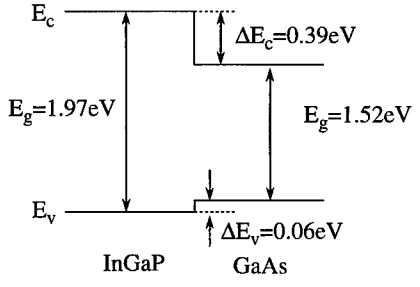


Fig. 9. Band lineup at InGaP/GaAs heterointerface with a large conduction band discontinuity value ΔE_C of 390 meV.

values of ΔE_C are also shown. As shown here, the observed dependence of the PL peak position on the QW thickness agrees reasonably with the theoretical curve calculated using a large conduction band discontinuity value ΔE_C of 390 meV as schematically illustrated in Fig. 9. This large value of ΔE_C was reported by Kodama et al. [11] by transport measurements for InGaP/GaAs single heterostructures grown by chloride VPE.

However, it should also be noted that an extremely large difference of 30–390 meV exists in the reported values of ΔE_C [11–16] and our data in this study agrees with the largest value exist in these reports. At present, we do not know why such a large variation of ΔE_C can take place in this system. One possibility is that ΔE_C of the InGaP/GaAs heterointerface is dependent on the growth method and conditions. Another possibility is that the incorporation of P into the GaAs QW layer may give rise to similar effects on the PL peak positions. Namely, it may shifts the PL peak energy to higher energy side either by forming GaAsP layer at the interface, which may reduce the effective QW thickness or by changing GaAs QW to GaAsP QW. (B) A third possibility is defect formation in the barrier layers during QW formation. Thus, a further detailed study is necessary to confirm the validity of the large ΔE_C value obtained here.

4. Conclusions

The feasibility of preparing device quality InGaP/GaAs heterostructures by GSMBE growth using TBP was investigated. The main conclusions are listed below:

1. In the GSMBE growth of InGaP using TBP, Si doping was possible up to the electron concentration of 2×10^{19} cm^{-3} without carrier saturation.
2. All of the undoped and doped InGaP layers showed electron mobilities comparable to those reported for InGaP layers grown by other methods.
3. The InGaP/GaAs QW samples showed intense and narrow PL emission lines, indicating the realization of high quality InGaP/GaAs heterointerface.
4. Observed PL peak positions of the QWs can be explained by assuming a large conduction band discontinuity value (ΔE_C) of 390 meV at the InGaP/GaAs interface. However, further work is necessary to confirm this.

References

[1] Mondry MJ, Kroemer H. IEEE Electron Device Lett 1985;6:175–7.
[2] Beam III EA, Henderson TS, Seabaugh AC, Yang JY. J Cryst Growth 1992;116:436–46.
[3] Kuo JM, Fitzgerald EA. J Vac Sci Technol B 1992;10:959–61.
[4] Ohba Y, Ishikawa M, Sugawara H, Yamamoto M, Nakanishi T. J Cryst Growth 1986;77:374–9.
[5] Sai H, Fujikura H, Hasegawa H. Jpn J Appl Phys 1999;38:151–8.
[6] Sai H, Fujikura H, Hasegawa H. Jpn J Appl Phys 1999;38:824–31.
[7] Besikci C, Razeghi M. IEEE Trans Electron Devices 1994;41:1066–9.
[8] Shitara T, Eberl K, Dickmann J, Wölk C. J Cryst Growth 1995;150:1261–5.
[9] Masselink WT, Zachau M, Hickmott TW, Hendrickson K. J Vac Sci Technol B 1992;10:966–8.
[10] Quigley JH, Hafitch MJ, Lee HY, Stave RE, Robinson GY. J Vac Sci Technol B 1989;7:358–60.
[11] Kodama K, Hoshino M, Kitahara K, Takikawa M, Ozeki M. Jpn J Appl Phys 1986;25:L127–L129.
[12] Garcia JC, Maurel P, Bove P, Hirtz JP. Jpn J Appl Phys 1990;30:1186–9.
[13] Nittono T, Sugitani S, Hyuga F. J Appl Phys 1995;78:5387–90.
[14] Hafich MJ, Quigley JH, Owens RE, Robinson GY. Appl Phys Lett 1989;54:2686–8.
[15] Kobayashi T, Taira K, Nakamura F, Kawai H. J Appl Phys 1989;65:4898–902.
[16] Rao MA, Caine EJ, Kroemer H, Long SI, Babic DI. J Appl Phys 1987;61:643–9.

PERGAMON

Solid-State Electronics 43 (1999) 1547–1553

SOLID-STATE ELECTRONICS

Potential profile measurement of GaAs MESFETs passivated with low-temperature grown GaAs layer by Kelvin probe force microscopy

K. Matsunami[a,*], T. Takeyama[a], T. Usunami[a], S. Kishimoto[a], K. Maezawa[a], T. Mizutani[a], M. Tomizawa[b], P. Schmid[c], K.M. Lipka[c], E. Kohn[c]

[a]*Department of Quantum Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan*
[b]*NTT System Electronics Laboratories, 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa, 243-0198, Japan*
[c]*Department of Electron Devices and Circuits, University of Ulm, Albert-Einstein-Allee 45, 89081, Ulm, Germany*

Abstract

Kelvin probe force microscopy (KFM) has proven to be an attractive method to measure the potential profile of the GaAs devices. We have applied the KFM technology to the GaAs MESFETs passivated with low-temperature (LT) grown GaAs cap layer, which was introduced to realize high breakdown voltage by lowering the electric field at the gate edge of the drain side. It was shown that the amplitude of the alternating voltage (V_{ac}) applied to the MESFETs to detect the electrostatic force between the devices and the tip gave little difference in the measured potential profile. High-field regions at the gate edge of the drain side was not so clear as the previous report on GaAs HEMTs. The role of the LT GaAs cap layer to relax the high-field at the gate edge was not confirmed because similar potential profile was obtained between those with and without LT GaAs cap layer. The obtained results were compared with the simulation results. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The shrinkage of semiconductor devices to the sub-micrometer level has led to the need for the direct measurement of two-dimensional potential profile with nanometer scale resolution. Recently Nonnenmacher et al. [1] have developed Kelvin probe force microscopy (KFM) for the surface potential measurement. The technology is based on the electrostatic force measurement between the tip and the sample surface. Contact potential difference was measured and reported on metal surfaces [1] and Si pn junctions [2]. Vatel applied this technique to the potential distribution measurement of a thin InGaAs resistor [3]. Two-dimensional dopant profiles of the Si pn junction were also reported [4,5]. Although KFM was proved to be effective in electrical characterization of the devices, there were few reports on the measurement of the cross-sectional potential profile of the GaAs devices under bias voltage. This information is very important for understanding both the transport properties and device physics, which are indispensable for the development of novel devices and improvement of the device performance.

Recently, we have reported that Kelvin probe force microscopy was successfully applied to the measurements of two-dimensional potential profile of the GaAs devices under bias voltage [6,7]. A high-field region was observed at the gate edge of the drain side.

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00102-1

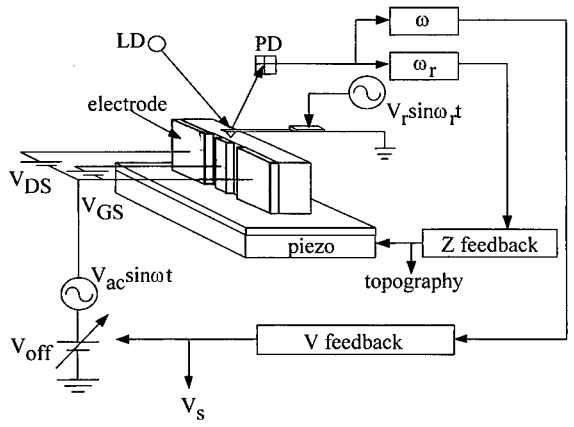


Fig. 1. Schematic diagram of the KFM measurement system.

The high-field region at the gate edge is responsible for the avalanche breakdown. A novel type of GaAs MESFETs using low-temperature (LT) grown GaAs cap layer as a low-conductive bypass has also been proposed to relax the high field at the gate edge, resulting in a high breakdown voltage [8].

In this paper, we report on the results of applying KFM technology to the GaAs MESFETs with LT GaAs cap layer. Two-dimensional potential profiles of the GaAs MESFETs with and without the LT GaAs cap layer will be compared.

2. KFM measurement system

Fig. 1 shows a schematic diagram of the KFM measurement system. The system has two feedback loops with deflected-light monitoring system: the Z feedback loop is used for the control of the z-axis position of the sample, and the other is used for the surface potential measurements. The operation principle of the KFM is based on the measurement of electrostatic force which arises when potential difference exists between the tip and the sample surface. An alternating voltage $V_{ac}\sin(\omega t)$ ($\omega = 8$ kHz) and a DC bias voltage V_{off} were applied to the sample holder to measure the surface potential V_S of the sample. The ω component of the electrostatic force is proportional to $V_{off} + V_S$. Then, when the DC bias voltage V_{off} is adjusted so that the amplitude of the ω component of the electrostatic force becomes zero, the sample surface potential can be determined as $-V_{off}$. The tip-to-sample distance is controlled using the ω_r force component similar to the noncontact-mode AFM. Hence a two-dimensional potential image and a topographic map can be obtained simultaneously. The measurements were carried out in an air ambient at room temperature. The measurement conditions are summarized in Table 1. A Au/Cr-coated Si cantilever with a tip radius of about 30 nm was used for the measurements. The tip-to-sample distance estimated from force curve was about 30 nm. The cantilever was driven by a piezoelectric bimorph transducer at a frequency of ω_r ($\omega_r = 20$–25 kHz depending on the cantilever), which is slightly above the resonant frequency of the cantilever.

Table 1
KFM measurement condition

Measurement temperature	Room temperature
Atmosphere	In air
Cantilever	Si tip (Au/Cr coat)
Tip radius	30 nm
Spring constant	1.5 N/m
Alternating voltage (V_{ac}, ω)	3 V, 8 kHz
Resonant frequency of the cantilever (ω_r)	23.6 kHz

3. Experimental

Fig. 2 shows a schematic cross-sectional view of the measured GaAs MESFETs with and without a LT GaAs cap layer (LT-MESFETs and C-MESFETs, hereafter). The epitaxial layer structure was grown by

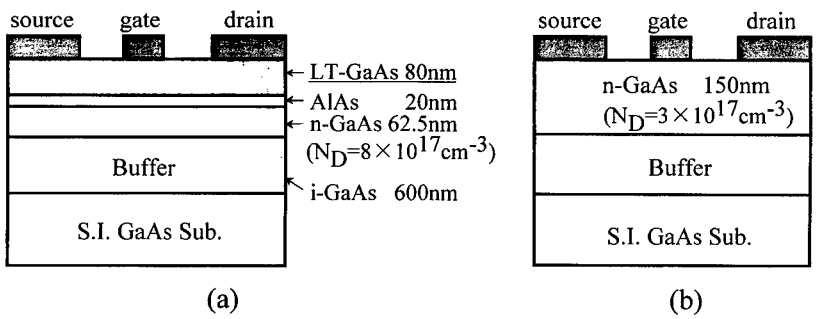


Fig. 2. Schematic cross section of the measured MESFETs (a) with and (b) without the LT GaAs cap layer.

Table 2
Breakdown voltages of GaAs MESFETs

	$V_{GS}=0$ V	$V_{GS}=V_{th}$
With LT	16 V	32 V
Without LT	12 V	25 V

MBE. The LT GaAs with 80 nm thickness was grown at 200°C. The post-growth anneal which is usually performed to obtain GaAs layer with high resistivity was not employed, resulting in a low-conductive layer. The layer is used as a low-conductive bypass with linear I-V characteristics [8]. It imposes an essentially constant lateral surface field, smoothing the lateral high-field in the channel. The AlAs layer with 20 nm thickness was grown to prevent the in-diffusion of excess As from the LT layer. The designed values of doping concentration and the thickness of the channel layer are 8×10^{17} cm^{-3} and 62.5 nm for the LT-MESFETs, and 3×10^{17} cm^{-3} and 150 nm for the C-MESFETs, respectively. The nominal gate length and gate-drain spacing were 1.2 and 2.4 μm, respectively. The source-drain breakdown voltages of the measured MESFETs are summarized in Table 2. They were 16 and 32 V at $V_{GS}=0$ V and $V_{GS}=V_{th}$ for the LT-MESFETs, which were larger than those for the C-MESFETs, 12 and 25 V, respectively. The larger breakdown voltage even with higher doping concentration in the channel of the LT-MESFETs may be due to the LT GaAs cap layer. KFM measurements were performed on cleaved surfaces of the GaAs MESFETs. Fig. 3 (a) and (b) show the drain and gate I-V characteristics of the LT- and C-MESFETs prepared by cleavage for the measure-

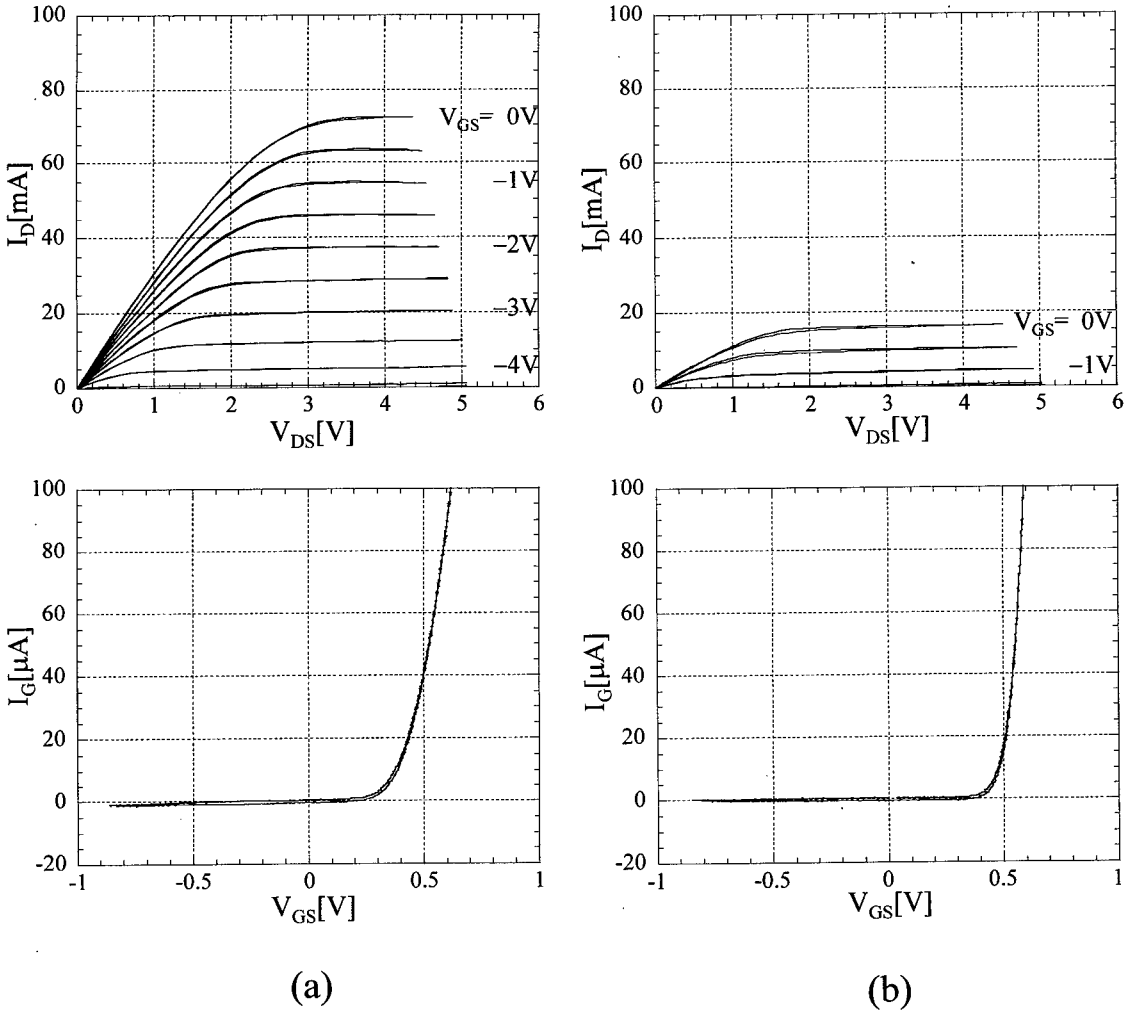


Fig. 3. Drain and gate I-V characteristics of the GaAs MESFETs (a) with and (b) without the LT GaAs cap layer.

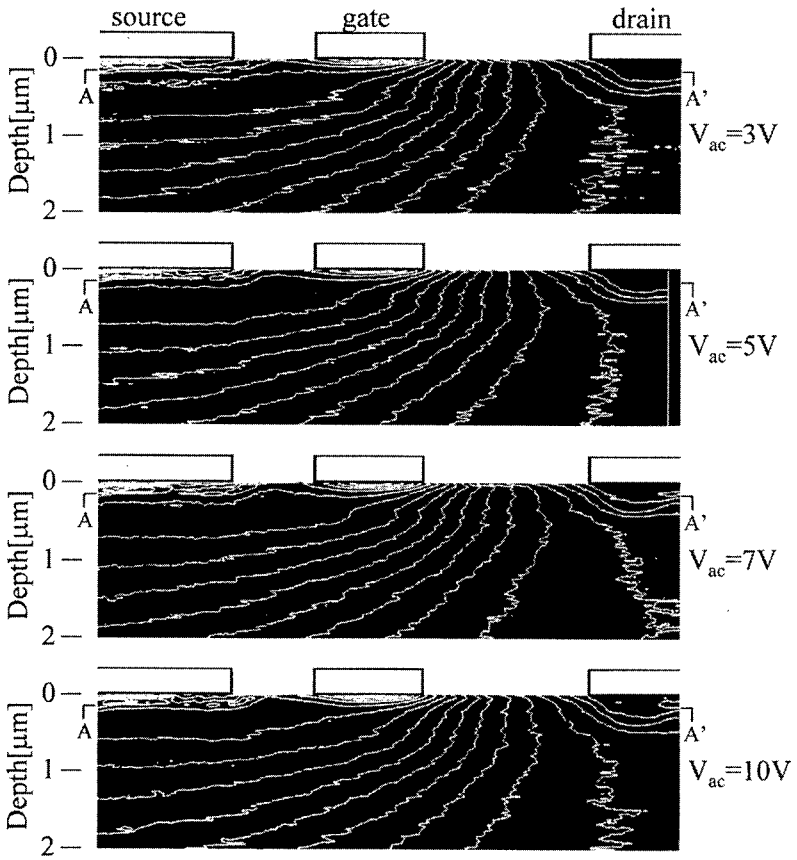


Fig. 4. The contour lines of the measured potential profile of the GaAs LT-MESFET for various V_{ac} of 3, 5, 7, and 10 V. Voltage step of contour line is 100 mV. The biased voltage are $V_{DS}=3$ V and $V_{GS}=-1$ V.

ments. The transconductances of the LT- and C-MESFETs were 144 and 83 mS/mm, respectively. Even though the reason for small g_m of the C-MESFETs is not so clear, the doping concentration and/or the channel thickness might be smaller than the designed values. It causes a large source resistance resulting in g_m decrease. The drain current ratios between those before and after the cleavage were 0.807 and 0.759 for the LT- and C-MESFETs, respectively. The ratios were almost equal to the ratios of the gate width before and after the cleavage, 0.793 and 0.732 for the LT- and C-MESFETs, respectively. The gate I-V curves do not show any extra leakage current, as shown in Fig. 3. These suggest that the cleavage did not cause any damage on the I-V characteristics of the MESFETs.

In order to study the effects of the amplitude of the alternating voltage V_{ac} on the measurements, the potential profile of the LT-MESFET was measured for various V_{ac}. The drain and gate voltages were 3 and -1 V, respectively. The measured two-dimensional potential image and the potential profile parallel to the device surface are shown in Figs. 4 and 5. Similar results are obtained even if the V_{ac} is changed from 3

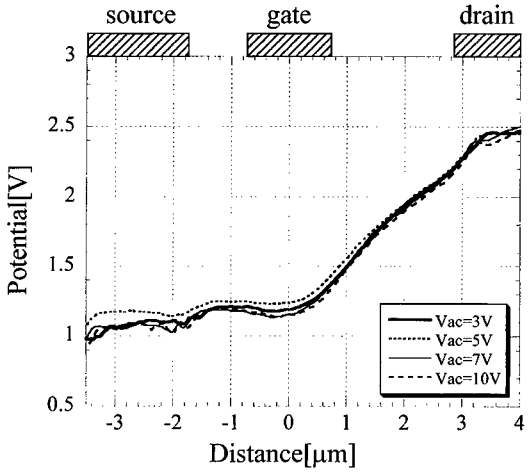


Fig. 5. Potential distribution parallel to the device surface along the AA′ line in Fig. 4 for various V_{ac}.

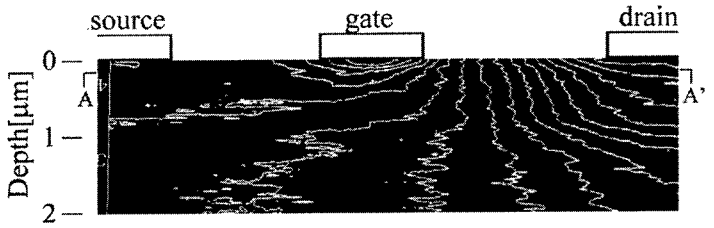


Fig. 6. The contour lines of the measured potential profile of the GaAs MESFET without the LT. Voltage step of contour line is 100 mV. The bias voltages are $V_{DS}=3$ V and $V_{GS}=-1$ V.

to 10 V, as shown in the figures. Then, the following experiments were performed at small V_{ac} of 3 V to avoid its effect on device operation. Even though high-density contour lines are observed at the drain side, the high field at the gate edge of the drain side is not so clear as the previous report on GaAs HEMTs [7]. The LT GaAs cap layer may contribute to relax the high field at the gate edge.

In order to confirm the role of the LT GaAs cap layer which was introduced to smooth lateral high field in the channel, a C-MESFET without LT cap layer was also measured for comparison. The measured potential image is shown in Fig. 6. V_{DS} and V_{GS} were 3 and -1 V, respectively. Similar image as that of the LT-MESFET is obtained. The potential distribution of the LT- and C-MESFETs parallel to the device surface is shown in Fig. 7. The electric field difference between the two is not so clear. According to the simulation, the maximum field of the LT-MESFET is lower by a factor of 3 than that of the C-MESFET [9]. This was explained by electron redistribution in the LT GaAs cap layer so as to linearize the lateral potential and lower the field at the gate edge of the drain side in the

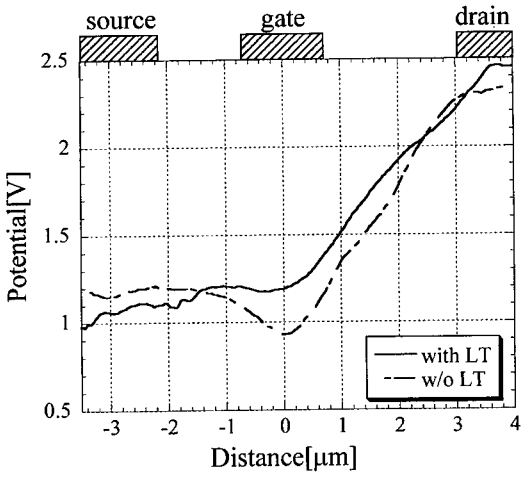


Fig. 7. Potential distribution along the surface of the MESFETs with and without the LT GaAs cap layer. The bias voltages are $V_{DS}=3$ V and $V_{GS}=-1$ V.

channel. The measured breakdown voltage of the LT-MESFETs was, on the other hand, not so high as expected from the simulation. It was about 30% higher than that of the C-MESFET. Then the difference in the maximum electric field between the LT- and C-MESFETs may not be so great. The fact that similar potential profiles parallel to the device surface were obtained for both types of devices means that it might be difficult to detect this amount of electric field difference by the present KFM.

In our previous study, the voltage resolution was less than 10 mV [6], which is sufficiently high to detect the electric field difference. The spatial resolution was estimated to be about 40 nm for the InAlAs/InGaAs heterostructures [10]. The resolution is dependent on the shape of the tip, cantilever and lateral inhomogeneity in the surface topography and potential [11]. This is because the V_{off} which is applied to minimize the electrostatic force, is the weighted average of the work function difference and capacitive components between the probe and the sample surface. Further improvement on the present KFM system will be possible by suppressing the effects of these parasitic coupling.

Fig. 8 shows cross-sectional potential images of the LT-MESFETs for different drain voltages of $V_{DS}=1$, 2, and 3 V. The gate voltage V_{GS} was -1 V. The MESFET was biased in a linear region at $V_{DS}=1$ and 2 V, and on a saturation point at $V_{DS}=3$ V. The density of contour lines increased gradually as the drain voltage was increased. Fig. 9 shows potential profiles along the surface of the MESFET. The increase in the drain voltage causes the electric field to increase at the drain side, in contrast with a small change in the potential profile at the source side. These phenomena qualitatively agree with the simulation results.

Fig. 10 shows the contour lines of the simulated potential profile of the GaAs MESFET without LT cap layer at saturation region ($V_{DS}=3$ V and $V_{GS}=-1$ V). If the high field is formed in a narrow region at the gate edge as shown in the figure, breakdown voltage does not strongly depend on the gate-drain length because almost all voltage drop occurs at the gate edge. However we often experience that the breakdown voltage is dependent on the gate-drain length. The po-

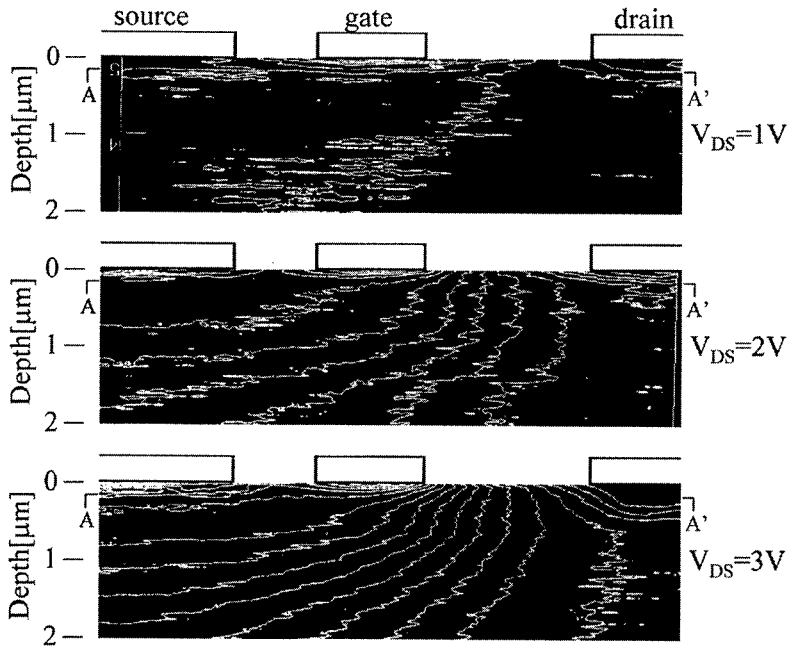


Fig. 8. The contour lines of the measured potential profile of the MESFETs with the LT GaAs cap layer for V_{DS} of 1, 2, and 3 V. Voltage step of contour line is 100 mV. The gate voltage is −1 V.

tential profile shown in Figs. 7 and 9 can explain the behavior of the breakdown voltage because the electric field is dependent on the gate-drain length. If minus charge exists at the device surface or the buffer layer, it will relax the high field at the gate edge.

The contour lines in Fig. 10 are perpendicular to the substrate surface because the simulation was performed under a boundary condition of floating substrate. The measured potential profile, on the other hand, shows different behavior. It shows a tendency to be parallel to the substrate surface, especially under the source electrode as shown in Figs. 4, 6 and 8. This is pronounced in the case of LT-MESFETs and may be related to the fact that the electrical characteristics of the buffer layer might be different between the LT- and C-MESFETs. Fig. 11 shows depth profiles of the measured potential under (a) source, (b) gate, and (c) drain electrodes. Under the drain electrode, the voltage decreases toward the substrate direction in agreement with the simulation results. Under the source and gate electrodes, on the other hand, it increases toward the substrate. This is equivalent to the fact that there is a downward electron flow, which is not allowed in

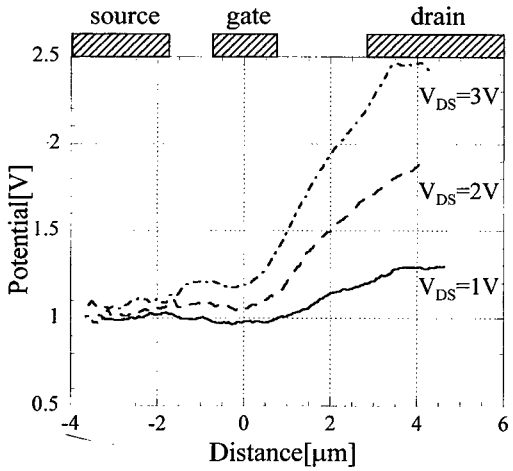


Fig. 9. Potential distribution along the AA′ line in Fig. 8 for various V_{DS}. The gate voltage is −1 V.

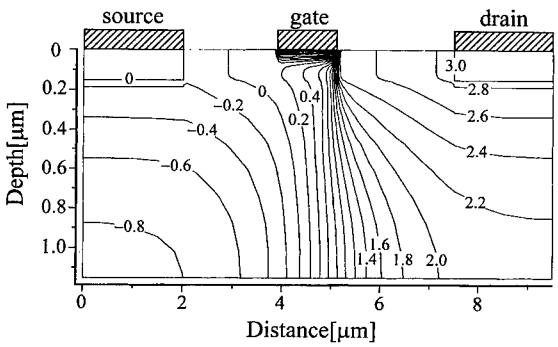


Fig. 10. The contour lines of the simulated potential profile of the MESFETs at saturation region (V_{DS}=3 V and V_{GS}=−1 V).

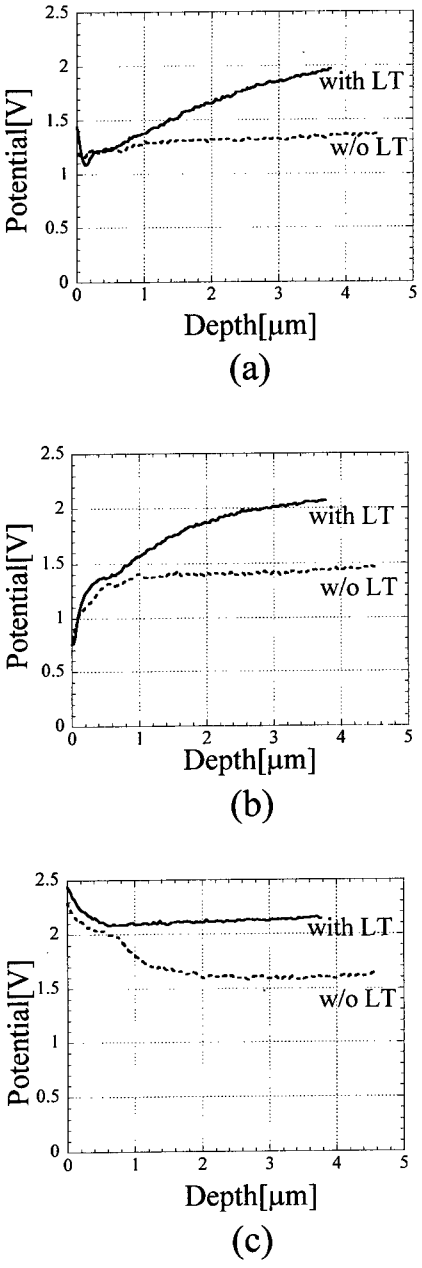


Fig. 11. Depth profile of the measured potential profile of the MESFETs with and without the LT GaAs cap layer under (a) source, (b) gate, and (c) drain electrodes. The bias voltages are $V_{DS}=3$ V and $V_{GS}=-1$ V.

devices without backside electrode. The tendency of the voltage increase toward the substrate is notable in the case of LT-MESFET. Lateral inhomogeneity in the surface potential might be responsible for this discrepancy in addition to the difference in the electrical characteristics of the buffer layer between the two devices. Further study is necessary to clarify this point.

4. Conclusion

Kelvin probe force microscopy (KFM) technology has been applied to the GaAs MESFETs passivated with low-temperature (LT) grown GaAs cap layer, which was introduced to realize high breakdown voltage by lowering the electric field at the gate edge of the drain side. It was shown that the amplitude of the alternating voltage (V_{ac}) applied to the MESFETs to detect the electrostatic force between the device and the tip, gave little difference in the measured potential profile. High-field regions at the gate edge of the drain side were not so clear as the previous report on GaAs HEMTs. Even though the LT GaAs cap layer may contribute to relax the high-field at the gate edge, the fact that a similar potential profile was obtained with and without LT GaAs cap layer suggests that further study is necessary to confirm these features. Comparison of the measured and simulated potential profiles suggests the existence of the minus charge at the device surface or the buffer layer which will relax the high field at the gate edge.

Acknowledgements

This work was supported in part by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Science and Culture, Murata Science Foundation, and the Grant of Hewlett Packard, Japan.

References

[1] Nonnenmacher M, O'Zboyle MP, Wickramasinghe HK. Appl Phys Lett 1991;58:2921.
[2] Nonnenmacher M, O'Zboyle MP, Wickramasinghe HK. Ultramicroscopy 1992;42(44):268.
[3] Vatel O, Tanimoto M. J Appl Phys 1995;77:2358.
[4] Henning AK, Hochwitz T. J Appl Phys 1995;77:1888.
[5] Tanimoto M, Vatel O. J Vac Sci Technol 1996;B14:1547.
[6] Arakawa M, Kishimoto S, Mizutani T. Jpn J Appl Phys 1997;36:1826.
[7] Mizutani T, Arakawa M, Kishimoto S IEDM Tech Digest, 31 1996.
[8] Lipka KM, Splingart B, Westphalen R, Kohn E, Theron D, Boudart B, Salmer G, Pond LL, Weitzel CE. In: IEEE Cornel Conference, Ithaca, August 7–9, 1995.
[9] De Jaeger JC ISSSE, San Francisco, CA, p. 111 1995.
[10] Usunami T, Arakawa M, Kishimoto S, Mizutani T, Kagawa T, Iwamura M. Jpn J Appl Phys 1998;37:1522.
[11] Hochwitz T, Henning AK, Levey C, Daghlian C, Slinkman J. In: Third International Workshop on the Measurement and Characterization of Ultra-Sallow Doping Profiles in Semiconductor, March 20–22, 1995.



PERGAMON

Solid-State Electronics 43 (1999) 1555–1560

SOLID-STATE ELECTRONICS

Photoemission studies on heterostructure bipolar transistors

Fritz Schuermeyer[a,*], Peter J. Zampardi[b], Peter M. Asbeck[c]

[a]*Air Force Research Laboratory, AFRL/SNDD, Wright Patterson AFB, Dayton, OH 45433-7322, USA*
[b]*Rockwell Science Center, 1049 Camino Dos Rios, Thousand Oaks, CA 91360, USA*
[c]*University of California, San Diego, Department of ECE 0407, 9500 Gilman Dr, La Jolla, CA 92093, USA*

Abstract

We apply in-situ photoemission techniques to characterize the band profile in heterostructure bipolar transistors. The measurements are performed on-wafer on fully fabricated InP based HBTs at room temperature. We demonstrate that this technique is sensitive to detect and analyze barriers between emitter and base as well as between base and collector. Furthermore, the photoresponse provides information on the doping characteristics of the active layers. Published by Elsevier Science Ltd.

1. Introduction

Heterostructure bipolar transistors (HBTs) are being developed for applications where speed, power, efficiency, and low noise are of importance. These transistors can also be used as fast photodetectors [1]. Especially, III-V compound semiconductor based HBTs are of interest due to their high electron velocities and mobilities. A multitude of III-V compounds exist which can be grown lattice matched or strained to form heterostructures. In these transistors, the base layers are chosen such that their bandgap is smaller than that of the emitter [2] to minimize carrier injection from the base to the emitter, which is a parasitic current. Conduction and valence band discontinuities exist at the heterointerface. The conduction band discontinuity creates an electron barrier between emitter and base. A similar barrier may exist between base and collector in double heterostructure HBTs (DHBTs). These barriers have negative effects on the device characteristics. The emitter base barrier increases the threshold voltage in n-p-n HBTs and reduces the uniformity. A barrier between base and collector leads to charge storage in the base, limiting the maximum output current. To minimize or eliminate the barriers, grading layers are often employed to obtain a smooth transition of the bands. These grading layers have to be very thin to retain low emitter or collector resistances. Frequently, one or both of these emitter, base, or collector materials are ternary compound semiconductors, and the grading layers need to be quaternary compounds. Furthermore, both the emitter and base layers are relatively heavily doped and the grading layer can only be lightly doped. Hence, diffusion of one or both of the dopants into the grading layer will negate its function. The growth of such a non-uniform, thin layer is difficult to accomplish and more analytical techniques are needed to assure that no barrier exists at the heterojunction.

Barriers to electronic carriers play an important role in most semiconductor devices and have been evaluated extensively [3]. For example, the Schottky contact to an n-type semiconductor represents such a barrier. Several techniques exist to analyze the height of the barriers, such as the evaluation of the temperature dependence of the current voltage behavior of diodes, capacitance voltage studies, and photoemission

* Corresponding author. Fax: +1-937-255-2306.

0038-1101/99/$ - see front matter Published by Elsevier Science Ltd.
PII: S0038-1101(99)00103-3

measurements. In the latter technique, the Schottky contact is illuminated with monochromatic light. The light is either incident on the backside of the structure or on the front side if the metal layer is thin. The light is absorbed in the Schottky metal, generating hot carriers. If the energy of these carriers is sufficiently large, they will be able to pass over the barrier and be collected, creating a photocurrent. At very low photon energies, the carriers have insufficient energy to cross the barrier. Hence, the threshold energy for the onset of the photocurrent gives a direct measure of the barrier height.

In this paper we apply the photoemission technique to evaluate heterojunctions in fully fabricated InP based HBTs to assess if barriers exist. Results obtained on transistors from two wafers are reported. The emitter base structure is nominally identical in both structures. One wafer contained a base collector heterojunction (DHBT) and the other a homojunction (SHBT). These were experimental wafers, stemming from a growth optimization study. Even though the base emitter junctions were designed to be identical for both wafers, their electrical behavior is quite different, suggesting that they possess dissimilar energy configuration due to variations in growth parameters.

2. Photoemission measurements

The photoemission studies are conducted on fully fabricated InP based HBTs. The measurements are performed on-wafer using backside illumination. Backside illumination is important since the front side is covered with metals, shading the active areas of the transistor. Backside illumination is feasible since the bandgap of the InP substrate is larger than that of the active transistor area. Hence, the substrate is transparent in a photon energy range of interest for these measurements. The measurements are performed on the base emitter diodes and the base collector diodes and the unused terminals are left floating.

Fig. 1 shows schematically the energy diagram of an emitter-base heterojunction. An InAlAs emitter and the InGaAs base layers have bandgaps of approximately 1.5 and 0.7 eV, respectively. Shown is also a barrier in the conduction band between the emitter and the base. Such a barrier occurs in a device without grading layer due to the conduction band discontinuity. Also, incorrect doping profiles at the junction or improper growth of the grading layer may lead to such a barrier. Monochromatic light is incident on the backside of the transistor and reaches the diode if the photon energy is less than 1.3 eV, the bandgap of the InP substrate. Since the bandgap of the emitter is larger than that of the substrate, band to band photoexcitation cannot take place in this layer. Light is absorbed

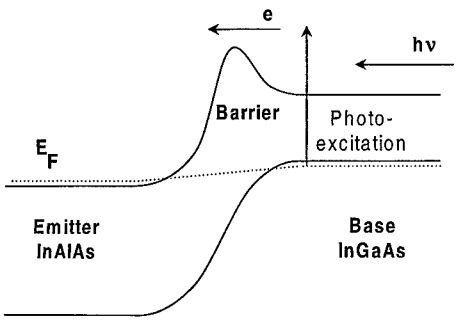


Fig. 1. Band profile of emitter base junction of an InGaAs/InAlAs/InP HBT.

in the base layer and the photogenerated electrons may be emitted into the emitter layer. Consequently, a photocurrent flows between the base and emitter contacts. As indicated in Fig. 1, only photogenerated electrons with sufficient kinetic energy may be emitted. The threshold photon energy is a measure of the energy of the barrier, referenced to the top of the valence band in the base layer. If this barrier is successfully eliminated by the use of a grading layer, the threshold photon energy will be equal to the InGaAs bandgap (0.7 eV). Similar considerations hold for the base collector junction.

The experimental setup has been described previously [4,5]. The measurements are performed on a probe station at room temperature. An Acton 150 monochromator provides chopped, monochromatic light to the backside of the tested device via an optical fiber cable. One end of the optical fiber is mounted in a hole, located in the center of the wafer chuck. The frequency of the mechanical chopper can be varied between 5 and 3000 Hz. In our measurements, the base of the HBT is kept at ground potential. The emitter or collector terminals are connected to the input of a Stanford Research System SR570 low noise current preamplifier. This amplifier provides the bias voltage to the electrode. The input impedance of the amplifier depends on the sensitivity setting and is less than 100 Ω. The output of the amplifier is fed to a lock-in-amplifier (Stanford Research System SR530), tuned to the chopper frequency and to a digital voltmeter. The SR530 measures the ac current due to the chopped light and the digital voltmeter the dc current. All instruments are computer controlled.

3. Experimental results on DHBT

Measurements were performed on double heterostructure bipolar transistors (DHBTs). The device structure is shown in Table 1. The collector consists of layers 2–5. Layers 7, 8, and 9 make up the base, and

Table 1
DHBT structure

Layer no.	Layer description	Concentration cm^{-3}	Thickness nm
13	n+InGaAs	4×10^{19}	120
12	n+InAlGaAs grading	5×10^{18}	7.5
11	n-AlInAs	5×10^{17}	25
10	i/n-InGaAlAs grading	5×10^{17}	10
9	i InGaAs spacer	–	15
8	p+ InGaAs	5×10^{19}	80
7	n-InGaAs spacer	5×10^{16}	20
6	n-InAlGaAs grading	5×10^{16}	10
5	n-AlInAs	5×10^{16}	300
4	n-AlInAs	5×10^{16}	250
3	n+AlInAs	5×10^{18}	50
2	n+InGaAs	2×10^{19}	400
1	i InGaAs	–	50
0	i InP	–	–

the emitter is composed of layers 11–13. Emitter, base, and collector consist of multiple layers to provide the desired diode characteristics as well as low sheet resistances. Layers 7 and 9 are undoped spacers to minimize the diffusion of the p+doping from the heavily doped base into the grading layers. Of importance are the AlInGaAs grading layers 10 and 6 at the emitter base junction and at the base collector junction respectively. The purpose of these layers is to eliminate the barrier due to the conduction band discontinuities at the AlInAs/GaInAs heterojunctions.

Fig. 2 shows the current voltage characteristics of a DHBT with an emitter area of 5000 μm^2. The transistors have excellent performance with an off-set voltage of 0.1 V and a gain of approximately 20. The base current is stepped in increments of 100 μA. The base to emitter voltage is 0.65 V at a base current of 400 μA, consistent with the narrow bandgap base layer and the successful implementation of the emitter to base grading layer. The Gummel plots reveal ideality factors of 1.1 and 1.2 for the collector and base currents, respectively.

Fig. 3 presents the photoyield of the base emitter and base collector junctions as a function of photon energy. The photoyield represents the photocurrent divided by the number of incident photons. The monochromator was calibrated using a Hilger and Watts bolometer and Ge and Si detectors. As expected, the photoyield is negligible for photon energies above 1.3 eV since the InP substrate absorbs light in this region. The substrate is transparent for smaller photon energies and the spectral characteristics are obtained. The measurements are performed at zero applied bias voltage. Due to the built in-voltage (see Fig. 1), photogenerated electrons flow from the base to the emitter or collector. Photogeneration does not take place in the emitter or collector layers since their bandgaps are

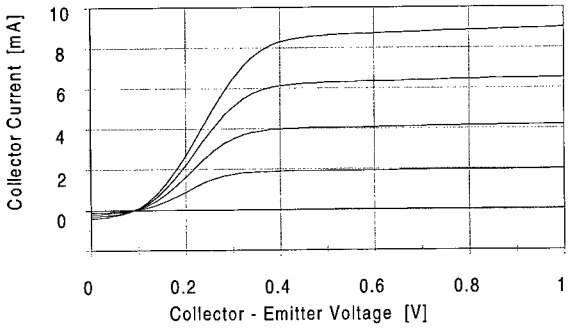


Fig. 2. Current voltage characteristics of DHBT with 100 μA base current steps.

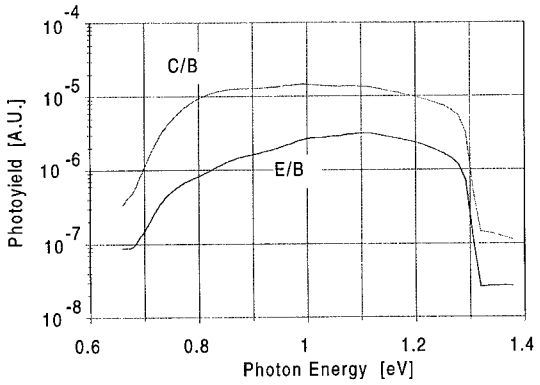


Fig. 3. Photoyield for emitter base and base collector junctions of a DHBT vs photon energy. The yield is shown in logarithmic format.

too large. Holes, photogenerated in the base cannot escape towards the emitter or collector due to the built-in voltage. One observes a threshold photon energy of approximately 0.7 eV for both the emitter base and the base collector junctions. This value is consistent with photoexcitation in the InGaAs base with a bandgap of 0.7 eV. These results are proof that no conduction band barrier exists either at the emitter base or the base collector junction. Such a barrier would result in a larger threshold photon energy. The result is consistent with the low base to emitter voltage, reported above and the excellent Gummel ideality factors. The absolute values and the shapes of the two curves are different. One reason for the observed difference is that the area of the base collector junction is approximately three times as large as the emitter base junction area. Also, the light intensity close to the emitter base junction is smaller than at the base collector junction since the light travels through the absorbing base layer. Finally, photoelectrons are not only generated in the p^+ base layer but also in the adjacent grading layers. These layers are not identical for the two junctions. Of interest is the drop-off at the band-edge energy. A drop-off value of approximately 60 meV/decade for both junctions is observed. A similar drop-off is seen in the optical absorption characteristics of III-V semiconductors [6]. This similarity is expected since the optical absorption in these materials causes photoexcitation. This drop-off depends strongly on the purity of the material. High purity GaAs and InAs have drop-off values of approximately 10–15 meV/decade at room temperature while this value becomes larger for heavily doped material. The larger drop-off value in photoyield in Fig. 3 is consistent with photoemission from the heavily doped base layer [7].

Fig. 4 presents the emitter base photocurrent as a function of emitter base voltage (V_{eb}) for two photon energies, 0.9 and 1.24 eV. At the 0.9 eV photon energy, one observes constant photocurrent over a wide range of applied voltages. The current remains constant without change in polarity even at $V_{eb} = -0.1$ V. This result is expected since photoexcited electrons are emitted from the base to the emitter. The built-in voltage exceeds the applied voltage. The base is heavily doped and the thickness of the depletion layer is insignificant. At an emitter base voltage of approximately −0.2 V, the photocurrent changes polarity. We assume that this current stems from photoexcitation of majority carriers in the emitter area, which are emitted into the base layer. At $V_{eb} < -0.5$ V, the photocurrent changes polarity again and becomes positive. We assume that the photocurrent in this regime stems from the modulation of the forward diode current. At the photon energy of 1.24 eV, the photocurrent changes by approximately a factor of two when the voltage is increased from 0 to 1 V. The reason for this effect is that photoexcitation takes place also in the grading layer. The energy configuration in this layer changes with bias voltage, causing the change in photocurrent with applied bias voltage.

Also shown in this figure are the dc currents, measured simultaneously with the ac photocurrents. As expected, the dc current is approximately half of the ac current. At negative V_{eb} the diode is forward biased and the magnitude of the dc current increases rapidly.

Fig. 5 presents similar results for the collector base junction. At 0.9 eV photon energy, the photocurrent is independent of applied bias voltage over a large voltage regime. This finding is consistent with photogeneration in the base and emission of photogenerated electrons into the collector. At a forward bias of $V_{cb} = -0.3$ V the current starts to decrease. This value of the forward bias is even larger than that found in the emitter base junction since the collector doping is significantly smaller than the emitter doping. Also, the

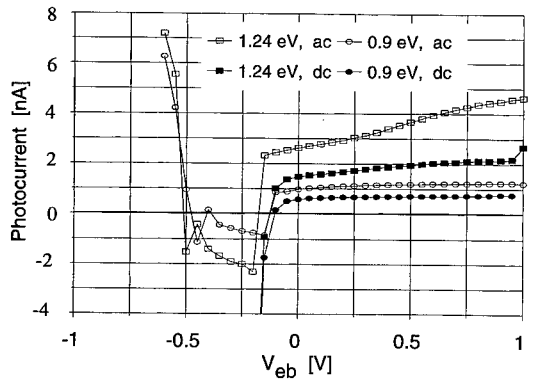


Fig. 4. ac photocurrent and dc current of an emitter base junction of DHBT vs emitter to base voltage.

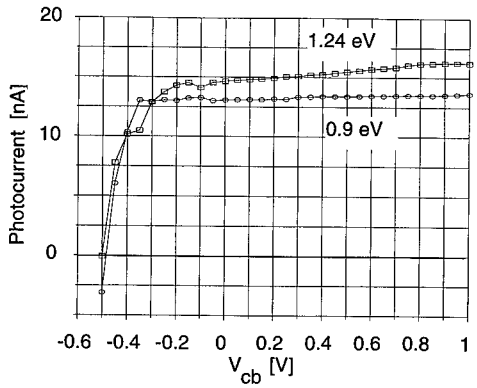


Fig. 5. Photocurrent of a collector base junction of DHBT vs collector to base voltage.

photocurrent, measured at 1.24 eV depends much less on applied voltage for the base collector junction than for emitter base junction. Again, the cause for this effect is the much lower collector doping in comparison to the emitter doping. The dc dark current is orders of magnitude larger than the ac photocurrent and is not shown in this graph.

In studies on HEMTs, it was observed that the photoresponse depended strongly on chopper frequency due to trapping of carriers [8]. Hence, it is important to investigate if similar charge trapping occurs in HBTs. The dependence of the photocurrent on chopper frequency is studied over a frequency range from 5 Hz to 2.4 kHz. The experiments show that in both the emitter base and the base collector junctions the photocurrent is independent of the chopper frequency. This is an indication that trapping effects do not play a role in the photomeasurements at this frequency range. All data presented in this paper are obtained at an 80 Hz chopper frequency.

4. HBTs with a barrier between emitter and base

In this section we describe photoemission results obtained on HBTs which contain a barrier at the emitter base junction. The devices studied are SHBTs. The base emitter voltages of these devices exceed those of similar DHBTs by more than 250 mV, measured under the same bias conditions. Also, the Gummel plot shows less ideal behavior in comparison with the DHBT with ideality factors of 2.10 and 2.11 for the collector and base currents, respectively. This finding suggests that a barrier of approximately 250 meV exists in these HBTs. Even though the emitter base junction has the same layer structure as the DHBT, differences in growth conditions may have resulted in an incomplete elimination of the barrier. Fig. 6 presents the photoemission characteristics for the emitter base junction of both the SHBT and the DHBT in linear representation. Extrapolation of the photoyield to zero results in energies of approximately 900 and 700 meV for the SHBT and DHBT, respectively. To obtain the exact threshold energy, the correct dependence of photoyield on photon energy needs to be established. The linear presentation gives only an approximation of this energy. The SHBT result indicates that a barrier exists at the emitter base junction with a height of 200 meV relative to the bottom of the InGaAs conductance band in the base, in good agreement with the electrical result. The values of this barrier obtained by the two methods differ slightly since the measurements are made at different bias voltages and the barrier depends on this voltage.

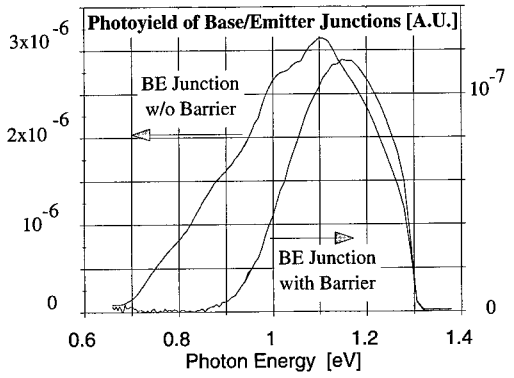


Fig. 6. Photoyield spectra of SHBT and DHBT emitter base junctions. The difference in threshold energy for the two devices indicates that a 200 meV barrier exists in the SHBT junction.

5. Photocurrents in homojunctions

The base collector junction of an SHBT is an InGaAs homojunction. The junction consists of the heavily p-doped base, a lightly n-doped collector area followed by an n^+-layer for low access resistance. Due to the built-in and applied voltages, a depletion layer exists. This layer is located mainly in the low-doped part of the collector. Carriers, photogenerated in this layer are separated due to the electric field and are detected as a photocurrent. In addition, minority carriers, photogenerated in the base and collector may move across the depletion layer and contribute to the photocurrent. Fig. 7 presents the spectral characteristics of an SHBT base collector junction, presented in logarithmic format. We observe that the on-set of the photoyield at the bandedge is very steep, approxi-

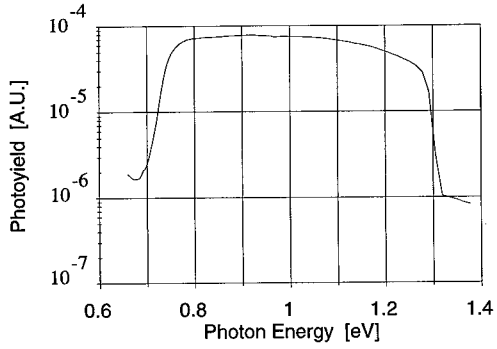


Fig. 7. Photoyield spectrum of a base collector homojunction. The yield is graphed on a logarithmic scale.

mately 30 meV/decade, consistent with photoexcitation in a low doped, high quality material. An approximation of the bandgap is obtained by extrapolating the photoyield in linear format to a value of zero. A bandgap of approximately 0.7 eV is obtained in good agreement with the DHBT data.

6. Summary and conclusions

The in-situ photoemission technique is ideally suited to study the existence of barriers in HBTs. These barriers may exist due to band discontinuities. These barriers degrade the performance of the transistors and an effort is made to eliminate such barriers. We have applied photoemission techniques to the evaluation of the band profile of fully fabricated InP based HBTs. The measurements are performed on-wafer, using backside illumination. Employing this technique, we have successfully demonstrated the existence of a barrier between emitter and collector in one type of device. We also proved that such barriers were successfully eliminated both in the emitter base junction and in the base collector junction in another set of devices. The measurements also provide bandgap and material quality information.

References

[1] Pedrotti KD, Pierson Jr RL, Sheng NH, Nubling RB, Farley CW, Chang MF. High-Bandwidth OEIC Receivers using Heterojunction Bipolar Transistors: Design and Demonstration. IEEE J Lightwave Technology 1993;11:1601–14.

[2] Shur M and references therein. In: Physics of semiconductor devices. NJ: Prentice Hall, 1990. p. 311–9.

[3] Sze SM. In: Physics of semiconductor devices. New York: Wiley-Interscience, 1969. p. 404–9.

[4] Schuermeyer FL, Loehr JP, Sherriff RE, Cerny C, Shur M. Photoelectric measurements of interband transitions in fully fabricated pseudomorphic high electron mobility transistors. In: Inst Phys Conf Ser No 145: Chapter 5, 1996. p. 791–6.

[5] Schuermeyer FL, Cerny C, Shur M. In-situ characterization of lateral and vertical band structure profiles and hole storage effects in PHEMTs by the photoconduction technique. In: Inst Phys Conf Ser No 155: Chapter 6, 1997. p. 495–8.

[6] Levinshtein M, Rumyantsev S, Shur M. In: Handbook series on semiconductor materials, vol. 1. Singapore: World Scientific Publishing, 1996. p. 96.

[7] Schuermeyer F, Zampardi PJ, Asbeck PM. InP-based HBTs for optical signal detection. Electrochemical Society Proceedings 1998;98-2:19–28.

[8] Schuermeyer FL, Cerny C, Bozada C, Fang Z-Q, Look DC. Charge storage effects in pseudomorphic high electron mobility transistors. Jpn J Appl Phys 1997;36:1330–4.



PERGAMON

Solid-State Electronics 43 (1999) 1561–1570

SOLID-STATE ELECTRONICS

In-situ characterization technique of compound semiconductor heterostructure growth and device processing steps based on UHV contactless capacitance-voltage measurement

Hiroshi Takahashi[a,*], Toshiyuki Yoshida[a], Morimichi Mutoh[a], Takamasa Sakai[b], Hideki Hasegawa[a]

[a]*Research Center for Interface Quantum Electronics and Graduate School of Electronics and Information Engineering, Hokkaido University, Sapporo, 060-8628, Japan*

[b]*Dainippon Screen Manufacturing Co Ltd, Kyoto, 612, Japan*

Abstract

In this paper, the applicability of an ultra-high vacuum (UHV) contactless capacitance-voltage (*C-V*) measurement method is evaluated for use in non-destructive characterization of crystal growth and device processing steps for the compound semiconductor heterostructure microelectronics. The UHV gap length between the sample surface and the field plate was made by an optical measurement using the Goos–Haenchen effect, and its accuracy was discussed by a theoretical calculation.

Then, zero-bias position of the surface Fermi level, conduction type and doping were successfully determined on free surfaces of GaAs and InP by basic *C-V* measurements.

Finally, the technique was applied to characterize each step of the UHV-based surface passivation process for InP utilizing the Si interface control layer (Si ICL). Change of the surface state distribution after each step was successfully detected without breaking UHV, demonstrating the powerful capability of the method. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

A great deal of effort is now being made to push the semiconductor device fabrication technology into the nanometer range, not only for Si CMOS gigabit to terabit ultra large scale integrated circuits (ULSIs), but also for advanced III-V compound semiconductor devices used in heterostructure microelectronics. This is mainly because only continued scaling down of the device feature sizes provides chances for III-V high-speed or high-frequency devices to compete with or dominate Si counterparts in terms of the speed, *rf* performance, noise and power consumption. Another drive for nanofabrication is the creation of quantum devices including quantum wave devices and single electron devices. Here, realization of perfect quantum structures such as quantum wires and dots with nanometer feature sizes is becoming increasingly important.

In the compound semiconductor heterostructure

* Corresponding author. Tel.: +81-11-706-7172; fax: +81-11-716-6004.

E-mail address: takahasi@rciqe.hokudai.ac.jp (H. Takahashi)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00104-5

microelectronics, devices including conventional and quantum devices normally consist of multitudes of interfaces. They include lattice-matched or coherently strained heterointerfaces utilizing material systems such as AlGaAs/GaAs and InAlAs/InGaAs/InP. Furthermore, they include various metal-semiconductor interfaces for Schottky and ohmic contacts, and insulator-semiconductor interfaces for surface passivation.

As the device feature sizes are reduced into the nanometer range, properties of each interface, as well as the effect of each processing step applied to the multi-interface structure, become extremely important. Because of this, in-situ non-destructive characterization of the effect of each heterostructure growth and processing steps on the electronic properties of the device structure becomes indispensable for optimization of device fabrication.

Recently, various surface/interface characterization techniques such as X-ray photoelectron spectroscopy (XPS), Auger electron spectroscopy (AES), scanning tunneling microscope (STM), ion scattering spectroscopy (ISS) have made great progress. However, there exists no well-established method for non-destructive characterization of the electronic properties of the 'grown' and 'processed' semiconductor surfaces. Conventional electronic characterization techniques are destructive methods because they need the preparation of Schottky or MIS (gate) structure to cause electrical modulation and ohmic electrodes for current supply.

The purpose of this paper is to describe a novel, non-destructive characterization technique and system based on the ultra-high vacuum (UHV) contactless capacitance-voltage (C-V) measurement. This technique seems to be extremely powerful for characterization of compound semiconductor growth and device processing steps. In this technique, a constant UHV gap is maintained between the sample surface and the field plate, using a piezoelectric mechanism based on capacitance measurement and feedback, and a standard C-V measurement is made, changing the field plate voltage. The length of the UHV gap is measured optically using the Goos–Haenchen effect.

After the description of the measurement method and system, a theoretical calculation was made for the optical measurement of the UHV gap for GaAs and InP surfaces. Then, basic C-V measurements were made on free surfaces of GaAs and InP, confirming the validity of the optical gap measurement as well as demonstrating the capabilities of the system for non-destructive determination of conduction type, doping level and zero-bias position of the surface Fermi level.

Finally, this technique was applied to characterize each step of the UHV-based surface passivation process for InP utilizing the Si interface control layer (Si ICL). Change of the surface state distribution after each step was successfully detected without breaking UHV, demonstrating the powerful capability of the method.

The present technique is also very powerful for characterization of UHV processing of Si wafers as discussed elsewhere [1,2].

2. Description of UHV contactless C-V measurement system

2.1. UHV contactless C-V system

The principle of the UHV contactless C-V method is to carry out C-V measurements from a field plate that is separated from the sample surface by a constant UHV gap. In the present study, extensive efforts were made to modify an air-gap-based commerical C-V measurement system (CV-8000, Dainippon Screen Mfg. Co. Ltd) [3] into a UHV-compatible system with a suitable UHV sample transfer mechanism. A schematic view of the completed UHV C-V chamber is shown in Fig. 1(a). It contains a movable sample stage and a C-V sensor head. The C-V sensor head is the most important part of the system, and described in the next section. It is connected to a computer for

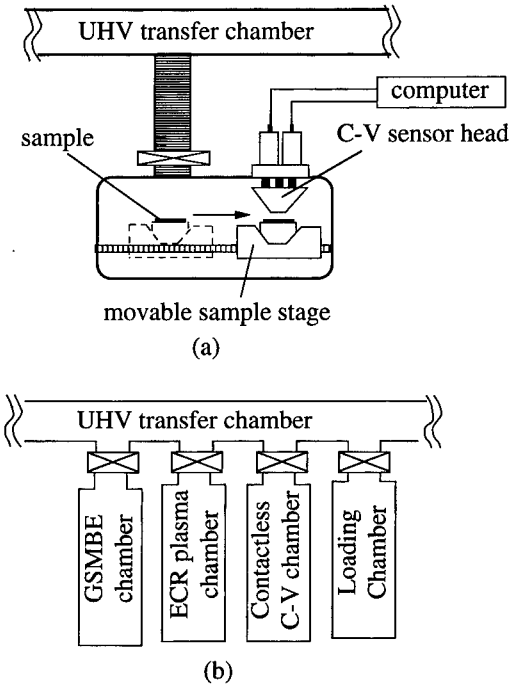


Fig. 1. (a) A schematic view of contactless C-V measurement system connected to the UHV formation/analysis system; (b) UHV-based formation/analysis system used for fabrication and characterization of samples.

feedback control and measurement. A special UHV transfer mechanism including a movable sample stage transfer was developed to transfer the grown or processed sample from other chambers into the UHV *C-V* chamber, and to place exactly under the *C-V* sensor head. In this study, the sample was attached on a standard molybdenum sample holder usually used for molecular beam epitaxy (MBE) growth. Care was taken to make all the parts 'bakable' to achieve the base pressure of 8×10^{-10} torr. The system was carefully protected from the vibration because variations of the UHV gap caused by external vibrations directly result in noises in the *C-V* measurements. For this purpose, the *C-V* chamber was placed on an anti-vibration support stage which had a weight of several hundred kg and a resonance frequency of several Hz. The system developed in this study can be attached to various growth and processing equipment for electronic characterization of each processing step.

In this study, the completed UHV contactless *C-V* system was connected to a UHV-based interface formation facility equipped in the super-clean room of the Research Center for Interface Quantum Electronics, Hokkaido University. Here, a gas-source molecular beam epitaxy (GSMBE) growth chamber, an electron cyclotron resonance (ECR) plasma chamber and another ten chambers are connected by a UHV transfer chamber as partly schematically shown in Fig. 1(b).

2.2. C-V sensor head

Fig. 2(a) shows a schematic view of the actual set-up used in this study for the UHV *C-V* measurement using a specially developed *C-V* sensor head. In the *C-V* sensor head, a ring shaped field plate is attached to an optical prism for the gap length measurement explained later, and is protected by a thick SiO_2 film. The electrode pattern at the bottom of the sensor head is shown in Fig. 2(b). In addition to the central field plate electrode of a ring shape, three surrounding 'parallelism' electrodes were formed to provide three additional parallelism capacitors with respect to the sample surface. The field plate had a diameter of about 1 mm ϕ and the bottom surface of the entire sensor probe had a diameter of about 2.5 mm ϕ. Chromium was used as the electrode material and the exact area of the field plate electrode was 7.5×10^{-3} cm^2.

During *C-V* measurement, parallelism between the field plate and a sample surface should be maintained at a constant UHV gap for a few 100 nm. This was accomplished through capacitance measurements of three parallelism capacitors and subsequent mechanical adjustment by a piezoelectric mechanism. Namely, the sensing probe was mechanically supported by three

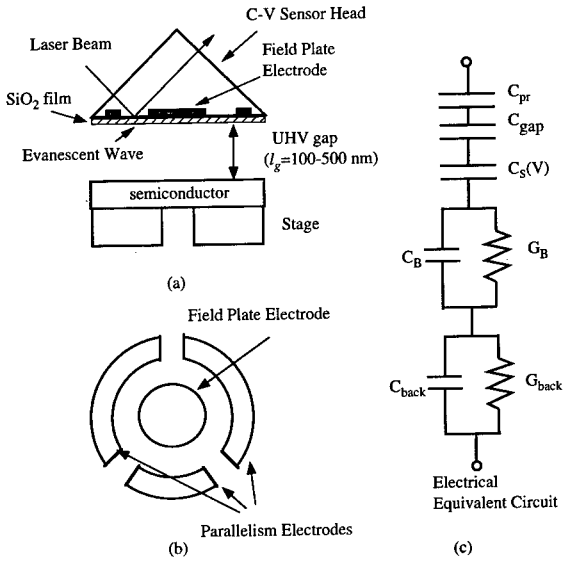


Fig. 2. Schematic illustration of (a) the sensor head, (b) field plate and parallelism electrodes and (c) electrical equivalent circuit of the system.

piezoelectric actuators and three mechanical screw actuators driven by the stepping motors. The stepping motors and the actuators were automatically controlled by the computer which provided suitable feedback on the basis of the capacitance measurement of three parallelism capacitors so that three capacitance values became equal and took an appropriate value. For achievement of planar gaps of a few 100 nm, the flatness values of the sample surface and the prism can become crucial. In this study, *rms* (root mean square) flatness values of the wafer and the prism used here were typically 3–4 nm and 6–7 nm over the entire field plate area.

2.3. Equivalent circuit and optical gap length measurement

The capacitance was measured at 500 kHz, using an *rf* bridge circuit. The electrical equivalent circuit of the measurement system is shown in Fig. 2(c). The circuit contains the capacitance of the protecting SiO_2 layer, C_{pr}, the UHV gap capacitance, C_{gap}, the voltage dependent semiconductor surface capacitance, $C_s(V)$, the semiconductor bulk capacitance and conductance, C_B and G_B, and the capacitance and conductance of the back contact, C_{back}, and G_{back}. In order to perform *C-V* measurements that can be analyzed by the standard metal-insulator-semiconductor (MIS) theory, a further simplification of the equivalent circuit is necessary. First, in doped semiconductors with the doping

levels larger than 10^{15} cm^{-3}, C_B and G_B/ω (ω: angular frequency of measurement, $2\pi \times 5 \times 10^5$ rad/s) are much larger than C_s as in the conventional MIS *C-V* measurements and can be ignored. As for C_{back} and G_{back} at the bottom of the sample, C_{back} and G_{back}/ω are again normally much larger than C_s, basically because the back surface area of the sample wafer is usually much larger than the field plate area. In this study In or Ga mount technique was used at the back surface of the wafer to ensure the validity of this condition.

Under the above conditions, the capacitance *C(V)* measured as a function of the field plate voltage by this system is given by

$$C(V) = [C_{pr}^{-1} + C_{gap}^{-1} + C_s^{-1}(V)]^{-1}. \quad (1)$$

Thus, for analysis of the *C-V* data, the value of C_{gap} or the length of the UHV gap, l_g, should be known. Direct electrical measurement of C_{gap} is, however, difficult, because it is always serially connected to unknown C_s which depends on various semiconductor parameters and voltage.

To solve this difficulty, the UHV gap length, l_g, was measured optically in this study, using the Goos–Haenchen effect in optics. Namely, l_g was determined by measuring the reflectivity of a totally reflecting prism equipped in the *C-V* sensor head as schematically shown in Fig. 2(a). In this set-up, the reflectivity becomes less than unity by the presence of the sample surface due to disturbance of the evanescent wave. To do this, a laser beam was emitted into the prism of the sensing probe and the reflected laser light (a semiconductor laser diode with a wavelength $\lambda = 0.780$ μm) was detected by a photodiode.

3. Calculation of laser reflectivity vs UHV gap length for GaAs and InP surfaces

To determine the UHV gap length between the sensor and the sample, the reflectivity was calculated as a function of the UHV gap length, solving Maxwell's equations. In the case of the simplest three-layer model shown in Fig. 3(a), the reflectivity R for the P-polarized light used in the present UHV *C-V* system is calculated as

$$R = \left| \frac{r_0 + r_1 \exp(i\delta)}{1 + r_0 r_1 \exp(i\delta)} \right|^2 \quad (2)$$

where r_0 and r_1 are reflectivity coefficients of the upper and lower interfaces in Fig. 3 and δ is the phase rotation angle; r_0, r_1 and δ are given respectively by

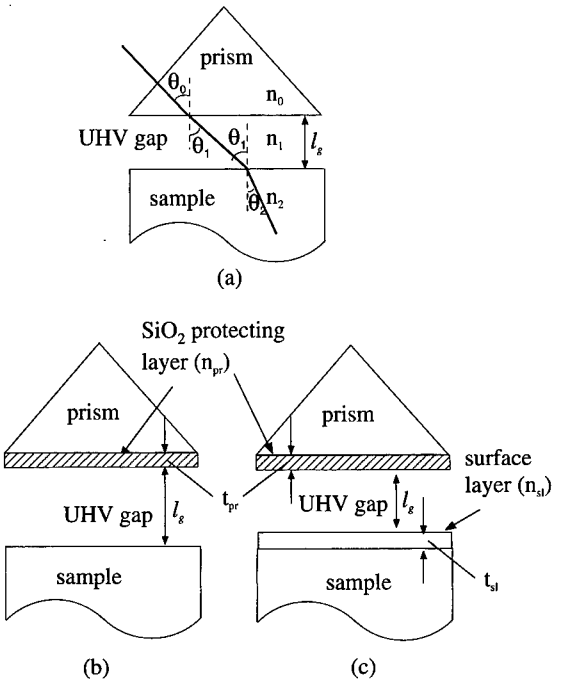


Fig. 3. Structures for the calculation of the reflectivity, (a) three-layer model, (b) four-layer model and (c) five-layer model.

$$r_0 = \frac{n_1 \cos\theta_0 - n_0 \cos\theta_1}{n_1 \cos\theta_0 + n_0 \cos\theta_1}, \quad (3)$$

$$r_1 = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2}, \quad (4)$$

$$\delta = \frac{4\pi}{\lambda} l_g n_1 \cos\theta_1. \quad (5)$$

Here, n_0, n_1 and n_2 are the refractive indices, θ_0, θ_1 and θ_2 are the angles in Fig. 3(a), λ is the wavelength of laser, and l_g is the length of the UHV gap. The angles θ_0, θ_1 and θ_2 satisfy the following relationship according to Snell's law.

$$n_0 \sin\theta_0 = n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad (6)$$

The above equations are general ones applicable for normal reflection of non-evanescent normal waves in a three-layered system. However, they can be used for the present case of reflection of evanescent waves by allowing θ_1, θ_2, $\sin\theta_1$, $\sin\theta_2$, $\cos\theta_1$, $\cos\theta_2$ to be complex variables and complex functions. It should be noted in this case that $\exp(i\delta)$ in Eq. (2) represents exponential decay of the evanescent wave rather than the phase angle rotation. By using Eqs. (1)–(6) on computers, R can be determined as a function of l_g.

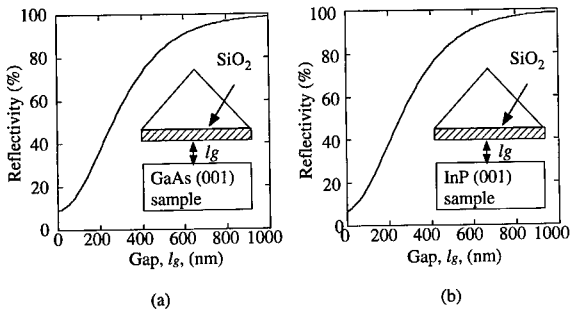


Fig. 4. Calculated reflectivity as a function of the UHV gap length for the (a) GaAs and (b) InP free surfaces.

In a more realistic and general case of a four-layer model shown in Fig. 3(b), where the prism has a SiO_2 protecting layer with a thickness t_{pr} and a refractive index n_{pr}, or of a five-layer model where, additionally, the sample has a surface layer with a thickness t_{sl} and refractive index n_{sl}, the total reflectivity can be calculated by repeated use of the above equations, starting from the bottom three layers first.

Fig. 4(a) and (b), shows the calculated reflectivity of the laser beam as a function of the UHV gap length for the free surfaces of GaAs and InP without any surface layer, respectively. The presence of the SiO_2 protecting layer was taken into account, and the parameters used here include the laser incident angle of $\theta_0 = 45°$, wavelength of $\lambda = 0.780$ μm, and the SiO_2 protecting film thickness of $t_{pr} = 3700$ Å. The refractive indices used here were $n_0 = 1.511$, $n_{pr} = 1.453$ and $n_1 = 1.0$ for prism glass (BK7 glass), SiO_2 and UHV gap, respectively. The refractive indices for semiconductor, n_2, were taken to be 3.82–0.19i and 3.45–0.28i for GaAs and InP, respectively [5,6]. As seen in Fig. 4, the reflectivity R becomes close to 100% as the UHV gap becomes sufficiently larger than the wavelength λ, corresponding to the total reflection. The reflectivity shows large variation for $l_g = 100$–600 nm in both cases of GaAs and InP; thus accurate measurements of the UHV gap length should be possible in this range.

Similarly, in the cases of samples having surface layers, the calculation can be made in a straightforward manner, if the thickness t_{sl} and refractive index n_{sl} of the surface layer are known. In practical cases, however, some unknown ultrathin oxide layer, adsorbed layer or surface reconstructed layer may be present, and the optical reflectivity modification by such a layer may interfere with the accurate determination of the gap length.

In order to provide ideas on the magnitudes of error caused by presence of such a layer, calculated results are shown in Fig. 5(a) and (b) for the cases that fairly thick layer with $t_{sl} = 5$ nm thickness and refractive index values of $n_{sl} = 1.0$ and 4.0 are present on GaAs and InP layers respectively. The curve for $t_{sl} = 5$ nm and $n_{sl} = 4$ almost overlaps with that for $t_{sl} = 0$. Approximately speaking, presence of the surface layer leads to overestimate the gap length by t_{sl}/n_{sl}. However, presence of the surface layer also reduces the effective insulator capacitance. Thus, the error of the

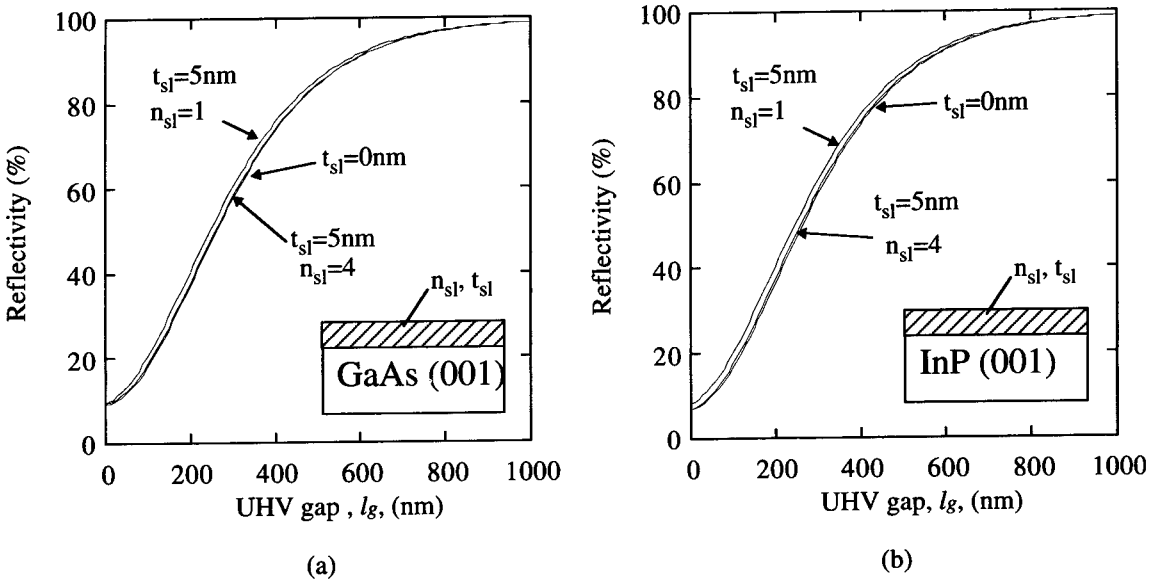


Fig. 5. Calculated reflectivity as a function of the UHV gap length for the (a) GaAs and (b) InP surfaces with fairly thick surface layers ($t_{sl} = 5.0$ nm), refractive indices of $n_{sl} = 1$ and 4.

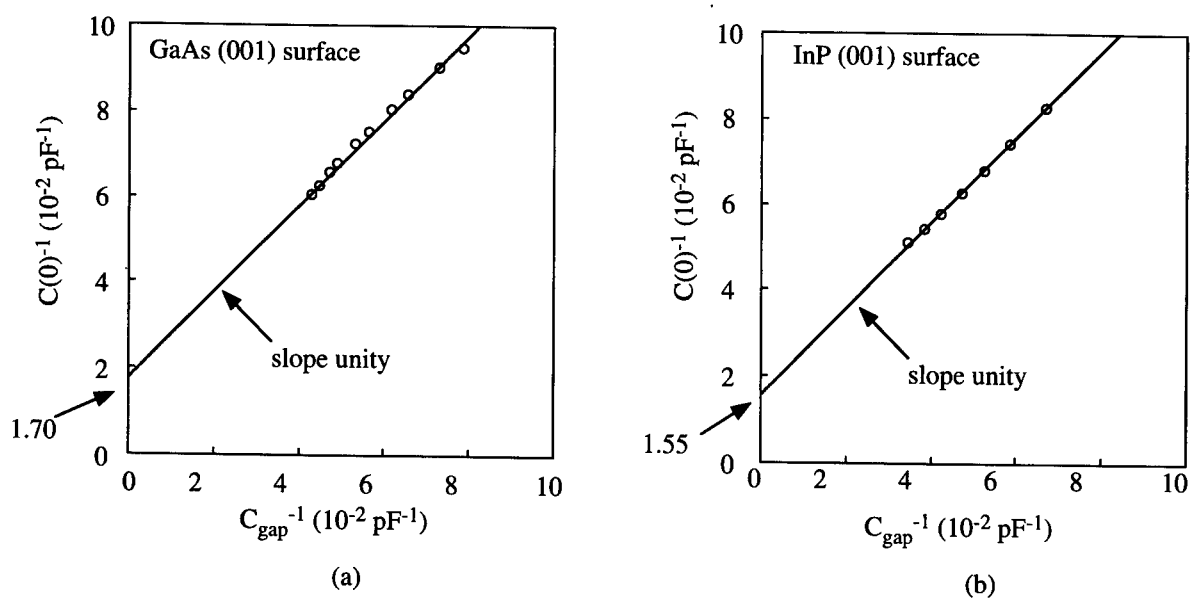


Fig. 6. Measured $C(0)^{-1}$ vs C_{gap}^{-1} plots for MBE-grown free surfaces of (a) GaAs and (b) InP.

insulating capacitance C_I caused by using the overestimated gap length is approximately given by

$$\left|\frac{\Delta C_I}{C_I}\right| \simeq \frac{t_{sl}}{n_{sl}\left(1_g + \frac{t_{pr}}{\varepsilon_{pr}}\right)}\left(1 - \frac{n_{sl}}{\varepsilon_{sl}}\right) \tag{7}$$

where ε_{sl} is the relative permittivity of the surface layer and ε_{pr}, that of the protecting SiO_2 layer. Thus, for a surface layer with $t_{sl} = 1$ nm and any value of refractive index, the error in the insulator capacitance is of the order of 0.1% or well below for measurements with l_g larger than 100 nm.

4. Applications to free surfaces of GaAs and InP

4.1. Determination of zero-bias position of surface Fermi level

The special feature of the present UHV contactless C-V measurement system is that the UHV gap length can be varied unlike the conventional MIS C-V measurements. Thus, if the zero bias capacitance $C(0)$ is measured for various values of the gap length, the following relation should be observed

$$C(0)^{-1} = C_{gap}^{-1} + (C_{pr}^{-1} + C_s^{-1}(0)) \tag{8}$$

where C_{gap} is the UHV gap capacitance obtained using the gap length by the laser reflectivity measurement. Here, it should be mentioned that variations of $C_s(0)$ with gap length variation due to contact potential difference are ignored, since the field plate is separated about 100–500 nm from the sample surface. With such a large gap, typical work function difference values between Cr field plate of 0.5–1.0 eV cannot affect the surface potential appreciably even in the most sensitive case of non-existence of surface Fermi level pinning. According to more accurate estimates in this worst case, $C_s^{-1}(0)$ in Eq. (8) should show variations of 2.5% for GaAs and 0.50% for InP, respectively. Furthermore, this effect should become even smaller due to the surface Fermi level pinning. In practice, this effect is much smaller or comparable as compared with the error of UHV gap measurement of 2–3% even in the worst case.

Fig. 6(a) and (b) shows the measured relationships of $C(0)^{-1}$ and C_{gap}^{-1} at different UHV gaps for free surfaces of MBE grown GaAs and InP. As seen in Fig. 6, the plots have good linearity with a slope of unity. This result proves, first of all, the validity of the UHV gap measurement using the laser beam.

In addition, the surface Fermi level position at zero bias can be determined from the intercepts of these plots which should be equal to $(C_{pr}^{-1} + C_s^{-1}(0))$. From these intercept values, the surface Fermi level positions at zero bias were determined to be $E_C - E_F = 0.96 \pm 0.1$ eV for the GaAs surface and $E_C - E_F = 0.43 \pm 0.1$ eV for the InP surface, respectively. The error of the obtained $E_C - E_F$ value of about 0.1 eV was mainly due to the error of UHV gap measurement of about 2–3 nm. These measured positions are consistent with the usually accepted pinning positions for strong Fermi level pinning [4].

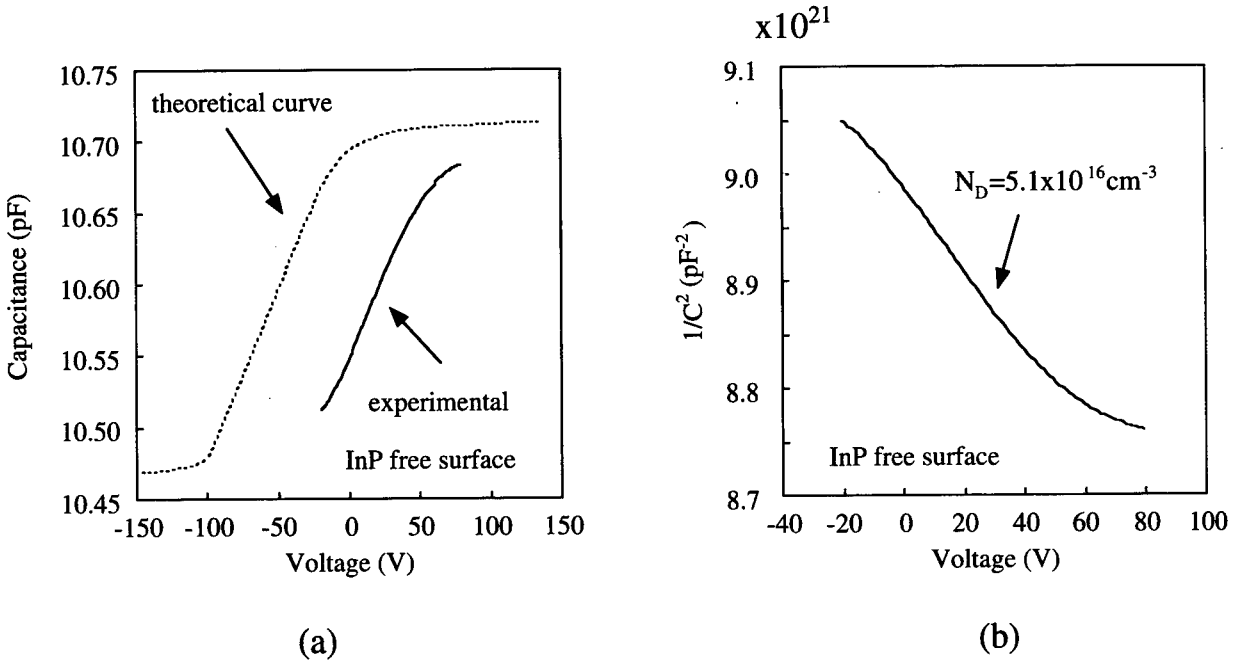


Fig. 7. (a) *C-V* curve obtained from InP surface under pulsed bias mode; (b) $1/C^2$ vs gate bias voltage (V_G).

4.2. Determination of conduction type and doping

By using the UHV *C-V* system, the conduction type and doping of the semiconductor layer can be determined by using the *C-V* data under deep-depletion condition. The deep-depletion condition was realized in this study by using a pulsed bias mode, where the gate voltage was supplied by such a pulse wave form that the generation rate of minority carrier could not follow the sweep rate of the pulsed mode bias.

An example of the *C-V* curve obtained in the pulsed mode is shown in Fig. 7(a) for a free surface of InP. The bias dependence of the curve shows that the sample is of *n*-type. A plot of $1/C^2$ vs gate bias voltage (V_G) obtained from the *C-V* curve in Fig. 7(a) is shown in Fig. 7(b). The carrier concentration determined from the slope of the curve was 5.1×10^{16} cm^{-3}. This was in excellent agreement with that of the Hall measurement of 5.0×10^{16} cm^{-3}. The reason why a correct doping value is obtained by the *C-V* measurement in spite of the existence of Fermi level pinning, is because the bias voltage is changed sufficiently quickly in the pulsed *C-V* measurement so that the surface states cannot change during capacitance measurement. As a result, the effect of surface states becomes effectively a d.c. flat-band voltage shift.

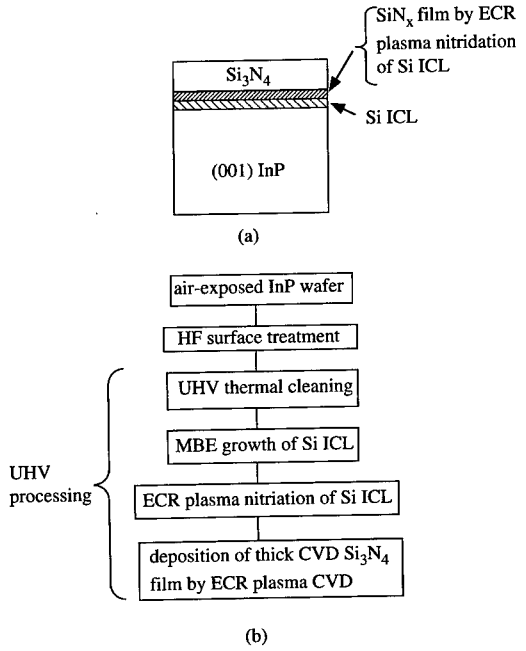


Fig. 8. (a) InP passivation structure using Si ICL studied here; (b) passivation process.

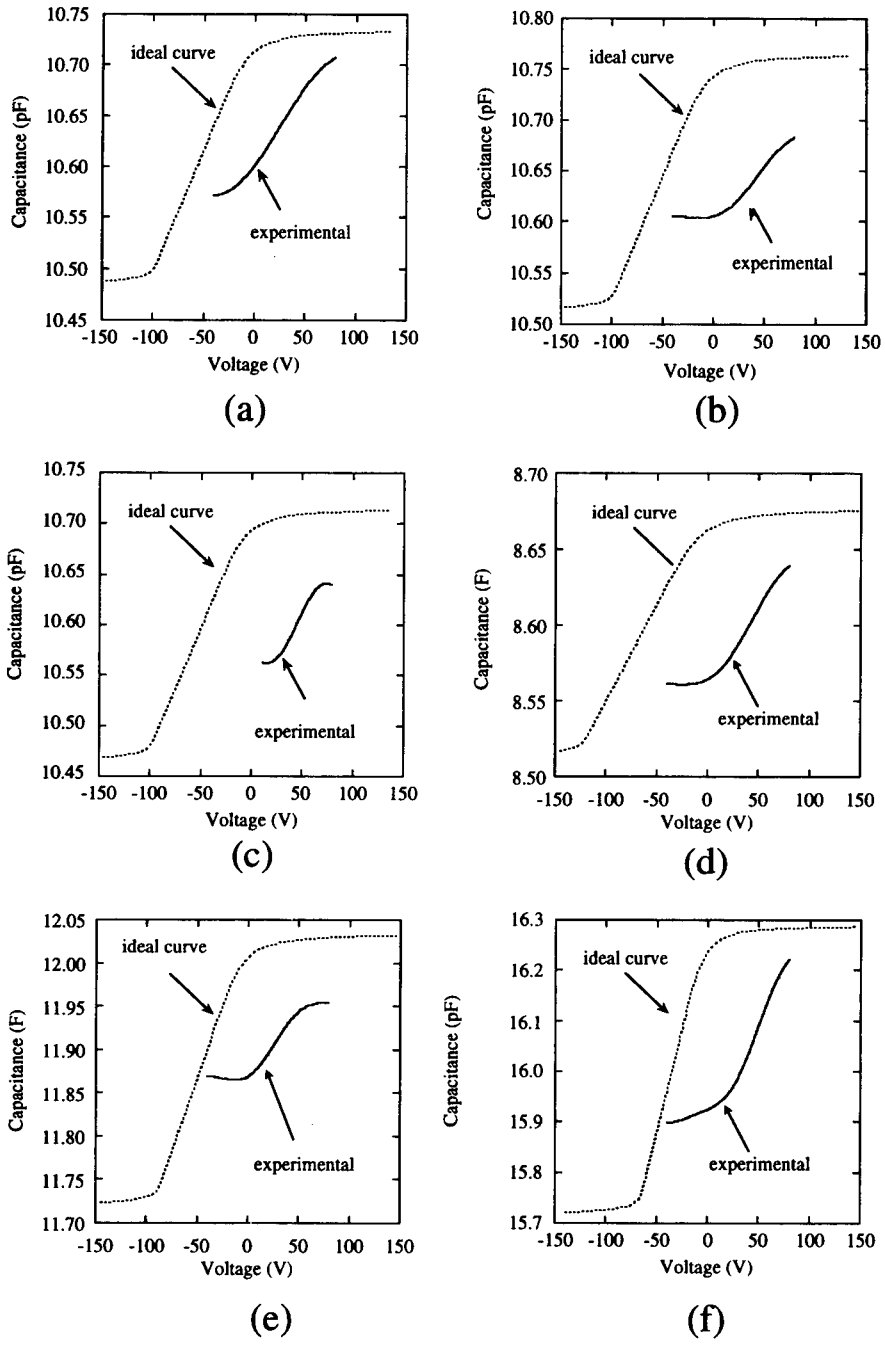


Fig. 9. UHV contactless *C-V* curves of InP surfaces: (a) after HF surface treatment; (b) after UHV thermal cleaning; (c) after growth of the Si ICL; (d) after 30 s nitridation of the Si ICL; (e) after 10 min nitridation of the Si ICL; (f) after deposition of SiN_x on the Si ICL.

5. Characterization of UHV-based surface passivation process using Si ICL

In order further to investigate the capability of the present method, in-situ UHV contactless *C-V* measurements were performed to characterize each processing step of the UHV-based passivation process utilizing the silicon interface control layer (Si ICL) originally proposed for GaAs and InGaAs [7,8]. In this study, the process applied to surface passivation in InP where the surface quantum well (SQW) nature of the Si ICL plays an important role. The basic concept for such a SQW effect and processing details for InP passivation were presented elsewhere [9,10]. Only the passivation structure and the processing sequence as shown in Fig. 8(a) and (b), respectively. An air exposed InP was used to see the effect of the passivation scheme on the commercially available wafers. The experiments were done using the system shown in Fig. 1(b).

Fig. 9(a) and (b) shows the ramped UHV contactless *C-V* curves of the InP surfaces after the HF treatment and after the UHV thermal cleaning. The sweep rate was 1 V/s. In spite of the removal of native oxide in UHV thermal cleaning, both of the slope and the range of capacitance variation of the *C-V* curve became smaller than those of the surface after HF treatment.

Fig. 9(c) shows a UHV *C-V* curve after growth of the Si ICL (1 nm) on InP surface. Although the slope of the *C-V* curve became close to that of the calculated ideal curve, the range of capacitance variation became smaller.

Fig. 9(d) and (e) shows UHV *C-V* curves after 30 s and after 10 min nitridation of the Si ICL. The slope of the *C-V* curve became very close to that of the calculated ideal curve after 30 s nitridation of the Si layer, indicating that a remarkable reduction of the interface state density took place. However, after 10 min nitridation, *C-V* behavior again became poor.

Fig. 9(f) shows the *C-V* curve after growth of a thinner Si ICL (0.5 nm) followed by direct deposition of SiN_x on the Si ICL surface by ECR-excited plasma CVD process. Although a comparatively large capacitance variation was observed, the slope was not close to that of the ideal curve.

Fig. 10 shows the distributions of the interface state density (*Nss*) by applying Terman's method to the measured UHV *C-V* curves. All the curves are U-shaped, and the energy position for minimum *Nss* remained more or less at the same position of $E_C-0.35$ eV. These features are in good agreement with the DIGS model concerning the origin of interface states [4]. A small difference between this value and the previous obtained pinning position of $E_C-E_F=0.43$ eV on the free surface may again be due to error of the UHV gap measurement, which produces surface potential error of ± 0.1 eV.

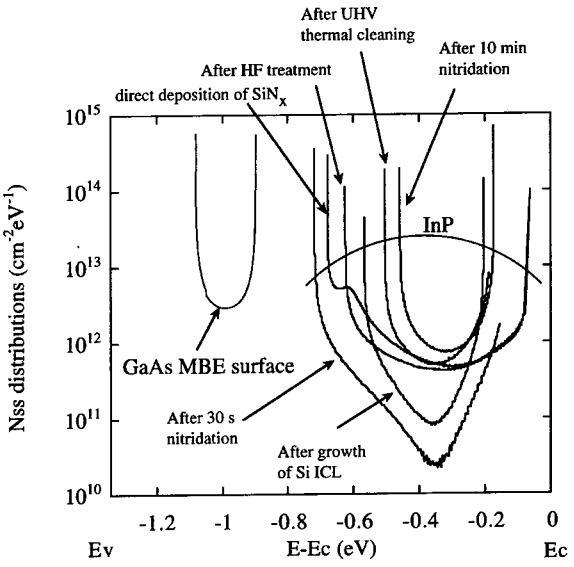


Fig. 10. *Nss* distributions calculated from the UHV *C-V* curves.

In order to compare the *Nss* distribution of the GaAs and the InP bare surfaces, a result of *Nss* determination by UHV contactless *C-V* measurement made on the MBE-grown GaAs surface [11] is also included in Fig. 10. As seen in Fig. 10, *Nss* minimum positions and state density values are very different between GaAs and InP, showing that the surface state density on InP is significantly smaller than that on GaAs.

Let us see the effect of each step of processing in more detail for InP passivation. After HF treatment, the surface was found to have the Nss_{min} value of 4×10^{11} cm^{-2} eV^{-1}, and slightly increased after UHV thermal treatment to 5.5×10^{11} cm^{-2} eV^{-1}. It decreased a great deal after growth of the Si layer down to 8×10^{10} cm^{-2} eV^{-1}. But in spite of the decrease of the Nss_{min}, the width of the *Nss* distribution became narrower. Then, the Nss_{min} remarkably decreased down to 2×10^{10} cm^{-2} eV^{-1} after nitridation of the Si layer for 30 s with a large increase of the distribution width. After prolonged nitridation of 10 min, Nss_{min} again increased up to the range of 10^{11} cm^{-2} eV^{-1}.

The fact that the Nss_{min} value slightly increased after UHV thermal cleaning indicates that simple removal of the native oxide from the surface did not cause a surface passivation effect. This is because the surface dangling bonds were not terminated properly. The reason why the *Nss* distribution became narrower after growth of the Si ICL on the InP surface can be explained by referring to the energy band diagram of

the strained Si_3N_4/Si/InP structure [9,10]. Namely, the bandgap of coherently strained Si in the Si_3N_4/Si/InP structure becomes very narrow. Because of this, the width of the *Nss* distribution became narrower. After 30 s nitridation of the Si ICL, the width of the *Nss* became wider to 600 meV and the Nss_{min} value became lower. By the nitridation process, the thickness of the Si ICL became narrower and this effectively increases the bandgap by the quantum confinement effect. Additionally, successful termination of interface bonds of the Si ICL reduces the density of gap states. On the other hand, after longer nitridation for 10 min of the Si ICL by the ECR plasma, the Nss_{min} value increased to a large value of 5×10^{11} cm^{-2} eV^{-1} because the prolonged nitridation damaged the interface structure.

In the case that direct CVD deposition of SiN_x onto a thinner Si ICL, the Nss_{min} value remained 5×10^{11} cm^{-2} eV^{-1}. This result is probably due to the large bond mismatch between the Si ICL and deposited SiN_x.

Thus, effect of each step of the UHV-based passivation could be sensitively detected with this technique, confirming the extremely powerful nature of the present method.

6. Conclusion

This paper has described an ultra-high vacuum (UHV) contactless capacitance-voltage (*C-V*) measurement method and system. The main conclusions are as follows.

1. The length of the UHV gap measured optically using the Goos–Haenchen effect is accurate enough for present type of system on the basis of a theoretical calculation made for GaAs and InP surfaces and actual measurements. The possible existence of a surface layer of a thickness of 1 nm does not cause large errors.
2. The system has the capabilities of non-destructive determination of conduction type, doping level and zero-bias position of the surface Fermi level.
3. On the basis of successful application of this technique to characterization of the UHV-based surface passivation process for InP utilizing the Si ICL, this technique is extremely powerful for use in non-destructive characterization of crystal growth and device processing steps for the compound semiconductor heterostructure microelectronics. Change of the surface state distribution after each step can be successfully detected without breaking UHV.

Acknowledgements

The present work was supported financially in part by a Grant-in-aid for scientific research B (No. 10555098).

References

[1] Chakraborty S, Yoshida T, Hashizume T, Hasegawa H. J Vac Si Technol 1998;B16(4):2159.
[2] Yoshida T, Hashizume T, Hasegawa Hideki, Sakai Takamasa. Presentation at 1998 Solid State Devices and Materials Conference, Extended Abstracts 128.
[3] Sakai T, Kohno M, Hirae S, Nakatani I, Kusuda T. Jpn J Appl Phys 1993;32:4005.
[4] Hasegawa H, Ohno H. J Vac Sci Technol 1986;4-4:1130.
[5] Properties of gallium arsenide. 2nd ed. An inspec publication, 1990.
[6] Properties of indium phosphide. An inspec publication 1991.
[7] Hasegawa H, Akazawa M, Matsuzaki K, Ishii H, Ohno H. Jpn J Appl Phys 1998;27:L2265.
[8] Hasegawa H, Akazawa M, Ishii H, Matsuzaki K. J Vac Sci Technol 1989;7-4:870.
[9] Takahashi H, Hashizume T, Hasegawa H. Appl Surf Sci 1998;123/124:615.
[10] Takahashi H, Hashizume T, Hasegawa H. Jpn J Appl Phys 1999;38:1128.
[11] Mutoh M, Tsurumi N, Hashizume T, Hasegawa H. Presentation at 1998 Solid State Devices and Materials Conference, Extended Abstracts. Jpn J Appl Phys 1999;38:2538.



PERGAMON

Solid-State Electronics 43 (1999) 1571–1576

SOLID-STATE ELECTRONICS

A preliminary study of MIS diodes with nm-thin GaAs-oxide layers

Tomoyuki Sugimura*, Tatsutoshi Tsuzuku, Takayuki Katsui, Yuki Kasai, Takao Inokuma, Shoushin Hashimoto, Kouichi Iiyama, Saburo Takamiya

Graduate School of Natural Science, Faculty of Engineering, Kanazawa University, 2-40-20 Kodatsuno, Kanazawa, Ishikawa, 920-8667, Japan

Abstract

Direct oxidation of GaAs wafers, by using UV and ozone oxidation system, was performed to form nm-thin GaAs-oxide layers. Composition of the oxidized layer, studied by XPS analysis, shows depth-dependent atomic percentage of the Ga, As, and O. Surface flatness and thickness of the oxide were investigated with AFM. The oxide thickness is nearly proportional to square root of the UV and ozone process time. Using the oxidized layer as the insulator, Ni/insulator/n-GaAs MIS diodes were fabricated, and compared with Ni/n-GaAs Schottky diodes of their gate leakage currents. Very clear leakage current suppression effect of the insulating layer is observed, which is controllable by controlling the UV and ozone process time. However, the suppression effect is much smaller than that theoretically estimated for a MIS diode with perfectly uniform SiO_2. Theoretical equations to estimate effect of a nm-thin insulating layer on current-voltage relation of a MIS diode are shown in the appendix. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Schottky junction has long been used as the gate structure of compound semiconductor Field Effect Transistors (FETs) and High Electron Mobility Transistors (HEMTs). One of their critical problems is that they are encountering tunnel limit [1–3] as a consequence of scaling down, which have been pursued in order to realize higher speed devices. From a view point of preventing the tunnel current, MIS structure is more desirable than the Schottky junction, because the former generally has much higher barrier than the latter, and makes a thinner barrier layer usable. However, the MIS structure compound semiconductor devices ever reported are not successful, due to high density of interface states, difficulty in achieving a very thin insulating layer without pin-holes etc., and no such devices with nm-thin insulating layers have been reported. UV and ozone process [4], which has been used to decompose organic contamination of semiconductor wafer surfaces etc., recently attracts attention as potential means to form very thin and uniform oxide layer on semiconductor surfaces. Driad et al. [5] improved noise performance of InGaAs/InP HBTs by passivating them by UV and ozone treatment.

We fabricated GaAs MIS diodes with nm-thin oxidized GaAs layers, by using the UV and ozone process, in order to study quality and effects of the GaAs-oxide layers. In the following, theoretical current-voltage characteristics of MIS diodes, which are being calculated from equations in the appendix, are explained in section 2, thicknesses and compositions of the oxidized GaAs layers are described in section 3, fabrica-

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00105-7

tion of the MIS diodes and effects of the insulating layer are described in section 4, and discussions and conclusions are described in section 5.

2. Theoretical current-voltage curves

Electrons flowing through a reverse biased MIS junction are illustrated by arrows in an energy band diagram in Fig. 1, in which J_T is a thermionic current, and J_{t1}–J_{t3} are tunnel current components. Theoretical equations to do a rough estimation of these currents are shown in the Appendix. As both of the barrier height at the MI-junction, Φ_o, and the dielectric constant, ε_o, of the GaAs-oxide are not known, we use for calculating the current-voltage characteristics, $\Phi_o = 3.7$ eV and $\varepsilon_o = 3.9$ times permittivity of the vacuum (referring from SiO_2), and a known donor density of the semiconductor of $N_D = 3 \times 10^{17}$ cm^{-3}, a dielectric constant of GaAs $\varepsilon_s = 10.9$, and a Schottky barrier height of $\phi_s = 0.7$ eV. Fig. 2 shows calculated current-voltage curves with the thickness of the insulating layer as a parameter. The current change by the increase of the insulating layer is somewhat larger in the forward or low negative voltage range, compared with high negative voltage region. However, the curves do not remarkably change their shapes, but shift depending upon the insulator thickness. This is due to the high barrier height of the MI-junction and so thin thickness of the insulator as compared with the depletion layer of the semiconductor that most of the applied voltage is shared by the depletion layer. The

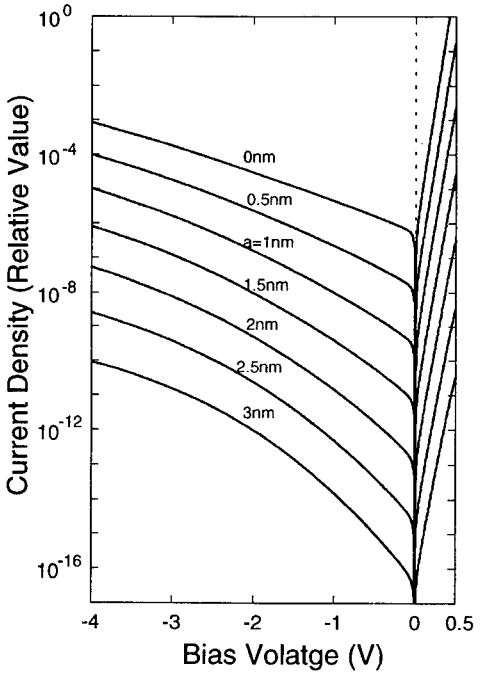


Fig. 2. Calculated current-voltage characteristics of MIS diodes with nm-thin insulating layers. Parameter a is thickness of SiO_2.

voltage dependence in the reverse current reflects dominance of the tunnel current in the original (a = 0 nm) Schottky diode due to the high donor density of the semiconductor [3].

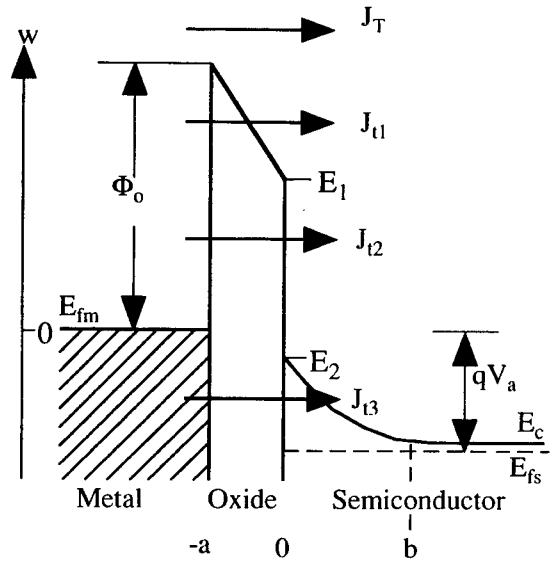


Fig. 1. Energy band diagram and electron flows in a reverse biased MIS diode.

3. Oxidation of GaAs surfaces

The authors used UV and ozone process (SAMCO UV and Ozone Cleaner) at 300°C in order to form very thin oxidized GaAs layers on the surfaces of n-GaAs(with a donor density of 3×10^{17}/cm^3, and a thickness of 0.5 μm)/Semi-Insulating-GaAs epitaxial wafers. Their native oxide was removed by buffered hydrofluoric acid etching. After rinsing by deionized water, they were directly oxidized by the UV and ozone process for 15–480 min. Then the oxidized layer was partially removed by photolithography and buffered hydrofluoric acid etching in order to measure thickness of the oxidized layer by using an Atomic Force Microscope (AFM). Fig. 3 is a three dimensional AFM image of the 480 min sample of around the oxide edge portion. Top surface of the oxidized GaAs has much larger roughness than the interface (surface of the oxide removed portion) suggesting a large nonuniformity of the oxide layer thickness. The average step height (average in single trace) from the

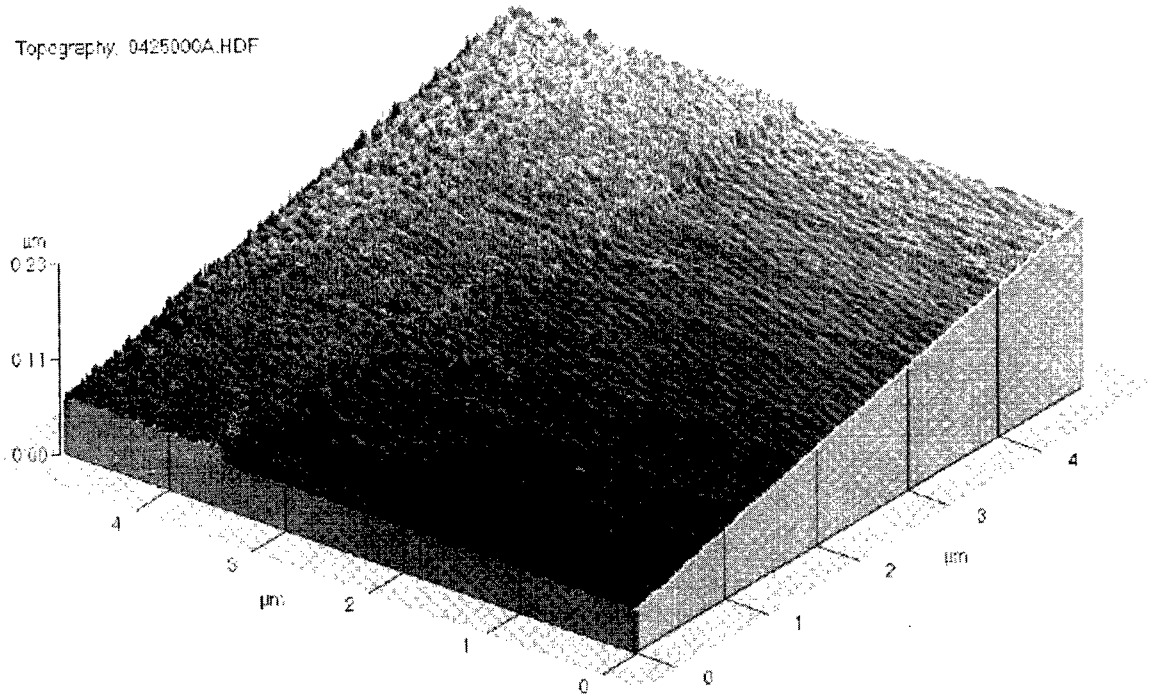


Fig. 3. Three dimensional AFM image of a wafer surface around the oxide edge portion. Left-up is the oxide surface, and right-down is the interface on which oxide layer was removed.

interface to the oxide surface was measured as the thickness of the oxidized GaAs layer. Fig. 4 shows thus measured oxide thickness of the samples. The wide spread of the data implies positional dependence of the data (between different traces), reflecting the large nonuniformity of the oxidized layers thickness. As a whole, the thickness is nearly proportional to square root of the process time. The thin solid line in Fig. 4 illustrates experimental data reported by Driad et al. [5].

X-ray Photoelectron-Spectroscopy (XPS) analysis was applied to investigate compositional informations of the oxide layer. Fig. 5 shows depth (Ar ion etching time) dependence of the composition. Our original data show existence of oxygen even after 15 min etching. We assumed that the oxygen becomes zero at a depth where it saturates. This is because that oxygen was detected even inside an untreated GaAs monitor wafer. Fig. 5 suggests a very gradual change of the composition and low As percentage near the surface region, although the measured data may be strongly affected by the large nonuniformity of the thickness.

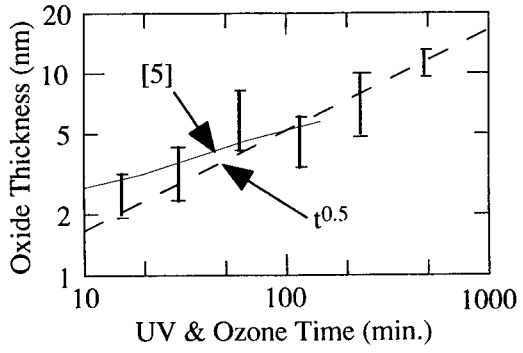


Fig. 4. UV and ozone process time dependence of the GaAs-oxide thickness. The thin line illustrates Driad et al.'s data [5].

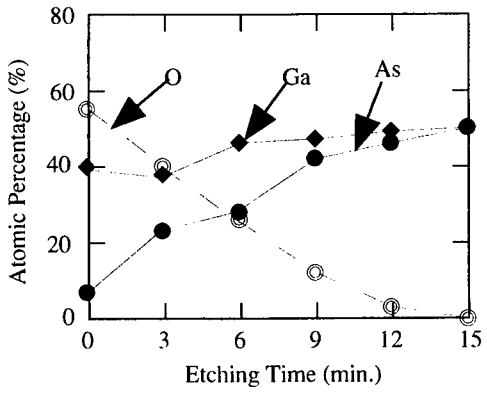


Fig. 5. Depth (etching time) dependence of composition of the oxide layer.

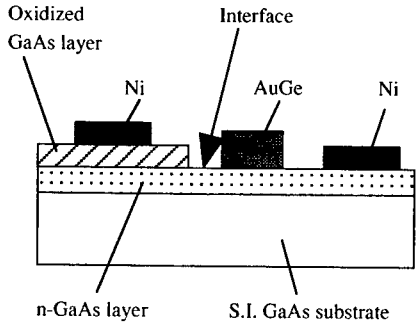


Fig. 6. Cross-sectional scheme of the MIS and the Schottky diodes sample.

4. MIS diode samples and current–voltage curves

A cross-sectional scheme of the MIS diode sample is shown in Fig. 6. After the oxidation, the oxide layer on half of the wafer was removed, and Ni was evaporated through a metal mask with 320 μm diameter holes on both the oxide surface and the interface (GaAs), in order to fabricate Schottky diodes and the MIS diodes on each wafer. Then, AuGe was evaporated as an ohmic contact on a part of the interface layer.

DC forward and reverse current-voltage characteristics of both of the Schottky and the MIS diodes were measured. Fig. 7 shows typical characteristics of four sample wafers. The Schottky diodes (open circle plots), as well as the MIS diodes (solid circle plots), show strong voltage dependences even in the reverse characteristics, due to dominance of the tunnel currents [3]. It is very clear that the nm-thin insulating layers effectively suppress the reverse leakage currents, but do not remarkably affect the shapes of the curves. Here, it should be stated that the currents of the Schottky diodes formed on the interfaces are larger than those ($\sim 10^{-6}$ A/cm^2 at -0.3 V and $\sim 10^{-3}$ A/cm^2 at -4 V) formed on an untreated wafer, which suggests that the interface is damaged by the oxidation process. The current suppression rate vs UV and ozone process time is summarized in Fig. 8, which demonstrates that one can control the current suppression rate by controlling the process time, although the data are widely spread this time due to the strong nonuniformity.

Fig. 9 summarizes the current suppression rate vs the oxide layer thickness relation. The dashed line curve shows as a reference the calculation result (at -1 V) of a Ni/Insulator/n-GaAs diode with perfectly uniform SiO_2. Fig. 9 suggests either the barrier height of the GaAs-oxide is much lower than that of SiO_2, or an effective thickness of the GaAs-oxide is much thinner than the measured average values.

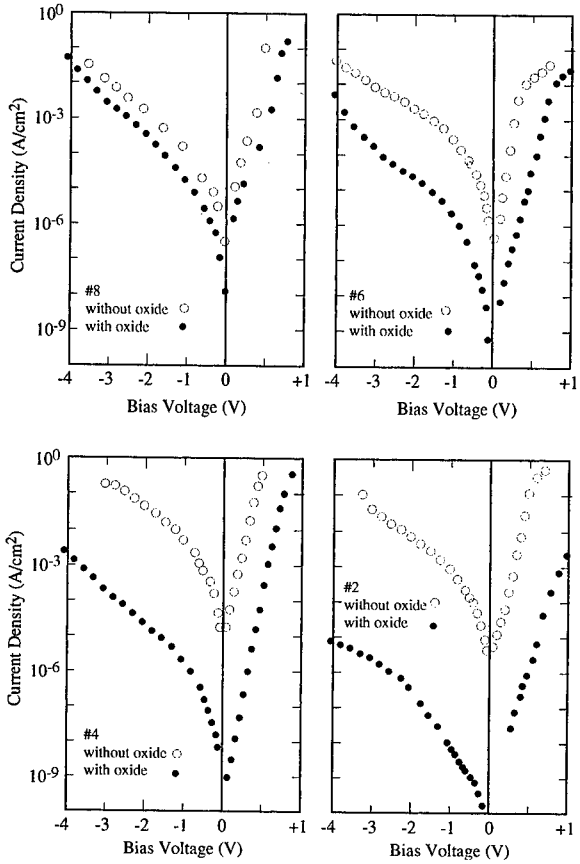


Fig. 7. Current voltage characteristics of the MIS (solid circle) and the Schottky diodes (open circles) from wafers processed for 30 min (♯♯8), 60 min (♯♯6), 120 min (♯♯4), and 240 min (♯♯2).

5. Conclusions and discussions

Theoretical current-voltage characteristics of a Ni/nm-thin-SiO_2/n-GaAs diode were calculated as a combination of a thermionic emission current and tunnel currents. Experimental MIS diodes, with GaAs-oxide as the insulator, showed clear and controllable current suppression effect as compared with Schottky diodes, which qualitatively agrees with the theory. However, the experiment showed much thicker oxide layers than the calculation for obtaining an equal current suppression effect. Moreover, the leakage currents of the Schottky junctions on the interface are larger than those formed on an unprocessed wafer. AFM and XPS analysis showed large nonuniformity of the oxide thicknesses and compositions. The former possibly results in current concentration at thinnest oxide point, and the latter possibly cause a difference between the measured and an actually effective oxide thickness, and both of them essentially give very strong influences to

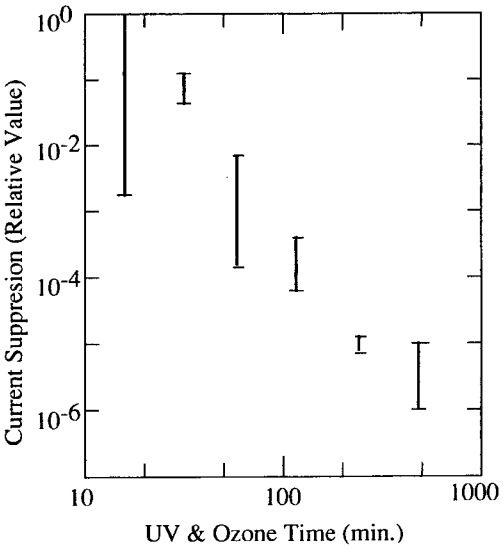


Fig. 8. Process time dependence of the current suppression rate.

the tunnel probability. The quantitative disagreement between the theory and the experiment can be attributed to this nonuniformity, and also to the used barrier height and the dielectric constant. In order to determine ε_o and Φ_o of the oxidized GaAs layer from comparison of theoretical and experimental C-V and I-V characteristics, high uniformities of the thickness and the composition of the oxidized layer are indispensable. Therefore, substantial improvement of the uniformity, as well as reduction of the process damage, are the main problems to be solved. The much better flatness of the interface than the surface of the oxide (Fig. 3) suggests a possibility of the improvement of the thickness uniformity by, for example, applying UV and ozone process multiple times.

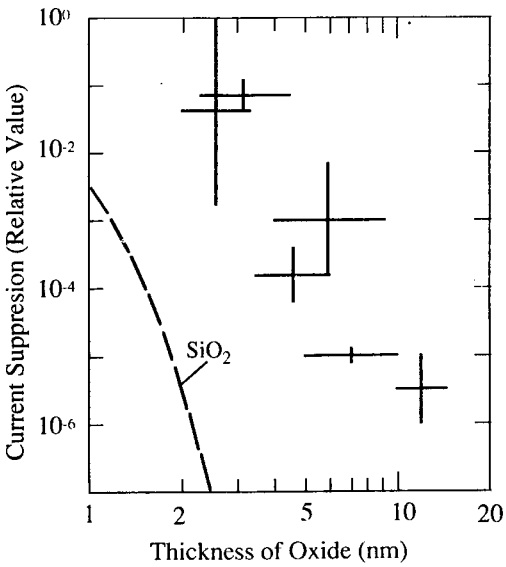


Fig. 9. Oxide thickness vs the current suppression rate. The dashed curve shows a calculation result for a $Ni/SiO_2/n$-GaAs diode.

Appendix A

Here, theoretical equations for a rough estimation of current-voltage curves of a MIS diode are described. Theoretical description for expressing effect of electron velocity in thermal equilibrium MIS system becomes too complicated to solve analytically, when tunnel is involved, due to dependence of tunnel probability on energy and its voltage dependence. However, as the electron velocity (proportional to square root of energy) gives rather minor effect as compared with a distribution function and a tunnel probability (both exponentially depend upon energy), we use constant effective velocities, v_{sx} and v_{mx}, and do a rough estimation.

The thermionic current (Fig. 1) is given by

$$J_T = \int_{\Phi_o}^{\infty} \{A_s D_s(w) E_s(w) - A_m D_m(w) E_m(w)\} dw, \quad (A.1)$$

$$A_s = 4\pi q (2m_s)^{3/2} v_{sx}/h^3 \quad (A.2)$$

$$A_m = 4\pi q (2m_m)^{3/2} v_{mx}/h^3 \quad (A.3)$$

$$D_s(w) = 1/[1 + \exp\{(w - qV_a)/kT\}] \quad (A.4)$$

$$D_m(w) = 1/\{1 + \exp(w/kT)\} \quad (A.5)$$

$$E_s(w) = \sqrt{\{w - qV_a - (E_c - E_{fs})\}} \quad (A.6)$$

$E_m(w)$: not known

where, suffix s implies the semiconductor and m implies the metal, the first term in Eq. (A.1) flows from the semiconductor to the metal and the second term flows in the reverse directions, respectively, A_s and A_m are velocity dependent constants, $D_s(w)$ and $D_m(w)$ are distribution functions, $E_s(w)$ and $E_m(w)$ are state densities, w is an energy level measured from the Fermi level of the metal, Φ_o is barrier height of the MI-junction, q is unit charge, v_{sx} and v_{mx} are above mentioned effective velocities of electrons in the x direction, h is Plank's constant, m_s and m_m are effective electron masses, V_a is applied voltage, k is

Boltzmann constant, T is absolute temperature, E_c is bottom of the conduction band, and E_{fs} is Fermi level of the semiconductor. While, the tunnel current is expressed by the following equations.

$$J_{tn} = \int_{E_c}^{\Phi_o} \{A_s D_s(w) E_s(w) - A_m D_m(w) E_m(w)\} P(w) dw \quad (A.7)$$

$$P(w) = \exp[-2\int_{-a}^{x_w} \{2\pi(2m_e(\Phi(x) - w))^{1/2}/h\} dx] \quad (A.8)$$

where, integration in Eq. (A.7) is done from E_1 to Φ_o for J_{t1}, from E_2 to E_1 for J_{t2}, and from E_c to E_2 for J_{t3} (Fig. 1), $P(w)$ is a tunnel probability for electrons at an energy level of w, x_w is a position where the w crosses the potential barrier, m_e is an effective electron mass in the barrier. Tunnel probability is strongly dependent on the effective mass of the electrons. We assumed $m_e = 0.3\ m_o$, as it is widely used in calculating a tunnel probability of SiO_2 [6]. $\Phi(x)$ is energy level of the bottom of the conduction band in the oxide

$$\Phi(x) = \Phi_o - q^2 N_D b(a + x)/\varepsilon_o, \quad (A.9)$$

and in the semiconductor

$$\Phi(x) = qV_a + (E_c - E_{fs}) + q^2 N_D (b - x)^2 / 2\varepsilon_s, \quad (A.10)$$

$$b = -(\varepsilon_s a/\varepsilon_o) + [(\varepsilon_s a/\varepsilon_o)^2 - 2\varepsilon_s (qV_a - \phi_s)/qN_D]^{1/2} \quad (A.11)$$

where, N_D is donor density, b is depletion layer thickness, a is the oxide thickness, x is position, ε_o and ε_s are dielectric constants of the oxide and the semiconductor, respectively, and ϕ_s is Schottky barrier height when the metal directly contacts onto the semiconductor. We use a constant state density for the metal and Eq. (A.6) for the semiconductor. This is because the former is not known but presumed to give very little effect, and the latter gives a strong effect when the voltage approaches a flat band condition. The forward/reverse ratio, A_s/A_m, is such determined that the current becomes zero at $V_a = 0$ V.

Calculations based upon the above equations and the assumptions yield current-voltage curves and their relative values dependent on the insulator thickness. It should be noted that temperature dependences of only the distribution functions but not of the velocities are explicitly included in the above equations.

References

[1] Bahl SR, del Alamo JA, Dickmann J, Schildberg S. Off-state breakdown in InAlAs/InGaAs MODFET's. IEEE Trans ED 1995;42(1):15.

[2] Putnum CS, Somerville MH, del Alamo JA, Chao PC, Duh KG. In: Proceedings of the International Conference on InP and Related Materials. 1997. p. 197–200.

[3] Takamiya S, Harayama M, Sugimura T, Tsuzuku T, Taya T, Iiyama K, Hashimoto S. Reverse currents of Schottky Gates of III-V MESFET/HEMTs: Field emission and tunnel currents. Solid-State Electronics 1998;42(3):447–51.

[4] Vig JR. UV/ozone cleaning of surfaces. J Vac Sci Technol 1985;A3(3):1027–34.

[5] Driad R, Lu ZH, Laframboise S, Scansen D, McKinnon WR, McAlister SP. Surface passivation of InGaAs/InP heterostructures using UV-irradiation and ozone. In: Proceedings of the Tenth International Conference on InP and Related Materials, 1998. p. 459–62.

[6] Matsuo N, Miura T, Urakami A, Miyoshi T. Analysis of direct tunneling for thin SiO_2 films. In: Extended Abstracts of the 1998 International Conference on Solid State Devices and Materials, 1998. p. 116–7.

SYSTEMS APPLICATION

PERGAMON

Solid-State Electronics 43 (1999) 1577–1589

SOLID-STATE ELECTRONICS

Applications of HEMT devices in space communication systems and equipment: a European perspective

Massimo Claudio Comparini, Marziale Feudale*, Andrea Suriani

Alenia Aerospazio, Space Division, Radio Frequency and Microwave Engineering, Via Saccomuro 24, 00131 Rome, Italy

Abstract

In the last years the satellite applications have seen a strong increase in the space hardware demand mainly for commercial communication services like fixed satellite service (FSS), broadcasting satellite service (BSS) and mobile satellite service (MSS). The new boundary will be the low cost, high production rate hardware needed for satellite constellations like Teledesic and for multimedia satellite-based systems (Euroskyway, Spaceway, Astrolink, Cyberstar). The design and development of microwave and millimeter wave equipment has been deeply modified in order to combine technical requirements and market needs (large productions, low cost and fast lead time). In this scenario the maturity of gallium arsenide MESFET and PHEMT devices and processes for space applications is a key factor and the design and integration of GaAs MMICs is mandatory to achieve the mentioned industrial drivers in the space companies. The paper will focus the application of GaAs MMIC technology in the European space industry taking as reference the space production done by Alenia Aerospazio in Italy and other European space companies and providing a future road map. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

As result of a fast evolution in the satellite commercial environment the payload complexity is growing with high number of channels either in transparent and regenerative architectures. While the payload complexity and the number of units per repeater is fast growing the required lead time is highly compressed, as consequence the technology solutions must address the miniaturization issue, the design for production approach and the performance repeatability issue required for relative high volume productions.

Only the massive use of MMIC technology allows integration of complex microwave functions like LNAs, front ends, receivers, frequency converters, and channel amplifiers. The combination with advanced packaging techniques like multi chip modules (MCM) and low temperature cofiring ceramic (LTCC) allows to achieve remarkable integration factors and impressive improvement with respect to the previous generation of space hardware. The definition and consolidation of a screening flow including accelerated life tests, burn in phases and lot acceptance tests has been very important in order to secure and assess, in the first application programs, satellite operating life of 15 years and more.

The modular approach is another key feature of the MMIC application in the space equipment. Simple functions like variable gain amplifiers (VGA), flatness correctors (FC), medium power amplifier (MPA), LNAs, mixers and VCOs have been developed as building blocks to be used in different configurations in the various modules and equipment.

* Corresponding author. Tel.: +39-6-4151-2515; fax: +39-6-4151-2507.

E-mail address: feudale@roma.alespazio.it (M. Feudale)

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00106-9

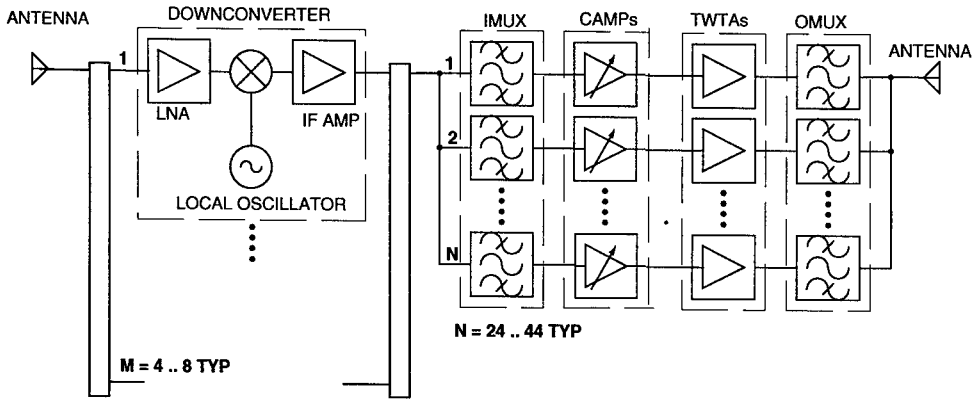


Fig. 1. Transparent payload architecture.

The mentioned approach requires the use of multi-project 'run' with well consolidated processes. At present only two fabrication processes are actually qualified for this purpose: low noise (medium power) 0.5 μm MESFET and low noise 0.25 μm PHEMT. MMIC based on these processes are currently flying in space equipment produced by European companies [1,2].

2. The present space applications

The market drivers in the design development and production of the space hardware in the last years can be summarized in the new paradigm: *better, faster and cheaper*. An impressive growing in the hardware demand has been observed in the last three years in the commercial telecommunication via satellite. Having as reference a simplified payload configuration (Fig. 1) it can be easily understood why the channel amplifiers (CAMP) have been the first unit in which a massive use of GaAs MMIC has been employed.

This unit, whose simplified block schematic is reported in Fig. 2, provides amplification within the Ku band portion of the spectrum (10.7–12.75 GHz). Selection of either gain (10–60 dB, fixed gain mode) or output level (−10–+4/+6 dBm, automatic level control mode) in discrete steps is also possible by means of an external command. The complete amplifier is composed of two sections: the microwave and the command/control section. The microwave section is fully MMIC based in order to achieve the best performances in terms of reduced weight, reduced size and integration factor. It is composed of eight MMICs

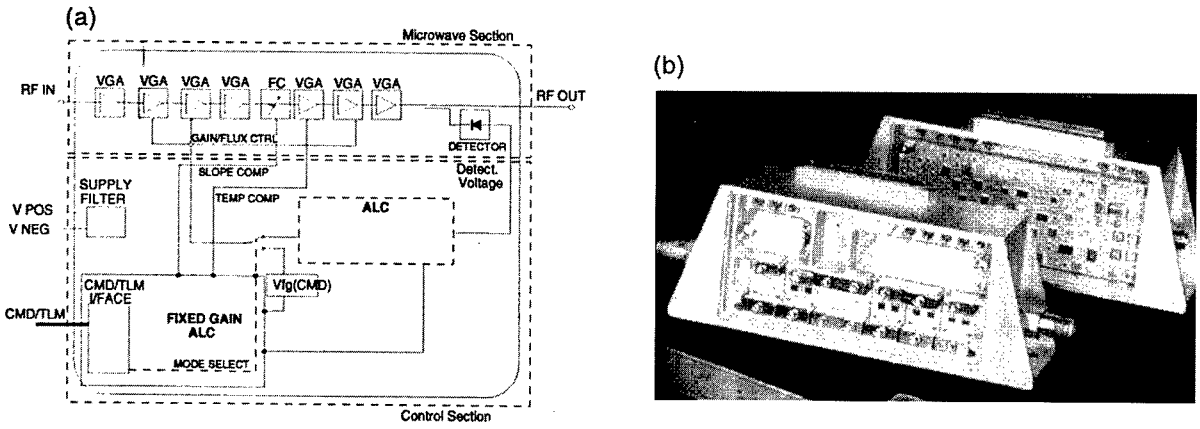


Fig. 2. (a) Ku band channel amplifier block diagram. (b) Ku band channel amplifier flight model.

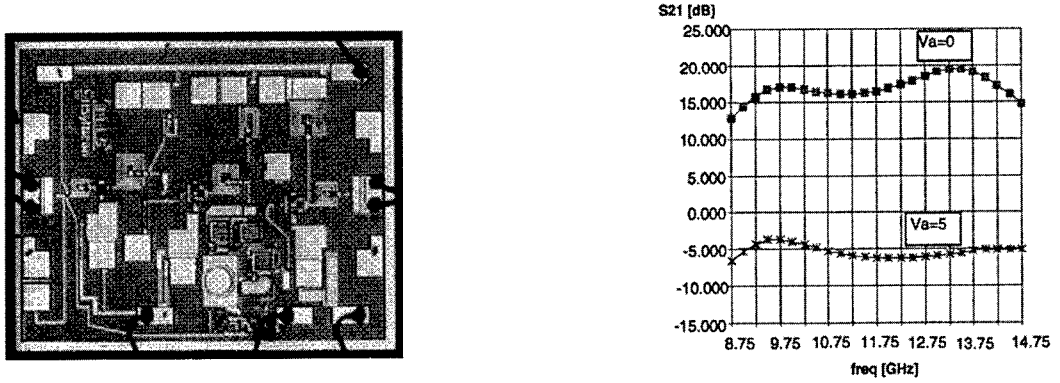


Fig. 3. MESFET variable gain amplifier: layout and measurements.

brazed on open carriers and integrated in a single hybrid hermetically sealed. A mass lower than 50 g is achieved with impressive reduction in the tuning effort with respect to a more traditional approach using discrete FET devices.

The basic building block of CAMP is a variable gain amplifier (VGA) whose layout is shown in Fig. 3. The MMIC, designed to work in the bandwidth 10.7–12.7 GHz, provides amplification in conjunction with gain control function. It is composed of three self-biased stages with an embedded analog attenuator. The analog attenuator is based on a broadband T structure. The driver for the necessary attenuator control is also implemented on the same chip. Typical gain is 14 dB over 2 GHz bandwidth (10.7–12.7 GHz) in combination with a gain control range of about 20 dB. The flatness corrector (FC) MMIC is a circuit that has been designed in order to control the channel slope and to recover slope variation at different gain settings and temperature conditions over the full bandwidth. The design is conceptually simple and allows the obtaining of both positive and negative slope control. Using semilumped elements and cold MESFET as varactor, a notch filter has been realized. The Q factor and resonant frequency are controlled by means of an external control voltage. Layout of this MMIC is reported in Fig. 4.

The communication receiver is another unit in which MMIC technology is extensively used. A conceptual block diagram of a typical communication receiver is shown in Fig. 5. The FSS or BSS receivers provide low noise amplification and down conversion of signals in a portion of the frequency range 12.7–14.7 GHz or 17–18 GHz down to 10.7–12.7 GHz, by means of a net subtraction of an internally generated local oscillator frequency. A hybrid hermetic module mostly

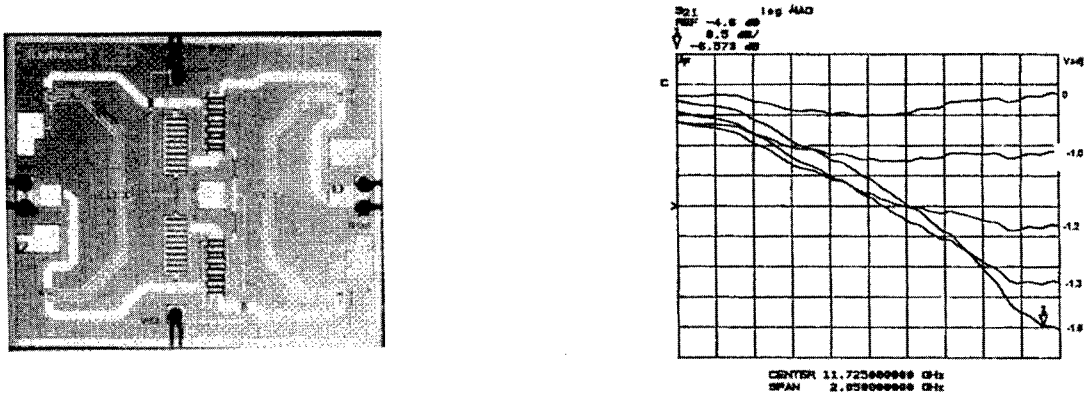


Fig. 4. MESFET flatness corrector.

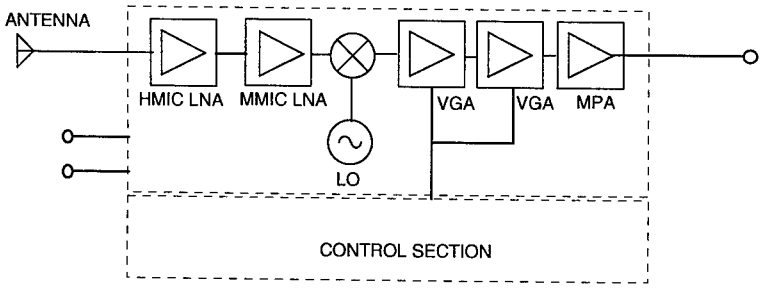


Fig. 5. Conceptual schematic of the receiver.

MMIC based has been designed to integrate low noise amplifiers (LNA), mixers (MIX) and the IF medium power amplifier (MPA) circuits. The first stage of the input section uses a hybrid MIC (HMIC) approach to achieve the best performances in terms of noise figure. The two following stages are done using a GaAs MMIC. This MMIC has been implemented using a 0.25 μm PHEMT process. The design employs a self-biased configuration and inductive source feedback on the first stage in order to achieve optimum noise figure associated to good input VSWR. A two-stage design is used at 14 GHz while three or four stages have been integrated in the 18 GHz MMIC. Layout of the 18 GHz LNA is shown in Fig. 6. Depending on the required characteristics in terms of spurious rejection, the mixing section of the receiver can be implemented using either MMIC or HMIC technology. Mixers have been developed both in the HMIC and in the MMIC form. In the first case, beamlead GaAs Schottky diodes were used, while, in the MMIC approach, active dual or single gate MESFET has been used to convert the RF signal from the frequency range 12.7–14.7 GHz or 17–18 GHz down to 10.7–12.7 GHz. The remaining part of the receiver (i.e. the IF section) is implemented cascading different MMIC circuits (the same VGA used in CAMP) which provide the required gain and gain control dynamic. Better linearity (when required) is reached using a medium power amplifier (MPA) whose layout is shown in Fig. 7.

Beyond the examples shown in the previous paragraphs, other MMICs have been developed for other applications at different frequencies, or to improve the performances and reduce the dimensions of the existing products. MMIC technology is beginning to be of interest in other equipment where discrete approach used to be the most suitable solution up to few years ago. As an example, the new generation C-band CAMP (Fig.8) will use only 3 MMIC (C band VGA) in the RF section and only a multi chip module for the control section. The housing will be as small as the one used for the Ku band CAMP allowing an even more evident reduction in weight, dimensions and lead time.

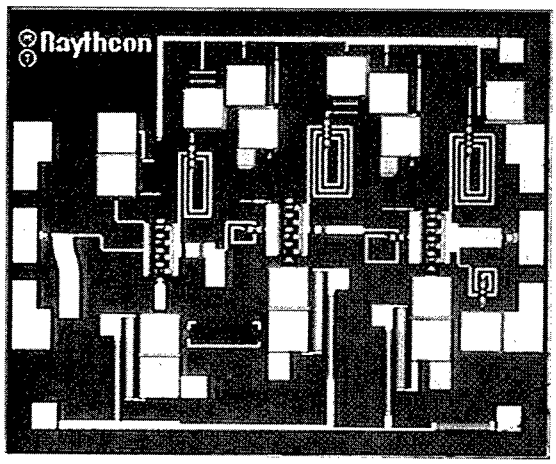


Fig. 6. Ku band PHEMT low noise amplifier.

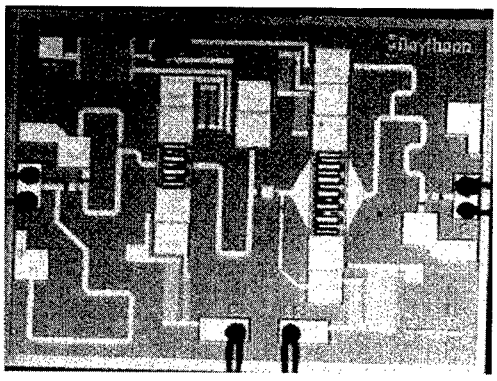


Fig. 7. Ku band medium power amplifier.

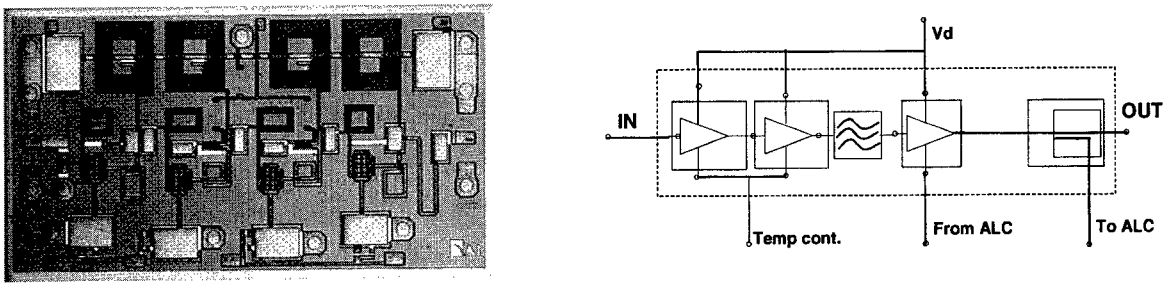


Fig. 8. C band VGA and conceptual diagram of C band CAMP.

The best proof of the successful application of GaAs MMIC technology and of advanced microwave hybrid techniques is summarized in Fig. 9. As shown in Fig. 9(a), the total number of MMICs that have been used in the last equipment production has grown from 100 in 1995 up to 3000 in 1997. Fig. 9(b) shows the number and the type of complex microwave hybrid circuits that have been produced since 1994 up to 1996. The strong increase in both MCM and HMIC circuits reveals how the confidence in MMIC technology circuit has completely modified both the integration factor and the number of produced equipment within one year.

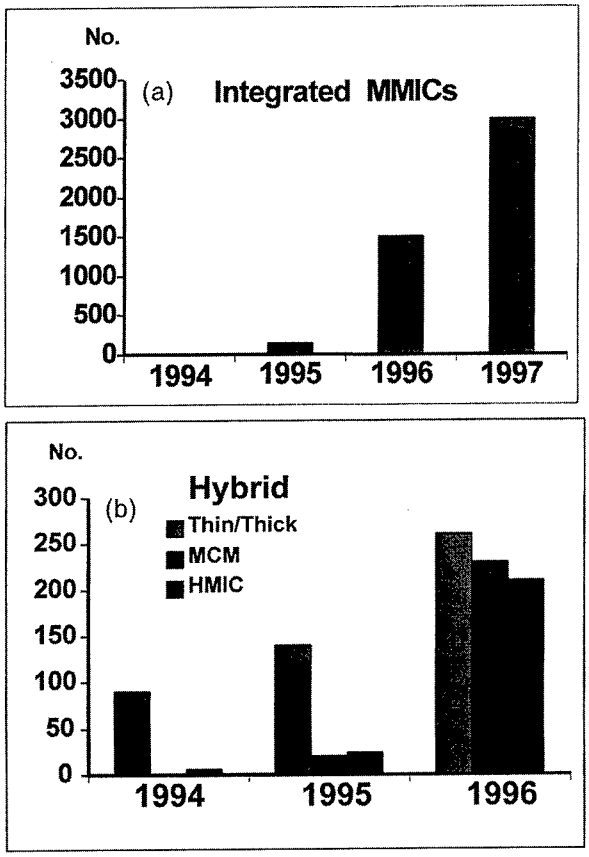


Fig. 9. (a) HMICs integration trend. (b) Flight MMICs integration trend.

3. Present developments

New satellite services (as multimedia transmission) require functions in the range from 20 to 30 GHz not only because of the amount of the transmitted data, but also to overcome the saturation of the channels in the Ku-band part of the spectrum. Moreover, a huge amount of equipment having small dimension and low weight is foreseen for the next years. In this frame the application of MMIC technology to Ka family circuits appears again to be the winning solution. As in Ku band applications the most attractive application for MMICs lies in those units which, depending on the payload complexity, are present in high quantities like receiver modules (Rx), frequency converters (DWC), LNA modules, channel amplifiers (CAMP) and in the linearisers (LIN). A number of equipment has already been developed and, as in the previous development for Ku band application, a modular approach was followed in the choice of the building block functions. A set of MMIC has been identified to be used in all the units where microwave sections are fully MMIC based. Fig. 10 shows how MMIC are used in a transparent payload configuration.

LNA: are extensively used in both DWC and CAMP. Nevertheless, receiver and down converters operate at 30 GHz while CAMPs operate at 20 GHz. Always using the modular approach concept, it was found more convenient to optimize two different MMICs at 20 and 30 GHz, respectively, rather than develop a single MMIC with such a large bandwidth. The 20 GHz LNA is also used in the lineariser equipment.

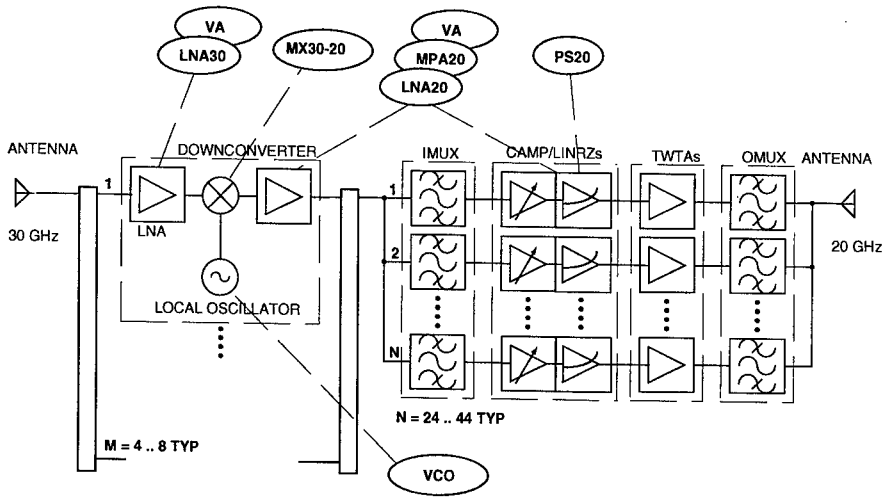


Fig. 10. Transparent payload configuration and MMIC usage plan for Ka band applications.

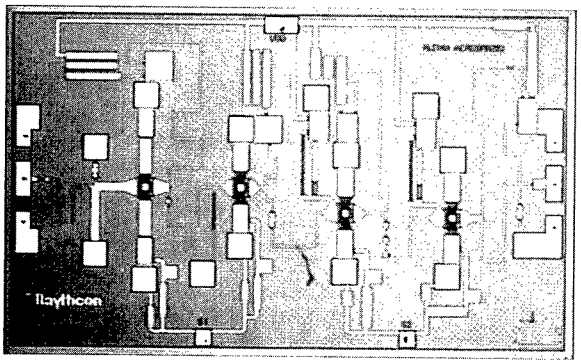
Voltage attenuator: it was found better to develop a single MMIC for this function rather than include it in a complete amplifier. VA alone can be reused as it is in several equipment and moreover a VGA function could give unstability at such high frequencies. VA is used in CAMP and DWC as a gain control element and in LIN as an element of the control function.

Phase shifter (PS): is used as a control element for LIN. It can be also used as a separate element where required.

MPA: is used in CAMP and DWC to improve the linearity of the overall circuit. The same MMIC is used in LIN as an element of the linearizing function.

Due to required performances in terms of operating frequency, bandwidth and noise figure, PHEMT represent the most attractive process to develop all the mentioned Ka band MMIC. All the defined functions have been developed using the same standard 0.25 μm low noise PHEMT process used for Ku band development. Again a multiproject foundry run has been used. In order to have the highest probability of success keeping the yield at acceptable values, more than one version for each circuit was developed increasing the number of processed wafers.

20 GHz low noise amplifier: source feedback was used to reach optimum noise figure and good return loss at the same time. A four stage configuration was necessary to get the expected gain-bandwidth performances. Each stage is self biased to allow the use of a single bias at equipment level and reduce the drift with



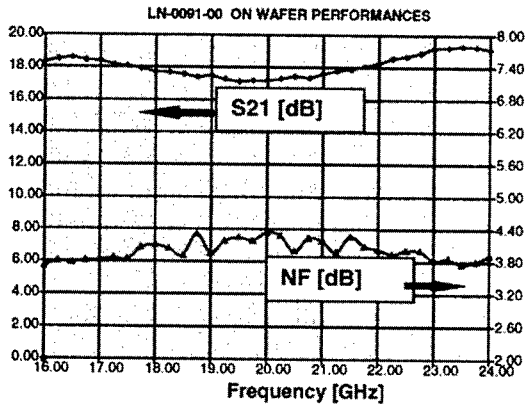
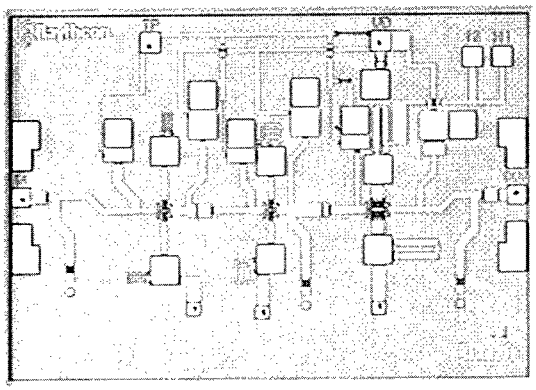


Fig. 11. 20 GHz LNA: layout and on wafer measurements.



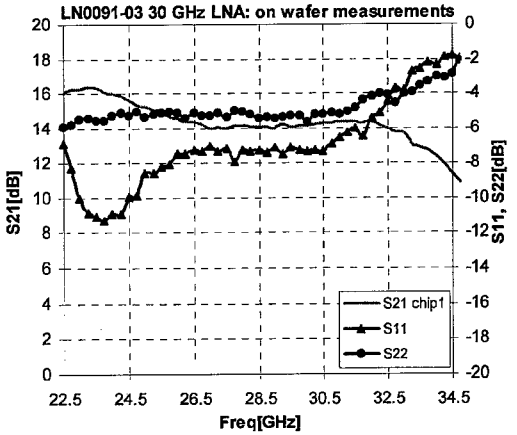


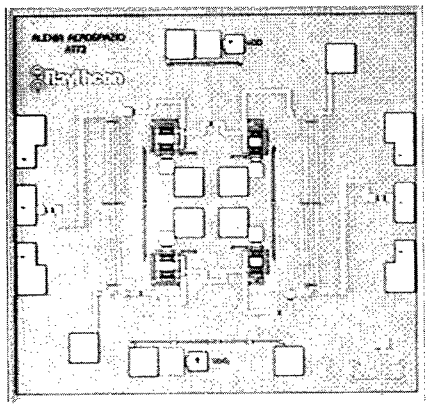
Fig. 12. 30 GHz LNA: layout and on wafer measurements.

temperature. In order to recover for the variations of the pinch-off voltage, selectable pads for proper bonding through ground have been implemented in the bottom of the MMIC. MMIC dimensions are 3×2 mm^2. Layout and on wafer measurements are reported in Fig. 11.

30 GHz low noise amplifier: the approach was similar to 20 GHz LNA. Again self biased configuration has been used. By optimizing the FET periphery of each stage, a three stages configuration was sufficient to get the expected gain keeping the power consumption within the specifications. Source feedback was used in the first stage to get at the same time good input matching and noise figure. To minimize the parasitic capacitance and inductance through ground, all the via holes have been put below the respective capacitors where possible. No selectable pads have been implemented in this MMIC to avoid further unwanted parasitic effects. Layout and measurements are reported in Fig. 12.

20 GHz variable attenuator: for reasons connected to the compatibility at equipment level, a simple configuration based on two cold FETs embedded within a pair of Lange couplers has been chosen. Enhancing a concept already published [3], the two FETs on each branch of the circuit have been chosen in order to have a more linear behavior of the attenuation versus control voltage. Layout was particularly studied to minimize parasitic effects. Fig. 13 shows the layout and the relative measurements.

20 GHz phase shifter: a simple reflective approach with cold FETS used as varactors, allowed to have both large phase range and linearity. A Lange coupler has been integrated on the chip itself as well as the net-



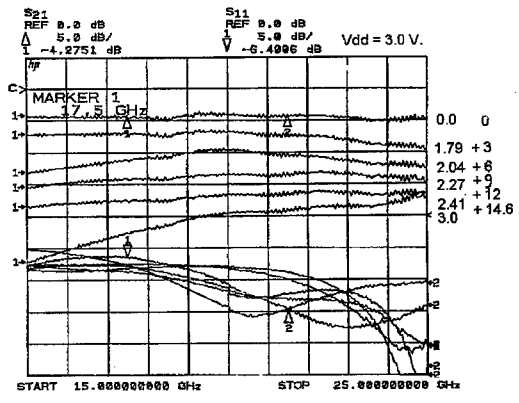


Fig. 13. 20 GHz variable attenuator.

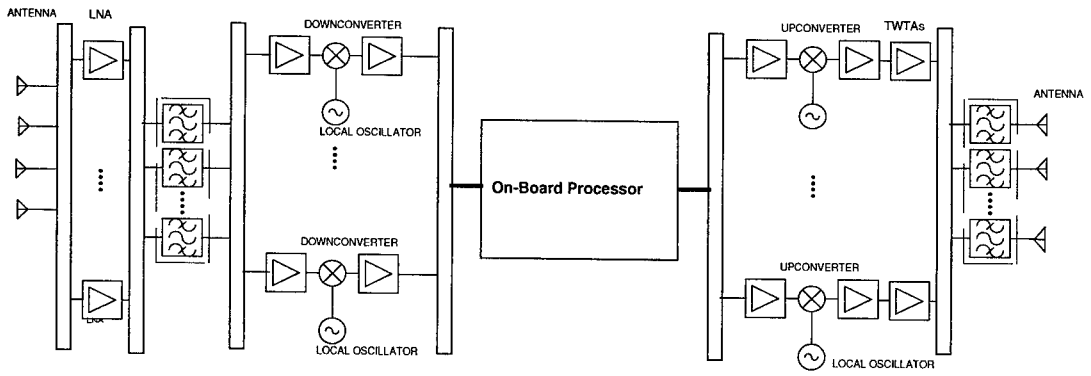


Fig. 14. Regenerative payload architecture.

work to provide the right control voltage to the diodes.

20 GHz medium power amplifier: a four stage configuration was necessary to obtain gain, power and bandwidth performances. The final stage uses a 1200 μm periphery FET and was designed to obtain required output power. The third stage was designed to drive the last stage while the first two stages were designed optimizing the gain. All the stages are self biased to keep performances with temperature variation. Selectable pads are also available to recover for process variations.

A conceptual block diagram of a regenerative payload is shown in Fig. 14. It is adopted, with some differences, in all new programs for multimedia satellite services in the constellation systems as well as in the geostationary (GEO) satellite. The modules and function required in regenerative payloads are easily derived from the transparent and other configuration. In the payload receiver section the integrated receiver usually is replaced by individual LNA functions and, through a redundancy ring, by a bench of frequency down converters. In the transmitter sections the channel amplifiers, standard or linearized, can be integrated with the up converter functions in a single module. Having as reference the transparent payload configuration (Fig. 10), it is evident where the MMIC technology is applied in this configuration.

Once the simple building blocks are defined, the concept of modular approach can be extended to all the required functions. In the payload receiver section the self-standing low noise function is easily derived starting from the integrated receiver configuration and the same approach is followed to obtain the receiver or down converter modules that in the regenerative configuration will have typical IF in C or Ku bandwidth (Fig. 15a). Exactly the same approach is followed in the transmitter section where the channel amplifier and the linearizer function are reused in combination with the frequency up converter to drive the TWTAs (Fig. 15b).

The number of satellites to be built for emerging sat-

Table 1
MMIC developed for Ku and Ka band equipment: amplifiers

MMIC	Bandwidth (GHz)	Gain (dB)	NF (dB)	1dBcp (dBm)	Power cons. (mW)	Process
Variable gain amplifier[a]	10.7–12.75	15	6	11	200	MESFET
Variable gain amplifier[b]	3.5–5.5	22	6	10	250	MESFET
14 GHz low noise amplifier	13.7–15.7	16	1.8	7	150	PHEMT
18 GHz low noise amplifier	17.0–19.0	16	2.5	7	180	PHEMT
20 GHz low noise amplifier	16.5–24.5	19	3	7	160	PHEMT
30 GHz low noise[b] amplifier	22.5–32.5	14	3	6.5	150	PHEMT
12 GHz medium power amplifier[b]	9.8–13.5	14	8	23	750	MESFET

[a] Courtesy of Alenia Research Division, Rome, Italy.
[b] Courtesy of Raytheon ADC, Boston, USA.

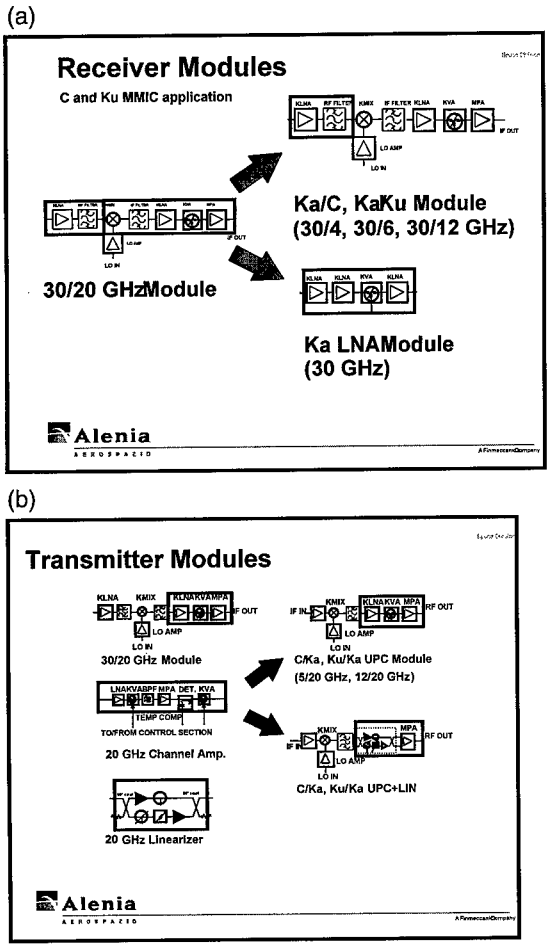


Fig. 15. (a) Payload receiver section: regenerative module configuration. (b) Payload transmitter section: regenerative module configuration.

ellite services and the payload complexity in terms of overall data throughput will support a remarkable trend in the number of modules and units to be produced with strongly reduced lead-times. For the European aerospace companies [4] the MMIC technology, as demonstrated in the consolidated Ku band payload applications, continues to be a key technology in the achievement of technical and industrial goals in the fast evolving commercial satellite world.

Tables 1–3 summarizes the main characteristics for all the MMICs developed in the last 5 years for the C-band, Ku-band and Ka band application, respectively. The chip dimensions range from 1×1 mm to 3.5×2 mm. Table 4 summarizes the MMIC usage plan as it was in early 1998. Fig. 16(a) and (b) reports the number of used MMIC per unit type and the number of equipment produced in the last two years.

4. Fabrication processes and qualification

All the mentioned MMIC have been developed using a multi project run and two well assessed fabrication processes: the 0.5 μm MESFET and the 0.25 μm PHEMT. Table 5 summarizes the main performances of these processes. The choice of the proper fabrication process is one of the key points for an effective application of the MMIC technology in space hardware. As the technology is evolving according to the market requirements, the question which arises is whether PHEMT is better than MESFET for this kind of application. The driving factor for this choice involves not only the technical matter but also other issues like cost, yield and reliability. From the technical point of view, it is obvious that for applications in the Ka band or higher frequencies, PHEMT is the most suitable process but, both MESFET and PHEMT still continue to be good for applications within the C band as well as the Ku band.

Up to a few years ago MESFET was much more characterized than PHEMT. Moreover, 4 inch wafers were already available for MESFET when PHEMT devices were produced on 3 inch wafers with an evident cost increase for the latter. Nowadays, PHEMT has reached the same maturity as MESFET since it is produced with the same rate and on 4 inch wafers showing a good yield also. As an example, Fig. 17 shows the dc and RF yield obtained over a production

Table 2
MMIC developed for Ku and Ka band equipment: mixers

MMIC	Bandwidth (GHz)	Conversion loss (dB)	LO-RF isolation (dB)	RF-IF isolation (dBm)	Power cons. (mW)	Process
18–12 Mixer	18–12	4	15	15	150	MESFET
14–12 Mixer	14–12	2	20	15	150	MESFET
30–20 Mixer	30–20	7	15	16	–[a]	PHEMT

[a] Use of resistive approach.

Table 3
MMIC developed for Ku and Ka band equipment: others

MMIC	Bandwidth	Slope (dB/GHz)	Out. Power (dBm)	Power cons. (mW)	Process
Flatness corrector	10.7–12.7	adjustable between 0 and 6	NA	< 2 mA	MESFET
		Sensitivity (GHz/V)			
VCO	9.5–11.5	1	10	100	MESFET
VCO	10.5–12	1	14	192	PHEMT
		Phase range (°)			
20 GHz phase shifter	18–20	120	–	5	PHEMT
		Attenuation range (dB)	1 dB cp(dBm)		
20 GHz variable attenuator	15–23	18	4	5	PHEMT

of 4 wafers for the mentioned developments in the Ka band. So, provided that the cost of the fabrication per chip is almost the same as the MESFET PHEMT is going to substitute the MESFET in the range between X and Ku band where the performances meet the technical requirements. For frequencies above the Ku band the use of PHEMT is almost mandatory. Fig. 18 (Table 6) shows what has been the percentage of use of the two mentioned processes in the production of the last years. It can be seen that the Ka band production is almost totally based on PHEMT MMIC, while MESFET technology is used where lower frequency or control functions are required. Another factor that will cause the PHEMT to replace the MESFET in the mentioned bandwidth is represented by the emergent power PHEMT process that allows high power in the Ka as well as in the X band. Several MMICs with an output power between 25 and 32 dBm and an efficiency greater than 30% are already available for this kind of applications. PHEMT is going to

Table 4
Equipment

DEVICE	C CAMP	Ku CAMP	Ka CAMP	FSS DWC	DBS DWC	Ka LIN	Ka LNA	Ka/Ku DWC	Ka/C DW
VGA 12 GHz		□		□	□			□	
VGA 5 GHz	□								□
MPA 12 GHz		□		□	□			□	
FC 10 - 12	□	□		□					
LNA 14 GHz				□					
LNA 18 GHz					□				
MIX 14 -12				□					
MIX 18 - 12					□				
VCO 9 - 12								□	
PH. SHIFT 20 GHz						□			
VA 20 GHz			□			□			
LNA 20			□			□	□	□	□
LNA 30			□				□	□	□
MIX 30 - 20								□	
MPA 20 GHz			□					□	

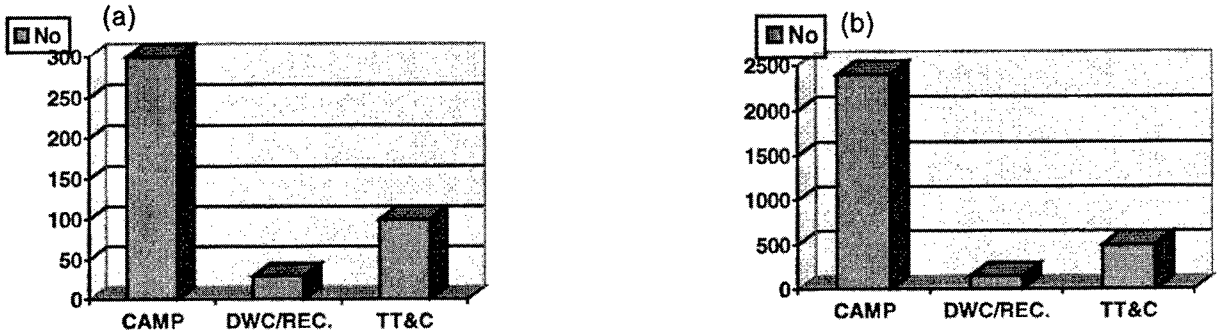


Fig. 16. (a) Number of equipment per unit type in the last two years. (b) Number of MMIC used per unit type in the last two years production.

Table 5

Parameter/process	MESFET	PHEMT
Gate length (μm)	0.5	0.25
Ft (GHz)	20	50
Gm(mS/mm)	190	350
Idss (mA/mm)	200	120
Vp (V)	−1.5	−0.6
Vbgd (V)	−6	−7

Table 6
Summary of life test performed on PHEMT 14 GHz LNA. Burn In condition: DC bias applied; Junction Temperature 150°C

Parameter	Pre B.I.	Post B.I. (240 h)	Post life (1500 h)	Delta
S21 [dB]	11.20	11.43	11.54	0.34
Idd [mA]	42.50	42.90	43.00	0.5
NF [dB]	2.99	3.05	3.06	0.07

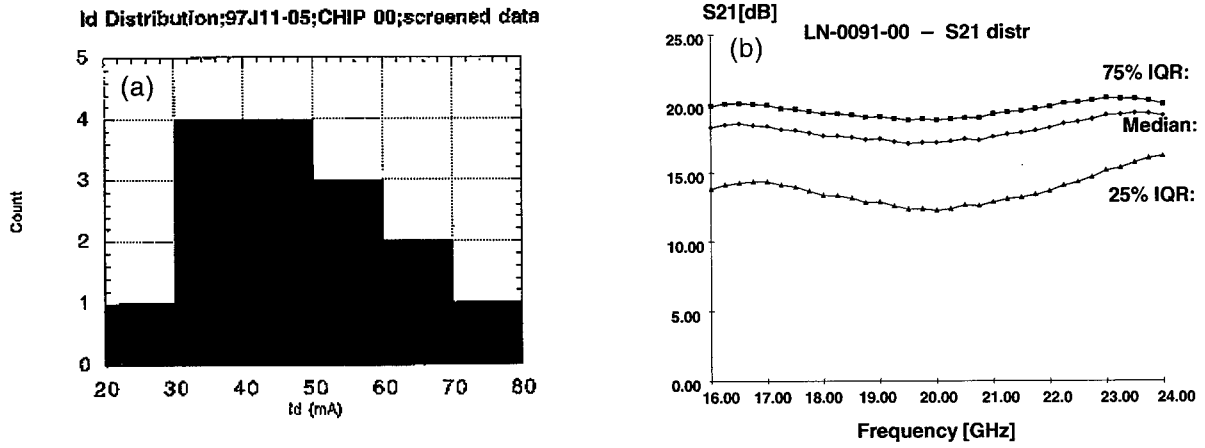


Fig. 17. (a, b) Dc and RF yield obtained over a production of 4 wafers for the mentioned developments in the Ka band.

extend also to the highest part of the spectrum (Q band, V band) through the assessment of the 0.15 μm and the InP process.

Qualification use to be a very big issue up to a few years ago when very few data were available and the selection of devices with a very narrow spread was expensive. Again, also in this case, the assessment of the processes made it possible to define a winning solution identifying a reduced number of tests to be performed at chip level to guarantee the performances over the entire satellite mission which is estimated around 15 years. Further optimization of the screening flow can be obtained if the multi-project approach is followed, identifying a set of critical structures that are representative of the implemented MMICs and adopted circuital solutions. In this way the time required for the

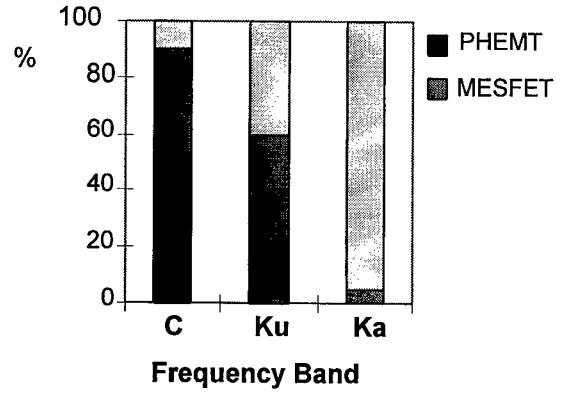


Fig. 18. Percentage use of MESFET and PHEMT technology.

qualification of one production lot can be reduced by up to 50 days in total. In Alenia Aerospazio, during the last production, more than 5000 h have been cumulated showing good reliability and a predicted lifetime of more than 1000 years at chip level. Fig. 18 shows the results of life tests and reliability analysis at MMIC level performed on PHEMT LNA developed for the Ku band production.

Beyond the example showed above, several new processes and new materials like GaN and SiC showed promising performances for high and low frequency power application at the prototype level, but they are still far from being ready for use with confidence in space hardware.

5. Future road maps

The maturity of gallium arsenide technology for space shapes in a complete different way the design approach, the manufacturing and the assembly techniques in design and development of flight units working in the microwave and millimeter wave frequency region. As shown in a previous paragraph a large number of MMIC have been integrated in space hardware. In parallel with MMIC, the assembly technique is also evolving toward new solutions likelLow temperature cofiring ceramic and flip chip mounting for high frequency and RF. A multi chip module will be the definitive assembly solution for low frequency and digital functions. The merge of these techniques will reduce the lead-time and the uncertainty of the bonding process. As far as the MMIC are concerned the trend for the next couple of years is to assess at 5000 integrated chips per year. In a long term forecast, as the use of MMIC increases, the technology itself is also evolving toward solutions like multifunction chips. So, the total number of integrated MMICs is foreseen to assess at very few chips per equipment. MCM technology will follow the same trend as for the MMIC. Very few chips mounted on alumina substrate using both traditional or LTCC technology will substitute every control section.

As far as the systems are concerned, the development of broadband satellite systems will represent a revolution in communication over the next few years [5]. The systems under development in the Ka band will require inter-satellite links to exchange information between the constellation spacecrafts or between GEO satellites in different orbital positions. Next generation systems will exploit the V band not only for ISL but also for the communication with ground. Receiver modules will integrate a low noise front end working at 60 GHz and a frequency conversion down to the regenerative payload IF (i.e. 5.5–7.5 GHz) while transmitter modules will integrate the up conversion function and the amplification to drive the power amplifier. The processes currently used in the flight production are the well stable and high yield 0.5 μm low noise MESFET and the 0.25 μm PHEMT. In the very next future two processes will be used for space applications: the 0.15 μm PHEMT for low noise functions and the 0.25 μm power PHEMT for solid state power amplification in constellation systems and large active antennas. The indium phosphide (InP) HEMT and HBT technology is beginning to mature for space application and research and development activities are planning to prepare the second generation constellation systems that could make use of the millimeter wave region (V-band) for link communication as well as inter-satellite link.

6. Conclusions

In this paper an overview of the latest technology in the microwave domain for space applications has been presented. The commercial telecommunication satellite application is a driving factor and the HEMT technology is massively used in the space hardware production. At the equipment level the trend is impressive in terms of weight, dimensions and recurring cost reduction. New families of components for the Ka band systems are presented and a future road map for next generation constellation and millimeter wave systems is presented.

Acknowledgements

Part of the PHEMT MMIC development is supported by the European Space Agency (ESA) with contract No. 12550/97/NL/SB. The authors would like to thank Mr. B. Van Rees from Raytheon ADC for the continuous support and the important cooperation

built with his company, Mr. P. Labasse and Mr. J. Primon of U.M.S. for their important role in the contract with ESA, Mr. G. Gatti from ESA for his suggestions and help, Mr. S. Rapisarda, Mr. A. Cetronio and Mrs. L. Marescialli for their job with us in the frame of development activity with Alenia GaAs foundry.

References

[1] Comparini MC, Farilla P, Feudale M, Suriani A. MMIC based modules for commercial satellite application. In: GAAS97 5th European Gallium Arsenide and Related III–V Compounds Application Symposium, Bologna, Italy, September 1997.

[2] Cazaux JL, Gregoris G, Pouysegur M, Soulard M. How MMIC technology is improving satellite transponders. In: 27th European Microwave Conference, Jerusalem, Israel, September 1997.

[3] Maoz B. Novel voltage variable MMIC attenuator. IEEE Trans on MTT 1990;38(11).

[4] Cazaux JL, et al. MCM building blocks for Ka-band telecommunication equipment. In: 2nd ESA Workshop on Millimetre Wave Technology and Applications, Espoo, Finland, May 1998.

[5] Nelson RA. V-Band expansion of the spectrum frontier. Via satellite, February 1998.

PERGAMON

Solid-State Electronics 43 (1999) 1591–1598

SOLID-STATE ELECTRONICS

RF device trends for mobile communications

Masahiro Muraguchi

NTT Wireless Systems Laboratories, 1-1 Hikarinooka, Yokosuka, Kanagawa, 239-0847, Japan

Abstract

This paper gives an overview of RF device trends for the next generation of mobile communication. First, the problems inherent in future wireless systems are described. Next, RF device technologies needed in the future wireless systems are presented. Finally, the close relationship between market size and RF device cost is discussed. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The younger consumers of today primarily desire smaller and lighter wireless terminals with longer talking and standby times. RF device technologies are of paramount importance as they represent the means by which these requirements can be met. Both GaAs and silicon device technologies have been pursued as solutions for use in wireless terminals. While GaAs devices have met the technical requirements, their cost is considered to be relatively high for many wireless applications. In contrast, silicon devices have found application in many low-tier products, but have failed to match the performance of GaAs devices. In fact, a historical pattern has emerged with GaAs devices winning the competition for initial high-tier product introductions, but losing to silicon devices for higher volume and lower cost products. It remains to be seen whether this will also hold true in the next generation.

2. Next-generation wireless systems

Next-generation wireless systems are classified according to mobility and information bit rate as shown in Fig. 1. The first class is IMT2000, which is characterized by high mobility and low capacity. IMT2000 stands for International Mobile Telecommunications 2000. The service bit rate of this system is 144 kbps at vehicular speed, 384 kbps at pedestrian speed, and 2 Mbps for indoor use with no mobility. The second class is the advanced wireless access (AWA), which is being developed by NTT [1]. The service bit rate of this system is 10 Mbps when the user terminal is fixed or moves very little. The third class is wireless local loop (WLL) or local multipoint distributed system (LMDS). These systems are characterized by no mobility and high capacity. The service bit rate of these systems is more than 20 Mbps.

Fig. 2 shows wireless system trends. The vertical axis is a measure of mobility, and the horizontal axis is the information rate. Analog cellular systems are called 1st-generation systems, and the present digital cellular and digital cordless are called 2nd-generation systems. The IMT2000 Phase 1 and the AWA, which are the next-generation systems, are called 3rd-generation systems. The 3rd-generation systems, however, only represent a midpoint in the planned development of mobile communication systems. Fourth generation systems, which are called the IMT2000 Phase 2 or the post-IMT2000, will provide high bit rates of more than 2 Mbps under high mobility conditions.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00107-0

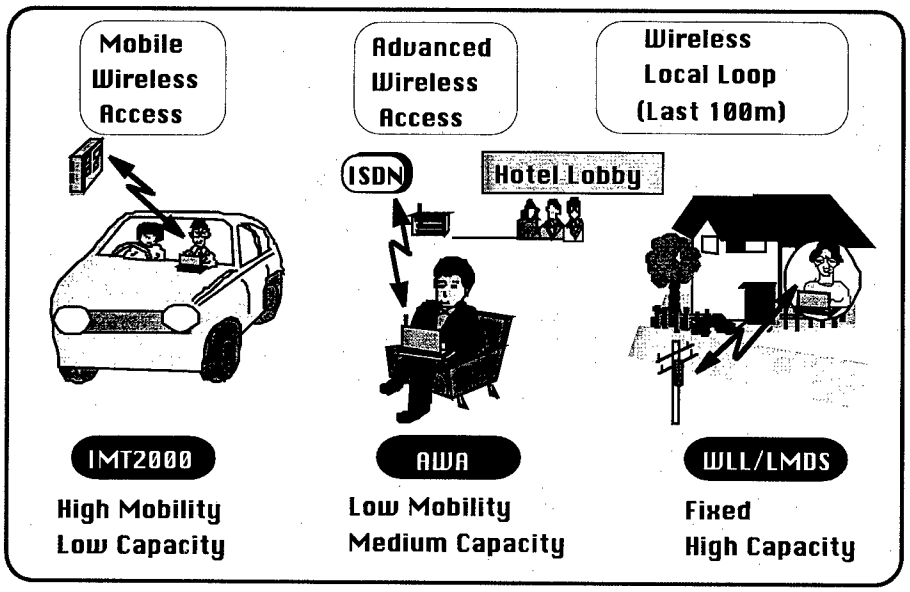


Fig. 1. Next generation wireless systems.

3. Problems in the next-generation wireless systems

Next-generation wireless system designs overcome three significant problems: propagation loss, shadowing, and multipath fading. Multipath fading is the most important issue. Because high bit rate signals occupy a wide frequency band, they are severely distorted by multipath fading.

In cellular systems, propagation loss is very different from free space propagation loss as shown in Table 1. Free space propagation loss is proportional to the square of the frequency and the square of the distance. However, terrestrial propagation loss is much greater than in free space because propagation wave undergoes refraction, reflection, and scattering. As a result, loss factor n is 4.35 in open areas such as rural environ-

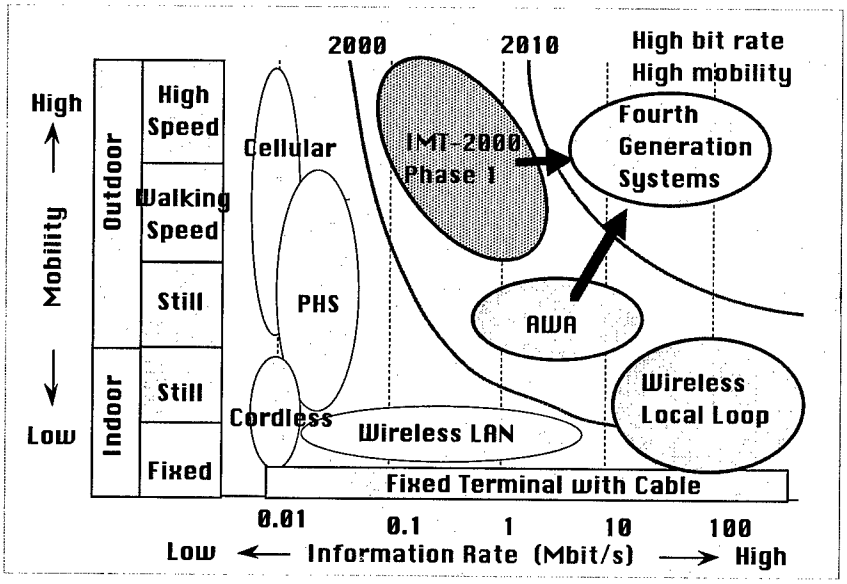


Fig. 2. Wireless system trends.

Table 1
Propagation loss in cellular systems[a]

Environment	Free space	Open (rural)	Suburban	Urban
Loss factor n	2	4.35	3.8	3–4.3

[a] Loss $\propto f^2 \times d^n$ (f: frequency, d: distance).

ments, 3.8 in suburban environments, and 3–4.3 in urban environments.

Shadowing is caused by objects that temporarily come between a base station and a mobile terminal (Fig. 3). These objects include buildings, automobiles, pedestrians, etc.

Multipath fading occurs when signals are received from multiple routes, as shown in Fig. 4. A mobile terminal receives a combination of signals of various phases and amplitudes through various routes. This situation causes a distortion in the frequency spectrum within own channel bandwidth. Multipath fading induces ripples within its own channel bandwidth, as shown in Fig. 4. This distortion causes a great increase in the bit-error-rate. This problem is compounded when the bit rate is increased. Therefore, some measures to combat this problem are required.

Equalization is the most primitive method to combat multipath fading, but it is very difficult to develop hardware and software that have effective equalization capabilities. Two other strong candidates have been proposed. One is orthogonal frequency division multiplexing (OFDM) and the other is code division multiple access (CDMA).

OFDM has been in the spotlight recently. It employs multi-carriers for one channel instead of a single carrier. The broad bandwidth of the original signal is divided into many narrow bandwidths, and then each narrow band signal is sent by a different carrier with a different frequency, as shown in Fig. 5. For example, a 5 GHz-band wireless LAN System employs 48 carriers in one channel, and each signal bandwidth is only 200 kHz, while the original signal bandwidth is about 20 MHz. Each signal is so narrow that its frequency spectrum looks almost flat. As a result, multipath fading does not increase the bit-error-rate in OFDM systems.

CDMA is a strong measure for combating multipath fading. CDMA cleverly combines frequency diversity and path diversity. Frequency diversity is achieved by using the spread spectrum scheme. Path diversity is achieved by using a RAKE receiver [2]. Here, the multipath signals are no problem because we can combine them back together into the original signal by using the RAKE receiver.

The key words in future wireless systems will therefore be OFDM and CDMA. For RF device engineers, the most important task with respect to these systems is the development of high-efficiency power amplifiers with high linearity.

4. Power amplifiers for the next-generation wireless systems

Fig. 6 shows input-output performances of a typical power amplifier and the operation point for each modulation scheme. Distortion becomes a very important issue in digital wireless systems. There are two main sources of distortion in power amplifiers: ampli-

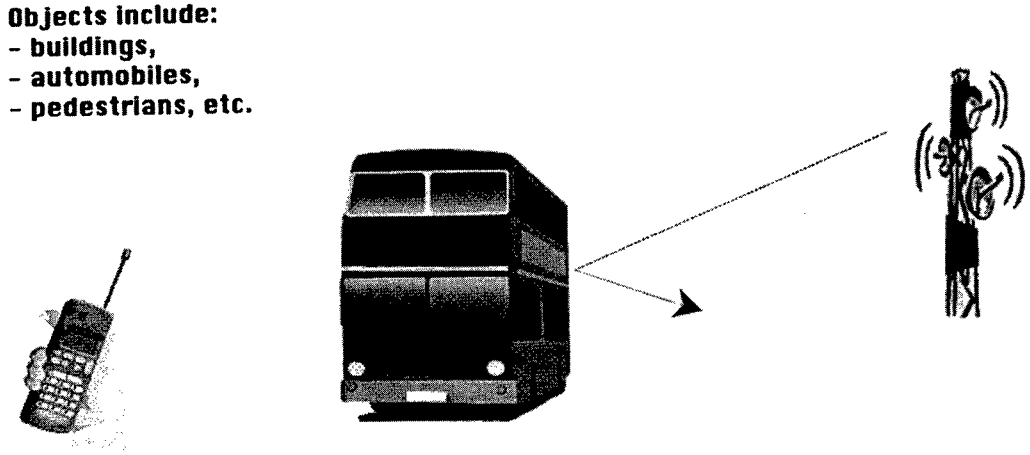


Fig. 3. Shadowing.

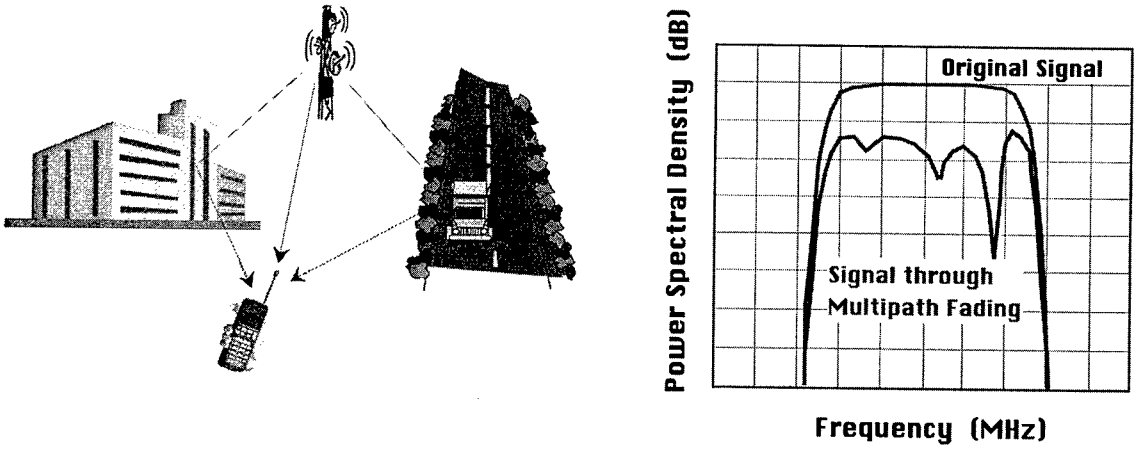


Fig. 4. Multipath fading.

tude distortion, or AM–AM distortion, and phase distortion, or AM–PM distortion.

In the linear region, the output-power curve is straight because the gain is constant. The distortion effect is very weak in this region. As input signal is increased further, the gain starts to decrease. The output power level at which the gain decreases by one dB from linear is called the 1-dB gain compression point. In general, the 1-dB gain compression point is used to describe the upper-limit of linear operation.

The power added efficiency (PAE) starts out at a very low level and rapidly increases around the 1-dB compression point. The PAE peak is at a point over the 1-dB gain compression point in the saturation region.

GMSK modulation, or Gaussian-filtered minimum shift keying, which is employed in the GSM system, can use the maximum point of the PAE because it is a constant envelope-type modulation. Power amplifiers that are designed for the GSM system can achieve more than 65% PAE. Although the constant envelope-type modulation scheme can be used with saturation amplifiers with high PAE, its wide frequency spectrum induces a penalty.

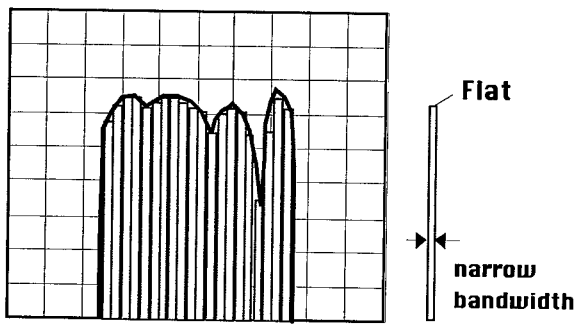


Fig. 5. OFDM system.

On the other hand, filtered linear modulations such as filtered QPSK modulations yield a narrow spectrum. However, the narrow spectrum loses its constant envelope property. Therefore, the filtered QPSK modulation, which is employed in the Japanese PDC and PHS systems, requires moderate linearity for power amplifiers. Amplifiers for the PDC and PHS can achieve more than 45% PAE, and this can be improved further by reducing the phase distortion around the 1-dB gain compression point.

Because the OFDM system employs a multi-carrier scheme, intermodulation distortion in the power amplifier is the most serious problem to be dealt with. A 5-GHz band wireless LAN system requires an output back-off, or OBO, of about 7 dB. Therefore, the PAE of a typical power amplifier is only 5% for GaAs MESFET PA and 10% for P-HEMT PA according to our experimental results.

The CDMA system requires precise output power control because of the near-far problem. The near-far problem arises at base stations. At the base station, the signal levels received from the mobile terminals are quite different due to the difference in the path lengths. The dynamic range of the received signal level becomes as large as 70 dB for path lengths between 100 m and 10 km. The adjacent channel interference becomes serious when the difference in the received signal levels is high. Therefore, some systems require an automatic transmit power control which controls the received signal level within a given range. CDMA systems must use this power control because spread spectrum signals are multiplexed on the same frequency using low cross correlation [3]. As a result, the typical operational output power for CDMA power amplifiers is between 15 and 20 dBm, while CDMA systems require a maxi-

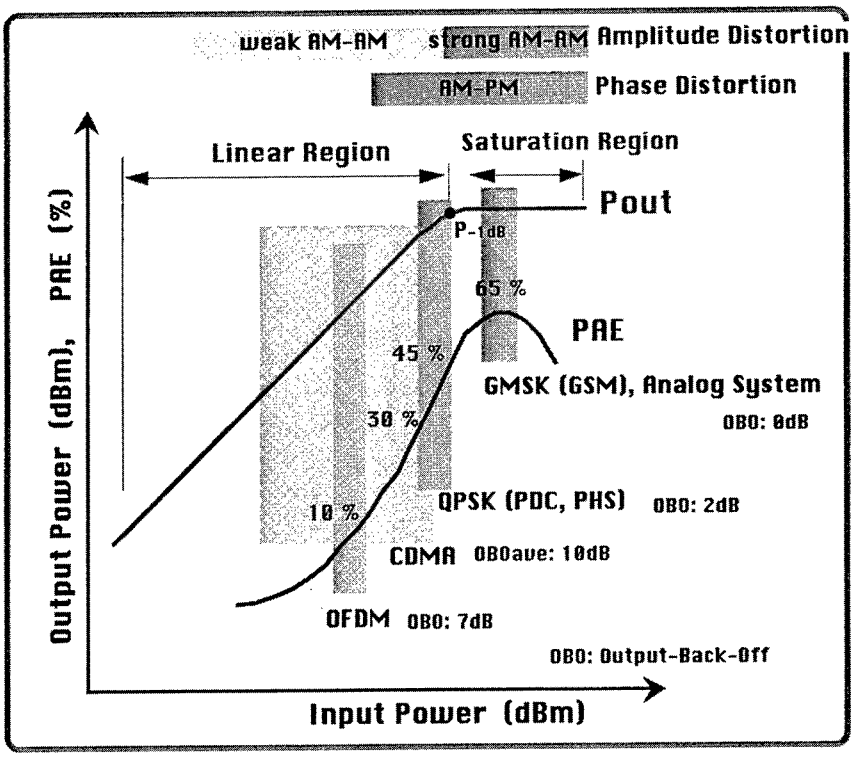


Fig. 6. Efficiency and linearity of power amplifiers.

mum output power of about 30 dBm. A large back off from the saturation power level, more than 10 dB for CDMA, can greatly reduce the average PAE of power amplifiers.

Here, the question is why current wireless systems require such linearity when it comes to power amplifiers. One answer is that the transmitted signal must satisfy the specifications of the adjacent channel interference. The specifications are defined by the power density of the adjacent channel and/or alternate channel, as shown in Fig. 7.

The original modulated signal does not have any sig-

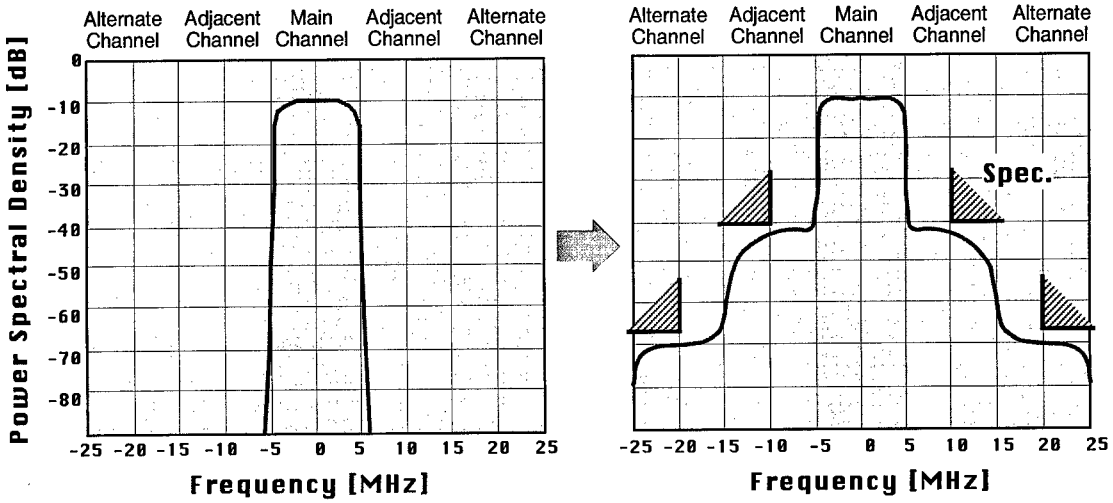


Fig. 7. Nonlinear distortion of modulated signal.

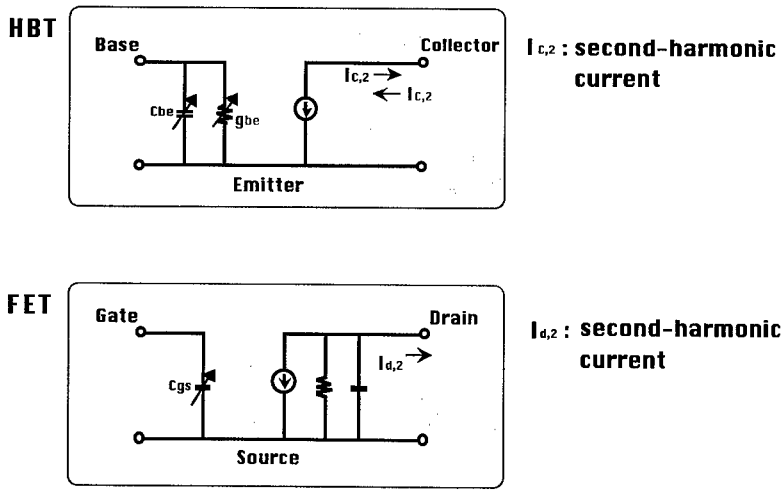


Fig. 8. Low intermodulation distortion of HBTs.

nal power outside of the main channel. However, the modulated signal amplified by a power amplifier has considerable power in the adjacent and alternate channel. The power leakage due to power amplifier non-linearity affects other channels and/or other systems. So, what is the origin of adjacent channel power leakage? Adjacent channel power leakage is caused by 3rd-order intermodulation distortion, which is intermodulation distortion between the fundamental and 2nd-harmonic signals. Alternate channel power leakage, on the other hand, is caused by 5th-order intermodulation distortion, which is intermodulation distortion between the 2nd- and 3rd-harmonic signals. Therefore, it is believed that the suppression of the 2nd-order harmonic power would be a very effective way of reducing the power leakage because the 2nd-order harmonic signal is closely related to both the 3rd- and 5th-order intermodulation distortions.

5. Linearity in HBTs and PHEMTs

It is commonly said that HBT amplifiers have good linearity. Fig. 8 shows the simplified equivalent circuits of an HBT and an FET. HBTs have two strong nonlinear components in the input side, i.e., the base-emitter capacitance, Cbe, and the base-emitter conductance, Gbe. The Cbe and Gbe generate the 2nd-harmonic currents in the output circuit. Here, those 2nd harmonic currents generated by the Cbe and Gbe have an out-of-phase difference and thus are canceled out in many cases. So, the good linearity of the HBTs results from this mechanism [4]. On the other hand, FETs have only one nonlinear component in the input side, i.e., the gate-source capacitance, Cgs. Therefore, the 2nd harmonic current in FETs is somewhat larger than that in HBTs.

The Gm and Cgs of a PHEMT are almost constant

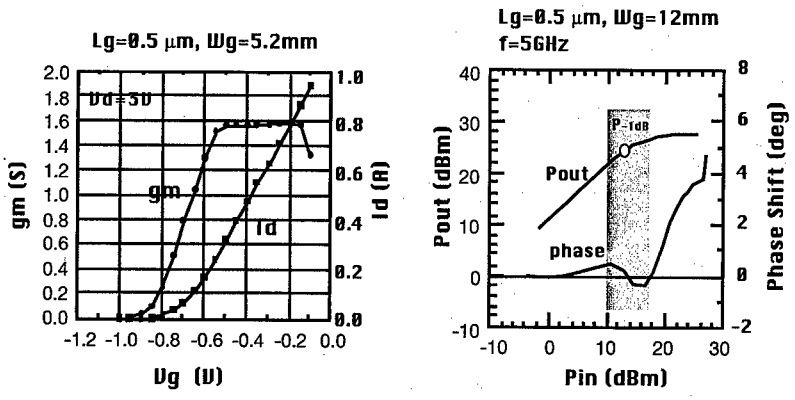


Fig. 9. Low phase distortion in PHEMT.

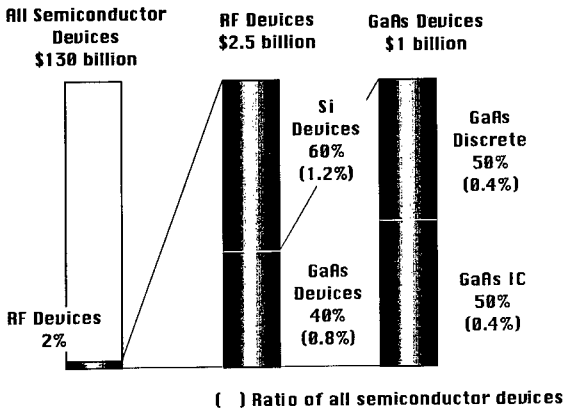


Fig. 10. RF device market in 1996.

over a wide range of gate biases. Therefore, PHEMT provides low phase distortion around the 1-dB gain compression as shown in Fig. 9. This low phase distortion performance is expected to increase the PAE of power amplifiers for QPSK systems such as the PDC and PHS.

6. Relationship between market size and RF device cost

The RF device market was about 2.5 billion dollars in 1996, which accounted for only 2% of the entire semiconductor market, as shown in Fig. 10. Silicon devices now occupy 60% of the RF device market, with the GaAs device market accounting for 40%. The total GaAs device market, except for opto-electronics devices, is about 1 billion dollars, and GaAs ICs take up just about half of the total [5]. The RF device market is considerable at 2.5 billion dollars, but it is a niche market in comparison with the whole semiconductor market.

Here, I would like to discuss why silicon RF devices are available at low cost. It is not simply a matter of silicon RF devices being cheaper to manufacture. Table 2 shows examples of the process lines at a certain company, which has had success in the cellular market. That company has three kinds of fabrication lines. Fab line A is a 0.35 μm-line which is used in producing advanced DRAMs and MPUs. Fab line B is a 0.5 μm-line that is used in producing custom digital LSIs and RF devices. Fab line C is a 1.3 μm-line that is used in producing analog devices such as power devices. Here, line-B and line-C are already depreciated. The company has a considerable market share in RF devices in the cellular market. However, only 10% of the line-B output is used for RF devices. The remaining 90% is used for custom LSIs such as SRAMs. Fortunately, since line-B operates at full fab capacity, the company provides RF devices at half the price of the corresponding GaAs devices. Therefore, the keys to being able to offer a low price are to use depreciated CMOS lines and to have some good partner products such as DRAMs. Here, process modifications for RF devices are not permitted because RF devices are usually minor products on the line.

Table 2
Examples of process lines at a successful company

Fab line	Product	Depreciation	RF devices
Line-A 0.35 μm	Digital	No	No
Line-B 0.5 μm	Digital/analog	Yes	Yes[a]
Line-C 1.3 μm	Analog	Yes	No

[a] About 10% of the line is used for RF devices.

Fig. 11 shows the relative cost per wafer vs the percentage of full fab capacity for MOSFETs. DRAMs are available at low cost because their lines operate at full fab capacity. Even a company with considerable market share in RF devices cannot operate a line at full capacity. For example, one successful company with a 60% share of the GSM handset market only operates at 10 or 20% of its full fab capacity. Therefore, some process partners such as DRAMs are indispensable for reducing RF device costs in the silicon process.

Fig. 12 shows relative cost per wafer vs the percentage of full fab capacity for GaAs MESFETs. The fab-line for GaAs MESFETs is much smaller than that of Si-CMOS and the wafer size of GaAs is 3 or 4 inches, while the wafer size of silicon is 6 or 8 inches. Therefore, if one has a considerable market share of

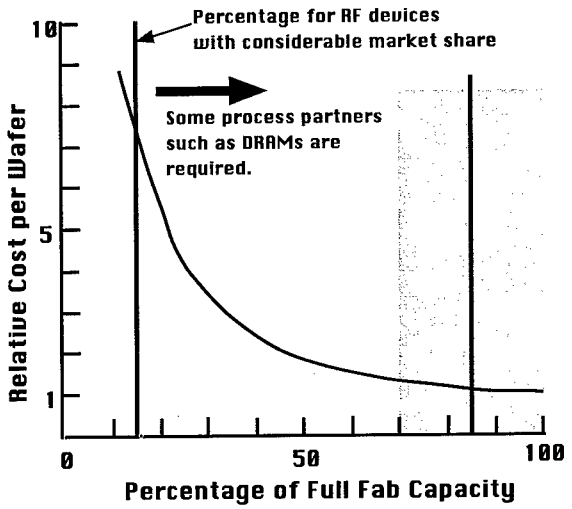


Fig. 11. Relative production cost versus percentage of full fab capacity. (In the case of MOSFETs).

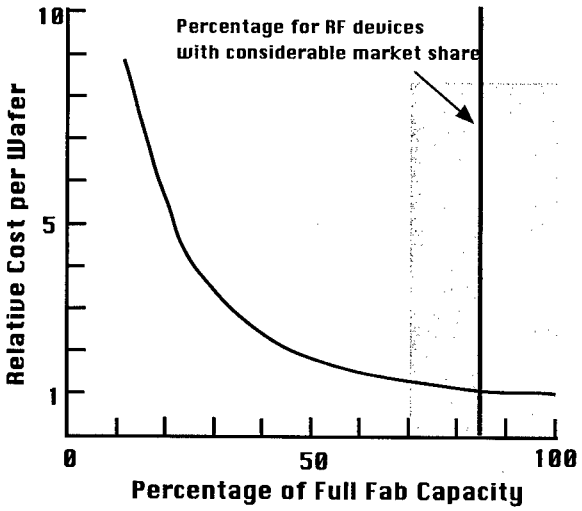


Fig. 12. Relative production cost versus perecentage of full fab capacity. (In the case of GaAs MESFETs).

the RF device market, the percentage of full fab capacity may reach 90%. So, I would say that GaAs fab lines fit the current RF device market better than Si fab lines.

7. Conclusion

The present RF device market is still small for silicon fab lines. In my opinion, GaAs fab lines fit the current RF device market better than Si fab lines. However, because the market for RF devices is limited, not all manufacturers who would like to join the market will find it feasible to do so. Although the device selection is important, it is more important to have your own original technology that distinguishes you from the competition.

References

[1] Ogawa K, Hashimoto A, Kohiyama K. Advanced wireless access system. In: Wireless personal communications. Netherlands: Kluwer Academic Publishers, 1997. p. 325–38.
[2] Akaiwa Y. In: Introduction to digital mobile communication. New York: Wiley-Interscience, 1997. p. 250–4.
[3] Akaiwa Y. In: Introduction to digital mobile communication. New York: Wiley-Interscience, 1997. p. 112–3.
[4] Maas SA, Nelson BL, Tait DL. Intermodulation in heterojunction bipolar transistors. IEEE Trans Microwave Theory Tech 1992;40(3):442–8.
[5] Compound Semiconductor (published six times yearly). Franklin Publishing, Saint Paul, MN.

PERGAMON

Solid-State Electronics 43 (1999) 1599–1606

SOLID-STATE ELECTRONICS

MM-wave integrated circuits and their applications to communication and automotive systems

P. Quentin*, C. Dourlens, M. Camiade, D. Pons

United Monolithic Semiconductors, Route departementale 128, 91400 Orsay, France

Abstract

After a short presentation of the market outlook of the commercial applications of millimetre-wave systems, we present three examples of a successful development of MMIC chip-sets: one chip-set for a full duplex Ka-band transceiver for a multipoint communication system; one chip-set for a 40 GHz MVDS subscriber terminal and, finally, one chip-set for a 76.5 GHz car radar system for ACC. In all three cases, a full MMIC architecture is used. In our view, this is the only solution to allow the high yield and high volume fabrication of the RF front-end modules at the very low costs which are required for market development. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

The wireless telecommunications are giving a formidable push to the GaAs MMIC industry. GaAs MMICs are now widely accepted by the consumer mobile phone market, the GaAs MMIC technology being used mainly for the power amplifier of the handsets. This application is at the low-end (below 2 GHz) of the frequency range of interest for the GaAs technology, which is engaged in a fierce price/performance competition with the Si or SiGe technologies.

However, the mobile phone is not the whole story for GaAs and other volume applications, especially at millimetre-wave, are blossoming rapidly and for which the GaAs MMIC technology will have no alternative. This includes short haul digital communication systems using wireless radio links, for interconnecting mobile phone base stations (point to point digital radio links) or wireless local loop systems (WLL) connecting subscribers directly to a base station for TV broadcast or for interactive multimedia distribution services (see Fig. 1).

Another extremely important millimetre-wave application which is under intense development phase, with strong investments involved, concerns the automotive industry: this is the development of forward looking car radars (FLR) for increased driving comfort and safety with, respectively, adaptive cruise control (ACC) and collision warning/avoidance (CW/A) systems.

In this paper we present a short outlook of the market perspectives for these rapidly developing millimeter-wave industrial applications, which will be crucial for the future of the GaAs MMIC industry. We will then describe with some detail the strong investments that we are devoting in UMS for the development of several families of standard MMICs designed to cover these applications.

* Corresponding author. Tel.: +33-169-33-0308; fax: +33-169-33-0309.

0038-1101/99/$ - see front matter © 1999 Published by Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00108-2

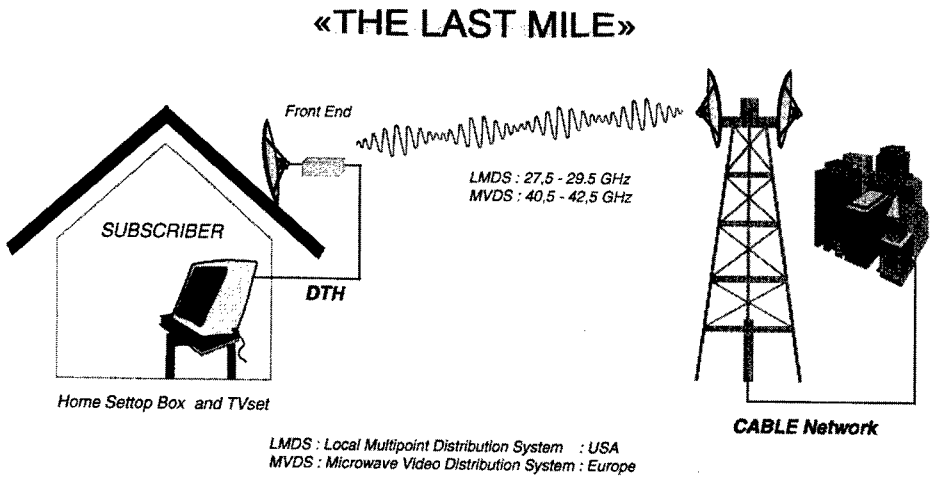


Fig. 1. Millimetre-wave multipoint distribution for interactive multimedia services (LMDS/MVDS).

2. Market outlook

2.1. Short haul millimetre-wave wireless communication systems

With respect to conventional wired high data rate communication systems, including optical fibres, the millimetre-wave wireless solutions offer the advantages (i) of being more cost effective, especially for the coverage of rural or suburban areas; (ii) of a much faster deployment, with potential rapid returns on investments and (iii) of being more easily scaleable to fit with the subscriber demand.

The market for *point-to-point digital radio links*, for mobile phone base station networks already exists with an estimated volume above 100,000 digital radios installed per year world-wide, with an annual growth rate of 25% [1]. These telecommunication systems are operating using various frequency bands from 7 to 18 GHz for long haul links and up to 60 GHz for short haul, with 38 GHz becoming a preferred choice. Originally developed in Europe, these systems are now being deployed world-wide and are especially economically interesting in the developing countries where a strong wired telephone network infrastructure is not existing. The market forecasts show that a majority of the installed radios links are and will be at Ka-band (23/26, 38 GHz) and that the market should exceed 200,000 millimetre-wave radios per year after 2000 [1]. The target prices for a complete transmit/receive module are now below US$1000 for reasonably large quantities (several 10^4 modules). This dictates the use of low-cost solutions with a highly integrated MMIC module technology, which will replace progressively the traditional hybrid approaches using a costly assembly and tuning of a large number of discrete components.

Millimeter-wave *multipoint distribution systems* are also already operating, for TV broadcast (LMDS: local multipoint distribution system), especially in the US, Canada, South America and Eastern Europe at 28 GHz [2]. Such local broadcast systems can use a simple analogue, one-way communication scheme, similar to the existing MMDS (multichannel multipoint distribution system, 'wireless cable TV system') operating at 2.5 GHz.

However, the available broader bandwidth at 28 GHz (1 GHz) opens the door for more sophisticated interactive multimedia services including telephone, video-on-demand, video conferencing and Internet/data communications, targeting residential family entertainment as well as business services. One specially interesting market segment concerns the small or home businesses (SOHO), whose demand could explode in the near future. Exactly similar interactive multimedia distribution services are also now being actively developed in Western Europe; it seems that most of these future services, using bi-directional digital radio links will be operated in Europe at a higher frequency, with a broader bandwidth: 40.5–42.5 GHz. These systems are usually called MVDS (multipoint video distribution system). In the European deregulation context for telecommunications, it is important that such wideband systems include also telephone services, allowing new operators to step in the field without depending on the installed infrastructure (with high rental costs) of their traditional competitors.

The market perspectives for such millimeter-wave multipoint distribution systems (LMDS/MVDS) are extremely exciting [2], with about 10 million subscribers world-wide expected by 2000 and about 40 million by 2005; the cost of the equipment will be of course one of the driving factors, with a target price below

US$700 (respectively, US$150) in 2000 (respectively, 2005) for the consumer market.

The competition for traditional wired systems will probably not come from LMDS/MVDS only, but also from space. The objective of the Teledesic(/Celestri) or Skybridge LEO *constellation satellite* systems is to build an 'Internet in the sky'; the GEO satellite systems Cyberstar, Astrolink or Spaceway have similar objectives. The launching of these multimedia satellite services opens however the same opportunities for the GaAs MMIC industry, for the expected large number of ground user terminals (VSAT) that these systems will generate. Most of these services will be operated at Ka-band with down-links at 19 GHz and up-links at 28 GHz.

2.2. Millimetre-wave forward looking car radars (FLR)

The market introduction of the first generation of *forward looking car radars* (FLR) for adaptive cruise control (ACC) is just starting this year. The FLR car application is even more challenging than millimetre-wave broadband telecommunications, due to the very high frequency which has been dedicated (76 GHz) and to the extremely low target prices. Most of these systems are today using a sensor based on a pulsed or FM-CW radar concept and are fabricated with discrete millimetre-wave diodes, with a Gunn diode for the emitter, Schottky diodes for immediate down conversion for the receiver and PIN diodes for signal switching. Once again, with the large volumes and the very low prices which are aimed at, cost effective MMIC solutions should rapidly take over, without talking about the specific reliability and supply problems for Gunn diodes. It is difficult to have reliable market forecasts for this just emerging application, considering the huge industrial competition issues for the car manufacturers and also all the technical, social and legal difficulties and risks associated with the introduction in the consumer market of such a sensitive new equipment.

It is however usually accepted that:

- The introduction of the ACC systems for increased driving comfort will start first, because they are much simpler to implement than collision warning/avoidance systems (CW/A).
- These ACC systems will be first introduced in Europe, starting this year, before US, because of the probable technical advance of the German car manufacturers; however the US market could follow very rapidly, because most of the US cars are already equipped with a cruise control which could be easily upgraded by an ACC system.
- The introduction of CW/A radars, bringing the desired feature of improved safety, will start a few years after (~2002–2003).
- The total number of equipped cars world-wide is expected to be about 400 K in 2000, increasing to above 5 M in 2005.

3. GaAs MMIC technology

In order to cover these millimetre-wave applications, UMS has developed several GaAs MMIC processes, which we now briefly describe:

- PH25, a pseudomorphic HEMT (P-HEMT) process with 0.25 μm Al gates, dedicated to very low-noise and small/medium signal circuits up to 40 GHz. The epitaxy material includes an etch-stop layer for improving the gate recess reproducibility and the on wafer uniformity of the electrical performances. The typical noise figure of PH25 transistors is 1.5 dB with 8 dB of associated gain at 40 GHz.
- PH15, a variant of the PH25 process with a shorter gate length (0.15 μm), for applications up to 80 GHz. The typical noise figure of PH15 transistors is 2.0 dB with 6 dB of associated gain at 60 GHz.
- PPH25, a power variant of PH25, also with 0.25 μm Al gates, with a double heterojunction/delta doping epitaxial sequence and a double gate recess for increased linear power amplification capability up to Ka-band. The PPH25 transistors exhibit a typical breakdown voltage $V_{bds} > 10$ V and about 600 mW/mm of power density at the 1 dB gain compression point.
- BES (buried epitaxy Schottky), a 1 μm Schottky diode MMIC process dedicated to millimetre-wave low conversion loss mixer circuits. The extrinsic cut-off frequency is about 3 THz.

All these processes are fully documented with design

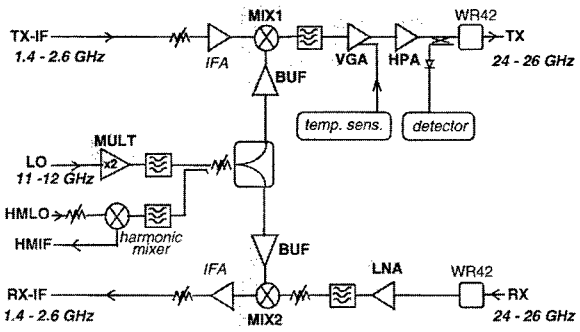


Fig. 2. Block diagram of a Ka-band point to multipoint transceiver module, stressing the integration of 8 millimetre-wave MMICs from UMS.

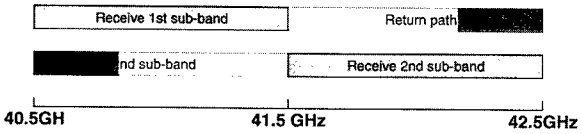


Fig. 3. The two possible frequency mappings of the MVDS 40.5–42.5 GHz bandwidth.

rules and small and large signal electrical models for MMIC design and simulation.

4. MMICs for millimetre-wave wireless communications

We have reported elsewhere the design of a highly integrated Ka-band (24–26 GHz) full duplex transceiver module for a high volume multipoint distribution system [3]. This system is developed by P.COM (one US industry leader for point-to-point and point-to-multipoint radio link systems), with the modules designed and manufactured by Thomson-CSF. The module uses a full MMIC solution in order to reduce drastically the module assembly cost and manufacturing time. The block diagram of the transceiver is given in Fig.2.

The chip-set is composed of 8 different MMICs, one LNA, one HPA, two LO buffer amplifiers, one variable-gain driver amplifier (VGA), one up- and one down-converter mixers and a frequency multiplier; all these MMICs are designed and manufactured by UMS using the PH25 process, except for the mixers which use the BES process and for the HPA, which uses the power PPH25 process. The description of the architecture and of the technology of the transceiver module and the main performances of the MMIC chip-set can be found in the proceedings of the conference mentioned above [3].

The performances of other MMICs, developed for millimetre-wave transceiver modules for another wireless telecommunication system are also given in another contribution in the same conference [4].

5. MMICs for a 40.5–42.5 GHz MVDS application

The MMIC chip-set that we are now describing in this section is developed for 40.5–42.5 GHz MVDS subscriber terminals. The RF front-end is composed of one receive circuit and one transmit circuit for the return path. It is designed to use the same satellite intermediate frequency band as for TV broadcast (0.95–2.175 GHz), in order to be fully compatible with the direct-to-home standard format (DTH) and use the same settop boxes.

Considering this available 1 GHz bandwidth, the full MVDS band is split into two 1 GHz receive subbands; the bandwidth of the transmit channel (return path) has been fixed to 300 MHz. The two possible frequency mappings are shown in Fig. 3. With the two possible polarisations for propagation, this gives 4 different signal combinations which will be used to prevent a possible interference with signals coming from adjacent cells.

The block diagrams of the receive and transmit RF circuits are given in Figs. 4 and 5. Figs. 4 and 5 also display the first generation of the MMIC chip-set that we have developed for this application. As can be seen, a full MMIC solution has been chosen in order to be

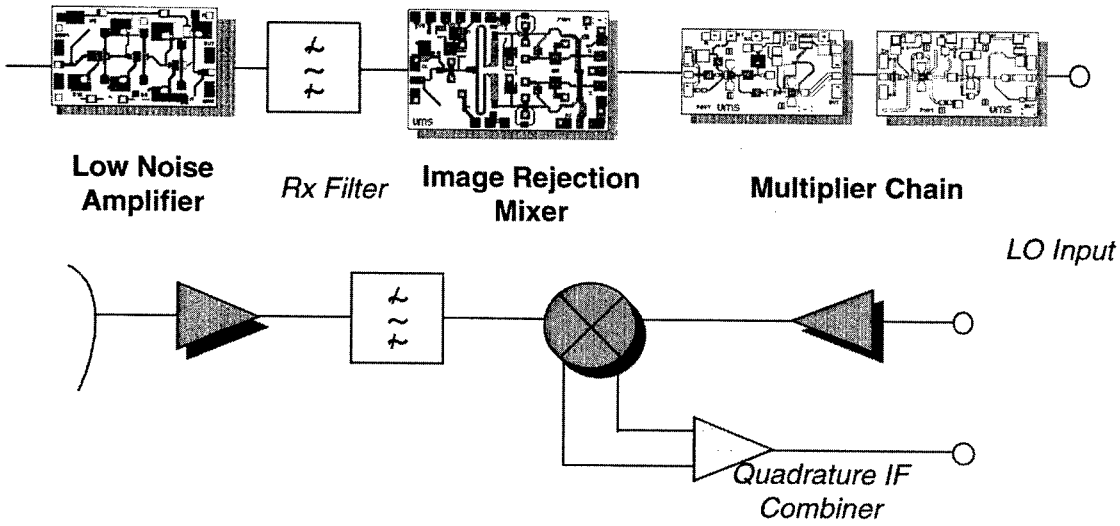


Fig. 4. Block diagram of the receive circuit of the MVDS terminal, integrating 4 MMICs: one LNA, one image rejection mixer and two frequency multipliers.

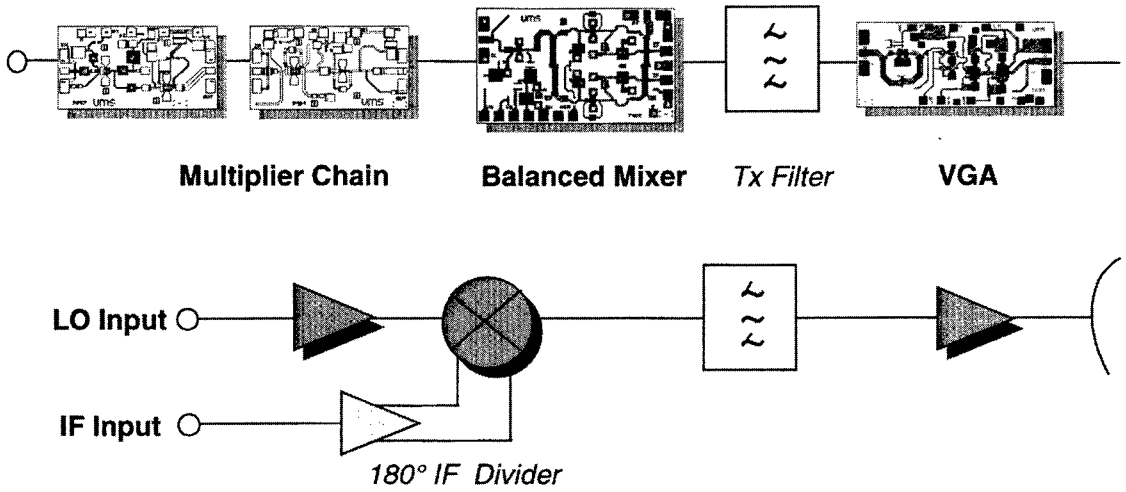


Fig. 5. Block diagram of the transmit circuit (return path) of the MVDS terminal, integrating 4 MMICs: two frequency multipliers (identical to those of the receive circuit), one balanced mixer and one variable gain amplifier.

compatible with the extremely low target prices requested for this high volume consumer application. All the MMICs of the chip-set are fabricated by using the PH25 process.

5.1. Low noise amplifier (LNA)

The present version of the chip-set uses a 35–40 GHz 3-stage LNA which is a UMS standard product (CHA2094) developed for 38 GHz point to point radio links[1]. Although it has been designed for applications at lower frequencies, its typical noise figure is below 4 dB with 19 dB of associated gain in the MVDS bandwidth. The input and output return losses are, however, insufficient and a redesign of this LNA is probably necessary to better match the MVDS bandwidth.

5.2. Down-converter mixer

This down-converter is an image rejection mixer (IRM). It is composed of two cold FETs biased at pinch-off. The LO signal is first amplified by a one-stage buffer amplifier and then split by a Wilkinson divider. The in-phase and in-quadrature RF signals are obtained using a Lange coupler. The LO and RF signals feed the drains of the two cold FETs. The image frequency rejection is achieved by an off-chip recombination of the I/Q intermediate frequency (IF) signals. This IRM is designed for an infradyne LO at 39.5/40.5 GHz (depending on the receive sub-band) for the down-conversion of the RF signal in the DTH bandwidth. The chip size of this MMIC is 2.1×1.3 mm^2.

Fig. 6 gives the on-wafer measurements of the conversion losses at the I and Q IF outputs, before recombination. The unexpected gap (1.5 dB) between the I and Q conversion losses should be suppressed in the next version of this chip; the first lot with a corrected design is under processing.

5.3. Frequency multipliers

The same frequency multiplier chain is used in the receive and transmit channels to convert the 10 GHz LO signal to 40 GHz. The chain is composed of two cascaded MMICs, each one being a frequency multiplier by 2. These MMICs are displayed twice in Figs. 3 and 4. The design of these two MMICs is similar, with lumped inductors for the first multiplier (from 10 to 20

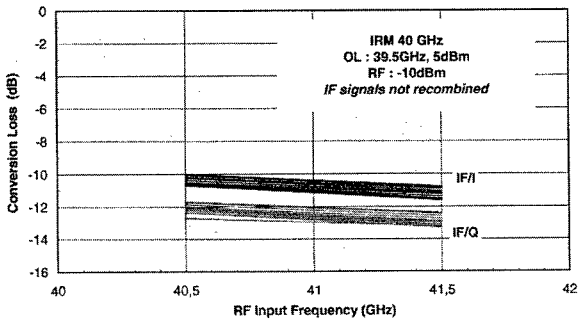


Fig. 6. On-wafer measurement of the conversion loss of the image rejection mixer on its two IF outputs. The IF in-phase and in-quadrature signals are not recombined during these measurements.

[1] The data sheets of our standard products are available on http://www.ums-gaas.com.

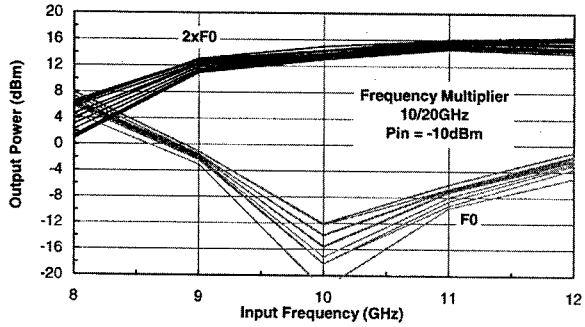


Fig. 7. On-wafer measurements of the performances of the first frequency multiplier MMIC (from 10 to 20 GHz).

GHz) and stubs for the second one (from 20 to 40 GHz). In both MMICs, the frequency multiplication is performed in the first stage of the circuit, the second stage being a buffer amplifier at the doubled frequency. The multiplication is achieved by biasing the active FETs at pinch-off. The chip size is 1.7×1 mm^2 for both MMICs.

The performances of these MMICs are displayed in Figs. 7 and 8, showing, as a function of the input frequency (F0), the output power at the doubled frequency ($2 \times$ F0) and the output leakage at F0. For the first multiplier (from 10 to 20 GHz) the gap between the output power levels at $2 \times$ F0 and F0 is better than 24 dB. For the second multiplier (from 20 to 40 GHz), it is better than 30 dB.

5.4. Up-converter mixer

The up-converter MMIC is a balanced mixer composed of two cold FETs biased at pinch-off between two Lange couplers for splitting the LO signal and recombining the RF signal at the output port, rejecting the LO frequency. Both the in-phase and inverted IF signals are required at two separated input ports; the

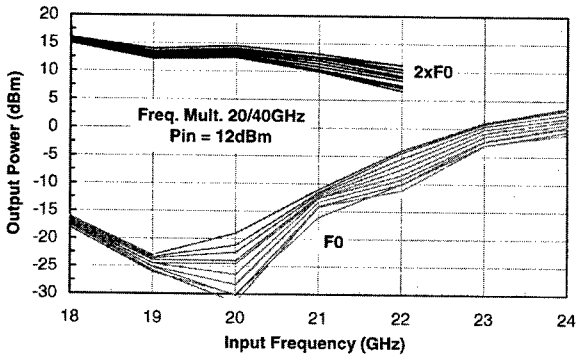


Fig. 8. On-wafer measurements of the performances of the second frequency multiplier MMIC (from 20 to 40 GHz).

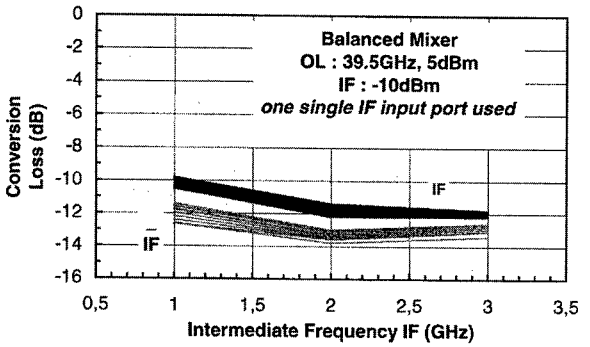


Fig. 9. On-wafer measurement of the conversion loss of the balanced up-converter mixer. The IF signal is injected in only one single input port.

inverted IF signal has to be generated off-chip in this version of the mixer. The chip size of this MMIC is 2.1×1.3 mm^2.

The measured conversion losses as a function of the intermediate frequency are given in Fig. 9, with the IF signal injected in one single input port. As for the down-converter mixer described in Section 5.2, an unexpected gap of about 2 dB is measured between the two IF channels, which should be corrected in the next version of this MMIC, which is also under processing.

5.5. Variable gain amplifier

This variable gain amplifier (VGA) MMIC controls the transmitted output power of the return path signal. The principle of the gain control is based on a double cold FET reflective attenuator using a Lange coupler. This attenuator stage is followed by a three-stage power amplifier. The chip size is 2.1×1 mm^2.

The measured small signal gain, input and return losses are given in Fig. 10. The gain control at 41 GHz is given in Fig. 11; the gain control dynamic is better

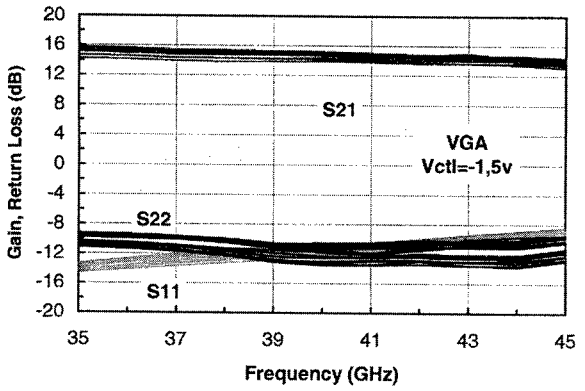


Fig. 10. On-wafer measurement of the small signal gain and return losses of the 40 GHz VGA.

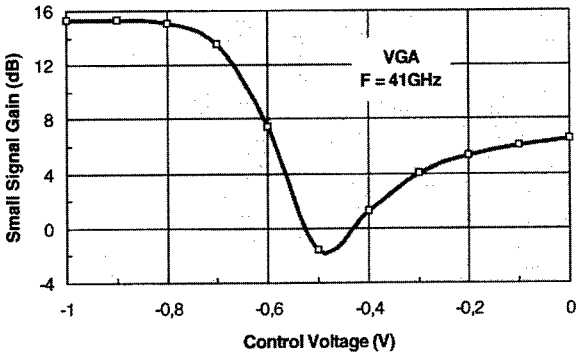


Fig. 11. On-wafer measurement of the gain control of the 40 GHz VGA.

than 15 dB. The power performances at 40 GHz are given in Fig. 12. The output power at the 1 dB compression point is in the range 16–17 dBm.

5.6. New generation of the MVDS MMIC chip-set

In order to reach the price target of the RF front-end of the MVDS subscriber terminal, we are designing the second generation of the MMIC chip-set, with the objectives to reduce the module manufacturing cost by simplifying the assembly and tuning. This will be obtained thanks to an increased level of monolithic integration:

- The two frequency multipliers will be integrated in one single MMIC, which will also integrate an X-band VCO.
- The design of a new LNA is also underway, better matched to the MVDS bandwidth; it will be monolithically integrated with the image rejection mixer. The Rx filter, displayed in Fig. 4 will be suppressed, thanks to the image frequency rejection. Also, the IRM will not integrate on-chip the IF signal combiner.

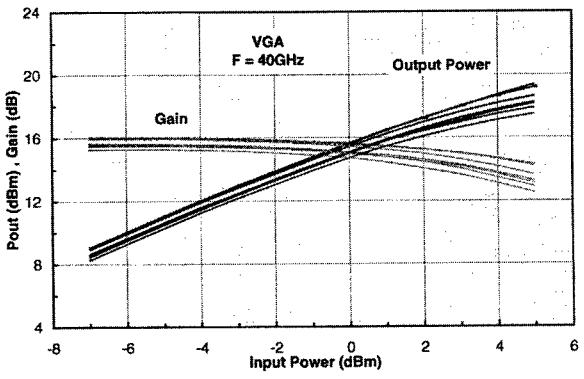


Fig. 12. On-wafer measurement of the power performances of the 40 GHz VGA.

- The up-converter mixer MMIC will be also redesigned, in order to integrate the generation of inverted IF signal on-chip. It will also integrate a buffer amplifier at the RF output.

This new generation will reduce the chip set count from 8 to 5 and will also reduce the number of passive components.

6. Forward looking car radars

The last example in this paper is a chip-set for a 76.5 GHz adaptive cruise control (ACC) car radar sensor; this radar is developed in co-operation with Thomson-CSF and Lucas Varity [5]. The radar is based on a FSK approach (frequency shift keying signal waveform). Again, a full MMIC solution has been developed. Fig. 13 gives the block diagram of the RF front-end of the sensor.

This RF front-end is fabricated with only 3 MMICs: one 38 GHz oscillator (PH25), one frequency doubler/power amplifier at 76.5 GHz (PH15) and one double mixer for the receive down-conversion (BES). This chip-set is described in more detail elsewhere [6]. The MMIC partitioning and the choice of MMIC technology for each circuit result from an extensive effort for optimising the overall RF front-end performances (and specially its noise performances, which determine directly the radar sensitivity) and the module fabrication yield, with minimum wire bonding and hybrid circuits at V-band.

The oscillator circuit is perhaps the most critical component of the front-end. The frequency is stabilised by an external resonator. With a high-Q resonator (DRO) a phase noise of -100 dBc/Hz has been measured in a test jig at 100 kHz off-carrier (38.25 GHz) with a 10 MHz tuning range. This limited bandwidth is compatible with the FSK approach of the radar.

The BES process has been selected for the down-converter mixer to minimise the conversion loss and the low frequency noise of the receiver. 7 dB and -160 dBm/Hz have been obtained, respectively, for the conversion loss and the IF noise at 100 kHz, which is a very significant improvement with respect to a former version using a PH15 cold FET approach.

7. Conclusion

Several commercial millimetre-wave applications, either for wireless communications or car radars are promised to very high volumes in the very near future, if the manufacturing costs, including the costs of the RF front-ends, can be drastically reduced. In our view,

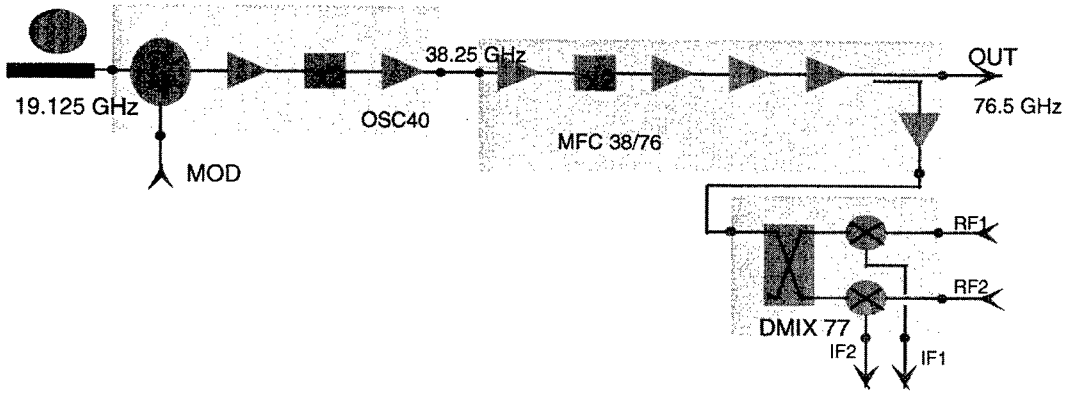


Fig. 13. Block diagram of the ACC radar RF front-end for a quasi bistatic system with two receiver channels.

this will be possible only if full MMIC solutions are followed, allowing to miniaturise and simplify considerably the RF module assembly and tuning.

We have demonstrated the successful development of high performance MMIC chip-sets for several of these millimetre-wave commercial applications. This has been obtained after strong R&D efforts for the development of the adequate GaAs MMIC technologies and the optimisation of several generations of MMICs. This optimisation has been performed in tight partnership with the system developers, our customers.

Acknowledgements

The MMIC chip-set for the 40 GHz MVDS application described in this paper is performed in the frame of a Eureka European collaboration, in partnership with Philips, ST and Thomson Multimedia (Medea A101). In this project, the MMIC development works of UMS are partly funded by the French Ministry of Industry (SERICS).

References

[1] Raffaelli. Microwave J 1997;40(10):92–6.
[2] Duncan. Microwave Eng Europe 1998:27–32.
[3] Viaud J, Levy D, Lajugie M, Vaudescal O, Tranchant S, Quentin P, Pons D. In: Proceedings of GAAS '98, Amsterdam, October, 1998. p. 709–14.
[4] Adelseck B, Schroth J, Guhl R, Braun P, Quentin P, Koblowski S. In: Proceedings of GAAS '98, Amsterdam, October, 1998. p. 725–30.
[5] Basten et al. Proceedings of Vehicle Electronics Conference, Baden-Baden, October 1998, submitted for publication.
[6] Camiade M, Dämbkes H. Proceedings of the 1998 IEEE International Conference on Intelligent Vehicles, Stuttgart, October 1998, submitted for publication.

PERGAMON

Solid-State Electronics 43 (1999) 1607–1612

SOLID-STATE ELECTRONICS

MM-wave HEMT based circuits and their system applications

Jürgen Dickmann*, Michael Berg, Volker Ziegler, C. Wölk, R. Deufel

Daimler-Chrysler AG, Research Center Ulm, Wilhelm-Runge-Strasse 11, D-89081 Ulm, Germany

Received 17 March 1999; accepted 25 March 1999

Abstract

The development status of millimeter-wave HEMTs on GaAs has improved to a status where widespread use has become normal in modern information systems for communication and sensor application. The remaining challenge for MMICs is to meet the cost margin in order to be economically attractive for commercial markets. For commercial applications it is generally preferred to offer performance identical to competing technologies at a lower cost than to offer improved performance at the same cost. Although it seems enigmatic, for certain rapidly evolving system applications, alternative InP-based generations of HEMT devices either on InP- or GaAs-substrate may become attractive. This paper focuses on this topic. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

While mobile communication in the low GHz regime is promoting the use of microwave systems on a large market scale millimeter-wave systems are now also being developed for use in civil systems with potentially high production volumes. This especially holds for radar-based sensor systems to be used in cars for intelligent cruise control or collision avoidance systems, but also for advanced communication systems. With increasing requirements concerning reduced power consumption and, forced by environmental considerations, reduced RF-power level and higher frequency bands it becomes reasonable to consider supplements to the well-established GaAs technology. InP-HEMT based components, either on InP- or GaAs-substrate (metamorphic approach) will have the potential to improve the performance of present systems in terms of power consumption, sensitivity, required chip area per function and since they show comparable electrical performance to their GaAs counterparts at larger gate geometry offer reduced fabrication costs. Moreover, they open higher frequencies to be used for advanced and novel systems.

2. Application scenario

The traditional applications for very high frequency HEMT-electronics are military systems for communication and radar-systems like battlefield communication, airborne active array radars, smart munitions, electronic warfare and missiles operating in the higher mm-wave regime like 50–94 GHz. More and more civil applications promote a market for millimeter-wave systems. This especially holds for radar-based sensor systems to be used in cars, but also for advanced communication systems. Beside this larger market scale applications, high-performance applications at much

* Corresponding author. Tel.: +49-731-505-2091; Fax: +49-731-505-4102.

E-mail address: juergen.dickmann@daimlerchrysler.com (J. Dickmann)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00109-4

smaller market scale like earth observation systems are still emerging. The most relevant applications that can benefit from InP-based HEMTs from our point of view are the following:

3. Wireless access applications

Microwave radio links play an important role in telecommunications as well as in cellar and other networks (GSM, PSTN). Thus, microwave links are used for radio links between L- and V-Band in long-haul transmission for interchange carrier networks, utility communication networks, local exchanges and several private networks. The fastest growth is foreseen in short-haul links such as between base station and controller in mobile telephone networks or between buildings in local area networks. In these systems the higher frequencies are being used i.e. 38 and 55 GHz. Point-to-Point and Point-to-Multipoint links are allowed for these higher frequencies.

Indoor local area networks take advantage of higher mm-wave (e.g. 60 GHz) frequencies such as wide frequency band width and ample number of radio channels, small equipment size and high sensitive receivers coupled to a very high gain antenna allow for low output power systems.

Frequencies in the 40 GHz-regime are used in high data-rate wireless transmission services for information transfer of video and other multimedia information. One example is MVDS (Microwave video distribution system in Europe) at Q-Band, which especially in the bi-directional version is considered to be an attractive alternative to wired cable or satellite delivery. In all above systems typical components required are up/down-converters, transmitters, receiver circuits (LNAs), frequency sources (oscillators), power control and switching circuits. First system generations have already been introduced or they are in an advanced development status. The present experience with these systems is that with respect to InP-based HEMTs enhancement is required in manufacturability and noise figure, which allows alternative technologies to be taken under consideration [1]. The new broadband satellite programs such as Teledesic will also need MMICs at 20 GHz and above. In order to reduce the power budget, low noise high gain receivers are necessary.

4. Automotive applications

In order to distinguish oneself from other competitors, passenger comfort and passenger safety have been and will be the major driving forces for the automotive industry to consider innovative technologies. In the past, RF-systems have been technically interesting, but commercially not affordable for automotive applications. Nowadays, due to the dramatic progress in RF-technology, including maturity of MMIC fabrication and module assembly, radar sensors become increasingly attractive for automotive applications. The technical advantages are obvious; virtual enlargement of the driver's visibility and hence passenger's safety under adverse weather- or traffic conditions. Exact measurement of distance and relative velocity of objects adjacent to the car, even of optically hidden objects in blind areas allow the introduction of driver assistant-systems into the car. All over the world car manufacturers are going to introduce forward looking radars for obstacle warning or intelligent cruse control functions operating at 76–77 GHz. Although many of the developed systems approaching market introduction very soon, there still remain challenges. The most pronounced once are low-cost high-volume production that forces the use of MMIC technology and very likely multifunctional approaches. From the technical point of view oscillators with high enough output power and sufficiently low phase noise and at the same time low-noise receiver circuits (e.g. LNAs) are required.

In order to make the vision of autonomous vehicle a reality, the first step is that future novel radar systems will become imaging radars in order to combine electronically scanned parallel beams with high spatial resolution. This allows for high image update rates and can be used for real time measurements while driving [2]. These systems may need much higher frequencies than 77 GHz. Higher frequencies up to at least 140 GHz are envisaged for such systems in order to achieve higher resolution, imaging capabilities, smaller size of the antenna. These applications will clearly favor InP-based components as the technology of choice.

5. High performance applications

Typically satellite-based sensor-systems normally require extreme technical performance. Earth observation systems e.g. for mapping the cosmic background radiation or identification of special species (e.g. oxygen lines) consist of a large number of detection receivers. Typically designed as direct detection receivers operating in the high frequency spectrum between 30 GHz up to several 100 GHz. In order to achieve the required sensitivity extremely low-noise amplifiers are required. In some cases cooling of the components is foreseen, which can be used to improve the low-noise performance of the LNA, which is advantageous for use with InP-based HEMTs. Other environmental observation systems use higher mm-wave frequency radars, e.g. for cloud profiling. These sys-

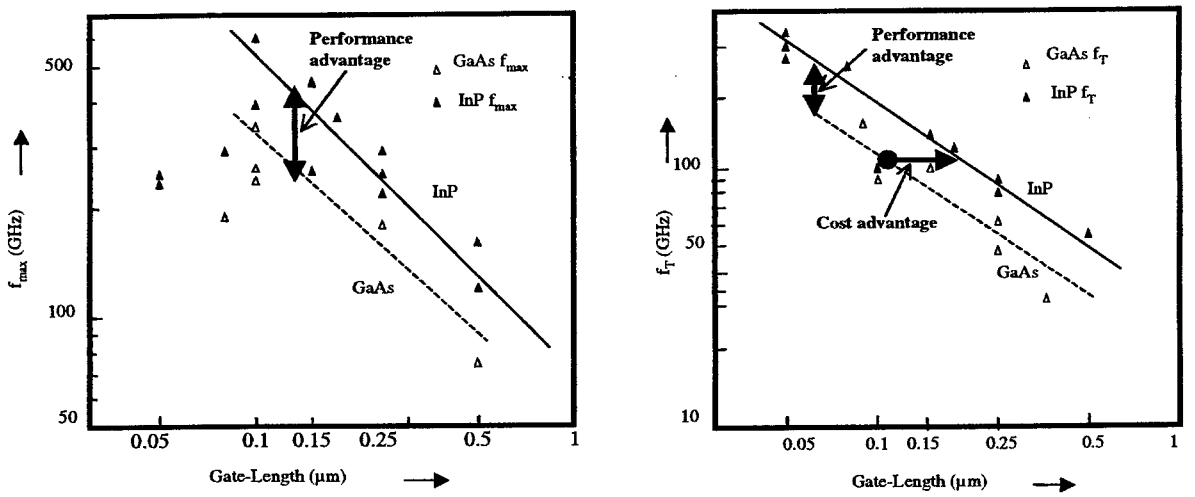


Fig. 1. (a) Power gain cut-off frequency and (b) current gain cut-off frequency as a function of gate-length for published HEMT devices.

tems require MMICs with sufficient noise- and gain-margin.

The component which is the key element in all above discussed systems and can benefit most from any kind of InP-HEMT technology (either on InP- or GaAs-substrate) is the LNA.

6. Why InP-based HEMTs?

Presently GaAs MMICs or even more discrete devices are used for first systems to be market-introduced. MMICs operating at the above-mentioned frequencies between 30–140 GHz will find increasing use only if they can be manufactured and assembled at reasonable costs. Thus, in order for these systems to be made in low-cost volumes in general two ways are possible. Cost optimization of existing fabrication lines, which becomes difficult the more mature and effective the line is still operating. The second way will help out of this dilemma is considering alternative/supplementary technologies which offer cost advantages combined with the potential to open the way for new functionality or new applications, e.g. increased receiver sensitivity and/or the use of higher frequencies. In this respect InP-based HEMT technology becomes attractive for (very) low noise and/or high gain amplifiers with increasing relevance the higher the frequency of operation is. With InP-based HEMTs the lowest noise figures and highest cut-off frequencies of any HEMT device have been achieved [3–9]. Fig. 1 and 2 summarize cut-off frequencies and minimum noise figures published in the literature so far from industrial research groups from TRW, Hughes, Lockheed Sanders, Raytheon, HP, NEC and Daimler Chrysler for GaAs and InP-HEMT devices. From this figure two potential advantageous of InP-based HEMTs become evident.

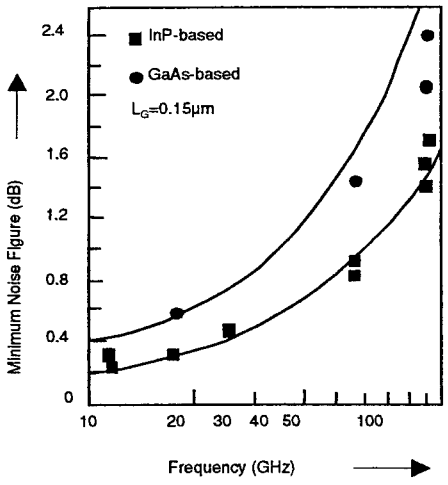


Fig. 2. Minimum noise figure for InP-and GaAs-based HEMTs. Data is collected from literature considering only results from industrial research centers.

First, performance advantages. At the same gate geometry compared to GaAs-based HEMTs InP-based HEMTs can either operate at higher frequencies or offer higher gain and lower noise per device at the same frequency. Moreover, they offer the lowest minimum-noise figure possible for any three-terminal device. This advantage in performance gives several options to the systems designer. The use of higher frequencies in new applications or improved receiver performance compared to established technologies, which

Table 1
Device performance of lattice matched InP-based HEMTs on InP-substrate

	L_G (μm)		
	0.5 Δ	0.25 T	0.20 T
G_m (mS/mm)	600	700	900
I_{Dsmax} (mA/mm)	500	500	600
f_T (GHz)	55	120	130
f_{max} (GHz)	150	250	> 300
NF_{min} (dB)/f (GHz)	1.1/@18	0.6/@18	2.3/@60
G_{ass} (dB)/f (GHz)	13/@18	11/@18	6/@60

gives several options with respect to the layout of the rf-system.

Second, cost advantage. For a given application, InP-based HEMTs offer the same gain, noise and high speed-performance at larger gate-geometry. This gives the opportunity of using 0.25 μm InP-based HEMTs instead of using 0.15 μm GaAs-devices. This in turn could mean using easier and lower cost stepper- instead of e-beam lithography combined with higher yield for the larger geometry. Another advantage, which can be deduced from this consideration, is reduced MMIC area-consumption for amplifier functions e.g. of a LNA. Two examples: For the identical two-stage V-band explorative amplifier design we achieved with InP-based HEMTs about 7 dB higher gain with the same noise figure compared to the GaAs-based version [10]. Especially for this application the amplifier could be reduced by one stage, which for this type would lead to a reduction of LNA-MMIC-area of about 40%. Another example, for a LNA operating in the MVDS frequency range a minimum noise figure of 3.5 dB with 20 dB gain is required. Calculations show, that using commercially available first class GaAs-based HEMT devices (PHEMTs) a four stage layout would be required while with our InP-based HEMTs an improved minimum noise figure of less than 2.5 dB with 20 dB gain using only a three stage version seems to be possible [11,12]. These facts definitely have an impact on cost and manufacturability of key functions for the above mentioned applications.

What kind of devices is required? We have developed a family of HEMT devices on InP substrates. Table 1 summarizes the key parameters for our workhorse, a 2 × 40 μm U-layout [10–12]. This data is sufficient for LNA-designs for the above mentioned applications. Since the gate length can further be reduced InP-based devices offer potential for future challenges. Other advantages of InP-based HEMTs compared to their GaAs-counterparts are the lower power consumption and lower bias voltages. For example, the bias for low noise is $V_{DS}=1$ V, $I_{DS}=80$ mA/mm, and for high gain $V_{DS}=2.5$ V, $I_{DS}=250$ mA/mm.

While in the past mostly microstrip lines (MSLs) have been used in MMICs as the transmission media, coplanar waveguides (CPWs) are now becoming an interesting alternative due to the dramatic advantages with respect to MMIC-fabrication (no substrate thinning, no via holes, no backside metallization hence much faster through-put) and design simplicity. Moreover, they show better electrical behavior (lower grounding parasitics, lower dispersion). Figs. 3 and 4 show some examples of low noise amplifier MMICs realized using CPW technology and $L_G=0.25$ μm InP-HEMTs from Daimler-Chrysler [10–12]. For the 77 GHz-amplifier a minimum noise figure of NF = 5 dB and a gain of 15 dB have been measured (Fig. 3). For the broad band amplifier a minimum noise figure of

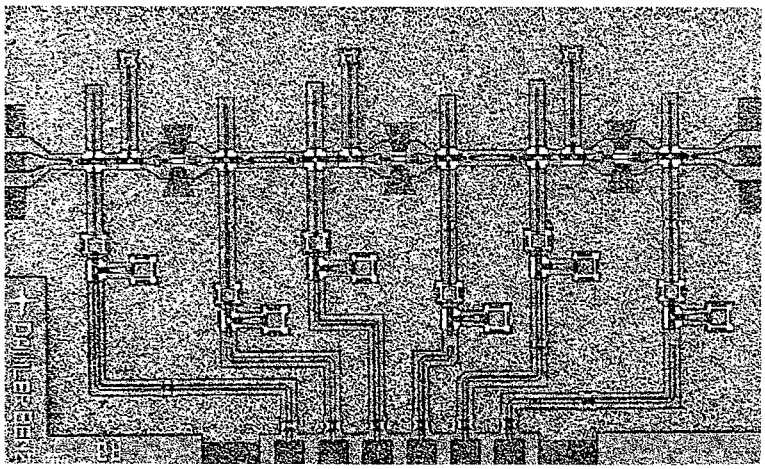


Fig. 3. 3-stage amplifier for 77 GHz operation using 0.25 pm InP-HEMTs. NF = 5.8 dB, Ga = 15 dB.

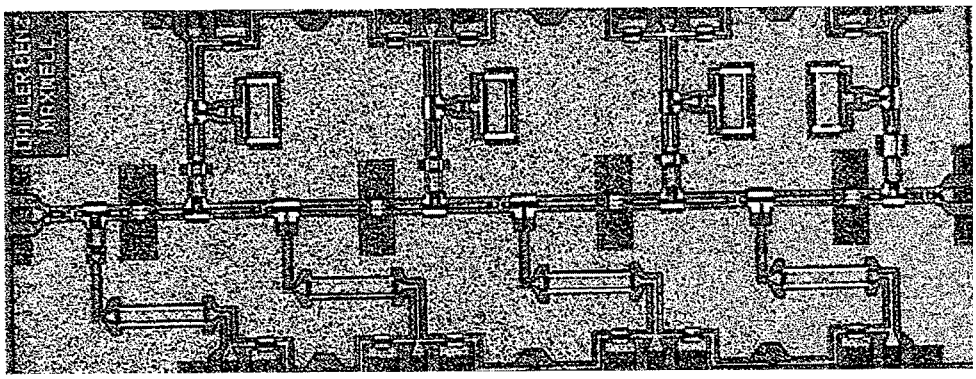


Fig. 4. 4-stage amplifier for 80–100 GHz operation using 0.25pm InP-HEMTs. NF = 7 dB, Ga = 3 dB.

NF = 7 dB with 15 dB gain within a 3 dB bandwidth of 80–100 GHz was measured. Biased for maximum gain, the amplifier achieved 19 dB.

Another function that could benefit from InP-based HEMTs is signal pathing. Concerning signal-pathing one finds different approaches like circulators or switches. Using our InP-based HEMTs as quasi-active circulators is an interesting approach for fast signal pathing. At Ka-band we have realized an active circulator using InP-HEMTs with an insertion loss of 5 dB and an isolation of −30 dB which was about 7 dB in isolation compared their GaAs counterparts. The reason is the drastically lower feedback capacitance C_{GD} by a factor of more than two [13].

7. Outlook and summary

Possible issues against a widely spread introduction of InP-HEMT on InP substrate for production are the following: higher breakage of 5–10% for GaAs and smaller wafer size which has a direct impact on manufacturing efficiency and presently higher substrate costs of InP substrates. All issues are related to the InP-substrate. In the future these problems will certainly become solved with reference to cost and wafer size along with the development of the InP-substrate market driven by the optoelectronic market. On the other hand, re-engineering of the HEMT's layer structure allows thinner buffer layers and hence cheaper epitaxial wafers [10–12].

Another approach is to move from InP- to GaAs-substrate (metamorphic approach) while maintaining the device electron physics and hence device performance [14]. Our own first experiments also support the experience reported in the open literature that InP-based HEMTs on GaAs-substrate show the same high speed performance as HEMTs with the same active layer structure on InP-substrate. Up to now it is not fully answered whether the same is true for all the other performances like e.g. power performance.

As mentioned above, the multifunctional approach will become relevant the more cost issues are of dominant concern. In order to allow the use of InP-based HEMTs in multifunctional MMICs, we presently evaluate the concept of heterointegration. A vertical non-planar technique allows the integration of different types of devices with different layer structures [15]. With this technique it will be possible to integrate metamorphic HEMTs and/or other InP-based functions via metamorphic technique with other standard GaAs functions on one chip.

The progress in both fields in the near future will decide which technology will make the race, but it is no doubt that one of both has to be considered for future systems.

References

[1] Raffaelli L. In: Workshop on The Emergence of Millimeterwave Video-on Demand Systems, 1998. p. 1998.
[2] Wenger, J. Dig. Conf. GaAs Manfacturing Technology, 1998, pp. 3.
[3] Nguyen LD, Brown A, Thomson MA, Jelloian LM. Microwave Journal. 1996. p. 96.
[4] Smith P. In: 7th Conf. on Indium Phosphide and Related Materials, May 9–13, Sapporo, Japan, 1995. p. 68.
[5] Duh KHG, Chao PC, Liu SMJ, Ho P, Kao MY, Ballingall JM. IEEE Microwave and Guided Wave Letters 1991;1:114.
[6] Streit DC, Tan KL, Liu PH, Chow PD. In: 4th Conf. on Indium Phosphide and Related Materials, Newport, Rhode Island, April 21–24, 1992, 1992. p. 682.
[7] Kao MY, Smith PM, Chao PC, Ho P. In: Proc. IEEE/Cornell Conf. on Advanced Concepts in High Speed Semiconductor Devices and Circuits, 1991. p. 469.
[8] Rhodin H, Nagy A, Robbins V, Su C, Maddan C, Wakita A, Raggio J, Seeger J. In: 7th Conf. on Indium

Phosphide and Related Materials, May 9–13, Sapporo, Japan, 1995, 1995. p. 73.
[9] Smith PM. In: Digest of IEEE Microwave Theory and Techn. Symposium, MTT-S, 1996. p. 5.
[10] Berg M, Dickmann J, Guehl R, Bischof W. Microwave and Optical Technology Letters 1996;11:139.
[11] Berg M, et al. In: Proc. 8th Int. Conf. on Indium Phosphide and Related Materials, 1996. p. 72.
[12] Berg M, et al. In: Proc. 9th Int. Conf. on Indium Phosphide and Related Materials, 1997. p. 245.
[13] Berg M, et al. In: Proc. 8th Int. Conf. on Indium Phosphide and Related Materials, 1996. p. 68–71.
[14] Wakita AS, et al. In: Proc. 10th Int. Conf. On Indium Phosphide and Related Materials, 1998. p. 223.
[15] Ziegler V, et al. Proc. GaAs'98, Amsterdam, The Netherlands, 1998, to be published.

HIGH-SPEED ICS FOR OPTICAL COMMUNICATION SYSTEMS

PERGAMON

Solid-State Electronics 43 (1999) 1613–1618

SOLID-STATE ELECTRONICS

Technologies for making full use of high-speed IC performance in the development of 40-Gb/s optical receivers

Katsumi Emura*

C&C Media Research Laboratories, NEC Corporation, 4-1-1 Miyazaki, Miyamae-ku, Kawasaki, 216-8555, Japan

Abstract

Key technologies for achieving high-speed IC modules for 40-Gb/s optical communication systems are summarized. Not only the development of high-speed devices, but also circuit and packaging designs are important to fully utilize high-speed transistor performance. By adopting a distributed amplifier circuit design, and high-speed package including newly designed, low reflection 50 Ω RF signal lines, high-performance IC modules were developed. A high-sensitivity 40-Gb/s optical receiver was fabricated with GaAs HBT ICs. An optical frontend, which consists of a waveguide pin-photodiode and HBT preamplifier IC, exhibited a transimpedance gain of 43 dB Ω and a bandwidth 31 GHz. For a distributed amplifier IC module, a 9-dB gain and a 39-GHz bandwidth were obtained. A D-type flip-flop IC module was also developed which acts as a demultiplexer for a 40-Gb/s signal. A high receiver sensitivity of −28.2 dBm was obtained for a 40-Gb/s optical return-to-zero (RZ) signal. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

As multimedia services become more sophisticated, there is an increasing demand for greater transmission capacity. High-speed optical fiber communication systems are being intensively developed to cope with these demands. A 40-Gb/s system is one research target and some experimental results have already been reported in this area, with the straightforward electrical time division multiplexing (ETDM) approaches [1–5]. At these very high speeds, technologies that make full use of device performance are desired. This paper introduces the key technologies for obtaining high-performance high-speed IC modules, such as (1) the development of a high-speed transistor, (2) broadband, flat-gain IC design, and (3) high-speed package design with optimized internal radio frequency (RF) signal lines. By adopting these technologies, the high-performance 40-Gb/s AlGaAs/InGaAs hetero-bipolar transistor (HBT) IC modules, such as an optical frontend module, a distributed amplifier module, and a D-type flip-flop (D-F/F) module, have been developed. High-sensitivity characteristics of the 40-Gb/s receiver, using these modules, are also reported.

2. AlGaAs/InGaAs HBT

The ICs were fabricated using AlGaAs/InGaAs HBT with p^+ regrown extrinsic base layers [6]. A heavily C-doped (2×10^{20} cm^{-3}) p^+-GaAs layer was regrown in the extrinsic base regions to reduce the

* Tel.: +81-44-856-8108; fax: +81-44-856-8072.
E-mail address: emura@ccm.cl.nec.co.jp (K. Emura)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00110-0

base contact resistance. The measured peaks of the cutoff frequency (f_T) and the maximum oscillation frequency (f_{max}) for the 1.6 μm × 4.6 μm emitter are 110 and 270 GHz at a collector-to-emitter bias voltage of 2 V. The transistor was specially designed for optical communications so it would simultaneously have high-speed operation (large f_T) and large output (high current density) characteristics.

3. IC design for high-speed operation

It is known that the transistor cutoff frequency f_T must be three to four times greater than the operating bit rate for broadband analog circuits (baseband amplifiers) and some digital circuits (such as D-F/Fs) [7]. Because the f_T obtained so far is around 100 GHz, there is a need for a circuit design that offers broader bandwidth operation for 40-Gb/s systems and beyond. As shown in Fig. 1, a distributed amplifier [8] is one of the attractive circuit configurations for extending the analog circuit bandwidth. In our distributed amplifier circuit design, in order to obtain a flat response from almost DC to 40 GHz, frequency dependent terminations [9] were applied in both the input and output transmission lines, as shown in Fig. 2. They prevent the degradation in gain characteristics (flatness and bandwidth) due to the mismatching between the transmission line and the termination load at high frequency. The frequency dependent termination consists of a capacitance and a resistor in parallel, designed so as to have a frequency dependence that is matched to the transmission line characteristic impedance. Stubs were inserted in both transmission lines as the tuning elements. Simulated S_{21} characteristics with and without the frequency dependent termination for a two-stage distributed amplifier are shown in Fig. 3. By having a matching in the high-frequency region, a broader bandwidth was achieved. It also contributes to having a flat response in the signal band. Moreover, in order to obtain a good performance after packaging, the effect of the outer bonding wires was taken into account when the circuit was being designed.

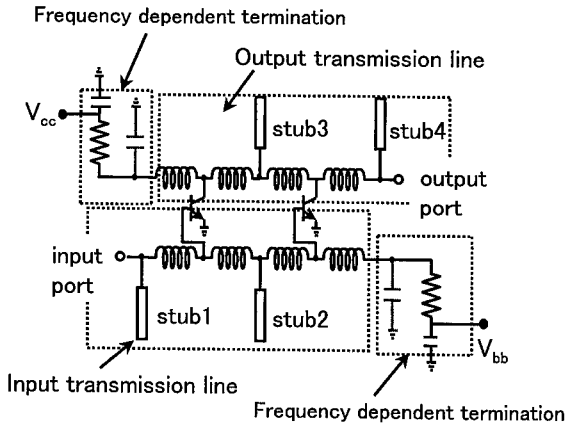


Fig. 2. Distributed amplifier circuit with frequency dependent terminations.

Using the AlGaAs/InGaAs HBTs, 2-stage and 3-stage distributed amplifiers were designed and completely packaged as shown in Fig. 4. Measured S_{21} characteristics at the connector interfaces are shown in Figs. 5 and 6. A 6-dB gain and a 44-GHz bandwidth (2-stage amplifier module), and a 9-dB gain and a 39-GHz bandwidth (3-stage amplifier module) were obtained. The module performance is in good agreement with the characteristics of the IC itself (measured by RF probes), which indicates that the degradation caused by module elements was sufficiently suppressed by the appropriate circuit design.

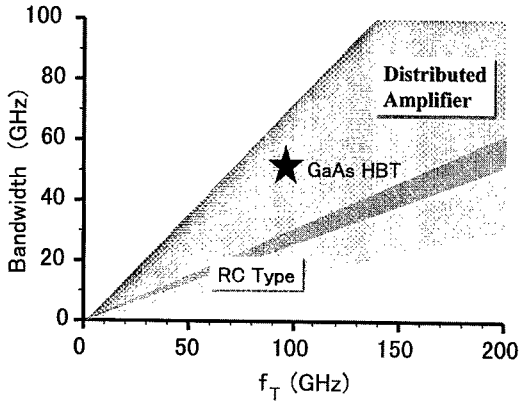


Fig. 1. Bandwidth expansion by distributed amplifier configuration.

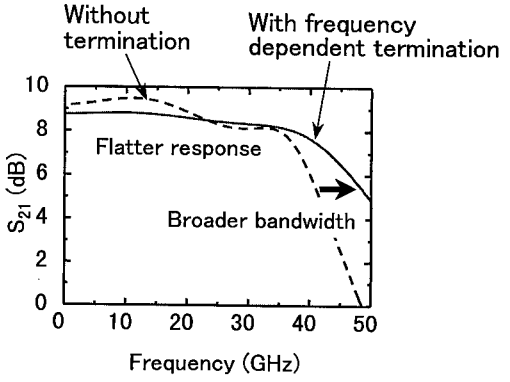


Fig. 3. Frequency response improvement for a 2-stage distributed amplifier by adopting frequency dependent terminations (simulation).

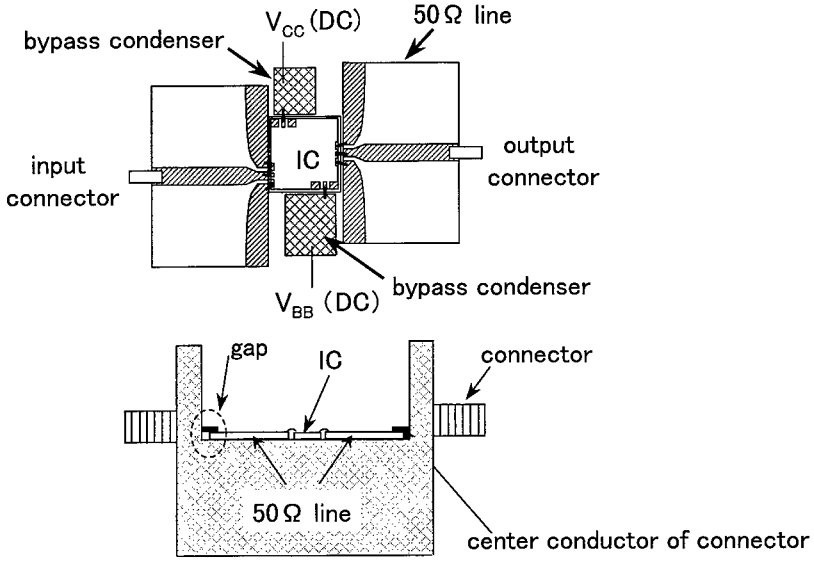


Fig. 4. Package configuration for the distributed amplifiers.

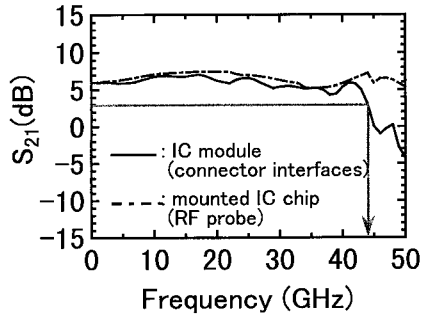


Fig. 5. Measured S_{21} characteristics for a 2-stage distributed amplifier.

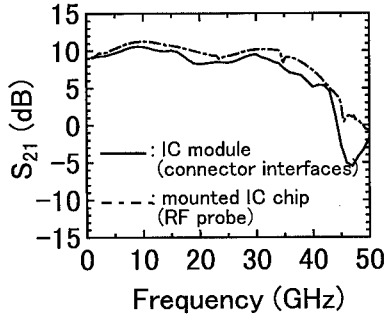


Fig. 6. Measured S_{21} characteristics for a 3-stage distributed amplifier.

4. Package design

In order to take full advantage of the performance of the high-speed IC chips, the package must be designed with care. The cavity influence must be taken into account in the three-dimensional package design. The RF signal feed line must also be designed so that a high-speed signal can be output without degradation. It is important to consider both interfaces of the IC side and the connector side so as to have a much better impedance matching. For the IC side, a conductor backed coplanar waveguide (CBCPW) was suitable, because:

- A 50 Ω line having an arbitrary signal line width can be designed without considering the substrate thickness.
- IC surface ground can be tightly connected to the module common ground via a CBCPW surface ground.

For a connector side, a microstrip line is selected, because:

- The frequency response of the module is less sensitive to any gap inside the package.

Considering these facts, a 50 Ω line with a section which transforms CBCPW into a microstrip line was designed by the two-dimensional electromagnetic analysis, as shown in Fig. 7. A broad bandwidth and a low reflection of less than −20dB was obtained, showing much better reflection characteristics compared with the all-microstrip line approach. This RF signal line is suitable for 40 Gb/s and above.

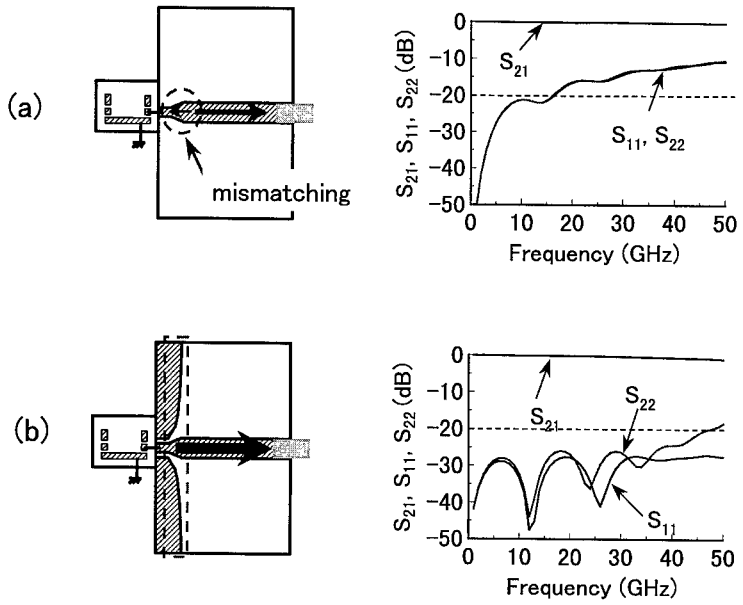


Fig. 7. Design and characteristics of RF signal lines: (a) the microstrip line approach; (b) the proposed transformational line.

5. Optical frontend design

For optical transmitters and receivers, circuit and module designs must be conducted considering optical device characteristics. For the receiver, transimpedance performance must be optimized with a photodetector (PD), which basically acts as a current source with stray capacitance. A good connection between the PD and IC is one of the most important issues. A short bonding wire may still be suitable for 40-Gb/s operation.

Using the S parameters of a preamplifier IC obtained by RF probe measurements, the overall frontend frequency response was simulated against the bonding-wire inductance. The simulation results are shown in Fig. 8. They indicate that the bandwidth broadens with the increase of wire inductance. However, the gain peaking, which may degrade the signal waveform, also increases gradually. Assuming that gain peaking of less than 1 dB is acceptable for keeping a good signal waveform, the bonding-wire inductance must be less than 0.1 nH. For this, a wire length of less than 130 μm is required, which is possible with proper packaging design.

The optical frontend module was fabricated using a waveguide pin photodiode (PD) [10] and a transimpedance-type GaAs-HBT amplifier IC [11]. The waveguide pin PD has waveguide width of 6 μm, the waveguide length of 6 μm and optical absorption layer is 0.4-μm thick. It has the junction capacitance of 30 fF and series resistance of 40 Ω, and is providing bandwidth of 41 GHz at the 50 Ω termination.

The frequency response of the fabricated optical frontend was measured by using an optical heterodyne

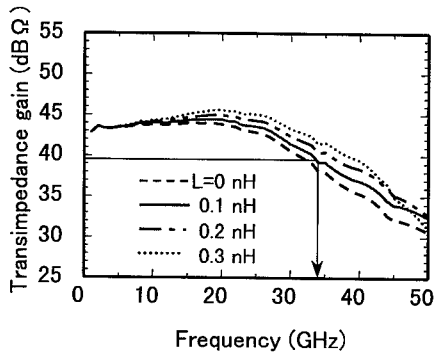


Fig. 8. Simulated frequency response of the optical frontend, taking into account the influence of the bonding wire between PD and IC.

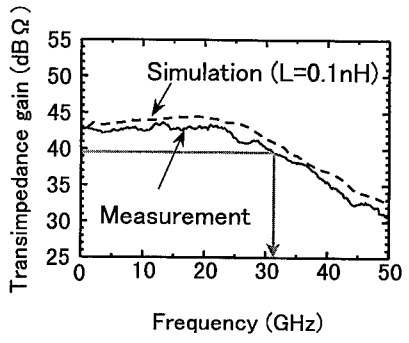


Fig. 9. Frequency response of the optical frontend module.

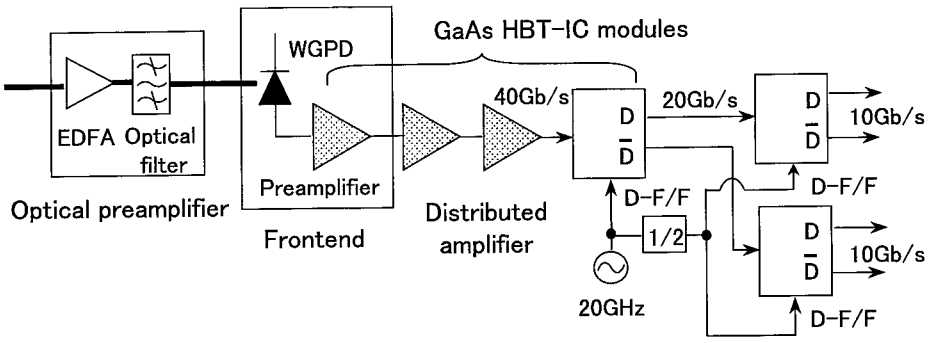


Fig. 10. Configuration of the developed 40-Gb/s optical receiver.

method, which uses two laser diodes to create a RF signal for bandwidth measurement, as follows. By receiving two beams, RF signal, corresponding to the optical frequency difference of the two laser diodes, is created through detection process. By changing the frequency difference, frequency response of the optical frontend can be measured including the PD characteristics. The result is shown in Fig. 9. Input light was injected into the waveguide pin-PD using a tapered optical fiber. Quantum efficiency was 40%. The frontend exhibits a 43-dB Ω transimpedance gain, 31-GHz bandwidth, and a flat response from DC to 25 GHz. A good eye-opening was obtained for the input photocurrent up to 2.0 $mA_{p\text{-}p}$, and a maximum output voltage amplitude of 200 $mV_{p\text{-}p}$ was obtained.

6. 40-Gb/s optical receiver performance

A 40-Gb/s optical receiver was constructed as shown in Fig. 10. The optical receiver consists of an optical preamplifier, the optical frontend, the base-band amplifiers, and demultiplexers. An optical preamplifier consists of an EDFA and an optical filter and has a noise figure of 3.9 dB and a 1.5-nm bandwidth. This amplifies the optical-signal level high enough to achieve high receiver sensitivity even with the relatively low-transimpedance optical frontend circuit. In the baseband amplifier section, the 2-stage distributed amplifier was placed before the 3-stage amplifier to obtain both broadband characteristic and high voltage output. The demultiplexer consisted of 40-Gb/s 1:2 demultiplexer

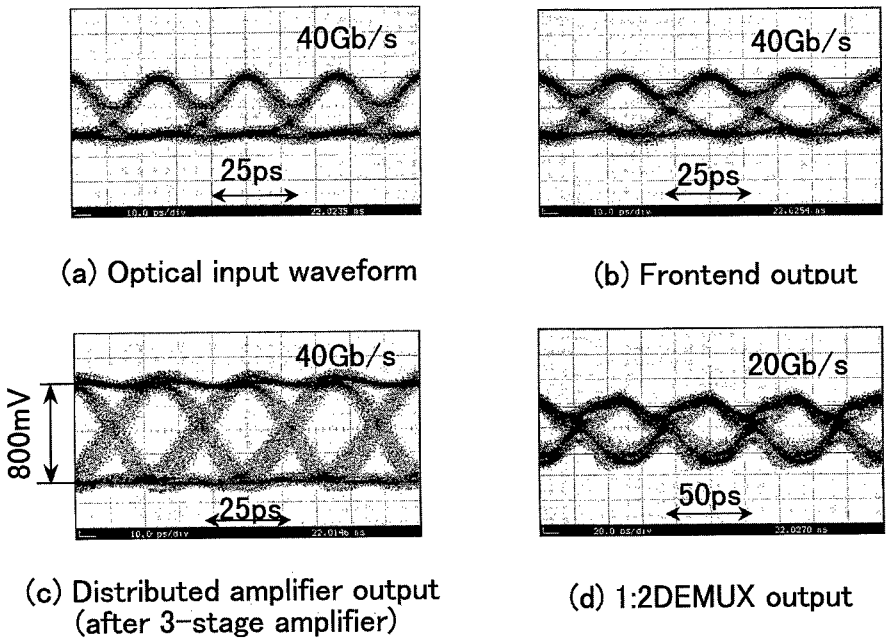


Fig. 11. Waveforms in the 40-Gb/s receiver.

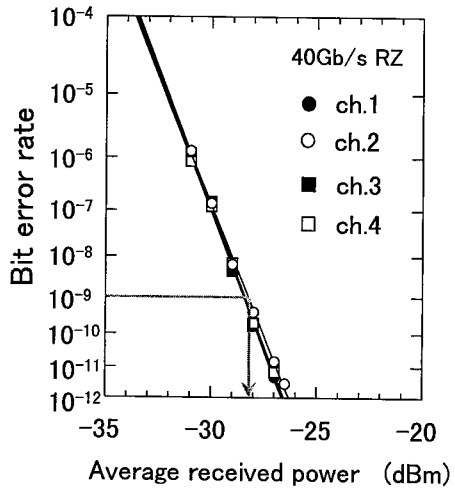


Fig. 12. Receiver sensitivity characteristics of the 40-Gb/s optical receiver.

using GaAs HBT and two 20-Gb/s 1:2 demultiplexers using SiGe HBT. The 40-Gb/s 1:2 demultiplexer was a simple static D-type flip-flop IC module. It worked as a selection switch with a 20-GHz clock signal. Thus the performance of four 10-Gb/s channels was evaluated in order by changing the clock phase.

The frequency response of the receiver was measured at the output of amplifier section by using the optical heterodyne method. A 24.5-GHz bandwidth with flat response was obtained.

Signal waveforms and receiver sensitivity for a 40-Gb/s return-to-zero (RZ) optical pulse train was measured. Fig. 11(a)–(d) shows the waveforms for various receiver positions. After the distributed amplifier, a clear eye-opening was obtained and the output voltage was as large as 800 $mV_{p\text{-}p}$ for 40 Gb/s. The reason why the output waveform seems like a non-return-to-zero (NRZ) signal is because the bandwidth of the receiver is about 60% of the bit-rate. The 40-Gb/s signal was successfully demultiplexed into two 20-Gb/s signals as shown in Fig. 11(d).

Fig. 12 shows the receiver sensitivity characteristics for four demultiplexed 10-Gb/s channels which had 2^7-1 pseudo random bit sequences. The sensitivity for the 10^{-9} bit error rate was −28.2 dBm with maximum difference of only 0.2 dB. This high receiver sensitivity was achieved by the flat frequency response of the entire receiver.

7. Conclusion

Practical technologies are being developed so that packaged high-speed IC modules can operate at the speed of IC itself. The IC design, which includes the influence of all module elements, resulted in the broad bandwidth and flat frequency responses after packaging. Optimum design of the RF signal lines in the module also contributes to the development of broad bandwidth and low reflection modules. Based on these design concepts, high-performance, high-speed IC modules for optical frontend, distributed amplifiers, and D-F/Fs are obtained. The 40-Gb/s optical receiver with these IC modules exhibited a high receiver sensitivity of −28.2 dBm for a 40-Gb/s optical RZ signal, confirming the feasibility of the developed technologies. These technological advances will accelerate the deployment of ultrahigh-speed optical communication for next-generation multimedia networks.

References

[1] Hagimoto K, Miyamoto Y, Yamabayashi Y. OFC'98 1998;WC3:114.
[2] Ohhira R, Amamiya Y, Niwa T, Nagano N, Takeuchi T, Kurioka C, Chuzenji T, Fukuchi K. OFC'98 1998;WI4:155.
[3] Yonenaga K, Matsuura A, Kuwahara S, Yoneyama M, Miyamoto Y, Yamabayashi Y, Hagimoto K, Noguchi K, Miyazawa H. OECC'98 1998;PD1-3:PD6.
[4] Yoneyama M, Sano A, Hagimoto K, Otsuji T, Murata K, Imai Y, Yamaguchi S, Enoki T, Sano E. IEEE MTT-S IMS 1997. p. 461.
[5] Washio K, Ohve E, Oda K, Tanabe M, Shimamoto H, Omai T. ISSCC98 1998;SA19.6:312.
[6] Shimawaki H, Amamiya Y, Furuhata N, Honjo K. IEEE Trans on Electron Devices 1995;42(10):1735.
[7] Sano E, Matsuoka Y, Ishibashi T. IEICE Trans Electron 1995;E78-C(9):1182.
[8] Beyer JB, Prasad SN, Becker RC, Nordman JE, Hohenwarter GK. IEEE Trans MTT 1984;32(3):268.
[9] Kimura S, Imai Y. APMC'94 1994. p. 249.
[10] Takeuchi T, Nakata T, Fukuchi K, Makita K, Taguchi K. OECC'98 1998;15B3-2:354.
[11] Suzuki Y, Shimawaki H, Amamiya Y, Nagano N, Yano H, Goto N, Honjo K. IEEE GaAs IC Symp. 1996. p. 143.



PERGAMON

Solid-State Electronics 43 (1999) 1619–1625

SOLID-STATE ELECTRONICS

SiGe HBTs and ICs for optical-fiber communication systems

Katsuyoshi Washio*

Central Research Laboratory, Hitachi, Ltd, 1-280 Higashi-koigakubo, Kokubunji, Tokyo 185-8601, Japan

Abstract

An ultra-high-speed selective-epitaxial-growth (SEG) SiGe-base heterojunction bipolar transistor (HBT) with self-aligned stacked metal/in-situ doped poly-Si (IDP) (referred to as SMI) electrodes has been developed. A 0.54-μm-wide SiGe base self-aligned to the 0.14-μm-wide emitter, which reduces collector capacitance, was selectively grown by using a UHV/CVD system. SMI electrode technology, which enables low parasitic resistance, allows the intrinsic base profile to be kept shallow, so it is well suited to a SiGe-base HBT. A 2-μm-wide BPSG/SiO_2 refilled trench was introduced to reduce substrate capacitance by reducing its sidewall element. This makes it possible to obtain a 95-GHz cut-off frequency and ultra-high-speed emitter-coupled-logic (ECL) circuit with an 8.0-ps gate-delay. As applications for these SiGe HBTs, various ICs for optical-fiber-link systems have been developed. These include a 1/8 static frequency divider with a maximum operating frequency of up to 50 GHz, a time-division multiplexer and a demultiplexer operating at 40 Gb/s, a preamplifier with a bandwidth of 35 GHz, an AGC amplifier core with a bandwidth of 32 GHz, and a decision circuit operating at 40 Gb/s. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Ultra-high-speed monolithic ICs will be key components in future optical-fiber communication systems that operate at over 10 Gb/s. Thus, ultra-high-speed transistors are needed and the feasibility of silicon bipolar technology that can operate at 20 or 40 Gb/s, has been investigated [1–3]. A SiGe-base bipolar transistor is a very attractive candidate to achieve a fast base transit time, and a below-15-ps emitter-coupled-logic (ECL) gate-delay [4,5] and a cut-off frequency of over 100 GHz [6,7] have been reported. To provide ultra-high-speed operation for 40-Gb/s systems, however, low parasitic capacitance and low parasitic resistance must be achieved simultaneously. Pure Si bipolar transistors with a self-aligned stacked metal/in-situ doped poly-Si (IDP) base electrode provide both low base resistance and low collector capacitance. In this way, high-speed circuits (for example, a 12-ps-delay ECL gate and a 45-GHz dynamic frequency divider) have been obtained even in an implanted base [8,9]. Both technologies are very suitable for obtaining high-speed performance; however, each standing alone is not sufficient to provide ultra-high-speed operation. This is because faster base transit time, lower base resistance, and lower parasitic capacitance need to be achieved simultaneously to improve the operating speed of analog and digital circuits.

To achieve this, a selective-epitaxial-growth (SEG) SiGe-base HBT with self-aligned stacked metal/in-situ doped poly-Si (SMI) electrodes has been developed [10,11]. In this HBT, a narrow graded-Ge-profile base fully self-aligned to the emitter provides a low collector capacitance. With SMI technology, tungsten films can

* Tel.: +81-42-323-1111; fax: +81-42-327-7764.
E-mail address: washio@crl.hitachi.co.jp (K. Washio)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00111-2

be selectively deposited on all poly-Si electrodes in a self-aligned manner and only emitter drive-in annealing is required after the SiGe-base formation. This means the intrinsic-base profile can be kept shallow. Furthermore, a 2-μm-wide BPSG (borophosphosilicate glass)/SiO_2 refilled trench was introduced to reduce the substrate capacitance. The low dielectric constant of BPSG/SiO_2 and the wide trench effectively reduce the sidewall element of the substrate capacitance. As a result, ultra-high-speed with a 95-GHz cut-off frequency and an ECL-gate-delay of 8 ps was achieved. Furthermore, as an application for these HBTs, ICs for optical-fiber-link systems that include a 1/8 static frequency divider with a maximum operating frequency of up to 50 GHz, a time-division multiplexer and a demultiplexer operating at 40 Gb/s, a preamplifier with a bandwidth of 35 GHz, an AGC amplifier core with a bandwidth of 32 GHz, and a decision circuit operating at 40 Gb/s, have been developed.

2. Device structure and fabrication process

A schematic cross-sectional view of an SEG SiGe-base HBT with SMI electrodes is shown in Fig. 1. To enable high speed, there are three key features in this transistor. First, to reduce the parasitic resistance of the base, emitter, and collector, tungsten films were selectively stacked in a self-aligned manner on in-situ boron-doped poly-Si (IBDP) as the base electrode and on in-situ phosphorous-doped poly-Si (IPDP) as the emitter and collector electrodes. Especially in the case of the narrow emitter, a metal film deposited at the bottom of the emitter poly-Si effectively reduces the emitter resistance. This area is shown by the SEM and schematic cross-sections of the main region in Fig. 2. The sheet resistance of the base electrode made from a tungsten/IBDP film was only 2 Ω/□, which is about 1/50 that of a p^+ poly-Si film of the same thickness. The contact resistance between the tungsten and IBDP was 20 Ω·μm^2. The base resistance can therefore be effectively reduced.

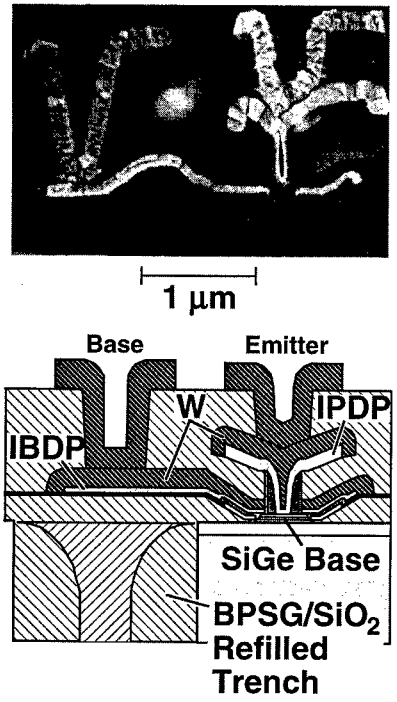


Fig. 2. An SEM cross-section of the main region of the SEG SiGe-base HBT. The emitter is 0.14 μm wide.

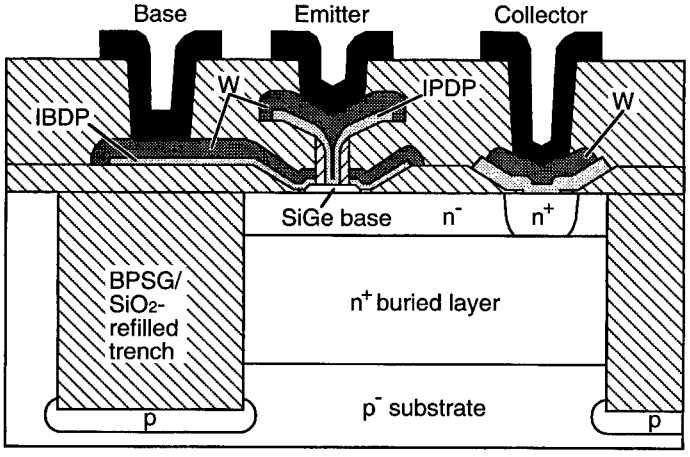


Fig. 1. A schematic cross-sectional view of an SEG SiGe-base HBT with SMI electrodes and a BPSG/SiO_2 refilled trench. Tungsten films were selectively stacked in a self-aligned manner on in-situ boron-doped poly-Si (IBDP) as the base electrode and on in-situ phosphorous-doped poly-Si (IPDP) as the emitter and collector electrodes.

Second, a 2-μm-wide and 4-μm-deep BPSG/SiO_2 refilled trench was introduced to reduce the substrate capacitance. (A wider trench can be used in the case of relatively low-integration ICs to obtain high-speed operation). This trench was filled with both CVD–SiO_2 formed at the sidewall and BPSG formed through reflow by annealing and etch back by a wet-etch process. The low dielectric constant of BPSG/SiO_2 and the wide trench reduces the substrate capacitance, especially that of the sidewall element. The effect of the wide BPSG/SiO_2 refilled trench on substrate capacitance is shown in Fig. 3. With a conventional 0.6-μm-wide poly-Si/SiO_2 refilled trench, the substrate capacitance was 3.8 fF at zero bias and reached a minimum of 3.3 fF at a reverse bias voltage of about 2 V. These capacitance values are a result of the high sidewall capacitance of 0.21 fF/μm caused by the narrow trench and the high dielectric constant of poly-Si (about three times that of SiO_2). On the other hand, in the case of the 2-μm-wide BPSG/SiO_2 refilled trench, the substrate capacitance was 1.5 fF at zero bias and fell to a minimum of only 0.6 fF. Here, the sidewall capacitance was decreased to 0.027 fF/μm, which was about 13% of the conventional value.

Last, to obtain both a high cut-off frequency and low parasitic capacitance, a SiGe-base self-aligned to the emitter was applied. A schematic cross-section of the intrinsic SEG Si/SiGe base is shown in Fig. 4. The process steps for fabricating the self-aligned SiGe base structure were as follows. After a 0.15-μm-thick epitaxial layer was formed, trench and wedge-shape isolation [12], seven films of thick CVD–SiO_2, undoped poly-Si, thin SiO_2, IBDP, Si_3N_4, thin poly-Si, and thin SiO_2 were deposited. The upper four films for the emitter area were patterned, then an 80-nm-thick CVD–SiO_2 side-spacer to isolate the emitter from the base was

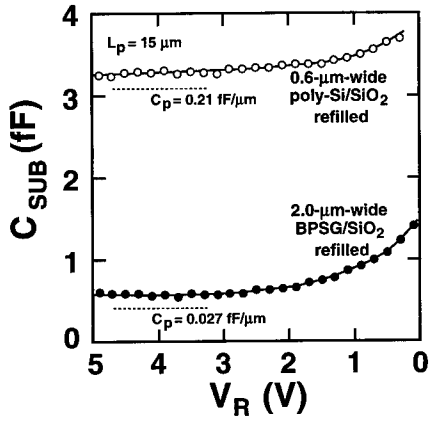


Fig. 3. The effect of a 2-μm-wide BPSG/SiO_2-refilled trench on the substrate capacitance compared with that of a conventional 0.6-μm-wide poly-Si/SiO_2 refilled trench.

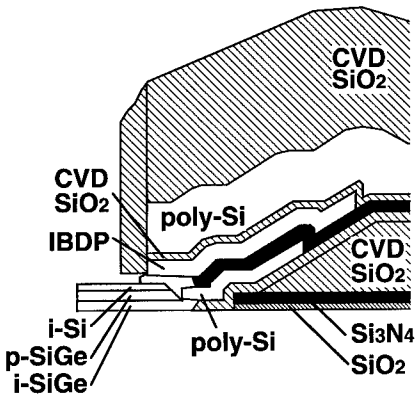


Fig. 4. A schematic cross-section of the intrinsic SEG Si/SiGe base just after the SEG was finished.

formed. After the lower three films were etched off, the wafer was inserted in a UHV/CVD chamber and any contamination on the Si surface was removed by H_2 cleaning at a partial H_2 pressure of 1300 Pa at 850°C [13]. The 0.54-μm-wide SiGe base and the Si-cap multilayer self-aligned to the 0.14-μm-wide emitter were selectively grown by using a UHV/CVD system with Si_2H_6, GeH_4 and B_2H_6 at 575°C for the SiGe and 600°C for the Si [14]. The whole surface except the base area was covered with SiO_2, because SiGe was grown selectively with respect to only SiO_2 (Fig. 5). The poly-Si/SiGe base contact formed simultaneously with the Si/SiGe intrinsic base was grown on the buffer poly-Si and beneath the IBDP. The intrinsic SEG Si/SiGe base was appropriately connected with the extrinsic IBDP base electrode; this self-aligned structure reduces collector capacitance. The buffer poly-Si between the IBDP and the single Si surface provides a good link between the intrinsic and extrinsic bases.

The impurity profile of the intrinsic region is shown in Fig. 6. The SEG layer consisted of a 20-nm-thick Si

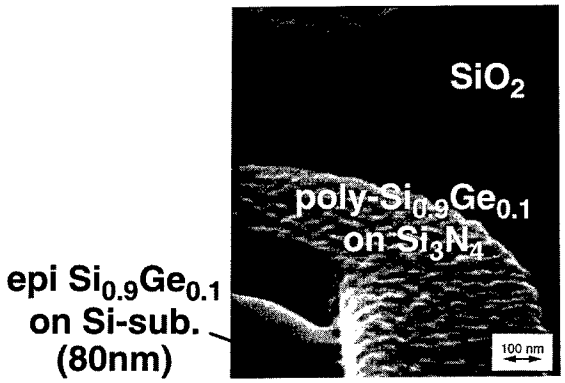


Fig. 5. An SEM bird's-eye view of SiGe selective growth. The SiGe layer was grown selectively with respect to only SiO_2.

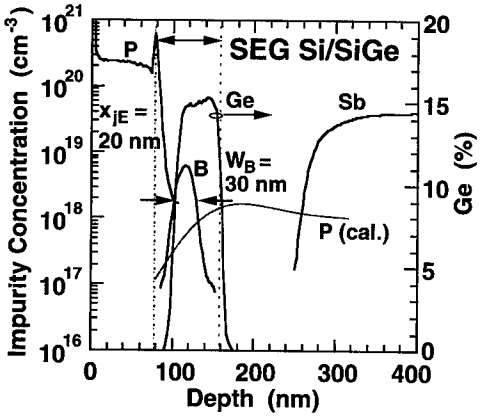


Fig. 6. Impurity profile of the intrinsic region.

cap, a 10-nm-thick two-step-ramped Ge-profile (from 0 to 10% over 5 nm and from 10 to 15% over 5 nm) $Si_{1-x}Ge_x$, a 40-nm-thick $Si_{0.85}Ge_{0.15}$ layer, and a 10-nm-thick Ge-retrograded $Si_{1-x}Ge_x$ layer. The 20-nm-thick 10^{19}-cm^{-3}-boron-doped $Si_{1-x}Ge_x$ layer formed the intrinsic base, which was 30 nm wide. Double selective phosphorous implantation in undoped SiGe and a 0.15-μm-thick Si layer increased the collector-doping level to about 10^{18} cm^{-3}. The shallow emitter junction (20 nm deep) was formed by out-diffusion from the IPDP into the Si cap at 900°C for 30 s.

3. HBT characteristics and applications

The developed transistors exhibit good I–V performance, as shown by the Gummel plot in Fig. 7 and the I_C–V_{CE} characteristics in Fig. 8, with an emitter area of 0.14 × 1.5 μm^2. A high current gain of 720 with a low base-recombination current of below 100 pA was obtained. This indicates there were no defects created in or relaxation of the strained Si/SiGe multilayer during the thermal cycle after low-temperature epitaxial growth. The high early voltage, V_A, of more than 100 V at a collector current of 1 mA indicates that the collector current was determined by the drift field created by bandgap grading. The observed negative resistance at a higher collector current was due to the self-heating effect, and is a typical characteristics of a genuine emitter/base heterojunction.

The cut-off frequency (f_T) and the maximum oscillation frequency (f_{max}) of the transistors, whose emitter area was 0.14 × 1.5 μm^2, were 95 and 97 GHz at a collector-to-emitter bias voltage of 2 V and a collector current of 2 mA, respectively (Fig. 9). These attractive high-frequency characteristics are attributed to the well-balanced transistor characteristics — a fast forward transit time of 1.18 ps and low collector capacitance of 3.6 fF — enabled by using the SEG SiGe base.

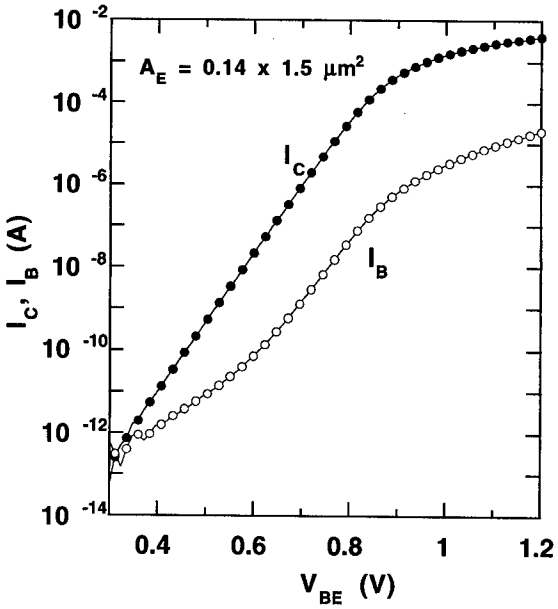


Fig. 7. Gummel plot for a typical transistor with an emitter area of 0.14 × 1.5 μm^2.

Typical transistor characteristics with an emitter area of 0.14 × 1.5 μm^2 are listed in Table 1. The emitter resistance was 50 Ω in spite of a very narrow 0.14-μm-wide emitter, and this low resistance is attributed to the stacked tungsten/IDP emitter electrode.

The dependence of the gate-delay time on the switching current measured in 45-stage differential ECL ring oscillators with a fan-in and a fan-out of 1 at a single-ended logic-swing voltage of 250 mV and a supply voltage of 3.5 V is shown in Fig. 10. The

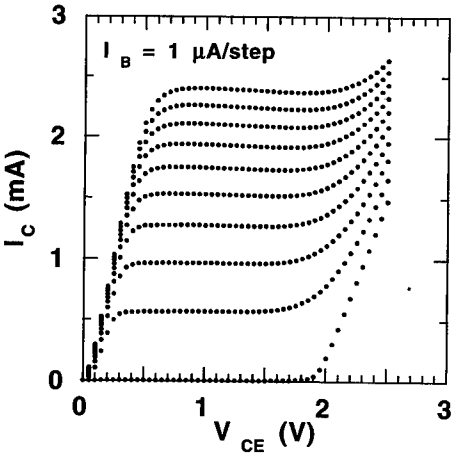


Fig. 8. I_C–V_{CE} characteristics for a typical transistor with an emitter area of 0.14 × 1.5 μm^2.

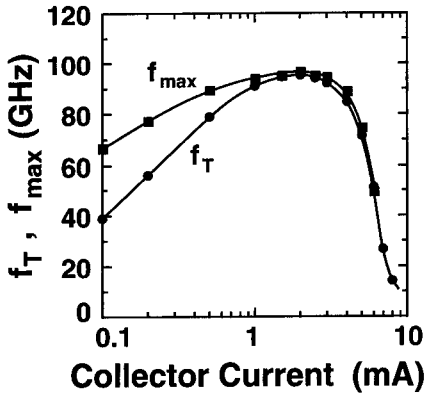


Fig. 9. Cut-off frequency (f_T) and maximum oscillation frequency (f_{max}) of the transistors with an emitter area of 0.14×1.5 μm^2 as a function of collector current.

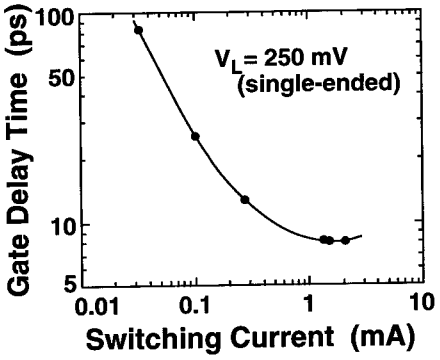


Fig. 10. Dependence of the gate-delay time on the switching current measured in 45-stage differential ECL ring oscillators with a fan-in and a fan-out of 1 at a single-ended logic swing voltage of 250 mV and a supply voltage of 3.5 V for the transistors with an emitter area of 0.14×1.5 μm^2.

Table 1
Typical transistor characteristics with an emitter area of 0.14×1.5 μm^2

A_E	0.14×1.5 μm^2
h_{FE}	720
BV_{CEO}	2.0 V
V_A	> 100 V
R_E	50 Ω
R_B	210 Ω
C_{jC}	3.6 fF
C_{SUB}	0.6 (5 V) fF

measured minimum gate-delay time for the transistors with an emitter area of 0.14×1.5 μm^2 was 8.0 ps at a switching current of 1.5 mA. This ultra-fast performance of the ECL gate was due to the fully-self-aligned SiGe base structure with a fast forward transit time and low collector capacitance, the SMI electrodes with low parasitic resistance, and the BPSG/SiO_2-refilled trench with low substrate capacitance.

As applications for these SiGe HBTs, various ICs for optical-fiber-link systems have been developed (Table 2) [11,15]. They include both digital ICs of a static frequency divider and a time-division multiplexer (MUX)/demultiplexer (DEMUX), and analog ICs of a preamplifier, an AGC amplifier core and a decision circuit. The maximum operating frequency of up to 50 GHz for a 1/8 static frequency divider is the highest reported to date for any semiconductor technology. A 2:1 time-division MUX and a 1:2 DEMUX built from basic circuit core modules operated at 40 Gb/s. A measured data rate of 40 Gb/s for a MUX with a DFF for retiming by a clock at 40 GHz was achieved for the first time. In a preamplifier with an input stage consisting of a common base transistor, a bandwidth of 35 GHz was achieved. In an AGC amplifier core, a bandwidth of about 32 GHz with a dynamic range of 19 dB was obtained by using a transimpedance amplifier as an active load circuit and a peaking capacitor. A well-opened eye diagram was obtained from a decision circuit with a 2:1 MUX to generate a bit stream at a data rate of 40 Gb/s. A block diagram of a transmitter and a receiver for an optical-fiber-link communication system using the fabricated IC chipset is shown in Fig. 11.

These excellent results indicate that self-aligned SiGe-base HBT technology, which offers high reliability and cost-effectiveness, will play an important

Table 2
ICs for optical-fiber communication systems fabricated by using SiGe HBTs

Circuit	Max. speed/bandwidth	Remarks
Multiplexer	40 Gb/s	2:1, 40 GHz DFF
Preamplifier	35.1 GHz	$Z_T = 48.7$ dB Ω
AGC amplifier	31.6–32.7 GHz	Dynamic range = 19 dB
Decision circuit	40 Gb/s	
Demultiplexer	40 Gb/s	1:2, 40 Gb/s DFF
Static frequency divider	50 GHz	

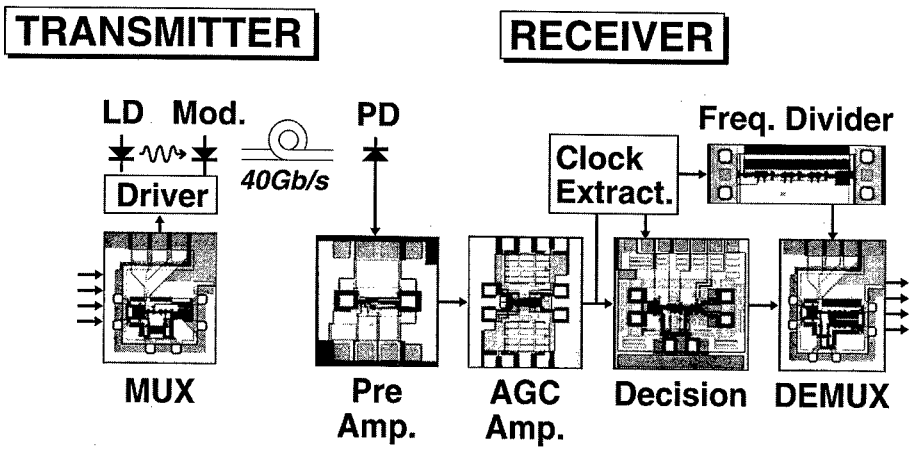


Fig. 11. A schematic block diagram of a transmitter and a receiver for an optical-fiber-link communication system operating at a data rate of 40 Gb/s with the fabricated SiGe HBT IC chipset.

role in future optical-fiber-link systems operating at a data rate of 40 Gb/s for global communications applications.

4. Summary

An ultra-high-speed selective-epitaxial-growth SiGe-base HBT with self-aligned stacked metal/IDP electrodes has been developed. The SiGe base, selectively grown by using a UHV/CVD system and self-aligned to the 0.14-μm-wide emitter, provides low collector capacitance. SMI technology provides low parasitic resistance and enables a shallow intrinsic base profile, so it is well suited to SiGe-base HBTs. A 2-μm-wide BPSG/SiO_2 refilled trench was introduced to reduce the substrate capacitance. This technology makes it possible to obtain ultra-high-speed operation with f_T and f_{max} of close to 100 GHz and an 8-ps gate-delay ECL circuit. An IC chipset has also been developed that includes a static frequency divider with the highest maximum operating frequency (up to 50 GHz) yet reported for any semiconductor technology, a MUX/DEMUX/decision circuit operating at 40 Gb/s, a preamplifier with a 35-GHz bandwidth, and an AGC amplifier core with a 32-GHz bandwidth. These excellent results show that Si bipolar technology, which offers high reliability and cost-effectiveness, will play an important role in future 40-Gb/s optical transmission systems for global communication.

Acknowledgements

I wish to express my sincere thanks to Dr. M. Miyao and Dr. O. Kanehisa at Hitachi's Central Research Laboratory (CRL) for their encouragement. I also thank Dr. T. Kure and the process staff members at the CRL pilot line for performing the wafer processing. Last, I thank those members in my research group at the CRL who were part of this work.

References

[1] Soda M, Tezuka H, Sato F, Hashimoto T, Nakamura S, Tatsumi T, Suzuki T, Tashiro T. Si-analog ICs for 20 Gb/s optical receiver. ISSCC Dig Tech Pap 1994:170–1.

[2] Felder A, Möller M, Popp J, Böck J, Rein H-M. 46 Gb/s DEMUX, 50 Gb/s MUX, and 30 GHz static frequency divider in silicon bipolar technology. IEEE J Solid-State Circuits 1996;31(4):481–6.

[3] Rein H-M, Möller M. Design consideration for very-high-speed Si-bipolar IC's operating up to 50 Gb/s. IEEE J Solid-State Circuits 1996;31(8):1076–90.

[4] Meister TF, Schäfer H, Franosch M, Molzer W, Aufinger K, Schler U, Walz C, Stolz M, Boguth S, Böck J. SiGe base bipolar technology with 74 GHz fmax and 11 ps gate-delay. IEDM Tech Dig 1995:739–42.

[5] Pruijmboom A, Terpstra D, Timmering CE, De Boer WB, Theunissen MJJ, Slotboom JW, Hueting RJE, Hageraats JJEM. Selective-epitaxial base technology with 14 ps ECL-gate-delay, for low power wide-band communication systems. IEDM Tech Dig 1995:747–50.

[6] Crabbé EF, Meyerson BS, Stork JMC, Harame DL. Vertical profile optimization of very high frequency epitaxial Si- and SiGe-base bipolar transistors. IEDM Tech Dig 1993:83–6.

[7] Oda K, Ohue E, Tanabe M, Shimamoto H, Onai T, Washio K. 130-GHz f_T SiGe HBT technology. IEDM Tech Dig 1997:791–4.

[8] Onai T, Ohue E, Tanabe M, Washio K. Self-aligned metal/IDP bipolar technology featuring 14 ps/70 GHz. IEDM Tech Dig 1995:699–702.

[9] Washio K, Ohue E, Tanabe M, Onai T. Self-aligned metal/IDP Si bipolar technology with 12-ps ECL and 45-

GHz dynamic frequency divider. In: Proc. ESSDERC '96, 1996. p. 807–10.

[10] Washio K, Ohue E, Oda K, Tanabe M, Shimamoto H, Onai T. A selective-epitaxial SiGe HBT with SMI electrodes featuring 9.3-ps ECL-gate-delay. IEDM Tech Dig 1997:795–8.

[11] Washio K, Ohue E, Oda K, Tanabe M, Shimamoto H, Onai T. 95 GHz f_T self-aligned selective epitaxial SiGe HBT with SMI electrodes. ISSCC Dig. Tech Papers 1998:312–3.

[12] Kondo M, Oda K, Ohue E, Shimamoto H, Tanabe M, Onai T, Washio K. Ultra-low-power and high-speed SiGe base bipolar transistors for wireless telecommunication systems. IEEE Trans Electron Devices 1998;45(6):1287–94.

[13] Oda K, Kiyota Y. H_2 cleaning of silicon wafers before low-temperature epitaxial growth by ultrahigh vacuum chemical vapor deposition. J Electrochem Soc 1996;143:2361.

[14] Oda K, Kiyota Y. Selectivity dependence on Ge composition in $Si_{1-x}Ge_x$ low-temperature epitaxial growth. MRS 1997 Spring Meeting.

[15] Masuda T, Ohhata K, Ohue E, Oda K, Tanabe M, Shimamoto H, Onai T, Washio K. 40 Gb/s analog IC chipset for optical receiver using SiGe HBTs. ISSCC Digest of Tech. Papers 1998:314–5.

PERGAMON

Solid-State Electronics 43 (1999) 1627–1632

SOLID-STATE ELECTRONICS

Heterostructure circuit applications in optical communications

Leda Lunardi

AT&T Labs, Newman Springs Laboratory, RedBank, NJ 07701, USA

Received 18 March 1999; accepted 25 March 1999

Abstract

Remarkable progress has been made in practical use of high-speed heterostructure device technologies for communications. To further expand the possible applications, devices with higher performance or able to perform new functions are going to be key elements for future transmission systems. In this article, the status of high-speed integrated circuits targeted for either wavelength division multiplexing (WDM) or time division multiplexing (TDM) is reviewed. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

In the last few years, the evolution of large capacity optical networks, due to the ever increasing demand for more bandwidth, has been basically progressing in two directions: high-speed time division multiplexing (TDM) and wavelength division multiplexing (WDM). In TDM systems, high-speed optoelectronic processing exists between source and destination of the optical signals. The complete optical repeater contains components that perform reshaping, time extraction and signal regeneration (the so-called 3R-repeater for reshaping, retiming and regenerating). In the design of such optical repeater, the optimum performance depends on maximising the speed of each individual circuit, which is primarily limited by the transistor technology used.

On the other hand, WDM systems have become a great successful story for solving the bandwidth crunch. Since WDM uses multiple channels, the systems provide aggregated speeds much higher than those obtained for single channel TDM with fewer demands on the electronic components. However, while the optical signals may propagate through either ring or mesh network architectures without the need for optical repeaters, there are some limitations that impair the ultimate capacity. These impairments arise mainly from optical phenomena, imposing a limitation on the network capacity and requiring dispersion compensation techniques.

With the development of broadband erbium doped fiber amplifiers, simultaneous amplification of both high-speed TDM and WDM signals are now possible. So future optical networks can be a hybrid mixture of TDM and WDM technologies with fiber dispersion management and depending on the architecture, may include interfaces with high-speed cross-connects or electrical crosspoint switches.

This paper will review the state-of-the-art of integrated circuit technologies, focusing mainly on heterostructure transistors-based circuits that have potential applications for TDM and WDM networks.

2. System applications

The progress reported in the last few years on transmission systems beyond 10 Gb/s capacity is shown in Fig. 1. While single channel transmission experiments have steadily reached longer distances and higher

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00112-4

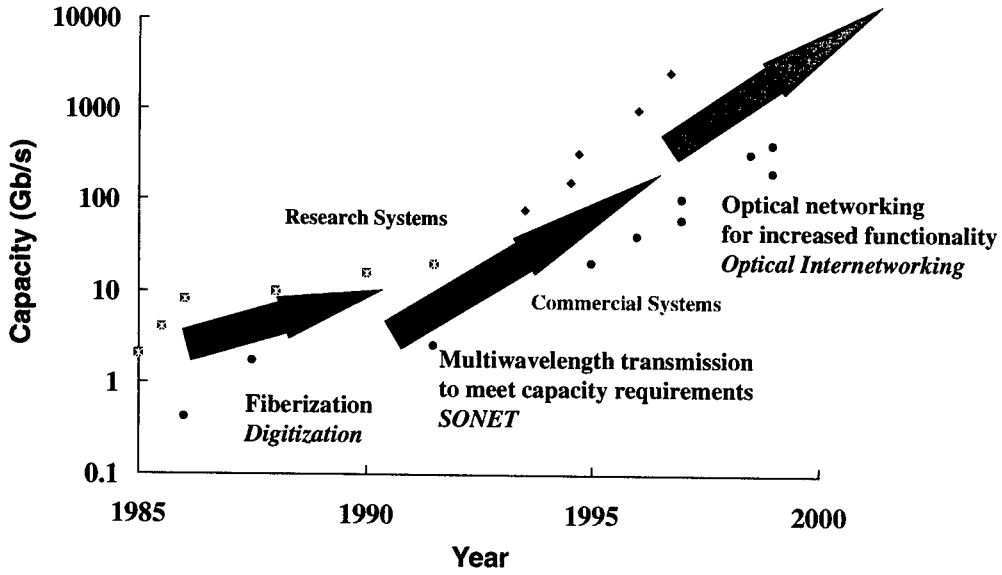


Fig. 1. Chronology on the progress of single fiber capacity from research (•) and commercial systems.

speeds, multi-channel carrier channels have demonstrated a total capacity beyond 2 Tb/s using 132 different wavelengths at 20 Gb/s per channel [1].

The major advantage of wavelength-division multiplexing (WDM) is that signals can propagate from source to their destination without optoelectronic conversion. It is format independent (each wavelength in a single network may carry analog, digital or any modulation scheme). The capacity (aggregated bit-rate) can be increased enormously without increasing the single channel bit-rate and can provide cost effective network upgrades. However, WDM uses signal propagation at high optical power levels through longer spans, which accumulates degradation in the transmission. These effects include chromatic dispersion, fiber non-linearities, crosstalk between channels, losses in filters, multi path interference besides incorporating the features of optical fiber amplifiers (gain flatness, noise figure). Another disadvantage is that transparent networks must meet standards like number of wavelengths, number of channels or even special component [2].

On the other hand, in time division multiplexing (TDM) the optical signal travels the network and is electrically regenerated after relatively short distances (50–60 km). The regeneration with advanced digital electronic processing provides also easy operation, maintenance and administration for the network. Moreover, the capacity of the network is given by the operating speed of the electronics, which depends directly on the transistor technology. TDM offers also another possibility for traffic demand over a shared network, since delay is reduced and throughput increased for a single channel when compared to parallel channels. However, the upgrading of TDM-based systems is not cost effective. Usually these systems are referred as opaque due to their nature of bringing the signal from the optical to the electrical domain before the signals arrive at their final destination. However, what looks like a limitation actually is an advantage considering that an opaque network employs the best technology available in each component that is added, without any barrier to interconnecting systems with different technologies. Actually, an opaque system is transparent to the adoption of new technologies. Eventually networks will have a combination of TDM and WDM technologies with either ring or mesh-like systems for flexibility of upgrading and cost effectiveness.

3. Multiwavelength networks

An opaque configurable network, displayed in Fig. 2, has some interesting features [3]. Although the immediate application is for 2.5 Gb/s, it can be expanded easily to any line rate. The core of this architecture relies on either electrical or optical switch with port count N. Another feature is the use of numerous optoelectronic converters (transponders) that transform long-reach 1.5 μm WDM signals to inter-office 1.3 μm and vice-versa. Since these transponders are bit-rate dependent, they can effectively restore the signal and provide performance monitoring at the network management layer. While at first this scheme can be readily

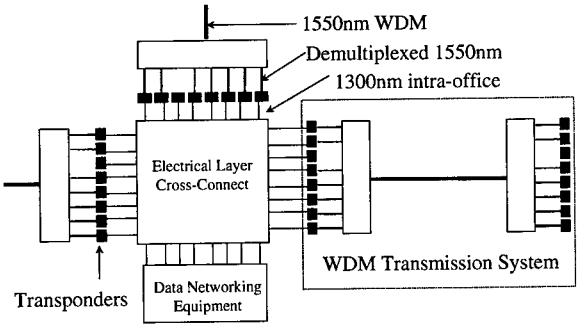


Fig. 2. Multiwavelength opaque crossconnect network architecture.

implemented at 2.5 Gb/s, for 10 Gb/s operation it would require optolectronic components and electrical cross point switch upgrades.

Two different laboratories have demonstrated HBT-based electrical cross point switches. Lowe [4] has demonstrated a 16 × 16 cross point switch using 2 μm emitter size GaAs HBT technology with a cut-off frequency of 80 GHz. The 160 Gb/s aggregate capacity switch has a total reconfigurable time of 4.3 ns and a data output jitter deviation of 7.7 ps.

Also benefiting from the large-scale integration capability of GaAs-based HBTs, Pedrotti et al [5] have demonstrated a 12 × 12-crosspoint switch. The chip, totalling 4500 transistors, operates at a maximum rate of 10 Gb/s with a 2.5 ps jitter in the data output for single channel with an increase of 0.15 to 0.25 ps per added channel. The latter switch has been used in the WEST project (WDM with electronic switching technology), sponsored by the defence agency for advanced research programs (DARPA), for connection of fibers handling up to 4 separate wavelength channels per fiber, at maximum line rate of 10 Gb/s per channel. The full data regeneration with retiming allows arbitrary reassignment of wavelengths to the data streams.

Moreover, the electronic approach in WDM architectures can drastically minimise the cumulative transmission impairments in addition to the performance monitoring and network management.

4. High-speed electronics

For very high-speed transmission systems either TDM or WDM, electronic devices continue to have some advantages over optical devices such as smaller size, lower cost, higher reliability and higher functionality. The state-of-the-art performance of compound semiconductor integrated circuit modules for optical communications is summarised in Table 1, [6–22]. Basically two compound semiconductor technologies have been used for those high-speed modules: heterojunction bipolar transistors (HBTs) and heterostructure field effect transistors (HFETs).

The performance of heterojunction bipolar transistors (HBTs) has steadily increased as a result of the maturity in the epitaxial growth and optimisation of the layer structures. For compound semiconductor based HBTs, the processing technology has evolved from mesa structure to be like silicon, with reduced emitter widths in the submicron size range [23,24]. Fig. 3 compares the pair f_T/f_{max} for different compound semiconductor HBT technologies for emitter areas below 10 μm^2. Notwithstanding are the results from Agrawal, where the combination of reduced parasitic capacitances and submicron emitter widths in a novel transferred substrate HBT technology has yielded maximum available f_{max} of 500 GHz [11]. From figure 3, the results from InGaP-based HBTs indicate that the suppression of the surface recombination can contribute to a more scalable technology [23].

In the case of HFETs, the steady improvement in performance has primarily been to the reduced gate lengths dimensions. Fig. 4 displays the dependence of current gain cut-off frequency for different FET technologies as function of the gate length. The departure of the scalability at reduced dimensions is related to short channel effects, which can be minimised by the use of heterostructure. As a matter of fact, InP-based HFETs benefiting from the larger electron saturation velocity have been reported with cut-off frequency of

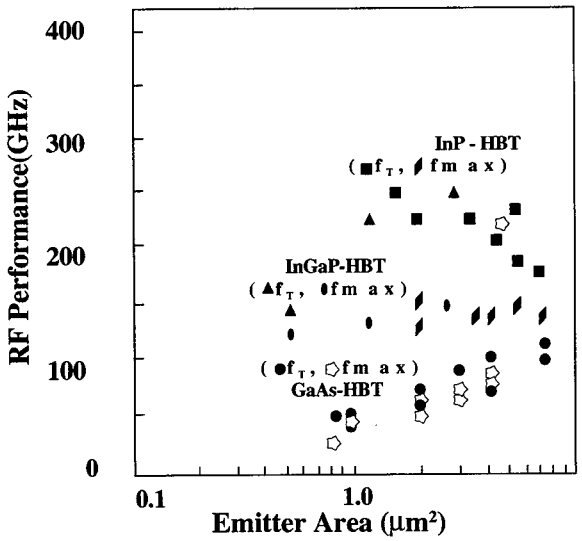


Fig. 3. Measured current gain cutoff frequency (f_T) and maximum available gain (f_{max}) for different technologies as function of emitter area.

Table 1
High-speed integrated circuits for optical communications in compound semiconductor technologies for state-of-the-art published results over 20 Gb/s

Circuit	Technology	Operating speed	Remarks	Reference
Multiplexer	InP-HBT	20 Gb/s	2:1	[6]
	GaAs-HBT	40 Gb/s	4:1	[7]
	InP-HFET	80 Gb/s	2:1	[8]
Laser driver	GaAs-HBT	20 Gb/s		[28]
Amplifier	GaAs-HFET	29.3 GHz	$A_v = 17$ dB	[9]
	InP-HBT	30 GHz		[6]
	GaAs-HBT	40 GHz	$A_v = 12$ dB	[10]
	InP-HBT	50 GHz	$A_v = 13$ dB	[11]
	InP-HFET	90 GHz	$A_v = 10$ dB	[12]
Decision circuit	GaAs-HFET	25 Gb/s		[13]
	InP-HFET	46 Gb/s		[14]
Demultiplexer	GaAs-HFET	20 Gb/s	1:4	[13]
	InP-HBT	25 Gb/s	1:2	[6]
	GaAs-HBT	30Gb/s	1:4	[15]
	InP-HFET	40 Gb/s	1:2	[14,19]
Digital divider	InP-HBT	39.5 GHz	1:2	[6]
	InP-HBT	44 GHz	1:2	[20]
	InP-HFET	46.5 GHz	1:2	[17]
Clock recovery	GaAs-HBT	20 Gb/s		[15]
	GaAs-HFET	20 Gb/s		[9,18]
	InP-HFET	40 Gb/s		[19,29]
Pin/preamp	InP-HBT	20 Gb/s	−17.0 dBm	[21]
	InP-HFET	20 Gb/s	N/A	[22]

over 300 GHz for a 50-nm gate length [25] and high breakdown voltage [26].

The design of digital circuits with clock frequencies as high as the transistor operating speed involve not only the optimisation of transistor operating points, but individual design of all transistors in the circuit as well as the optimisation of on-chip and off-chip wiring. There is a trade off in improving technologies with respect to higher operating speeds and the breakdown voltage. For digital integrated circuits with a logic swing of 1 V, a minimum of 3 V is needed for proper operation. However, this value can increase considerably for circuits that require larger output voltage swing, for example like a laser or optical modulator drivers. Fig. 5 displays two figures of merit, current gain cut-off frequency (f_T) and maximum available gain (f_{max}) as function of the breakdown voltage for the different high-speed technologies.

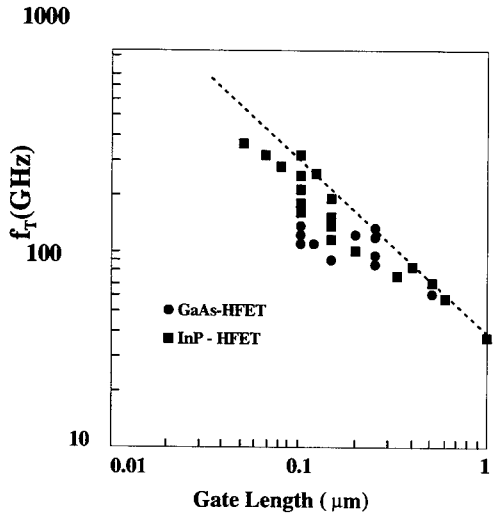


Fig. 4. Current gain cutoff frequency (f_T) as function of gate length for heterostructure FETs.

Another fact in the potential of these high-speed technologies is to observe the circuit operating speed as function of the individual transistor speed, given by f_r and f_{max} In the case of bipolar-based digital circuits, where emitter coupled logic (ECL) is normally used, the delay time of an ECL inverter varies as roughly to $2/(1/f_T + 1/f_{max})$, while for source coupled FET logic (SCFL), the inverter delay time is proportional to the inverse of f_T [27]. As seen in Fig. 6, the trend given by HBT-based circuits indicate that higher operating speeds are possible to be achieved. As the technology

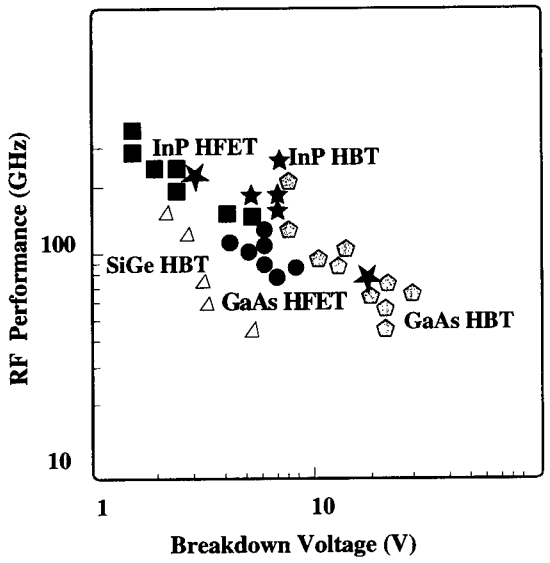


Fig. 5. Comparison of measured current gain cutoff frequency (f_T) and maximum available gain (f_{max}) for different technologies as function of the breakdown voltage.

matures, better models for the transistors and circuit parasitics will permit optimised circuit designs. InP-based transistors seem to be the potential candidates for over 40 Gb/s operating speeds in circuits.

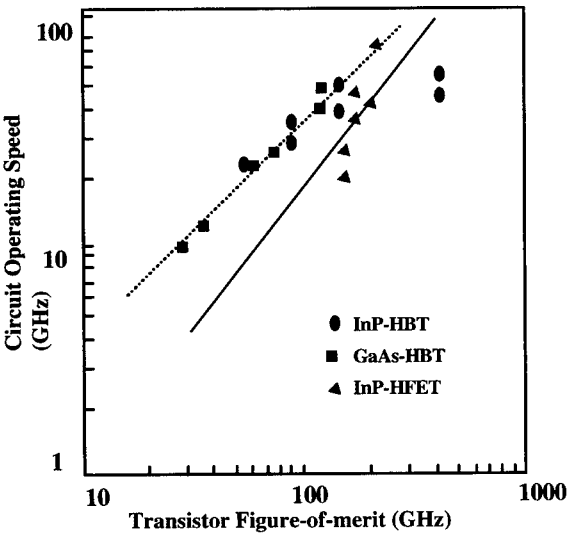


Fig. 6. Relationship between operating speed of D-type flip-flop integrated circuits based on ECL inverters for bipolar transistors and SCFL for FETs. The transistor figure-of-merit is simply proportional to measured f_T for FETs but it varies with $2/(f_T^{-1}+f_{max}^{-1})$ for bipolar transistors [27].

5. Conclusions

The demand for high performance modules in optical communications will continue to increase in the near future. Devices that allow novel network architectures certainly will include sophisticated heterostructures. GaAs-based HBTs have already reached medium scale of integration (MSI) at an aggregated bit rate of 160 Gb/s. The performance of InP-based HFETs and HBTs will certainly be further optimised with higher breakdown voltages while keeping the high-speed performance. Practical aspects such as cost, reproducibility and reliability should be addressed in order to bring high-speed technologies from research to manufacturing.

Acknowledgements

I would like to thank E.L. Goldstein, J.M. Wiesenfeld and R.W. Tkach for providing material for the first two figures. Special thanks to A.M.M. Saleh for supporting this work.

References

[1] Yano Y, Ono T, Fukuchi K, Ito T, Yamazaki H, Yamaguchi M, Emura K. 2.6 Tb/s WDM transmission experiment using optical duobinary coding. In: Proc European Conference on Optical Communication, Th B 3, 1996. p. 1.

[2] Tkach RW, Goldstein EL, Nagel JA, Strand JL. Fundamental limits of optical transparency. Optical Fiber Conference Technical Digest 1998: 161–162.

[3] Goldstein EL, Nagel JA, Strand JL, Tkach RW. National-scale networks likely to be opaque. Lightwave 1998;2:92.

[4] Lowe KE. A GaAs HBT 16 × 16 10-Gb/s/channel cross-point switch. IEEE J Solid State Circuits 1997;32:1263.

[5] Pedrotti KD, Chang CE, Price A, Beccue SM, Campana AD, Gutierrez G, Meeker D, Wu D, Wang KC, Metger A, Asbeck PM, Huff D, Kwong N, Swass MN, Zjamg SZ, Bowers J. WEST 120-Gbit/s 3 × 3 wavelength-division multiplexed cross-connect. OFC Technical Digest 1998;3:66.

[6] Mokhtari M, Swahn T, Walden RH, Stanchina WE, Kardos M, Juhola T, Schuppener G, Tenhunen H, Lewin T. InP-HBT chip-set for 40 Gb/s fiber optical communications systems operational at 3 V. IEEE J Solid-State Circuits 1997;32:1371.

[7] Runge K, Pierson RL, Zampardi PJ, Thomas PB, Yu J, Wang KC. 40 Gbit/s AlGaAs/GaAs HBT 4:1 multiplexer IC. IEE Electronics Lett 1995;31:876.

[8] Otsuji T, Murata K, Enoki T, Umeda Y. 80 Gbit/s multiplexer IC using InAlAs/InP HEMTs. IEE Electronics Lett 1998;34:113.

[9] Lang M, Wang ZG, Lao Z, Schlechtweg M, Thiede A,

Rieger-Motzer M, Sedler M, Bronner W, Kaufel G, Koehler K, Huelsmann A, Raynor B. 20–40 Gb/s 0.2 μm GaAs HEMT chip set for optical data receiver. IEEE J Solid State Circuits 1997;32:1384.

[10] Sano E, Yamahata S, Matsuoka Y. 40 GHz amplifier IC using AlGaAs/GaAs ballistic collector transistors with carbon-doped base. IEE Electronics Lett 1994;30:635.

[11] Agarwal B, Mensa D, Lee Q, Pullela R, Guthrie J, Samoska L, Rodwell MJW. A 50 GHz feedback amplifier with AlInAs/GaInAs transferred-substrate HBT. IEEE IEDM Technical Digest 1997;743:746.

[12] Kimura S, Imai Y, Umeda Y, Enoki T. 0–90 GHz InAlAs/InGaAs/InP HEMT distributed baseband amplifier IC. IEE Electron Lett 1995;31:1430.

[13] Lao Z, Berroth M, Thiede A, Rieger-Motzer M, Kaufel G, Seibel J, Bronner W, Huelsmann A, Schneider J, Raynor B. Low power 20 Gbit/s data decision and 17 GHz static frequency divider ICs with 1.5 V supply voltage. IEE Electronics Lett 1997;33:289.

[14] Yoneyama M, Otsuji T, Imai Y, Yamaguchi S, Enoki T, Umeda Y, Hagimoto K. 46 Gbit/s super-dynamic decision circuit module using InAlAs/InGaAs HEMTs. IEE Electronic Lett 1997;33:1472.

[15] Runge K, Yu RY, Zampardi PJ, Pierson RL, Wang KC, Blixt P, Petersen AK. Packaged 30 Gb/s data demultiplexing and clock extraction IC fabricated in a AlGaAs/GaAs HBT technology. IEE Electronic Lett 1996;32:588.

[16] Jensen JF, Hafizi M, Stanchina WE, Metzger RA, Rensch DB. 39.5 GHz static frequency divider implementation in AlInAs/GaInAs HBT technology. IEEE GaAs IC Symp 1992;101:104.

[17] Otsuji T, Yoneyama M, Murata K, Imai Y, Enoki T, Umeda Y. 46.5 GHz quasi-static 2:1 frequency divider IC using InAlAs/InGaAs/InP HEMTs. IEE Electronic Lett 1997;33:1376.

[18] Wang ZG, Berroth M, Thiede A, Rieger Motzer M, Hofmann P, Huelsmann A, Koehler K, Raynor B, Schneider J. 40 and 20 Gbit/s monolithic integrated clock recovery using a fully-balanced narrow band regenerative frequency divider with 0.2 μm AlGaAs/GaAs HEMTs. IEE Electronic Lett 1996;32:2081.

[19] Otsuji T, Imai Y, Sano E, Kimura S, Yamagushi S, Yoneyama M, Enoki T, Umeda Y. 40-Gb/s IC's for future lightwave communications systems. IEEE J Solid State Circuits 1997;32:1363.

[20] Masuda H, Ouchi K, Terano A, Suzuki H, Watanabe K, Oka T, Matsubara H, Tanoue T. Device technology of InP/InGaAs HBTs for 40-GB/s optical transmission application. IEEE GaAs IC Symp 1997;19:139.

[21] Lunardi LM, Chandrasekhar C, Gnauck AH, Burrus CA, Hamm RA. (20-Gb/s monolithic p-I-n/HBT photoreceiver module for 1.55 μm applications. IEEE Photonic Technol Lett 1995;10:1201.

[22] Umback A, van Waasen S, Auer U, Bach HG, Bertenburg RM, Breuer V, Ebert W, Janssen G, Mekonnen GG, Passenberg W, Schlaak W, Scramm C, Seeger A, Tegude F-J, Unterbosch G. Monolithic pin-HEMT 1.55 μm photoreceiver on InP with 27 GHz bandwidth. IEE Electron Lett 1996;32:214.

[23] Oka T, Hirata K, Ouchi K, Uchiyama H, Mochizuki K, Nakamura T. InGaP/GaAs HBT's with high-speed and low-current operation fabricated using WSi/Ti as the base electrode and burying SiO2 in the extrinsic collector. IEEE IEDM Technical Digest 1997;739:742.

[24] Yanagihara M, Sakai H, Ota Y, Tanabe M, Inoue K, Tamura A. 253-GHz fmax AlGaAs/GaAs HBT with Ni/Ti/Pt/Ti/Pt-contact and L-shaped base electrode. IEEE IEDM Technical Digest 1995;807:810.

[25] Nguyen LD, Brown AS, Thompson MA, Jelloian LM. 50-nm self-aligned gate pseudomorphic AlInAs/GaInAs high electron mobility transistors. IEEE Trans Electron Devices 1992;39:2007–14.

[26] Zaknoune M, Bonte B, Gaquiere C, Cordier Y, Druelle Y, Theron D, Crosnier Y. InAlAs/InGaAs metamorphic HEMT with high current density and high breakdown voltage. IEEE Electron Dev Lett 1998;19:345.

[27] Sano E, Matsuoka Y, Ishibashi T. Device figure-of-merits for high-speed digital ICs and baseband amplifiers. IEICE Trans Electron 1995;E78:1182.

[28] Lao Z, Thiede A, Nowotmy U, Schlechtweg M, Hurm V, Bromner W, Hornung J, Rieger-Motzer M, Kaufel G, Kohler K, Hulsmann A. High power modulator driver ICs up to 30 Gbit/s with AlGaAs/GaAs HEMTs. IEEE GaAs IC Symp 1997;19:223.

[29] Murata K, Otsuji T, Enoki T, Umeda Y. Exclusive OR/NOR IC for >40 Gbit/s optical transmission systems. IEE Electron Lett 1998;34:764.

PERGAMON

Solid-State Electronics 43 (1999) 1633–1643

SOLID-STATE ELECTRONICS

Heterostructure-based high-speed/high-frequency electronic circuit applications

P.J. Zampardi[a,*], K. Runge[a], R.L. Pierson[a], J.A. Higgins[a], R. Yu[a], B.T. McDermott[a], N. Pan[b]

[a]*Electronic Devices Laboratory, Rockwell Science Center, 1049 Camino Dos Rios, Thousand Oaks, CA 91358, USA*
[b]*Noren Pan, Kopin Corporation, 695 Miles Standish Blvd., Taunton, MA 02780, USA*

Abstract

With the growth of wireless and lightwave technologies, heterostructure electronic devices are commodity items in the commercial marketplace [Browne J. Power-amplifier MMICs drive commercial circuits. Microwaves & RF, 1998. p. 116–24.]. In particular, HBTs are an attractive device for handset power amplifiers at 900 MHz and 1.9 GHz for CDMA applications [Lum E. GaAs technology rides the wireless wave. Proceedings of the 1997 GaAs IC Symposium, 1997. p. 11–13; "Rockwell Ramps Up". Compound Semiconductor, May/June 1997.]. At higher frequencies, both HBTs and p-HEMTs are expected to dominate the marketplace. For high-speed lightwave circuit applications, heterostructure based products on the market for OC-48 (2.5 Gb/s) and OC-192 (10 Gb/s) are emerging [http://www.nb.rockwell.com/platforms/network_access/nahome.html#5.; http://www.nortel.com/technology/opto/receivers/ptav2.html.]. Chips that operate at 40 Gb/ have been demonstrated in a number of research laboratories [Zampardi PJ, Pierson RL, Runge K, Yu R, Beccue SM, Yu J, Wang KC. hybrid digital/microwave HBTs for >30 Gb/s optical communications. IEDM Technical Digest, 1995. p. 803–6; Swahn T, Lewin T, Mokhtari M, Tenhunen H, Walden R, Stanchina W. 40 Gb/s 3 Volt InP HBT ICs for a fiber optic demonstrator system. Proceedings of the 1996 GaAs IC Symposium, 1996. p. 125–8; Suzuki H, Watanabe K, Ishikawa K, Masuda H, Ouchi K, Tanoue T, Takeyari R. InP/InGaAs HBT ICs for 40 Gbit/s optical transmission systems. Proceedings of the 1997 GaAs IC Symposium, 1997. p. 215–8]. In addition to these two markets, another area where heterostructure devices are having significant impact is for data conversion [Walden RH. Analog-to digital convertor technology comparison. Proceedings of the 1994 GaAs IC Symposium, 1994. p. 217–9; Poulton K, Knudsen K, Corcoran J, Wang KC, Nubling RB, Chang M-CF, Asbeck PM, Huang RT. A 6-b, 4 GSa/s GaAs HBT ADC. IEEE J Solid-State Circuits 1995;30:1109–18; Nary K, Nubling R, Beccue S, Colleran W, Penney J, Wang KC. An 8-bit, 2 gigasample per second analog to digital converter. Proceedings of the 1995 GaAs IC Symposium, 1995. p. 303–6.]. In this paper, we will discuss the present and future trends of heterostructure device applications to these areas. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The rapid growth of wireless and telecommunications markets has driven the commercialization of several heterojunction device technologies that have, to this point, existed to serve predominantly military ap-

* Corresponding author. IBM, M/S 861A, 1000 River Road, Essex Junction, VT 05452, USA. Tel.: +1-802-769-6719; fax: +1-802-769-7147.
E-mail address: pz@us.ibm.com (P.J. Zampardi)

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00113-6

plications. One example of this is the high volume production of heterojunction bipolar transistors (HBTs) at Rockwell, RF MicroDevices, and several other emerging companies. Silicon has long benefited (particularly in cost) from large-scale production. III–V's are just beginning to take advantage of similar leverage and are breaking ground in many commercial areas. The main thrust of heterojunction devices has been in handheld power amplifiers. This application, with its yield, cost, reliability and scheduling requirements (delivery schedules), provides a manufacturing base that is being leveraged by other applications such as lightwave telecommunications and data conversion. In this paper, we begin by discussing heterojunction devices in power amplifier applications and the prospects for their continued success. We then focus on lightwave applications from 2.5 to 40 Gb/s. Finally, we highlight the process by discussing the application of heterostructure devices to data conversion applications which often require transistor counts between 5 and 10,000.

2. Power amplifiers

Heterostructure devices are finding widespread use in commercial handheld applications at high volumes (> 1 million die per month by several vendors) [1]. For these applications, high-efficiency is essential (longer battery life), single polarity supply is preferred, traditional figures of merit (f_T and f_{max}) are irrelevant, breakdown and power density requirements are relaxed, and COST is the primary driver [2]. High-electron mobility transistors (HEMTs) and HBTs are both widely used and each has its own merits in terms of efficiency, linearity and output power. These parameters also strongly depend on the power supply voltage and operating frequency at which the devices are to be used. For the commercial wireless market, the emphasis on reducing battery voltage is strong and tends to favor FET based technologies. This drive for low-voltage/low-power operation is because the use of lower voltages (a) reduces the weight and volume of the product and (b) follows the trend of digital chip power supplies (so both the power amplifier (PA) and digital circuits can use the same supply). The choice of amplifier, in these applications, is not totally governed by the device performance, but by other factors such as total system cost and time to market.

The requirements for these handheld applications are contrasted by the needs of infrastructure or defense applications where the given elements of device performance are the dominant factor and there is stringent requirement for low-power operation. For example, CATV and base-station applications require high-efficiency, high-linearity and high-breakdown. HBTs are excellent candidates for these applications. For low-noise, high-frequency performance (10–100 GHz) HEMTs are the device of choice due to superior thermal properties (no current crunch) and higher gain at higher frequencies.

Rockwell Science Center has developed, and is supplying, several low-noise amplifier designs using 0.25 μm p-HEMT technology to deliver low-noise, high-performance solutions above 30 GHz. For example, amplifiers in the 33–36 GHz range with 22 and 26 dB nominal gain with 2.5 dB nominal noise figure and 8 dBm saturated power have been fabricated [3–5]. Amplifiers, such as the one shown in Fig. 1, for applications at W band (91–GHz) with 15 dB nominal gain with 5.5 dB with 3 dBm saturated power have also been demonstrated [6]. These circuits clearly demonstrate that HEMT is, and will continue to be, the dominant technology for applications above 30 GHz.

HBTs are also excellent devices for high-power microwave applications. They offer high-power densities, high-efficiency and high-linearity. The high-voltage and high-current density achievable with HBTs allow them to form compact amplifiers. The high-power added efficiency (PAE) achievable with HBTs was demonstrated at power densities of 2.83 mW/μm^2 at 10 GHz with 67.8% PAE [7]. Further improvements in power density have been achieved using a thermal shunting technique [8] that demonstrated power densities above 10 mW/μm^2 at 10 GHz with 7-dB gain and 60% PAE. The importance of thermal management and reliability for these applications cannot be understated.

For power amplifiers, the high volumes of product currently being shipped along with clearly defined higher frequency applications bode well for the future of heterostructure devices in power amplification. The high-volume production of these devices also has several important secondary benefits. The first benefit is that it provides a solid manufacturing base for the development of future high-performance GaAs chips for other applications. As an example, the production discipline necessary to deliver low-cost HBT power amplifiers (reliable, high-yields devices and simpler, tightly controlled processes) results in the ability to yield large digital circuits like those described later in this paper. The second benefit is that, as circuit designers become more familiar designing with these devices, the product cycle time for next generation chips is greatly reduced.

3. Lightwave circuits

With the explosion of internet use, the requirements for bandwidth globally are ever-increasing. As a result, there is a growing need for high-bit rate communi-

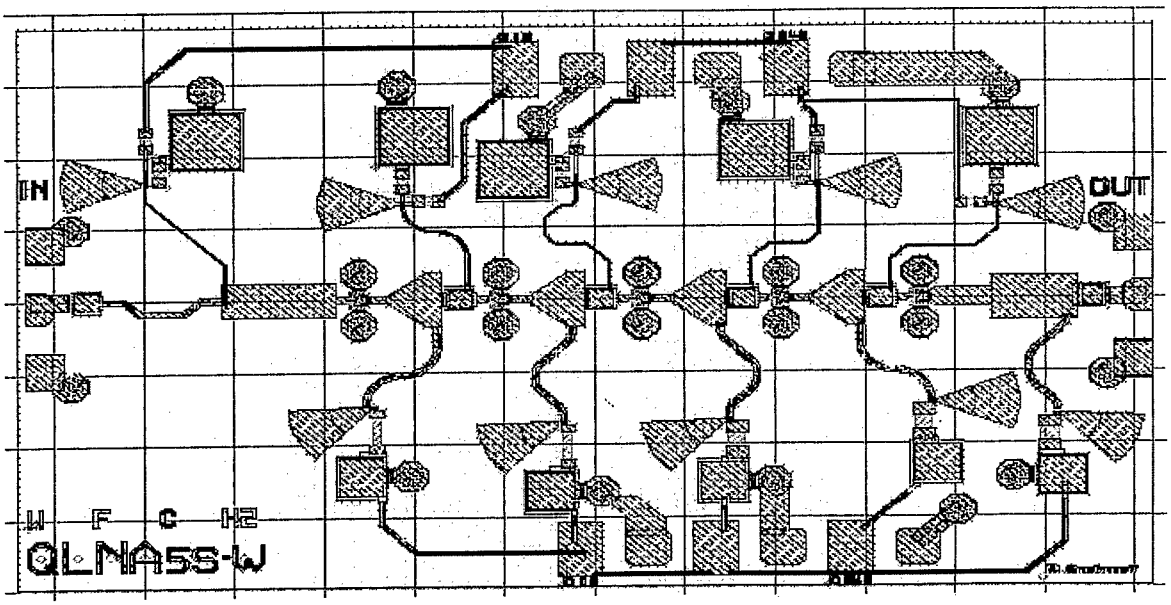


Fig. 1. Layout of W-band p-HEMT amplifier. The die is 2.5 × 1.67 mm and operates from a 4-V supply.

cations. While some question about wavelength division multiplexing (WDM) versus time division multiplexing (TDM) remains, clearly higher bit rate *channels* are necessary to meet the needs of future net surfers. High-bit-rate communication systems require very high-speed technologies for their implementation. Many different technologies, including heterojunction devices, are competing for the 'low bit-rate' (< 2.5 Gb/s) markets. At higher bit-rates (> 10 Gb/s) heterojunction devices are very capable of offering a full chip-set solution. These chip-sets typically include amplifiers, laser drivers, MUX/DEMUX, and clock and data recovery chips. Several companies already have commercially available products that use heterojunction devices at both the 2.5 and 10 Gb/s using HBTs [9–10]. HBTs (GaAs, SiGe and InP) and HEMTs are the leading candidates for higher bit rate operation [11–14]. We will now highlight some of the high-speed circuits demonstrated at Rockwell for use in 40 Gb/s optical systems.

3.1. 4:1 Multiplexer

High-speed MUX circuits are important for both time division multiplexing (TDM) and wavelength division multiplexing (WDM). We have demonstrated a 4:1 multiplexer that operates at 40 Gb/s using our hybrid digital/microwave process [15]. The architecture used requires only the last 2:1 MUX stage to operate at the maximum bit rate. The rest of the circuit operates at half the bit rate. Another outstanding feature of this design is the low devices count and low power consumption compared to multiplexing with four master–slave D flip–flops (MS-DFFs). The circuit was tested on-wafer, using RF probes. A 10 Gb/s bit-error-rate tester (BERT) generated test signals. The 4:1 multiplexer operated up to a maximum bit rate of 40 Gb/s, with data input levels of 125 mV peak-to-peak per channel and a single-ended clock input level of 500 mV peak-to-peak. A microphotograph of this 4:1 multiplexer is shown in Fig. 2(a), the output eye diagram, when operating at 40 Gb/s, is shown in Fig. 2(b). To our knowledge, this is the fastest 4:1 MUX reported. This circuit was converted to be compatible with Rockwell's non-self-aligned production process (with only minor changes to the transistor layout). Excellent yield was achieved at 30 Gb/s with a maximum operating frequency of 39 Gb/s. In this case, the major differences between the two technologies are the non-self-aligned emitter-base and lack of a C_{bc} reduction implant in the production technology.

3.2. 1:4 Demultiplexer

The 1:4 demultiplexer, shown in Fig. 3(a), used similar multifunctional circuits as the 4:1 MUX [15–16]. Testing was performed on-wafer using a 30 Gb/s test signal generated by connecting a 4:1 MUX and a 1:4 DMUX with 0.8 m of semi-rigid cable. The resulting eye diagrams and divided clock are shown in Fig. 3(b). The divided clock output was designed with a falling edge that is independent of data rate in its relative position to the output data channel eyes, and is used to time subsequent circuits. The MUX/DMUX pair were

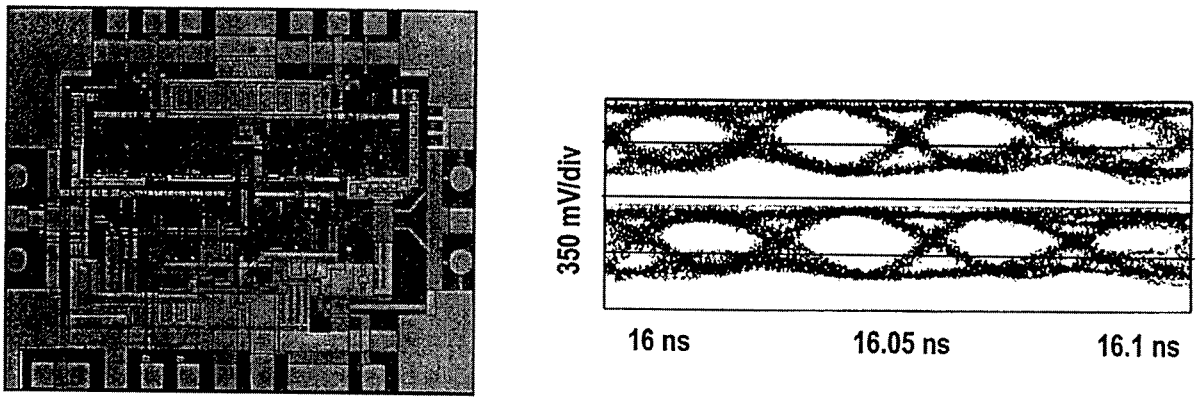


Fig. 2. (a) Microphotograph of 40 Gb/s 4:1 multiplexer. The die size is 1 × 2 mm^2 and the chip dissipates 3.5 W from a single −7.5 V supply. (b) Eye diagram at 40 Gb/s.

measured back-to-back with a BER test set to have error-free performance at 30 Gb/s. The same transistor geometrics were used in both the 1:4 and 4:1 circuits in our research HBT process.

3.3. Gain control amplifier

Broadband variable gain amplifiers are also key elements for optical communication systems. Receivers for such systems require linear channel response with little magnitude variation, constant group delay and frequency response down to dc to achieve good bit error rates. We have demonstrated a *packaged* amplifier with these properties from DC to 26 GHz [17]. A microphotograph of this chip is shown in Fig. 4(a). The circuit consists of an input buffer, two amplifier stages connected by emitter follower buffers and an output buffer. The input and output buffers provide good terminations for 50 Ω source and load. The basic amplifier in this circuit is a modified transadmittance/transimpedance (TA/TI) amplifier pair [18]. Variation of I_1 by an external gain controlled voltage changes the transimpedance (g_m) of the TA amplifier which in turn varies the overall gain of the amplifier. The transistors in the parallel feedback path of the TI amplifier buffer the output from the feedback network, which results in higher gain, wider bandwidth and constant output dc voltage independent of the gain-controlled voltages [19]. Emitter followers between the stages provide level shifting and improve the input/output impedance matching between the output TI portion of the first amplifier stage and the input TA portion of the second amplifier stage. The measured performance of this variable gain amplifier (VGA or AGC) is shown in Fig. 4(b). Note that this circuit has a gain variation of only ±1 dB, constant group delay within the passband, and a gain controlled range of 10–16 dB. This

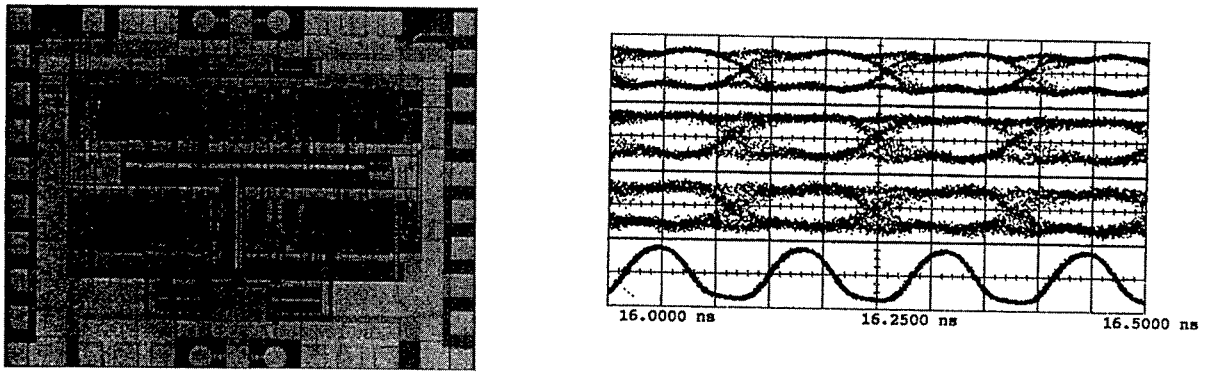
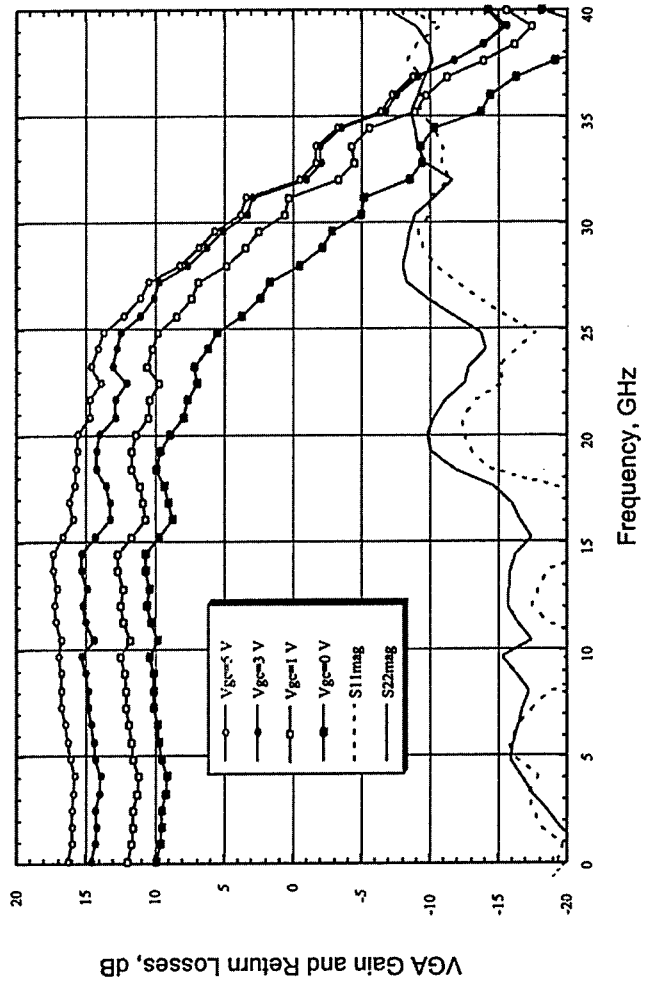


Fig. 3. (a) Microphotograph of 30 Gb/s 1:4 demultiplexer. The die size is 1 × 2 mm^2 and the chip dissipates 3.1 W from a single −7.7 V supply. (b) Output eye diagrams of channels 1–3, divided clock (7.5 GHz)(400 mV/div, 50 ps/div).



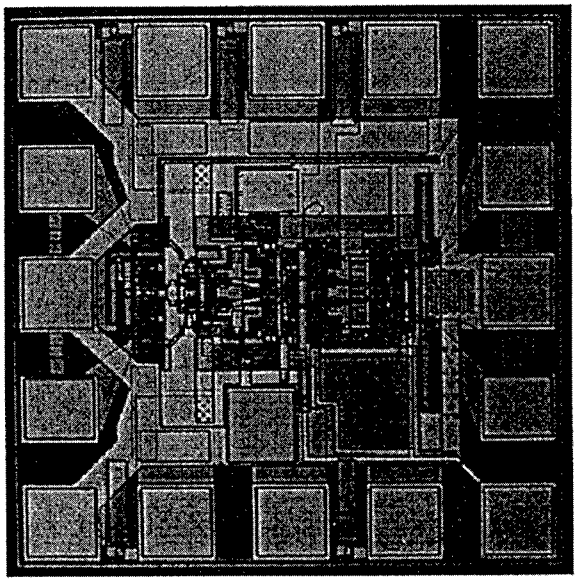


Fig. 4. (a) Microphotograph of packaged variable gain amplifier (VGA). The chip is 0.72×072 mm^2 and dissipates 930 mW into two supplies, $V_{cc} = 5$ V and $V_{ee} = -5.2$ V. (b) Measured packaged VGA performance.

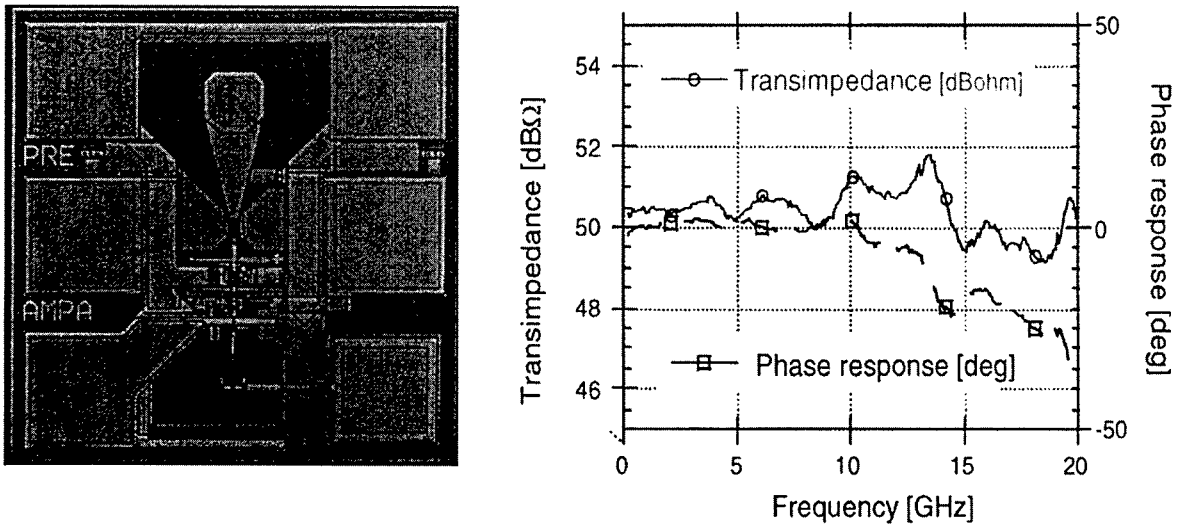


Fig. 5. (a) Microphotograph of preamplifier circuit. The die is 0.45×0.45 mm^2 and dissipated 0.3 W into ± 5 V supplies. (b) Measured performance using optical network analyzer.

performance is suitable for use in >30 Gb/s fiber-optic TDM transmission systems [20].

3.4. Preamplifier

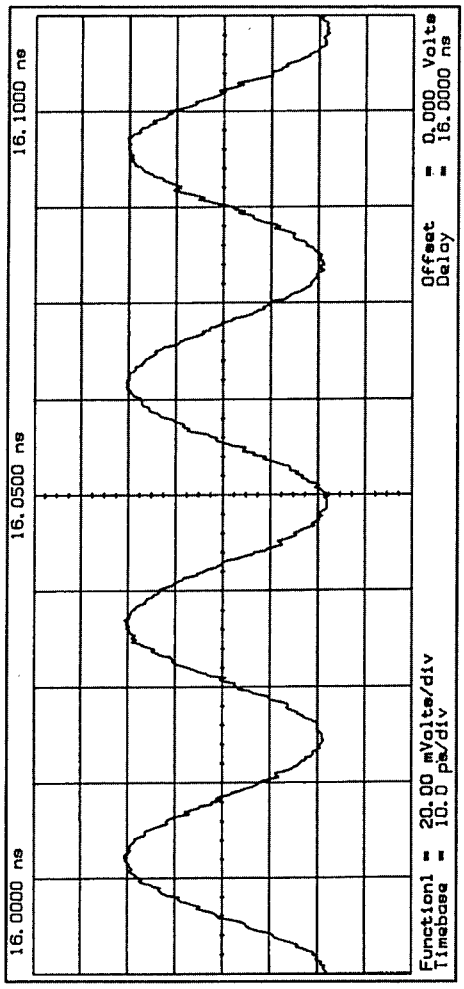
The preamplifier was a transimpedance amplifier designed to have maximum flatness (±0.5 dB) from 0 to 30 GHz. The chip presented here was an electronically probable version of this part [21]. Another two versions consisted of a flip-chip compatible transimpedance (T_z) amp and an integrated T_z amp + Schottky detector (for $\lambda = 0.86$ μm) circuit. A microphotograph of the 0.86 μm version is shown in Fig. 5(a). The chip was measured using an optical spectrum analyzer. The results are shown in Fig. 5(b). The 3-dB bandwidth of this chip was 25 GHz. Note the excellent gain flatness and good phase response. Several different chips on each wafer were tested. Preamplifiers on the same wafer showed very similar responses, indicating excellent process uniformity.

3.5. Clock and data recovery

One of the conventional ways to regenerate the clock signal from a NRZ data stream is to differentiate and rectify the NRZ data stream [22]. In this differentiate/rectify (DR) circuit, a capacitance–resistance (C–R) network achieves differentiation. The capacitor is implemented with base-emitter (BE) junction of a diode-connected HBT, and the regenerated clock power is maximized by proper V_{be} bias; rectification is achieved with a pair of common collector/emitter HBTs [23]. Because passive components and rectification perform differentiation by current switching in emitter followers, high-speed operation can be obtained with this circuit. The regenerated clock signal is high-pass filtered to remove the unwanted low-frequency signal/noise and to maintain proper dc bias for the DR circuit.

The delay/multiply (DM) circuit (Fig. 6(a)) consists of a Gilbert multiplier and a variable delay line. The recovered clock power can be maximized when the delay is half the bit period [24]. The variable delay line is implemented with a high-impedance transmission line periodically loaded with reverse-biased BE junctions of diode-connected HBTs. The delay of the line can be optimized by changing the BE junction bias of the diode connected HBTs to obtain maximum regenerated clock power. The nominal characteristic impedance of the delay line is 50 Ω and terminated at the end of the line with 50-Ω resistors (not shown) to minimize reflections. It can be shown that the delay per section of line is $\tau_o = \tau_l \times \chi$ and the overall impedance of the line is $Z_o = Z_l/\chi$, where τ_l and Z_l is, respectively, the delay per section and impedance of the unloaded transmission line, and $\chi = (1 + C_j/C_l)^{1/2} \geq 1$, with C_j being the junction capacitance of the HBTs and C_l the transmission line capacitance. Fig. 7 shows τ_o and Z_o as a function of χ with $\tau_l = 4$ ps and $Z_l = 90$ Ω. While τ_o depends linearly on χ, Z_o remains relatively constant. As a design example, to obtain τ_o variable from 8.3 to 6.3 ps (corresponding to optimum delay for 30 and 40 GBit/s input data), the line impedance only changes from 43 to 58 Ω; hence the return loss remains under −22 dB. Fig. 6(b) shows the single ended output waveform from the delay and multiply circuit.

To obtain high-speed operation, the phase detector (PD) circuit is implemented with a Gilbert-cell type



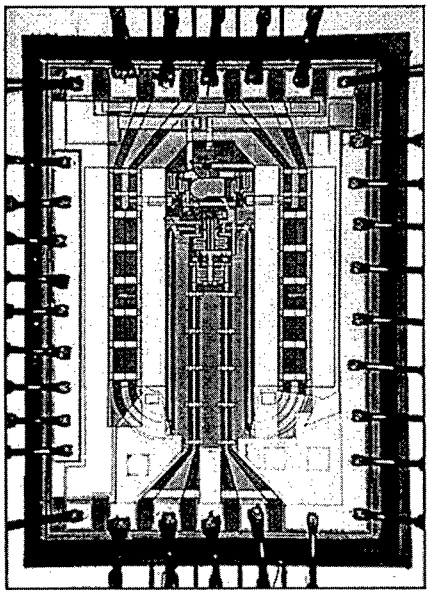


Fig. 6. (a) Microphotograph of delay and multiply (DM) circuit. (b) Recovered clock signal using DM circuit.

mixer. When two high-speed signals with similar frequencies are applied to the LO and RF ports of the mixer, the output dc (or low-frequency) signal is proportional to the input signal phase difference [25] and the mixer operates as a phase detector. With high-speed HBT devices, large phase-detector gain over wide bandwidth is attainable with proper transistor sizing and biasing. The output of the phase detector is low-pass filtered to reject high-frequency noise and help ease the output port impedance matching requirements.

Having discussed circuits that take advantage of the high-performance offered by heterostructure devices, we now discuss data conversion circuits that place great emphasis on not only the performance, but also on the maturity of a technology.

4. Data conversion

Data conversion circuits such as DAC's, ADC's and direct digital synthesizers are of great interest to both military and commercial markets. HBT based 12-bit 1 GHz DAC's are just one of the commercially available chips available for data conversion [26]. In addition, the large Early voltages (>500), high-speed ($f_T > 50$) and compatibility with high-performance Schottky diodes, make HBTs excellent devices for analog to digital conversion applications. The excellent device matching and high-yields achievable with HBTs are also critical for these applications. We have fabricated several high-speed ADCs that leverage the capabilities of our commercial fabrication facility. In addition to these applications, a direct digital synthesizer ($>10{,}000$ transistors) operating at speeds greater than 1.5 GHz has been demonstrated in this technology. We now briefly describe several of the analog to digital converters that have been demonstrated at Rockwell [27].

A joint Hewlett Packard/Rockwell team funded by Air Force Wright Laboratory [28] demonstrated a 6-bit ADC suitable for application in digitizing oscilloscopes. This ADC features an analog folding architecture, includes on-chip master–slave track-and-hold (T/H) circuit and provides Gray-encoded digital outputs. The ADC achieved 5.6 effective bits at 4 Gsps. It has a resolution bandwidth of 1.8 GHz at 4Gsps. The chip operates at up to 6.5 Gsps.

An 8-bit 2Gsps ADC, as shown in Fig. 8, has been fabricated using Rockwell's AlGaAs/GaAs HBT foundry process [29]. The ADC uses a folding and interpolation architecture to provide wide bandwidth with moderate device count. When sampling at 2 Gsps, the ADC demonstrated an effective resolution between 6.5 and 7.0 effective bits for analog inputs between dc and 1.5 GHz and a single-tone, spurious-free dynamic range of 48 dB at Nyquist. The ADC incorporates

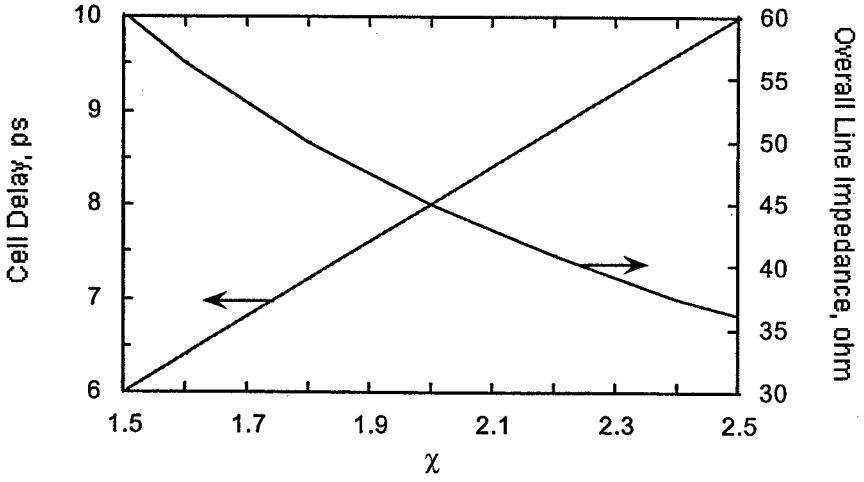


Fig. 7. (a) Delay line impedance and cell delay as a function of χ.

over 2500 transistors and dissipated 5.3 W from 5.0 and −5.2 V supplies. This chip is being used in several system demonstrators, including a laser altimeter for NASA.

A 10-bit 1 Gsps ADC currently under development [30] consists of 5500 HBTs and has 7-W power dissipation. Measurements on the first iteration of this chip show fully functional ADCs with 7.5–8.5 effective number of bits and a spurious-free dynamic range of 60 dB up to 500 MHz clock speed. Several areas of improvement have been identified and a second iteration of the chip is in fabrication.

Direct digital synthesizers are useful for a number of applications, from satellite receivers to high-performance test equipment. Rockwell has demonstrated the world's most advanced, fast-hopping frequency synthesizer chip for EHF-satellite terminals that operates from 10 MHz to 1.5 GHz clock with 4 Hz increments and a total power dissipation of 3.5 W. The spurious-free dynamic range is 60 dBc ($f_0 = 65 \times f_{clk}$). This chip consists of a phase accumulator, sine ROM look-up

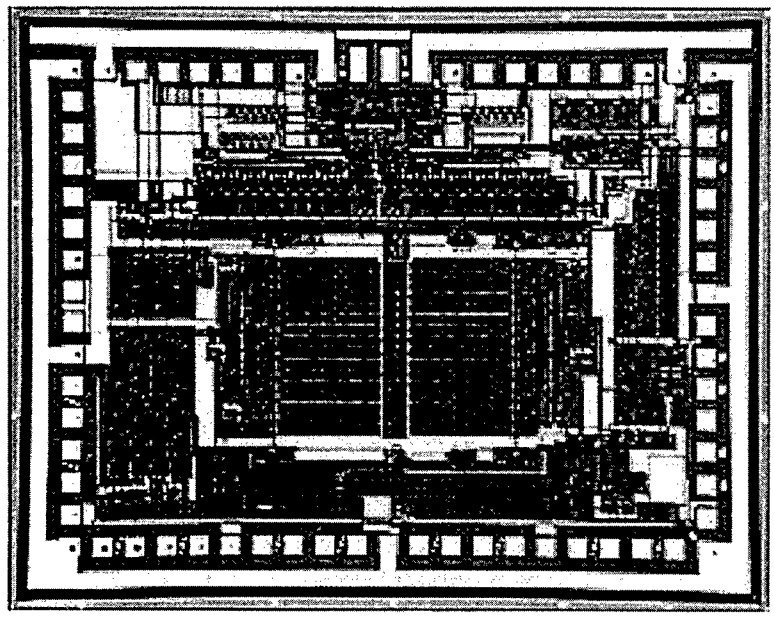
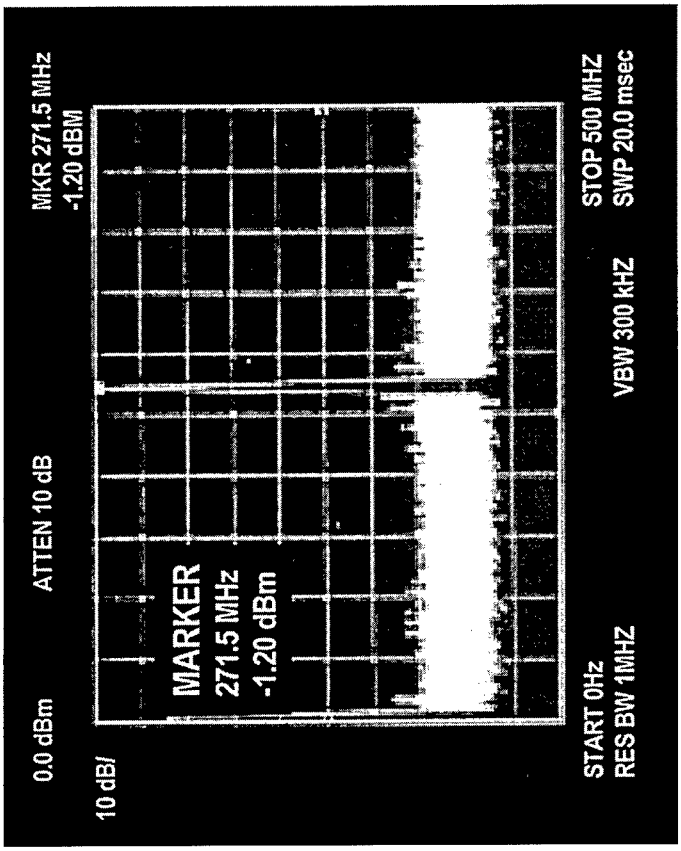


Fig. 8. (a) 8-Bit, 2Gs/s HBT analog to digital converter. The die is 2.775 × 3.525 mm^2 and dissipates 5 W into +5 V and −5.2 V supplies.



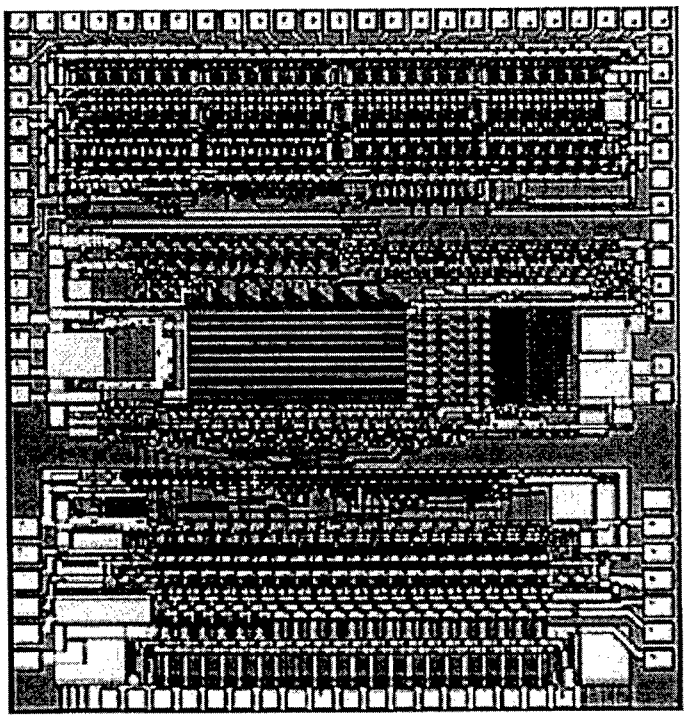


Fig. 9. (a) Microphotograph of direct digital synthesizer (DDS) chip. (b) Measure performance of DDS chip.

table and a DAC. The entire circuit uses over 10,000 transistors and is shown in Fig. 9(a). The resulting frequency spectrum is shown in Fig. 9(b).

It is clear, from the device counts and critical dependence of performance on device matching that these data conversion applications are taking full advantage of the growing infrastructure for HBTs. Thus, heterostructure devices are no longer limited to small demonstration circuits, but should be considered viable technologies for larger systems applications.

5. Summary

All of the above application areas stress a different aspect of the rapid expansion of heterostructure devices into commercial and military markets. Power amplifiers demonstrate that HBTs are cost-competitive with other technologies and are gaining a significant manufacturing base. HEMTs are currently the preeminent device in applications above 30 GHz. Telecommunications applications are taking advantage of the superior performance of heterostructure devices to ease the development and implementation of high-bit rate communication chips. The prospects for continued growth in this area are excellent. Finally, data conversion applications that require good matching of characteristics (and accurate device models), high-performance and high-yield for transistor counts >5000 transistors are readily achievable using current HBT production devices.

Acknowledgements

The authors acknowledge the support and guidance of JP Rode, J Bowers and M Melliar-Smith. We also gratefully acknowledge the technical contributions to this work of P Asbeck, K Poulton, R Nubling, K Nary, S Beccue, MF Chang, M Rodwell, M Miller, M Bloom, K Weller and KC Wang. We are indebted to the processing staffs of RSS Newbury Park and Rockwell Science Center, especially S Fitzsimmons for fabricating the circuits described in this work. The funding for the circuits described in this paper has come from a variety of sources including DARPA, WPAFB, NASA and Rockwell internal funds.

References

[1] Browne J. Power-amplifier MMICs drive commercial circuits. In: Microwaves and RF, 1998. p. 116–24.
[2] Golio M. Low-voltage microwave electronics. In: Proceedings of IEEE MTT-S IMS, paper TU4C, 1998.
[3] Rockwell Science Center Data Sheet KaLNA3S.03.
[4] Rockwell Science Center Data Sheet KaLNA3S.04.
[5] Rockwell Science Center Data Sheet KaLNA4S.03.
[6] Rockwell Science Center Data Sheet Wlna5ssp.
[7] Wang NL, Ho WJ, Higgins A. Ultrahigh-power efficiency operation of common-emitter and common-base HBT's at 10 GHz. IEEE Trans Microwave Theory Tech 1990;38:1381–9.
[8] Bayraktaroglu B, Barrette J, Kehias L, Huang CI, Fitch R, Neidhard R, Scherer R. Very high-power-density CW operation of GaAs/AlGaAs microwave heterojunction bipolar transistors. IEEE Electron Device Lett 1993;14(10).
[9] http://www.nb.rockwell.com/platforms/network_access/nahome.html#5.
[10] http://www.nortel.com/technology/opto/receivers/ptav2.html.
[11] Swahn T, Lewin T, Mokhtari M, Tenhunen H, Walden R, Stanchina W. 40 Gb/s 3 Volt InP HBT ICs for a fiber optic demonstrator system. In: Proceedings of the 1996 GaAs IC Symposium, 1996. p. 125–8.
[12] Suzuki H, Watanabe K, Ishikawa K, Masuda H, Ouchi K, Tanoue T, Takeyari R. InP/InGaAs HBT ICs for 40 Gbit/s optical transmission systems. In: Proceedings of the 1997 GaAs IC Symposium, 1997. p. 215–8.
[13] Runge K, Zampardi PJ, Pierson RL, Thomas RB, Beccue SM, Yu RT, Wang KC. High-speed AlGaAs/GaAs HBT circuits for up to 40 Gb/s optical communication. In: Proceedings of the 1997 GaAs IC Symposium, 1997. p. 211–4.
[14] Berroth M, Lang M, Wang Z-G, Lao Z, Thiede A, Motzer M, Bronner W, Kaufel G, Kohler K, Hulsmann A, Schneider J. 20–40 Gbit/s 0.2 μm GaAs HEMT chip set for optical data receiver. In: Proceedings of the 1996 GaAs IC Symposium, 1996. p. 133–6.
[15] Runge K, Pierson RL, Zampardi PJ, Thomas PB, Yu J, Wang KC. 40 Gbit/s AlGaAs/GaAs HBT 4:1 multiplexer IC. Electron Lett 1997;31(11):876–7.
[16] Runge K, Pierson RL, Zampardi PJ, Thomas PB, Yu J, Wang KC. 30 Gbit/s 1:4 demultiplexer IC using AlGaAs/GaAs HBTs. Electron Lett 1997;33(9):765–6.
[17] Yu R, Beccue S, Zampardi P, Pierson R, Petersen A, Wang KC, Bowers J. A packaged broadband monolithic variable gain amplifier implemented in AlGaAs/GaAs HBT technology. IEEE J Solid-State Circuits 1996;31(10):1380–7.
[18] Cherry BM, Hooper DE. The design of wide-band transistor feedback amplifiers. In: Proc. IEE, 1963. p. 375–89.
[19] Ishihara N, Nakajima O, Ichino H, Yamauchi Y. 9 GHz bandwidth, 8–20 dB controllable-gain monolithic amplifier using AlGaAs/GaAs HBT technology. Electron Lett 1989;25(19):1317–8.
[20] Petersen A, et al. 3 MHz-30 GHz traveling-wave optical front-end receiver. In: Proceedings of OFC '95, 1995. p. 157–8.
[21] Zampardi PJ, Pierson RL, Runge K, Yu R, Beccue SM, Y J, Wang KC. Hybrid digital/microwave HBTs for >30 Gb/s optical communications. In: IEDM technical digest, 1995. p. 803–6.
[22] Wang ZG, Langmann U, Bosch B. Multi-Gb/s silicon

bipolar clock recovery. IEEE J Selected Areas Commun 1991;9((5)):656–63.

[23] Yu R, Pierson R, Zampardi P, Runge K, Campana A, Meeker D, Wang KC, Petersen A, Bowers J. Packaged clock recovery integrated circuits for 40 Gbit/s optical communication links. In: Proceedings of the 1996 GaAs IC Symposium, 1996. p. 129–32.

[24] Imbeaux J. Performance of the delay-line multiplier circuits for clock and carrier synchronization in digital satellite communications. IEEE J Selected Areas Commun 1983;1(1):82–95.

[25] Gray P, Meyer R. In: Analysis and design of analog integrated circuits, 2nd ed. John Wiley and Sons, 1984. p. 604.

[26] http://www.rsc.rockwell.com/converters/12bit_data.pdf.

[27] http://www.rsc.rockwell.com/converters/index.htm.

[28] Poulton K, Knudsen K, Corcoran J, Wang KC, Nubling RB, Chang M-CF, Asbeck PM, Huang RT. A 6-b, 4 GSa/s GaAs HBT ADC. IEEE J Solid-State Circuits 1995;30:1109–98.

[29] Nary K, Nubling R, Beccue S, Colleran W, Penney J, Wang KC. An 8-bit, 2 gigasample per second analog to digital converter. In: Proceedings of the 1995 GaAs IC Symposium, 1995. p. 303–6.

[30] Yu R, Sheng NH, Cheng K, Gutierrez G, Wang KC, Chang MF. A 1 GS/s 11-b Track-and-Hold Amplifier with <0.1 dB Gain Loss. In: Proceedings of the 1997 GaAs IC Symposium, 1997. p. 87–90.

MANUFACTURING

PERGAMON

Solid-State Electronics 43 (1999) 1645–1654

SOLID-STATE ELECTRONICS

A 0.1-μm MHEMT millimeter-wave IC technology designed for manufacturability

Hans Rohdin*, Arlene Wakita, Avelina Nagy, Virginia Robbins, Nick Moll, Chung-yi Su [1]

Hewlett-Packard Laboratories, P.O. Box 10350, Palo Alto, CA 94303, USA

Abstract

We describe and discuss our approach to improved manufacturability of millimeter-wave InP-type HEMT IC technology. The main ingredients are the use of (1) GaAs (rather than InP) as substrate, with a buffer that grades the lattice constant to that of the high-mobility $Ga_{0.47}In_{0.53}As$ channel; (2) low-resistance non-alloyed ohmic contacts with good reliability and reproducibility; (3) an e-beam process that produces 0.1-μm T-gate fingers with one exposure and (4) two-step selective etching of the gate trough for good uniformity. The process is compatible with standard GaAs FET IC front and back end unit processes, including backside vias. The yield, uniformity and performance of devices and circuits made on this metamorphic (MHEMT) material are consistently as good as those we get using InP substrates. The reliability of our devices, whether made on GaAs or InP substrates, is limited by impact ionization in the channel, which has also been linked to R_d-degradation and burnout. At low-noise drain bias ($V_d = 1$ V), where the impact ionization is small, we have demonstrated reliability, i.e. extrapolated MTTF at 125°C channel temperature is $> 10^6$ h. This is as good as FETs made on InP substrates. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

The need for a low-noise, high-performance FET MMIC technology for future communication and instrument applications can be met with a short-gate InP HEMT process (e.g. Refs. [1–3]). However, before the full benefits of such a process can be realized, issues with manufacturability have to be worked out [2,3]. At HP there is a degree of incompatibility between existing GaAs-based III–V FET/MMIC processes in manufacturing (e.g. Ref. [4]) and a next-generation higher-performance FET process on InP. Relative to GaAs, InP substrates are more expensive and brittle, particularly for larger wafer sizes (≥3 inch). We have overcome this obstacle by implementing a buffer technology that allows the use of GaAs substrates [5,6]. There are additional features in our process that promote manufacturability. Our good experience with non-alloyed ohmic contacts to pseudomorphic $In_xGa_{1-x}As$ (x = 0.25, 0.35) on GaAs [7], made these particularly attractive in the present case of even higher In mole fractions. Other workers have also pursued this approach because of the reproducibly low and uniform contact resistance [1]. The 0.1-μm T-gates are presently defined by a one-step electron beam lithography process, which self-aligns the top of the T with the stem. A higher throughput deep-UV phase–shift mask approach is being developed as an alternative [8]. The need for threshold uniformity has been addressed

* Corresponding author.

[1] Present address: HP Wireless Semiconductor Division, San Jose, CA 95131, USA.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00114-8

by an approach that initially attempted to leverage off our dry-etch experience with earlier GaAs-based MODFET technology [9–13]. Ultimately, however, we settled on a two-step selective wet-etch approach. In the following we will discuss these approaches in more detail, show some circuit results, and finally discuss our experiences with device degradation and reliability. Most of the thinking that has gone into the process has been summarized in a simple tutorial way in Ref. [14].

Table 1
Typical FET parameters

$R_c = 0.15\ \Omega$ mm	$d_{gc} = 23$ nm
$R_S = 0.35\ \Omega$ mm	$L_g = 120$ nm
$V_{th} = -0.35$ V	$f_{Tx}^{(max)} = 190$ GHz
$\sigma_{V_{th}}^{(wfr)} = 60$ mV	$f_{max} = 350$ GHz
$\sigma_{V_{th}}^{(wfr\text{-}to\text{-}wfr)} = 80$ mV	$C_{fs} = 150$ fF/mm
$I_d^{(max)} = 700$ mA/mm	$P_{sat}^{(60\ GHz,\ 2V)} = 210$ mW/mm
$BV_{ds}^{(off)} = 6$ V	$NF_{min}^{(12\ GHz,\ on\text{-}wfr)} = 0.34$ dB
$g_{mx}^{(max)} = 950$ mS/mm	$G_a^{(12\ GHz,\ on\text{-}wfr)} = 16.8$ dB
$g_{dx}^{(sat)} = 5$ mS/mm	

2. Structure, growth and process

The epitaxial structure and gate definition of our basic FET is shown in Fig. 1. Typical characteristics are shown in Table 1. The structure is grown by MBE. Except for the GaAs substrate, and the first and the last three layers grown, the structure is conventional. The first layer is the linearly-graded low-temperature buffer (LGLTB) in which the lattice constant is varied from that of GaAs to that of InP by gradually replacing Ga in $Al_{0.48}Ga_{0.52}As$ with In over 1 μm of thickness, so that, at the top of the buffer, $Al_{0.48}In_{0.52}As$ is grown [5,6]. The vast majority of the misfit dislocations generated in this process remain confined to the LGLTB. The threading dislocation density in the device layers is typically below the limit observable in TEM ($< 10^7$ cm^{-2}) [14,15]. It is possible that the lower yield strength of material with less In tends to confine the dislocations closer to the GaAs substrate [16]. We have not observed any reduction in FET performance or yield as a result of the use of a GaAs substrate [17]. The approach allows some freedom in the choice of In mole fraction. The conduction-band offset ΔE_c for the 2DEG is maximum for ≈30% indium, which could

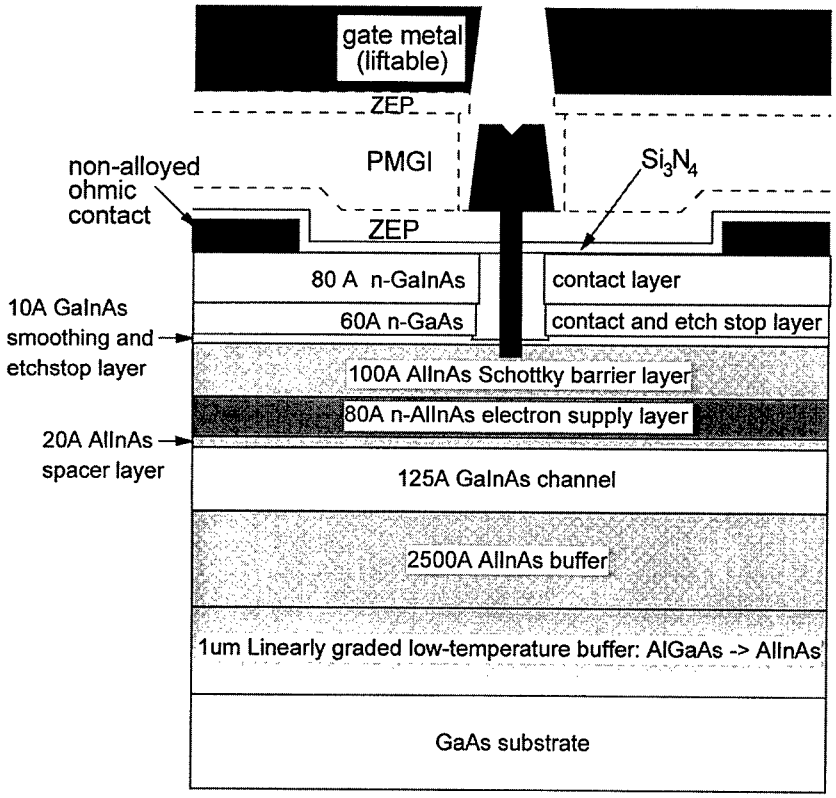


Fig. 1. Standard epitaxial structure and gate definition of the 0.1-μm MHEMT.

lead to larger maximum carrier concentration n_{so} and breakdown voltage BV_{ds} [18]. However, it also makes it harder, if not impossible, to achieve good contact resistance R_c with non-alloyed ohmic contacts, because of the $\approx$70% Al in the Schottky barrier layer.

The top two layers in Fig. 1 provide for low parasitic source and drain resistances (R_s, R_d) and a well-defined, uniform and reproducible threshold voltage V_{th}. Low R_c is accomplished by the choice of material (GaInAs and GaAs) and by high Si doping (6×10^{18}–3×10^{19} cm^{-3}). This results in a low barrier for the electrons and, thus, low tunneling resistance. After completion of the IC process, R_c is $\approx$0.15 Ω mm. The control and reproducibility of MBE and metal deposition are better than of typical alloying processes. Threshold control is accomplished by a two-step selective recess process. The first succinic-acid based etch removes the top GaInAs layer at a rate that is 10–20 times higher than in the underlying GaAs. This allows for sufficient lateral etch of the GaInAs before the GaAs layer is consumed. With GaAs (rather than AlAs) as an etch stop, the succinic etch had to be modified from that in Ref. [19]. The second etch (for the remaining GaAs) was initially intended to be RIE, since recipes with excellent selectivity between GaAs and AlGaAs have been developed by us and by others. It has been shown that such a recipe can be ported to InP-based HEMTs with the inclusion of a GaAs layer on top of the AlInAs Schottky barrier layer [20]. Early attempts with our CCl_2F_2/He process [10,11] resulted in very low breakdown voltage. Considering the challenges involved in not just tweaking an existing process, but actually having to develop from scratch a low-damage, environmentally acceptable RIE process, we instead decided to leverage off our experience with selectively etching GaAs on pseudomorphic InGaAs [7]. With the $\approx$50% indium in the present layers, the selectivity of NH_4OH:H_2O_2:H_2O is even better. It consumes what is left of the GaAs layer and literally stops on the underlying In-containing layer. This is key to threshold control. The good selectivity of the two-step wet etch is illustrated in Fig. 2. The amount of lateral etch strongly affects the maximum drain current $I_d^{(max)}$ and off-state breakdown voltage $BV_{ds}^{(off)}$ and determines where on the $I_d^{(max)}$–$BV_{ds}^{(off)}$ trade-off line the FETs will fall [14]. The lateral etch also affects the knee voltage of the FETs, since R_s and R_d will have a non-negligible component associated with these etched higher-resistance regions. Excessive lateral etch can introduce an unacceptably large kink [1]. There is no lateral etch stop, but as the succinic-acid based etch proceeds in the thin layers, the rate is reduced, eventually to zero at a final lateral extent of approximately 0.1 μm, which results in good overall device characteristics.

The thickness and quality of the GaAs etch layer is important. If it is too thick it will not be coherent with the layers below, and the doping efficiency will be reduced dramatically, preventing a low R_c. If it is too

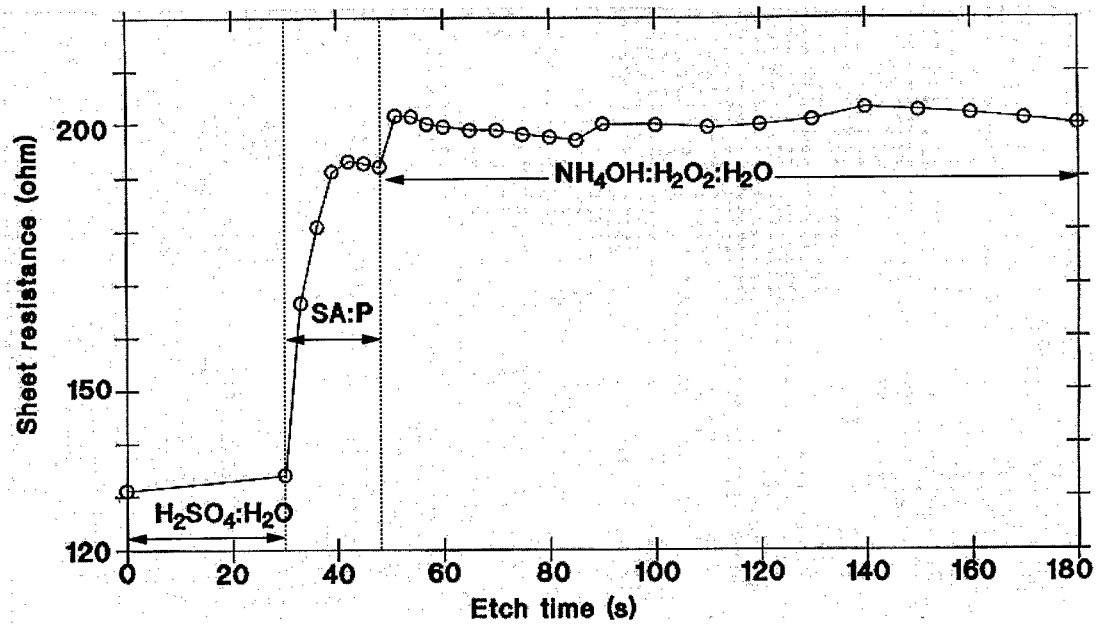
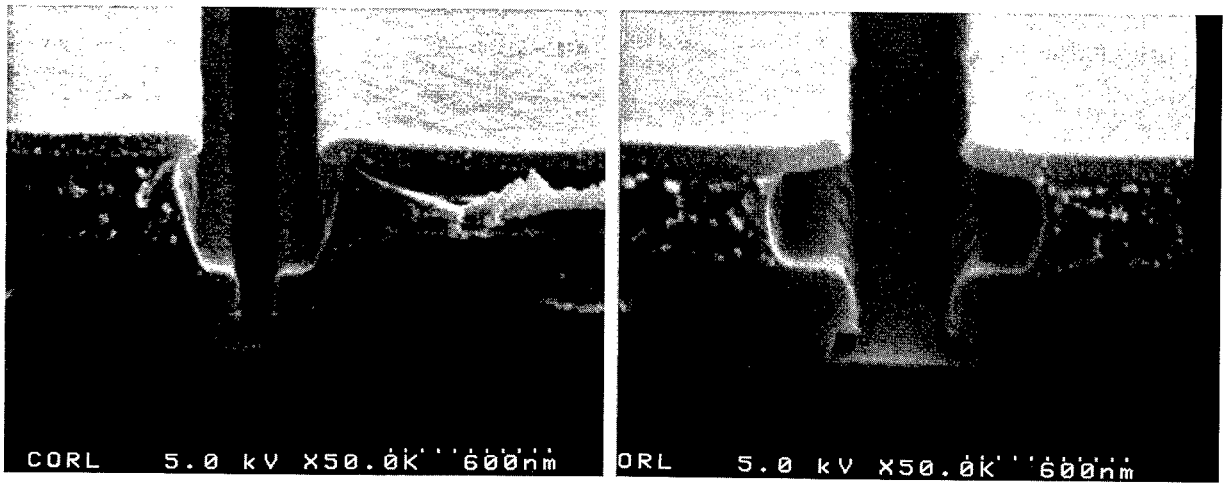


Fig. 2. Demonstration, using the contactless sheet resistance of a blank wafer, of the selectivity of the two-step wet gate recess. SA = succinic acid, P = H_2O_2.

thin, the semi-selective succinic-based first etch will consume it and some uncontrollable fraction of the underlying AlInAs Schottky-barrier layer, leading to a nonuniform V_{th}. Initially, 200-Å GaAs layers proved to not be efficiently doped, while 100-Å layers were. Although Matthews–Blakeslee theory [21] leads to a 40 Å critical thickness for GaAs cladded by GaInAs, experience suggests that the actual critical thickness can be significantly larger. This has recently been quantified experimentally for layers in tensile stress [22]. Together with our good results with 100 Å layers, this suggests that our choice of 60 Å is safe. We have successfully made FETs on material that was rapid-thermal annealed at 650°C for 5 s, which further suggests that our design of the cap structure is robust. The final layer of particular interest, the thin GaInAs layer over the AlInAs Schottky barrier layer in Fig. 1, was initially inserted to improve the growth of the strained GaAs layer. It was felt that starting the GaAs growth by simply closing the In shutter was better than starting it on AlInAs, particularly if a growth interruption, to increase the Si cell temperature for the cap, was used. The etch layer structure we settled on was in large part the result of the constraints imposed by the element sources initially available to us in the MBE reactor used. Alternative etch-stop layers with phosphorous have been successfully included by other workers in the field [1]. Being particularly concerned with reliability and contact resistance, our cap design is quite different from that in Ref. [20], where a 600 Å uncapped layer of GaAs was used, in conjunction with alloyed ohmic contacts and RIE recess. In such a thick GaAs cap, the first 300–400 Å of the GaAs contain misfit dislocations [20].

Prior to the recess, the gates have been defined by direct e-beam writing in a trilayer resist [23] as illustrated in Fig. 1. This results, after development and RIE, in a $\approx$0.12 μm gate cut in the bottom layer ZEP resist. Shorter gates can be defined with this approach, but the present performance and reproducibility requirements are met with 0.12 μm. Because of fringing fields that modulate the 2DEG (not included in the passive fringe capacitance $2C_{fs}$) the electrical gate length is $\approx$30 nm longer than the metallurgical length L_g in Table 1. The opening in the top layer resist (also ZEP) is larger because of a stronger (low contrast) developer. The middle layer resist (PMGI) develops out further to provide an overhang. Thus, a single e-beam exposure produces a resist cut for an easily liftable T-gate. The approach relies on the excellent selectivity of the developers [23].

Given the deep sub-micron dimension of the gate cut, controlling the uniformity and reproducibility of the parameters with wet chemistry is non-trivial, even with the built-in vertical and lateral etch control discussed above. The controlled reproducible initiation and quenching of wet etching in small semi-confined openings with poor aspect ratio require special care. Fig. 3 shows the SEM cross-sections of two *non-selectively* etched gate cuts in a GaAs test wafer. The 0.11 μm cut corresponds to our standard FET, while the 0.35 μm cut corresponds to gates used for level-shifting diodes. The very similar vertical and lateral etch in the two cuts suggests that that the initiation and quenching are quite controlled. The etch rate appears to be somewhat larger than in GaAs test pieces with large openings in nitride. This suggests that the quenching is slower than the initiation. The threshold control evi-



(a) (b)

Fig. 3. Cross-section scanning electron micrographs of two *non-selectively* etched troughs in a GaAs test wafer with different gate-cut sizes defined by one-step e-beam direct writing in tri-layer resist. (a) Standard 0.1-μm gates. (b) 0.35-μm gates used for diodes.

dent in Table 1 is based on whole or partial 2-inch wafers processed in HP Labs R&D environment. Better uniformity has been demonstrated on full 3 inch wafers at HP manufacturing divisions, where 20-mV standard deviation in V_{th} is the norm. This indicates that the etching techniques and recipes are effective and robust.

As with all good T-gate processes, the gate metallization access resistance is kept quite low ($r_{ga} \approx 100\ \Omega$/mm). In fact, it is too low to explain the value for the typical FET gate resistance R_g extracted from microwave measurements. In the course of our process development, it has become clear that the dominant component in R_g is a component R_{gi}, which scales inversely with gate width W_g [17]. By a theoretical and experimental process of elimination we have identified R_{gi} to be an ac interfacial layer resistance between the gate metal and the semiconductor surface states. This has rather profound consequences for device optimization and model scaling. Details on this topic can be found in Ref. [24]. A theoretical analysis suggests that even a pristine metal–semiconductor interface produces a small but possibly significant R_{gi} due to the resistance associated with tunneling across the interfacial dipole layer that exists even in an ideal Schottky barrier [25]. With a given lithography, we can do the following to approach this ultimate situation: (1) use an effective clean-up dip prior to evaporation; (2) minimize the delay to evaporation and (3) use a gate metal that can be controllably sintered into the semiconductor, while consuming remaining native oxide and/or contaminants. Only the third point warrants further discussion. From the standpoints of a clean interface and of barrier height, the best gate metal appears to be Pt, which has been studied and used by workers in the III–V FET field for quite some time (e.g. Ref. [26]). More recently it has been used for AlInAs/GaInAs MODFETs [27,28]. Because of the scaling between Pt thickness and sintering depth [26], the Pt thickness must be limited [27]. We use 25 Å Pt separated from the top Au by a Ti/Pt barrier much like in Ref. [28]. Interestingly, we have found, by analyzing V_{th} versus Pt thickness, the same 1.4 scaling factor between sintering depth and initial Pt thickness as was found in Ref. [28]. The 35-Å shift in metal–semiconductor interface position is reflected in the effective gate-to-channel distance d_{gc} in Table 1. The barrier height is $\Phi_b \approx 0.7$ eV, with better ideality than our earlier Ti gates. The undoped GaInAs smoothing and etch-stop layer discussed above should be thinner than the sintering depth, so that the Schottky barrier is formed on the wide band-gap AlInAs. The GaInAs area contacted from the side by the Pt is too small to significantly increase the gate leakage. The minimum reproducible normalized interfacial gate resistance $r_{gi} = R_{gi}L_gW_g$ associated with the Pt gate is $\approx 3 \times 10^{-7}\ \Omega\ \text{cm}^2$ [24].

3. Circuit examples

Several demonstration circuits, with emphasis on low noise and wide-band applications, have been designed and fabricated in HP's Microwave Technology Division. In a 2–50 GHz distributed amplifier a 3.1 dB noise figure was measured in the 36–42 GHz band, which is a 3 dB improvement over the existing 0.25-μm PHEMT process [29]. A distributed amplifier with 12 dB low-frequency gain and 8 dB gain at 110 GHz has also been demonstrated [30]. Results like these heighten the interest in getting this process into manufacturing. The major stumbling block is not manufacturability per se, but rather reliability at $V_d \geq 1.5$ V. We will discuss this in the next section. The two circuits shown in Figs. 4 and 5 were designed and fabricated at HP Laboratories. Fig. 4(a) shows a three-stage broadband MMIC amplifier designed with coplanar transmission-lines. The design is based on Ref. [31]. Fig. 4(b) shows the gain in V-band (50–75 GHz). The gain is 25 dB at 50 GHz, and has dropped 3 dB at ≈68 GHz. This further illustrates the potential of this technology for millimeter-wave applications.

Given the uncertainty in the number of dislocations that thread through the device layers we were interested in the yield of a more complex circuit. Fig. 5(a) shows a static divide-by-4 circuit with a total of 97 FETs, 49 of which have 0.1 μm gates. The others are used for level-shifting and have 0.35 μm gates (Fig. 3b). The circuit is of interest for potential digital applications, and the design is very similar to that in Ref. [13]. The differences are due to the present availability of only one type of FET (depletion mode). The average divider yield for four wafers, three on GaAs and one on InP, was 45%. Two of the three GaAs wafers actually had higher yield than the InP wafer, indicating that the LGLTB does not impact the yield negatively. In estimating a density of electrically active dislocations from the yield numbers, we assume (1) a uniform threading dislocation density, (2) that the yield loss is due exclusively to dislocations and (3) that only 0.1-μm switching FETs are affected by dislocations through the gate area. Based on the non-unity yield of the InP wafer and visual inspection, assumption (2) is not strictly true, although observable flaws, such as metal peeling and spitting, have been considered too infrequent to warrant quantification. Assumption (3) excludes the possibility that 0.1-μm current sources and 0.35-μm level-shifting diodes are also affected by dislocations through the gate area. The estimate $(2–7) \times 10^5\ \text{cm}^{-2}$ is thus an upper limit of an effective uniform density of harmful threading dislocations. This is similar to the $5 \times 10^5\ \text{cm}^{-2}$ arrived at by analyzing the yield of discrete FETs on a 3 inch wafer from HP's Wireless Semiconductor Division [14], where observa-

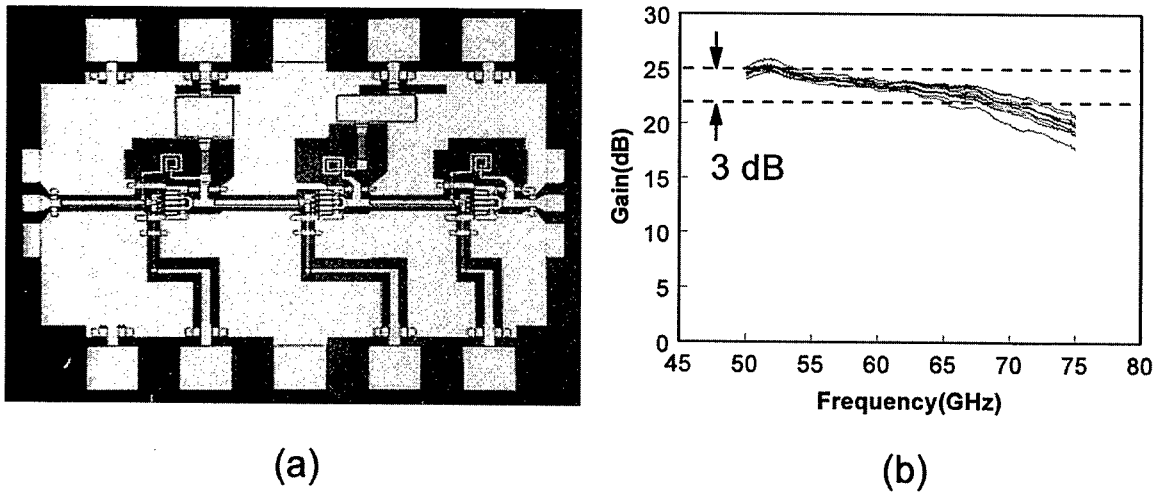


Fig. 4. Three-stage MHEMT feedback amplifier (a) and its V-band frequency response (b). The traces correspond to different circuits locations on the 2 inch wafer.

ble metallization flaws are very few. These similar values are indeed below the TEM detection limit and low enough to yield useful circuits. Division up to 31 GHz, with a sensitivity shown in Fig. 5(b), was measured on-wafer. Summarizing the device and circuit results, we have seen no reduction in performance, yield or uniformity by using the GaAs+LGLTB as a substrate for InP-type HEMTs.

4. Reliability, R_d-degradation and burnout

At low-noise bias, the reliability of the MHEMTs [15] is as good as that of similar HEMTs grown on InP [32]. Fig. 6 shows the MTTF Arrhenius plot for MHEMTs stressed at $V_d = 1$ V and $I_d = 200$ mA/mm, at channel temperatures of 200, 220 and 240°C. With a 15% drop in I_{dss} as the failure criterion, MTTF extrapolates to 7.5×10^6 h at a 125°C channel temperature. Stress for 3500 h at $T_{ch} = 220$°C shifts V_{th} by +65 mV (Fig. 3 in Ref. [15]), which is equivalent to 12 Å further Pt sintering. Comparing this result with those in Ref. [32], it appears that the reliability of our sintered Pt gate is as good as that of a refractory gate. A more serious reliability issue is the increase ΔR_d in drain resistance, previously observed on lattice-matched InP-based HEMTs [3,32]. At 220°C channel temperature even low-noise biasing causes a significant

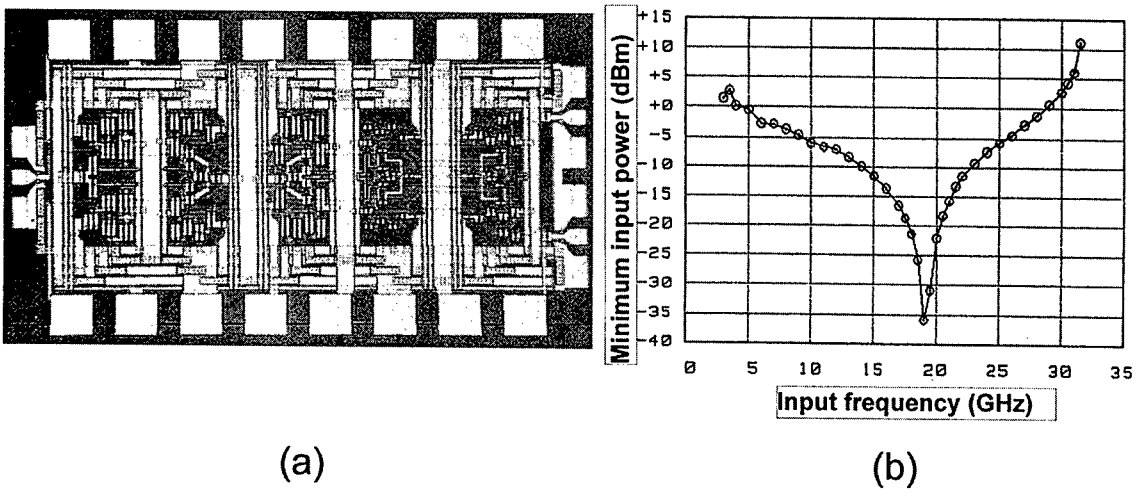


Fig. 5. Static MHEMT divide-by-4 circuit (a) and its input sensitivity curve (b).

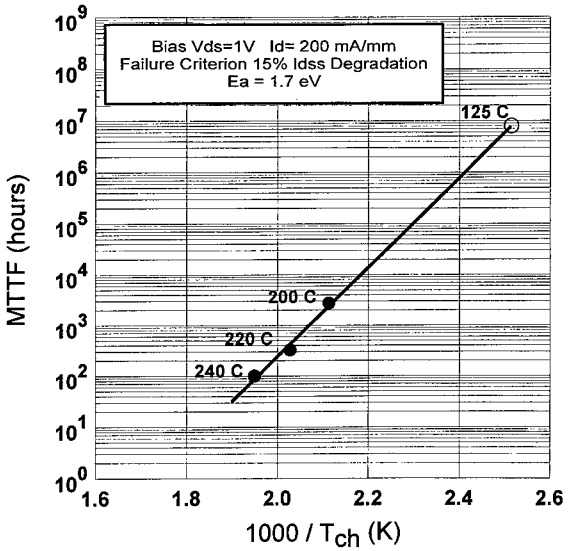


Fig. 6. MTTF Arrhenius plot for 0.1-μm MHEMTs. Solid circles are HTOL MTTF data for the failure criterion of 15% degradation in I_{dss}. The open circle is the extrapolated 7.5×10^6 h MTTF for 125°C channel temperature. With g_m peaking near $V_g = 0$ V, I_{dss} degrades more rapidly than $I_d^{(max)}$ [15].

(40–80%) ΔR_d (compared to 0–20% for ΔR_s) [15]. In our FETs at $V_d > 1.5$ V, R_d-degradation has been experimentally and theoretically correlated with impact ionization in the channel [33,34]. This is likely to be the case also at low-noise bias. While the impact ionization component in the gate current is typically too low to be discerned, impact ionization has been shown theoretically, in the context of modeling the kink phenomenon, to exist at $V_d < 1$ V [35]. Higher drain biases that are desirable for medium-power applications increase the impact ionization for both MHEMTs and standard HEMTs on InP. The *amount* of impact ionization will depend on the design of the channel and cap layers, the gate trough and passivation. The *sensitivity* of the device to impact ionization will depend on the mechanism responsible for the degradation, which is largely unknown. One can envision it to depend on general material quality, and the design of the buffer and Schottky barrier layers. The best reliability reported on AlInAs/GaInAs HEMTs on InP (at $V_d > 1$ V) [36] appears to have been achieved on a structure with depleted cap and Al-rich Schottky-barrier layer [37], both requiring alloyed ohmic contacts.

Even without considering R_d-degradaton and long-term reliability, impact ionization has an important limiting effect on these narrow-bandgap FETs, whether grown on GaAs or InP. It determines the on-state destructive breakdown (or burnout) voltage $BV_{ds}^{(on)}(I_d)$ and renders the off-state breakdown voltage $BV_{ds}^{(off)}$ essentially irrelevant [34]. This is illustrated in Fig. 7, where the experimental data are the burnout biases of 45 FETs on a single typical wafer. Each data point was generated by fixing the gate bias V_g and ramping up V_d until the device burned out. $BV_{ds}^{(on)}$ was recorded along with the associated drain current I_d. Varying V_g from device to device then generates the experimental dependence of $BV_{ds}^{(on)}$ on I_d. For low currents, this approaches the independently measured $BV_{ds}^{(off)}$ [38]. As a reference we superpose on the experimental data dotted curves corresponding to various power dissipations in the device. As described in Ref. [34] we also utilize a simple semi-analytical FET model for two purposes. First, we model the full-channel I_d versus V_d by fixing V_g at +0.75 V. Although not able to accurately predict the output conductance, the model predicts the measured larger burnout currents rather well. Second, we calculate the secondary drain current generated by impact ionization. This is done non-self-consistently by using the field dependent impact ionization coefficients [34]. We then locate the points in I_d–V_d space corresponding to a constant secondary current, and plot those with heavy solid lines in Fig. 7. These are different from those in Fig. 2 in Ref. [34], because of the use of more realistic impact ionization coefficients for the GaInAs channel [35,39]. The general shape of a locus of constant impact ionization is very different from that of constant power dissipation. In

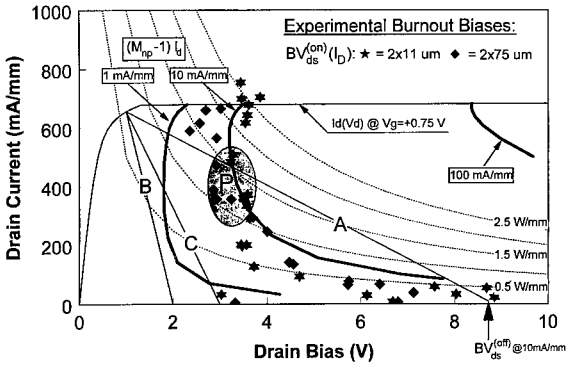


Fig. 7. Measured and modeled on-state breakdown voltage characteristics for 0.1-μm AlInAs/GaInAs/GaAs MHEMTs, modeled maximum current ($V_g = +0.75$ V), constant dc power loci (dotted lines), and three load-lines (A, B and C). The heavy solid lines are calculated loci of constant channel impact ionization current. Load-line A is based on the measured off-state breakdown voltage $BV_{ds}^{(off)}$ [38]. Load-lines B and C are for more realistic drain biases (1.5 and 2 V, respectively) which avoid the critical shaded area P of minimum on-state breakdown. The measured burnout data points were taken on FETs with 22 and 150 μm gate widths, on a representative wafer. V_g was held fixed for each measurement, but varied from FET to FET.

particular for larger currents, the former is much steeper, and actually bends over because of the lower fields in a full channel. Two very different FET sizes were used to be able to notice any differences attributable to heating effects [34]. Fig. 7 reveals that the experimental $BV_{ds}^{(on)}(I_d)$ data points for these $Ga_{0.47}In_{0.53}As$-channel FETs deviate markedly from a constant-power locus, which usually describes the similarly measured breakdown for GaAs MESFETs [40] and PHEMTs [41,42]. Instead, $BV_{ds}^{(on)}(I_d)$ follows a modeled locus for 2–12 mA/mm secondary drain current, independent of the primary drain current I_d. The impact ionization current corresponding to burnout in our present devices is then expected to be on the order of 1% of the full-channel current. The measured gate current at burnout varied between 1 and 6 mA/mm, suggesting that, for this device structure, a significant fraction of the holes generated by impact ionization in the high-field drain region of the channel are collected at the gate.

Fig. 7 also includes three potential load-lines for a power amplifier, all emanating from the knee voltage of the modeled full-channel $I_d(V_d)$ curve. The one marked A is the standard textbook load-line intersecting the V_d-axis at the off-state breakdown voltage $BV_{ds}^{(off)}$. Obviously, this is unacceptable for these FETs. Instead we have to move the load-line below the shaded area P in Fig. 7, with some margin for reliability, as illustrated by load-lines B and C. The area P corresponds to a flat minimum $BV_{ds}^{(on,min)}$ in the $BV_{ds}^{(on)}(I_d)$ curve. The slightly smaller $BV_{ds}^{(on)}$ for the larger FETs, at larger currents, may be a small thermal effect [34]. Since this will not affect the load-line we will instead focus on $BV_{ds}^{(on,min)}$ at P, which is near the point of maximum transconductance g_m and the $I_d^{(max)}/2$ class-A bias current. $BV_{ds}^{(on,min)}$ thus becomes a parameter worth focusing on for device structure optimization. Because of the flatness of the $BV_{ds}^{(on)}(I_d)$ minimum, the choice of I_d is not critical. For low I_d, $BV_{ds}^{(on)}$ increases because the primary drain current is low. For high I_d, as mentioned above, $BV_{ds}^{(on)}$ increases slightly because the channel field is reduced. $BV_{ds}^{(on,min)}$ is henceforth defined with V_g fixed at V_{gm} corresponding to maximum g_m at low drain bias ($V_d = 1$–1.5 V).

Fig. 8 shows a collection of measured $BV_{ds}^{(on,min)}$ plotted versus the drain current ($\approx I_d^{(max)}/2$) at which the device burns out, for a variety of MBE structures, unit processes, FET geometries, gate orientation and biasing schemes. Some of the variations in structure

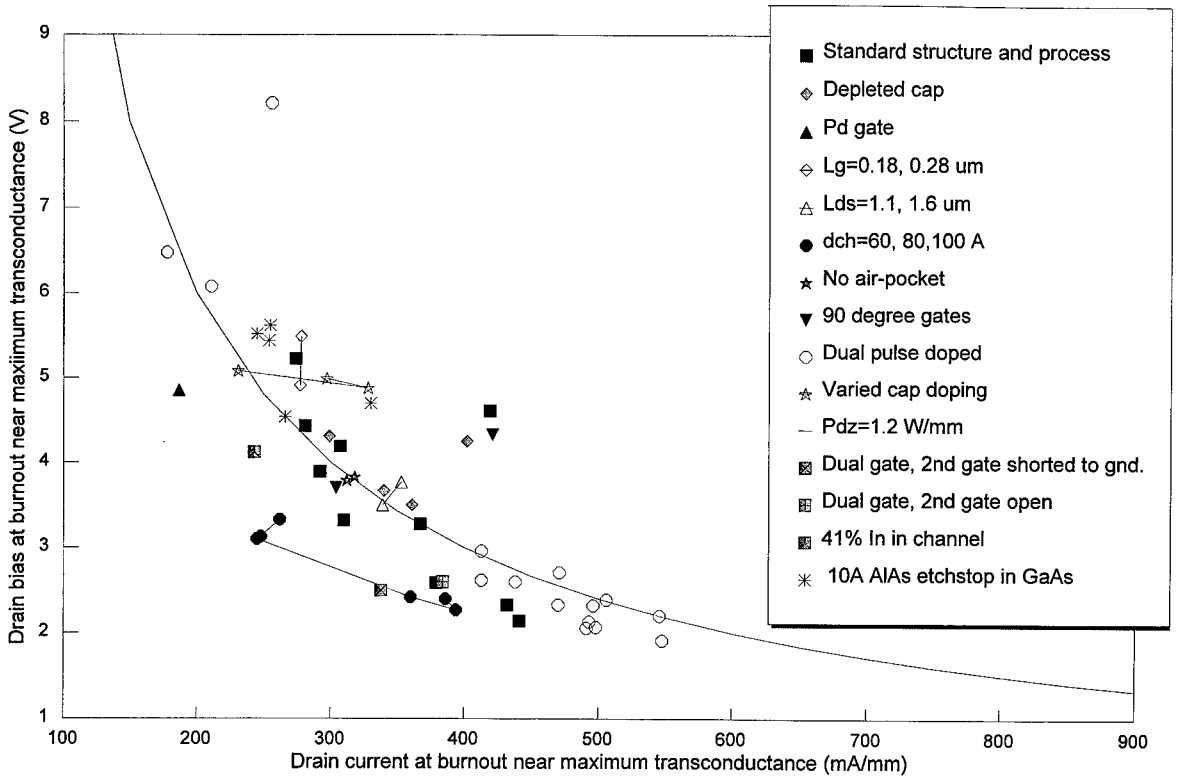


Fig. 8. On-state burnout points of various AlInAs/GaInAs/GaAs MHEMT wafers with the gate biased for maximum transconductance at $V_d = 1$–1.5 V. A variety of different FET structures and processing are plotted, as indicated in the legend. The solid line corresponds to $V_d I_d = 1.2$ W/mm.

and process were specifically made to improve the millimeter-wave output power that can be reliably delivered. The data follow a 1.2 W/mm $BV_{ds}^{(on,min)} \times I_d^{(max)}/2$ tradeoff curve, with a standard deviation of 0.3 W/mm. This is a very different trend than that in Fig. 7. Recall, however, the different parameters that are varied and looked at in the two cases. In Fig. 7, various burnout points were generated on a wafer by varying V_g and hence I_d, but obviously keeping the device design the same. In Fig. 8, on the other hand, we vary the device itself, fix V_g at V_{gm} for the particular device, then ramp up V_d until burnout occurs at $BV_{ds}^{(on,min)}$. The data in Fig. 7 established that the burnout is not a thermal effect. The apparent constant-power tradeoff in Fig. 8 is best considered a coincidence, until the physical origin can be established. Variations in the linear part of the parasitic access resistances R_s and R_d between the structures are too small to significantly affect the burnout voltage and current. Variations in the nonlinear part due to the unmetallized free-surface part of the gate trough, however, can have significant effect. In fact, variation in the lateral etch is one non-thermal physical mechanism that can produce a constant-power-like tradeoff [34]. This is unlikely, however, to explain more than a subset of the data points in Fig. 8, since, as discussed in Section 2, our recess is quite controlled. Some of the data on dual–doped FETs with low breakdown in Fig. 8 may be explained by this mechanism, since these FETs were marred by poor nitride adherence, which did significantly alter the shape of the etched trough. Some of the MBE structure variations in Fig. 8, such as the use of a depleted cap and thin quantized channels, may also be reconciled with a constant-power-like tradeoff, since they do tend to simultaneously produce lower currents and reduced impact ionization. Our thinner channel devices ($d_{ch} = 60$, 80 and 100 A), however, all fall rather significantly below the 1.2-W/mm tradeoff. So do our dual-gate FETs, where the lack of improvement is related to the redistribution of the channel field which reduces the current, but does not prevent a high-field region from ultimately appearing at one of the gates. For similar reasons the gate length does not strongly affect the on-state breakdown. The channel field is always concentrated at the drain-end of the gate, with the rest of the channel essentially acting as a relatively small low-field series resistance.

To summarize the observations in this section, Fig. 9 combines, qualitatively, the two trends in Figs. 7 and 8. Taking into account the finite knee-voltage, Fig. 9 implies that larger output power would be achieved with lower-current devices. The requirement for sufficient gain at the millimeter-wave frequencies of interest will severely limit the acceptable reduction in current, and alternative approaches to improved reliability at larger drain bias are therefore needed.

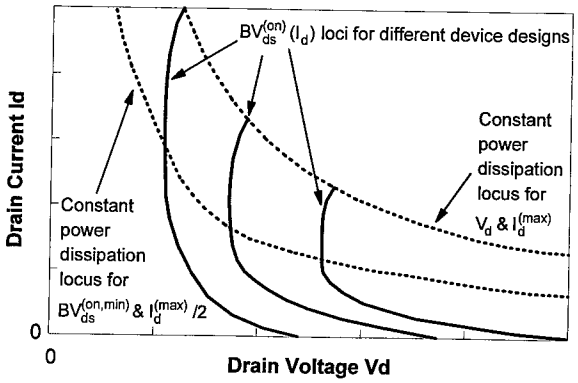


Fig. 9. Qualitative picture summarizing the lessons from Figs. 7 and 8. For different bias on similar FETs, the on-state breakdown $BV_{ds}^{(on)}(I_d)$ follows a locus of constant impact ionization (rather than one of constant power), with a minimum $BV_{ds}^{(on,min)}$ near maximum g_m and $I_d^{(max)}/2$. As the FET structure is varied, resulting in a different $I_d^{(max)}$, the $BV_{ds}^{(on)}(I_d)$-trace is stretched or compressed to fit under $I_d^{(max)}$ and tends to move laterally following a constant power tradeoff.

5. Conclusions

Low-cost rugged GaAs substrates can be used to fabricate InP-type HEMTs, with excellent performance, yield, uniformity, and reliability at low-noise-bias. The linearly-graded low-temperature buffer used to grade the lattice constant from GaAs to InP adds to the growth time, but subtracts substrate cost and wafer breakage. For us, the seamless fit of the process with existing GaAs IC manufacturing infrastructure is an overriding economic advantage of this technology. Non-alloyed ohmic contacts, etch stop layers and a one-step e-beam T-gate process also contribute to the manufacturability of these MHEMTs. The reliability issues at higher drain biases do not appear to be related to the GaAs substrate.

Acknowledgements

We gratefully acknowledge the contributions and support of many of our colleagues at Hewlett-Packard, including D. D'Avanzo, S. Bahl, G. Baldwin, D. Basile, D. Briscoe, P. Chye, B. Collins, A. Fischer-Colbrie, J. Gladstone, P. Gregory, S. Harris, W. Hargrave, D. Hornbuckle, K. Hur, T. Hwang, R. Jaeger, M. Kanemura, F. Kellert, C. Madden, D. Mars, R. Merchant, J. Miller, E. Par, G. Patterson, J. Perdomo, R. Powell, M. Pustorino, J. Raggio, D. Reed, D. Ritchey, D. Scherrer, K. Seaward, J. Seeger, L. Studebaker and R. Van Tuyl.

References

[1] Enoki T, Kobayashi T, Ishii Y. Tech Dig GaAs IC Symp. 1994, p. 337.

[2] Nguyen L, Le M, Delaney M, Lui M, Liu T, Brown J, Rhodes R, Thompson M, Hooper C. IEEE MTT-S Dig. 1993. p. 345.

[3] Sonoda T, Yamamoto Y, Hayafuji N, Yoshida H, Sasaki H, Kitano T, Takamiya S, Ostubo M. Solid-State Electron 1997;41:1621.

[4] D'Avanzo DC, Estreich DB, Taylor TW, Kondoh H, Orr J, Quach A, Wong EH, Yeats R, Makker VK. Tech Dig GaAs IC Symp. 1988. p. 317.

[5] Fischer-Colbrie A, Zhou GG, Hasnain G. Electron Mater Conf. 1993. Paper A4.

[6] Goldman RS, Chen J, Kavanagh KL, Weider HH, Robbins VM, Miller JN. Inst Phys Conf Ser 1995;141:313.

[7] Wakita A, Moll N, Fischer-Colbrie A, Stickle W. J Appl Phys 1990;68:2833.

[8] Wang JG, Hur KY, Studebaker LG, Keppeler BC, Quach AT. Tech Dig GaAs IC Symp. 1997. p. 74.

[9] Rohdin H, Nagy A. Tech Dig IEDM. 1992. p. 327.

[10] Lin BJF, Luechinger H, Kocot CP, Littau E, Stout C, McFarland B, Rohdin H, Kofol JS, Jaeger RP. Tech Dig GaAs IC Symp. 1988. p. 143.

[11] Seaward KL, Moll NJ, Stickle WF. J Vac Sci Technol B 1988;6:1645.

[12] Rohdin H, Nagy A, Fischer-Colbrie A, Kocot C, Jaeger R. Proc IEEE ISCAS. 1990. p. 2538.

[13] Rohdin H, Straznicky J, Jekat H, Nagy A, Fischer-Colbrie A, Mars DE, Jaeger R. IEEE J Solid-State Circ 1992;27:1353.

[14] Rohdin H, Nagy A, Robbins V, Su C-Y, Wakita AS, Seeger J, Hwang T, Chye P, Gregory PE, Bahl SR, Kellert FG, Studebaker LG, D'Avanzo DC, Johnsen S. Hewlett-Packard J 1998;49(1):37 [http:/www.hp.com/hpj/98feb/feb98a4.htm].

[15] Wakita AS, Rohdin H, Robbins VM, Moll N, Su C-Y, Nagy A, Basile DP. Proc InP and Relat Mater. 1998. p. 223.

[16] Krishnamoorthy V, Lin YW, Park RM. J Appl Phys 1992;72:1752.

[17] Rohdin H., Nagy A, Robbins V, Su C-Y, Madden C, Wakita A, Raggio J, Seeger J. Proc InP and Relat Mater. 1995. p. 73.

[18] Cappy A. Proc InP and Relat Mater. 1996. p. 3.

[19] Broekaert TPE, Fonstad CG. IEEE Trans Electron Devices 1992;39:533.

[20] Kuroda S, Harada N, Sasa S, Mimura T, Abe M. IEEE Electron Device Lett 1990;11:230.

[21] Matthews JW, Blakeslee AE. J Cryst Growth 1974;27:118.

[22] Pizani PS, Boschi TM, Lanciotti Jr. F, Groenen J, Carles R, Maigne P, Gendry M. Appl Phys Lett 1998;72:436.

[23] Wakita AS, Su CY, Rohdin H, Liu HY, Lee A, Seeger J, Robbins VM. J Vac Sci Technol B 1995;13:2725.

[24] Rohdin H, Moll N, Su C-Y, Lee G. IEEE Trans Electron Devices 1998;45:2407.

[25] Rohdin H, Moll N, Bratkovsky A, Su C-Y. Phys Rev B 1999;59:13102.

[26] Kumar V. J Phys Chem Solids 1975;36:535.

[27] Harada N, Kuroda S, Hikosaka K. IEICE Trans Electron 1992;E75-C:1165.

[28] Chen KJ, Enoki T, Maezawa K, Arai K, Yamamoto M. Proc InP and Relat Mater. 1995. p. 428.

[29] Perdomo J. Private communication.

[30] Scherrer D. Private communication.

[31] Madden CJ, Van Tuyl RL, Le MV, Nguyen LD. Dig Tech IEEE ISSCC. 1994. p. 178.

[32] Enoki T, Ito H, Ishii Y. Solid-State Electron 1997;41:1651.

[33] Wakita AS, Rohdin H, Su C-Y, Moll N, Nagy A, Robbins VM. Proc InP and Rel Mater. 1997. p. 376.

[34] Rohdin H, Su C-Y, Moll N, Wakita A, Nagy A, Robbins V, Kauffman M. Proc InP and Rel Mater. 1997. p. 357.

[35] Suemitsu T, Enoki T, Tomizawa M, Shigekawa N, Ishii Y. Proc InP and Rel Mater. 1997. p. 365.

[36] Nguyen LD, Le MV, Liu T, Lui M, Kaneko K, Holzman E, Delaney MJ. Solid-State Electron 1995;38:1575.

[37] Matloubian M, Brown AS, Nguyen LD, Melendes MA, Larson LE, Delaney MJ, Pence JE, Rhodes RA, Thompson MA, Henige JA. IEEE Electron Device Lett 1993;14:188.

[38] Bahl SR, del Alamo JA. IEEE Trans Electron Devices 1993;40:1558.

[39] Urquhart J, Robbins DJ, Taylor RI, Moseley AJ. Semicond Sci Technol 1990;5:789.

[40] Vashchenko VA, Kozlov NA, Sinkevitch VF, Martynov JB, Tager AS. Proc European Microwave Conf. 1995.

[41] Rohdin H, Studebaker L. unpublished.

[42] del Alamo JA, Somerville MH. Tech Dig GaAs IC Symp. 1998.

PERGAMON

Solid-State Electronics 43 (1999) 1655–1663

SOLID-STATE ELECTRONICS

Wafer bonding technology for optoelectronic integrated devices

Hiroshi Wada*, Hironori Sasaki, Takeshi Kamijoh

Optoelectronics Oki Laboratory, Real World Computing Partnership, c/o Opto-Electronics Laboratory, Oki Electric Industry Co. Ltd, 550-5 Higashiasakawa, Hachioji, Tokyo, 193, Japan

Abstract

Wafer bonding has been investigated as a key technology to integrate InP lasers on Si for optoelectronic integrated circuits. The bonding process has been optimized to allow the integration of InGaAsP/InP double-heterostructures (DHs) on Si with keeping the crystal qualities good enough to realize the lasers. As a result, room temperature continuous-wave (CW) operation of InP edge-emitting lasers has been achieved. In addition, as one of the building blocks to implement the optimal interconnections between Si LSIs, InP optical devices on Si integrated with the back-surface diffractive lenses have been demonstrated. A novel bonding process which allows an integration on structured wafers, such as Si LSI wafers, has also been proposed. The wafer bonding is thought to be a promising technique to implement optical interconnections between Si LSI chips. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Optical interconnections between Si LSI chips have been attracting increasing interest as a key technology to overcome communication bottlenecks in electrical interconnections of the future high-speed LSI system and/or to realize new functional LSI processors [1–3]. Fig. 1 schematically illustrates the proposed structure of the optically interconnected LSIs. InP lasers, preferably surface emitting lasers, are integrated on LSI chips and the optical signals are emitted through the Si substrate to another LSI chip. InP-based lasers are suitable for this application because Si is transparent at the emission wavelength. In addition, diffractive optical elements (DOEs) such as diffractive lenses and beam deflectors can be integrated on the back-surface of the Si wafer. In two-dimensional integration, the LSIs with lasers are mounted on a waveguide substrate such as glass or Si and are optically interconnected using DOEs fabricated on the Si back-surface and the waveguide substrates.

In order to realize such integrated structures, usually called as optoelectronic integrated circuits (OEICs), integrations of III-V optical devices on Si has been widely studied with several different approaches, including solder bonding [4], lattice-mismatched epitaxial growth [5,6] and wafer bonding [7,9]. Solder bonding with metals such as AuSn [4] is the most straightforward method, but it is not suitable for vertical optical interconnections (Fig. 1(a)) because the metals interrupt the vertical light propagation. Lattice-mismatched epitaxial growth has been widely investigated and room-temperature CW operation of the lasers on Si has been successfully demonstrated [5,6]. However, there are some problems in this growth technique, such as high dislocation densities ($> 5 \times 10^6$ cm^{-2}) and high growth temperatures (typically

* Corresponding author.

0038-1101/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0038-1101(99)00115-X

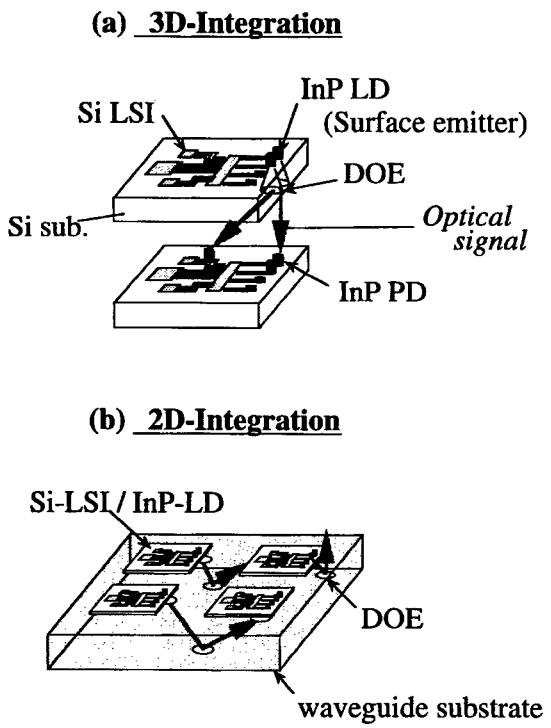


Fig. 1. Schematic pictures of proposed optically interconnected Si LSIs.

>900°C for thermal pre-cleaning of Si and >550°C for epitaxial growth). The high growth temperatures may cause a big problem in the device integration because the interconnection metals in LSIs, especially aluminum, will be degraded when the fully processed LSIs are exposed to the temperatures higher than 500°C [10,11].

Wafer bonding is another promising candidate. It is a technique to bond different wafers by contacting the flat surfaces face-to-face after appropriate surface cleaning without using any adhesives. Since the technique was first demonstrated in 1986 for Si–Si bonding [12] and the fabrication of silicon-on-insulator (SOI) structures [13], it has been extensively investigated mainly for Si-based materials, and has recently been applied to other materials including InP–GaAs [14–16], GaAs–Si [7,17], InP–Si [8,9,18,19], GaAs–$LiNbO_3$ [20] and $LiNbO_3$–$LiTaO_3$ [21]. With the wafer bonding technique, it is possible to bond wafers at low temperatures (even at room temperature) through Van der Waals force [7] and/or hydrogen bonding between surface OH-groups [12]. Therefore, we can integrate InP/InGaAsP materials on fully processed LSI wafers without degradation of the LSI characteristics. Additionally, optical devices can be fabricated using photolithography after wafer bonding because the bonds can be strong enough to endure following device fabrication processes. This will facilitate alignment of large number of small optical devices with Si circuits, resulting in a potentially inexpensive integrated system.

In this paper, we report on the procedures and properties of InP–Si bonding and the fabrication of InGaAsP/InP optical devices on Si. In Section 2, we review the bonding process and discuss the optimization of the process to realize the InP lasers with device characteristics as good as the conventional lasers fabricated on lattice-matched substrates. Section 3 describes the fabrication of the InP light emitting diodes (LEDs) on Si integrated with back-surface diffractive lenses, which will be one of the basic building blocks of the optical interconnections. In Section 4, a novel bonding process is proposed which allows the integration of optical devices on structured wafers such as actual LSI wafers, and conclusions are presented in Section 5.

2. Wafer bonding process

2.1. Bonding process and mechanism

The bonding process starts with chemical cleaning of the wafer surfaces with H_2SO_4:H_2O_2:H_2O solution, followed by rinsing in de-ionized (DI) water. After this cleaning, the surfaces show hydrophilic characteristics, which are reported to be preferable for achieving strong bonds at relatively low temperatures [22,23]. The wafers are then spin dried and their surfaces are placed in contact at room temperature under a pressure of 4 kg/cm^2 applied on the wafers to ensure the surface contact. The wafers adhered to each other at this stage, though the adhesion is not very strong. Finally, the wafers are annealed at elevated temperatures typically for 30 min in an H_2 atmosphere to increase the bonding strength.

The bonding mechanism has been extensively discussed for Si–Si wafer bonding [12,24] and a similar mechanism is also thought to be responsible for the InP–Si bonding. The proposed mechanism is schematically summarized in Fig. 2. First, the OH-groups are absorbed on the wafer surface during the cleaning process with H_2SO_4:H_2O_2:H_2O (Fig. 2(a)), which is thought to be an origin of the hydrophilicity. When the cleaned surfaces are brought into contact at room temperature, the wafers adhere to each other through the hydrogen bonding between the OH-groups (Fig. 2(b)). In the following heat treatment at moderate temperatures (200°C < T < 400°C), water molecules (H_2O) start to evaporate and escape from the interface and the hydrogen bonds are replaced by bonds like InP–O–Si (Fig. 2(c)). This bond is usually stronger than the hydrogen bond, which results in increased strength at the higher temperatures. When the wafers are annealed

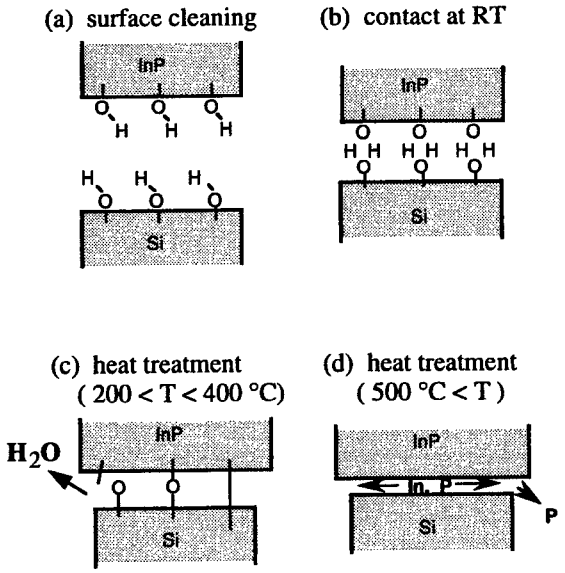


Fig. 2. Proposed bonding mechanism.

at even higher temperatures ($T > 500$°C), evaporation of phosphorus from InP and the migration of indium atoms become pronounced which can cause atomic rearrangement at the interface (Fig. 2(d)) and the two wafers can be atomically bonded to form stronger bonds [8,14,15].

A unique feature in the case of InP–Si bonding is that the thermal expansion coefficients are different between the materials and it causes a thermal stress which may produce cracks or even separate the wafers during annealing under some conditions [9,25]. Therefore, the actual bonding mechanism could be more complicated than that described above. In addition, the bonding properties may also depend on the surface morphology and the wafer size and thickness, which makes full understanding of the mechanism more difficult. Further investigations are needed to understand the mechanism more clearly.

2.2. Optimization of bonding temperature

As suggested by the model described above, the bonding temperature is a very important parameter. Therefore, first of all, we have optimized the bonding temperature, In this work, it has been optimized in terms of (i) quality of the bonded InP crystals, (ii) bonding strength and (iii) compatibility with the device processes.

The higher limit of the bonding temperature is imposed by the degradation of InP crystal quality and Si LSI circuits. The quality of the bonded InP materials has been evaluated by measuring photoluminescence (PL) intensity. InP/InGaAsP ($\lambda_{gap} = 1.3$ μm) double-heterostructure (DH) films were bonded on Si at several temperatures and the PL intensities from the InGaAsP active layers before and after the heat treatment were measured and compared. The thickness of the InP cladding layer which determines the distance from the bonded interface to the active layer is 1.5 μm. The PL was measured through the Si substrates using a 1.2-μm wavelength laser diode as the pump source. The measured PL peak intensity is shown in Fig. 3, normalized to the value obtained before the heat treatment. At temperatures above 600°C, the intensity is found to decrease drastically. This indicates that the treatment temperature must be lower than 550°C to maintain a good crystal quality. When DH wafers were annealed at 700°C without bonding, the PL intensity exhibited no significant change. Therefore, the loss of intensity is not due to the heat treatment itself, but originates in the bonding process and could be related to the threading dislocations caused by thermal stress introduced during the heat treatment. Actually, in the cross-sectional transmission electron microscope (TEM) observation of the sample bonded at 700°C, threading dislocations were clearly observed in the InP crystals near the bonded interface [26].

Additionally, bonding below 500°C is preferable in terms of compatibility with Si LSI circuits, as the previously patterned Al interconnects in LSIs may be degraded at above 500°C [10,11].

On the other hand, the lower temperature limit is imposed by the laser fabrication process following the wafer bonding. The formation of ohmic contacts on the InP lasers requires the annealing of the metals at about 400°C, which will be described later again. Additionally, higher bonding temperatures are preferable in terms of bonding strength as the bonds usually strengthen at higher temperatures. Fig. 4 shows the measured dependence of the bonding strength on the

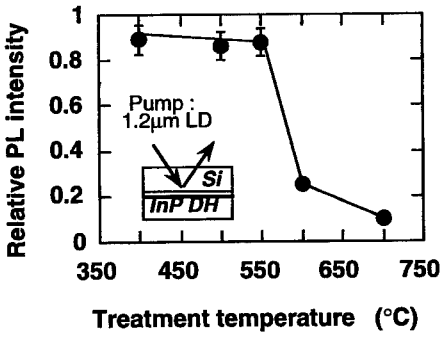


Fig. 3. Peak photoluminescence intensity from InP/InGaAsP DHs bonded on Si at different temperatures. The active layer was optically pumped through the Si substrate by a 1.2 μm wavelength laser and the photoluminescence emitted through the Si was measured.

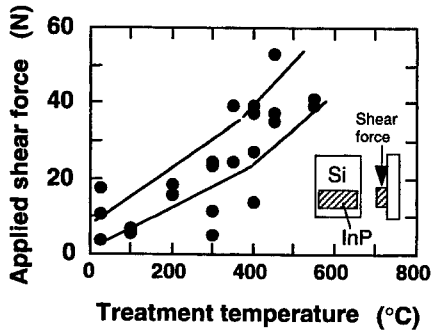


Fig. 4. Dependence of bonding strength on treatment temperature.

treatment temperature. In this measurement, 350-μm-thick InP substrates, 5 × 10 mm^2, were bonded on larger Si substrates and the strength was evaluated from a shear force required to separate the wafers. As shown in Fig. 4, the bonding strength basically increases as the temperature rises, which is consistent with the mechanism explained above, although large scattering of the data is observed. We speculate the data scattering to be mainly due to the unexpected formation of voids at the bonded interface. Incompleteness of the measurement technique, such as nonuniform distribution of the applied shear force, may be an additional reason for the data scattering. Because stronger bonding will facilitate the following device fabrication, a higher temperature is also preferable from this point of view.

On consideration of these factors, the bonding temperature has been determined at 400°C in this experiment.

2.3. Fabrication of InP edge-emitting lasers on Si

In order to examine if the bonding process is really applicable to the integration of the lasers on Si, we first fabricated the InGaAsP/InP edge-emitting lasers on Si. The laser fabrication process is summarized in Fig. 5. First of all, a DH structure was grown on a p-type (100) InP substrate by metalorganic vapor phase epitaxy (MOVPE) and was composed of a 0.2-μm InGaAs etch-stop layer, a 1.5-μm p-InP cladding layer, a 0.15-μm InGaAsP bulk active layer (λ_{gap} = 1.3 μm), a 15-μm n-InP cladding layer and a 0.1-μm n-InGaAs cap layer. The epitaxial wafer, with a typical size of 10 × 10 mm^2, was then stuck on a glass plate with wax for mechanical support (Fig. 5(a)). Next, the p-InP substrate and the InGaAs etch-stop layer were selectively etched to leave a thin DH film on a glass plate. The exposed surface of the p-InP cladding layer was then cleaned with N_2SO_4:H_2O_2:H_2O. The surface of a Si substrate was also cleaned with H_2SO_4:H_2O_2:H_2O. After a DI water rinse, the wafers were spin dried and their surfaces were placed in contact under a pressure of 4 kg/cm^2 at room temperature (Fig. 5(b)). The wafers adhered to each other at this stage through the hydrogen bonding between the surface OH-groups.

Then, the bonded wafers were immersed in warm organic solvent to dissolve the wax from the side completely and detach the glass plate from the wafers. The wafers still remained bonded after this wax dissolving. The sample was then annealed at 400°C for 30 min in an H_2 atmosphere (Fig. 5(c)). After this annealing step, the crystal quality of the bonded DH film was evaluated by measuring PL intensity, X-ray rocking curve and etch-pit density (EPD). All the measured results indicated that the quality is good enough to realize the lasers on Si—the PL intensity as high as 90% of that before bonding, the full-width-at-half-

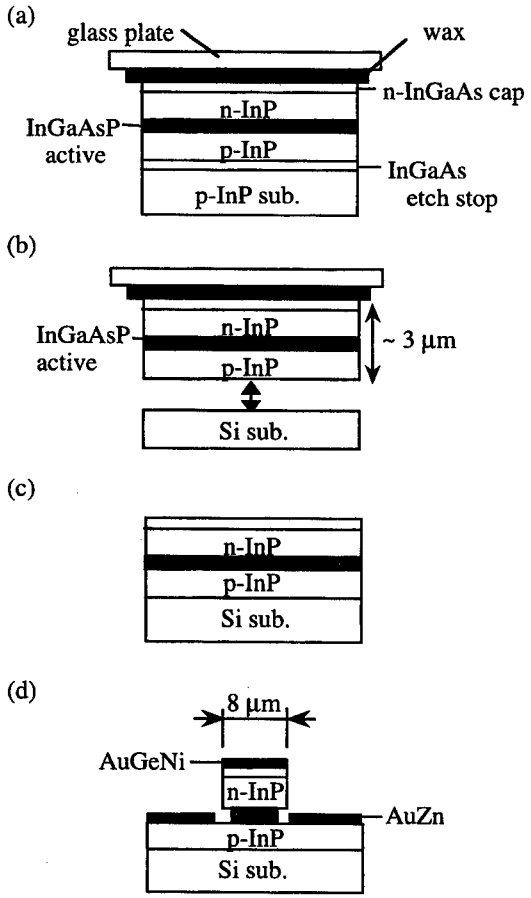


Fig. 5. Schematic diagram of the edge-emitting laser fabrication process: (a) DH wafer growth and sticking of the wafer on a glass plate for mechanical support, (b) substrate etching, surface cleaning and contact at room temperature, (c) heat treatment at 400°C after detachment of the glass plate and (d) mesa stripe and ohmic contact formation.

maximum (FWHM) of the (400) X-ray rocking curve as narrow as 35 arcsec, and the EPD of 2×10^4 cm^{-2}.

After bonding of the thin film, 8-μm-wide mesa stripes were formed by standard photolithography and wet etching. The active layer was slightly undercut to reduce the active width to 6 μm. Finally, AuGeNi and AuZn were evaporated and alloyed at 400°C to form n- and p-type ohmic contacts to complete the laser fabrication (Fig. 5(d)).

In this fabrication process, the 350-μm-thick InP substrate was etched away before the annealing step. When we tried to etch the substrate after annealing, the InP wafer often cracked during substrate etching, probably due to nonuniform redistribution of the thermal strain, while it hardly occurred when the wafer was bonded after substrate removal. Moreover, the flexibility of the thin film may allow it to conform to the undulations of the wafer surface, as proposed in the epitaxial liftoff process [7]. Therefore, thin film bonding is preferable and actually produces a higher yield than thick wafer bonding.

The yield of the bonding process seems to depend critically on the surface quality and preparation. For example, dust particles trapped at the interface during the bonding process lowers the yield greatly. Therefore, in this experiment, the wafers were dried and brought into contact on a clean bench with 0.3-μm particles of less than 30 counts/ft^3. Moreover, relatively small wafers (typically 10×10 mm^2) were used, to reduce the chances of trapping the particles between the wafers. However, the process yield was still much lower than that of conventional laser process. Other factors, such as impurities which vaporize at elevated temperatures, surface undulation and warp, can also cause void formation at the interface, which further degrades the yield. Therefore, more careful control of the wafer surface will be required to achieve a yield high enough for manufacturing, especially when larger wafers are required to be bonded.

After thinning the Si substrate down to about 100 μm, the lasers were cleaved into 300-μm-long cavities and mounted on Si heat sinks with junction-up configuration. For comparison, the same laser structure was fabricated on a p-InP substrate. The light–current (L–I) characteristics of the devices were measured at room temperature (RT) without any facet coatings. Room temperature CW operation has been achieved with these devices as shown in Fig. 6. In this figure, the L–I curves of three lasers on Si are shown, together with those of the conventional lasers on InP for comparison. The kinks in the L–I curves are due to multiple transverse-mode operation of the lasers with 6-μm-wide active regions. The threshold currents are about the same (I_{th} = 39 mA) for both the structures, which again demonstrates the high crystal quality of the bonded DH films. Moreover, maximum

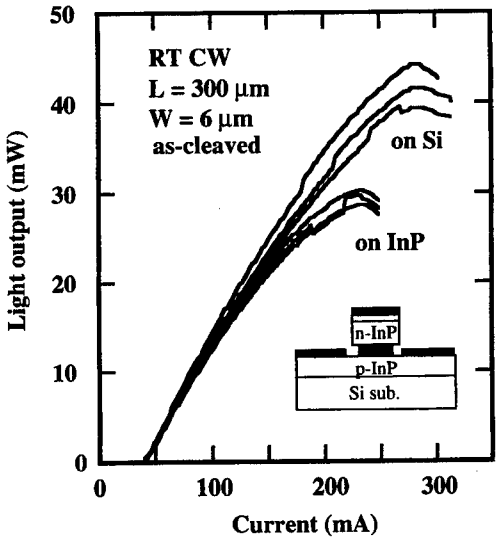


Fig. 6. L–I curves of the lasers fabricated on Si under room-temperature CW operation, together with those of lasers on InP for comparison. Higher output powers are obtained with lasers on Si.

output powers of lasers on Si are higher than those of conventional lasers on InP. This can be attributed to the lower thermal conductivity of the Si substrates, which is one of the advantages of fabricating lasers on Si. These results indicate that the wafer bonding is a promising technique to integrate high-quality InP lasers on Si wafers.

3. InP LEDs on Si integrated with back-surface diffractive lenses

3.1. Fabrication

In order to implement the optical interconnections between Si LSIs illustrated in Fig. 1, InP surface-emitting lasers (SELs) are required to be integrated on the LSI wafers. We are thus trying to fabricate the InP SELs on Si by wafer bonding and have so far succeeded to achieve room temperature lasing operation by photo-pumping [27]. However, it is still very difficult to achieve lasing by current-injection in the InP SELs because of complexivity in the device design [28]. Therefore, in this work, we have fabricated InP LEDs, instead of SELs, and demonstrated the integration of diffractive lenses on the back-surface of the Si wafers, which will be one of the basic building blocks in the interchip and/or intrachip optical interconnections.

The fabrication process is basically the same as that for edge-emitting lasers (Fig. 5) and briefly summarized in Fig. 7. First, the InGaAsP/InP double-heterostruc-

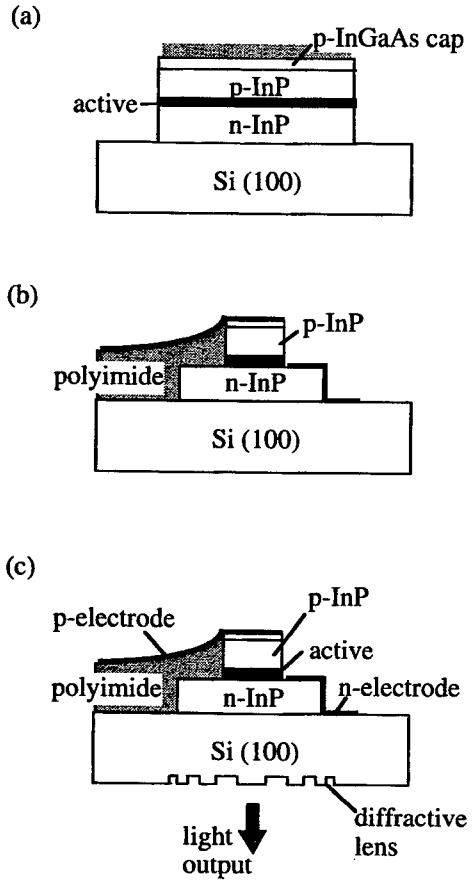


Fig. 7. Schematic fabrication process of the InP LEDs with back-surface diffractive lenses: (a) bonding of a DH thin film on a Si substrate, (b) fabrication of LED structures and (c) integration of diffractive lenses on the back-surface of the Si wafer.

ture (DH) was grown on an n-InP substrate by MOVPE and the thin DH films were directly bonded on Si substrates (Fig. 7(a)). The thickness of the Si substrate used in this work was 300 μm. After bonding, the active region of the LED was defined by etching the p-InP cladding and active layers to form mesa structure 20-μm in diameter. P- and n-type ohmic contacts were formed on InP with using polyimide as an insulator (Fig. 7(b)). Finally, the binary diffractive lenses were fabricated on the back-surface of the Si substrates by photolithography using double-view mask aligner and reactive ion etching of Si (Fig. 7(c)).

Fig. 8 shows (a) the picture of the fabricated pattern of the back-surface diffractive lens and (b) simultaneous view of the front LED and the back lens patterns observed with a double-view micro-scope. As shown in Fig. 8(b), the LED and lens patterns are well-aligned with an alignment error typically less than 1.5 μm.

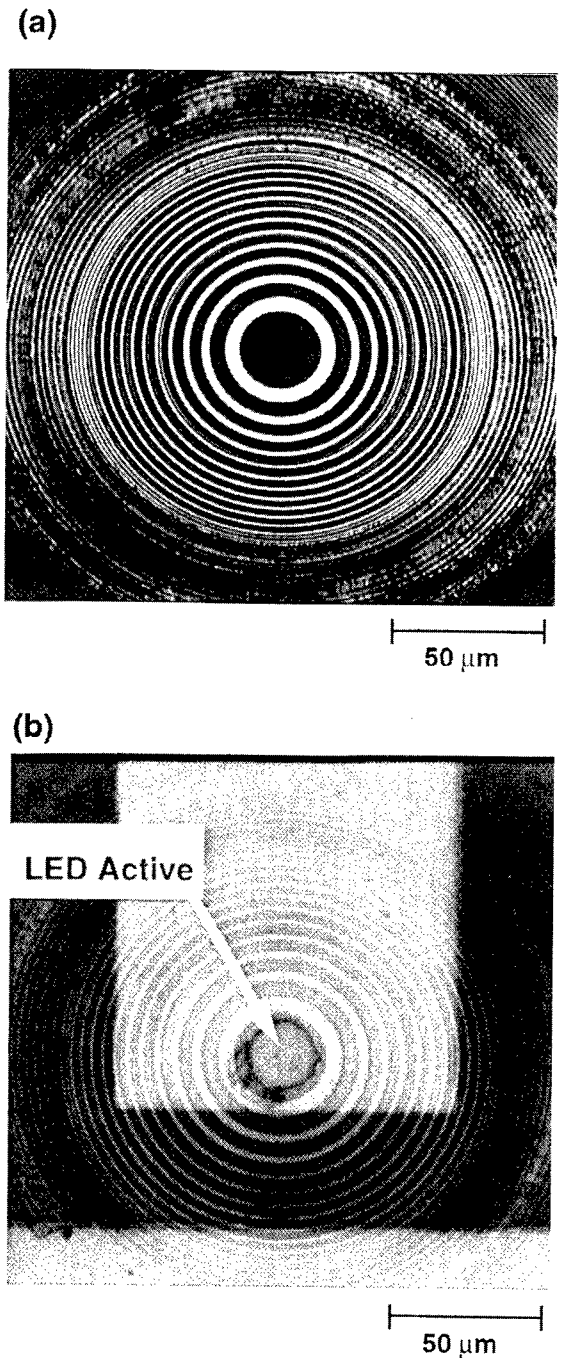


Fig. 8. Pictures of the LEDs with diffractive lenses: (a) patterns of the back-surface diffractive lens and (b) simultaneous view of the front LED and back lens patterns observed with double-view microscope.

3.2. Characteristics

The output power was measured under room temperature CW operation using large-area (10 mm in di-

ameter) Ge photodetector (PD). The total power emitted through the substrate without the diffractive lens was measured by placing the Ge detector as close to the LEDs as possible, as shown in Fig. 9(a). The light output higher than 200 μW was obtained. Fig. 9(b) and (c) show the power from the LEDs (b) without and (c) with the diffractive lens, which were detected by the PD placed at some distance away from the LEDs so as to cover the solid angle of only 20°, as illustrated in Fig. 7(b). Without the diffractive lens (Fig 7(b)), the detected power was only one tenth of the total power, which agrees well with the theoretical prediction. On the other hand, with the lens (Fig. 7(c)), about 35% of the total power was detected by the PD with the same configuration as Fig. 7(b), which indicates the LED light is collimated by the back-surface diffractive lens. Because the lens fabricated in this work is binary (two-phase) lens, the maximum diffraction efficiency is theoretically limited at about 40%. Higher collimated power will be available by using multi-level diffractive lenses, e.g. 80% with four-phase-level and 95% with eight-phase-level diffractive lenses.

Fig. 10 shows the beam size measured as a function of a distance from the diffractive lens. Although the beam is slightly divergent (with a divergent angle of about 4°), which is attributed to that the Si substrate used in this work was 20% thinner than the designed thickness, it is obvious that the directivity of the LED light is significantly improved by the diffractive lens. These results show the InP LEDs integrated with the back-surface diffractive lenses are promising devices for optical interconnects between Si LSI chips.

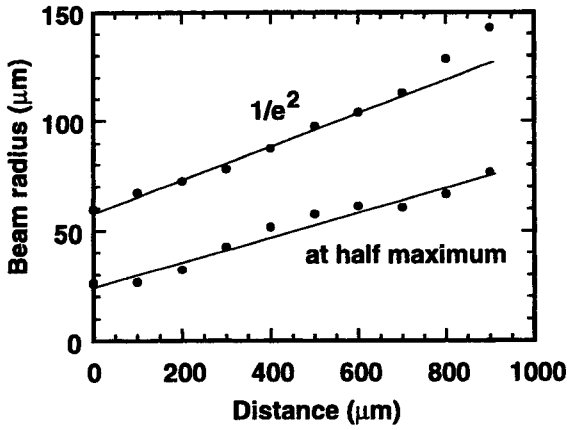


Fig. 10. Beam size measured as a function of the distance from the diffractive lens.

4. Selective-area wafer bonding

In the previous sections, we have shown that high-quality InP devices can be fabricated on pure Si substrates. However, actual Si LSI wafers, on which the devices are required to be integrated, have many structures and steps on their surfaces. In this section, we propose a bonding process which allows the integration on such structured surfaces.

Fig. 11 illustrates the proposed integration process. First, Si LSI wafers are designed and fabricated to have some vacant spaces allocated for optical devices, as shown in Fig. 11(a). Concurrently, InP DH thin

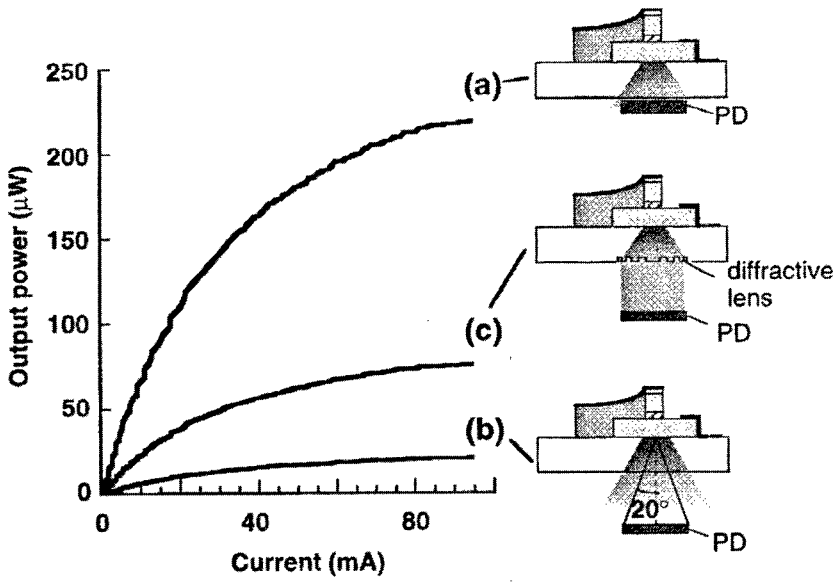


Fig. 9. Output power vs current characteristics of the fabricated LEDs with and without the back-surface diffractive lenses.

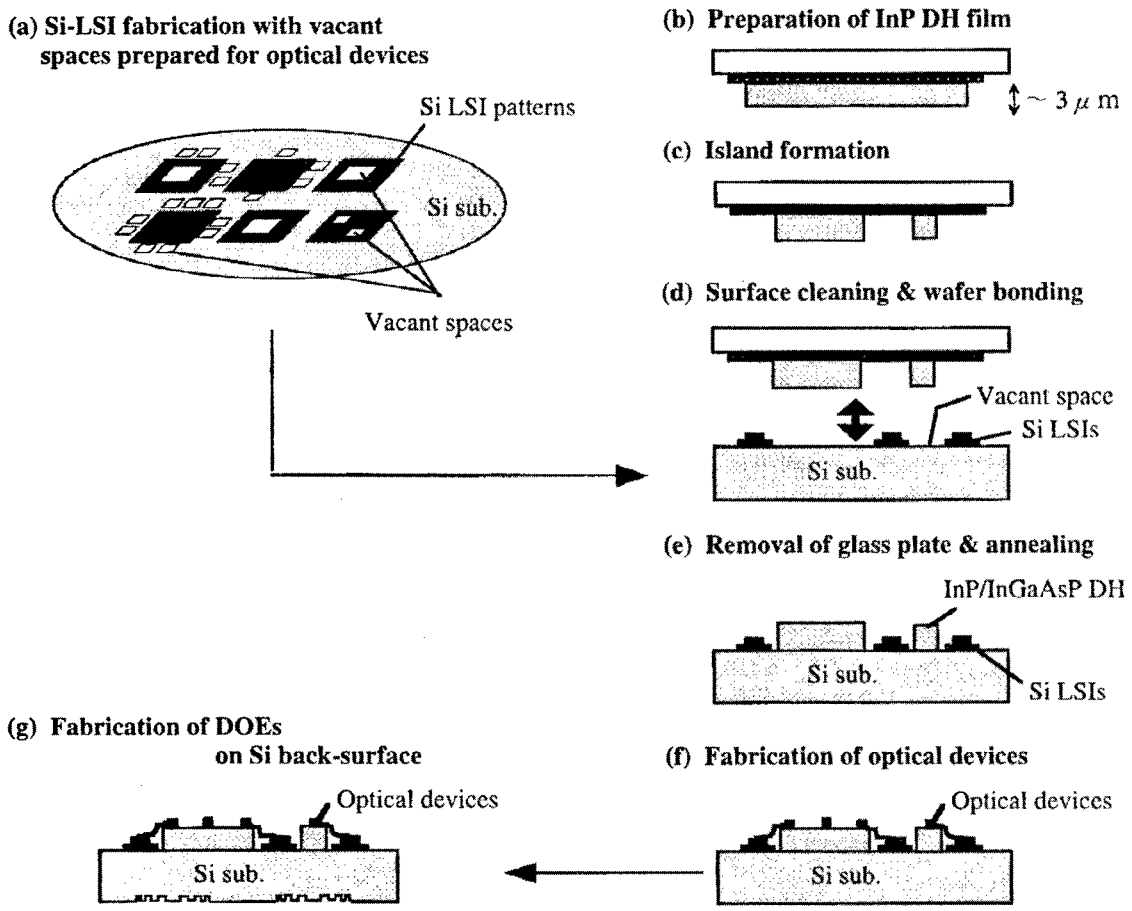


Fig. 11. Schematic illustration of the process which allows the integration of optical devices on Si LSI wafers.

films are prepared as described previously (Fig. 11(b)) and the islands are formed in the thin films by photolithography and etching with the same pattern as the vacant spaces on the LSI wafers (Fig. 11(c)). The surfaces of both the wafers are then cleaned and the islands are aligned to the vacant spaces and directly bonded at room temperature (Fig. 11(d)). After removing the glass plate and annealing at 400°C (Fig. 11(e)), optical devices are fabricated in the DH islands by lithography and the interconnection metals are deposited and patterned (Fig. 11(f)). Since small optical devices can be defined at this step, the DH islands and the vacant spaces may be much bigger than the optical devices, which greatly facilitates the alignment at bonding in the step of Fig. 11(d). Finally, the back-surface optical elements, such as diffractive lenses and beam deflectors, can be fabricated on the back-surface of the Si wafers (Fig. 11(g)).

By using this process, we can bond to DH islands even if there are LSI patterns on the Si wafers and align the optical devices to the LSI patterns by standard photolithography. Therefore, we can integrate large number of small optical devices well-aligned to the LSI patterns, which will allow the inexpensive optoelectronic integrated circuits. Actual design of the Si LSIs with some vacant spaces (Fig. 11(a)) and integration of the optical devices on the LSIs are now under way.

5. Conclusion

Wafer bonding of InP to Si has been investigated as a key technology to integrate the optical devices on Si for interchip and/or intrachip optical interconnections. The bonding process has been optimized to allow the integration of the InGaAsP/InP DH films on Si with keeping crystal qualities good enough to make the lasers. As a result, InP edge-emitting lasers have been successfully fabricated on Si and room temperature CW operation has been achieved. In addition, InP LEDS integrated with back-surface diffractive lenses have been fabricated on Si as one of the basic building blocks in the optical interconnections. A novel bonding

process which allows the integration on actual LSI wafers has also been proposed.

Finally, it would be worthwhile to summarize the advantages of the wafer bonding over other bonding techniques, such as bonding with adhesives and flip-chip bonding. Bonding with adhesives (e.g. polyimide) has been developed and GaAs SELs were successfully integrated on Si chips [29]. However, the high thermal resistance of the adhesives causes the undesirable temperature increase of the optical devices, which will deteriorate the device performance. On the contrary, as discussed in Fig. 6, the lasers fabricated on Si with the wafer bonding technique has shown higher output powers than the conventional lasers on InP, which indicates a good heat dissipation through the bonded interface. Although metals or solders may be other candidates as adhesives with lower thermal resistance, it is opaque and not applicable to vertical optical interconnections between stacked LSIs (Fig. 1(a)). Flip-chip bonding is another promising integration technique. However, when a large number of small optical devices are required to be integrated, alignment at flip-chip bonding may become stringent and expensive. In the wafer bonding, small devices can be fabricated with conventional photolithography after grafting the InP-based materials with larger area (Fig. 11(f)), which will make the alignment very easy and inexpensive. Therefore, we believe the wafer bonding described in this paper has advantages over other bonding techniques and is very promising to realize optical interconnection between LSI chips and other optoelectronic integrations.

References

[1] Goodman JW, Loenberger FI, Kung S-Y, Athale RA. Proc IEEE 1984;72:850.

[2] Koyanagi M, Takata H, Mori H, Iba J. IEEE J Solid-State Circuits 1990;SC-25:109.

[3] Hayashi I. Jpn J Appl Phys 1993;32:266.

[4] Shieh CL, Chi JY, Armiento CA, Haugsjaa PO, Negri A, Wang WI. Electron Lett 1991;27:850.

[5] Razeghi M, Defour M, Blondeau R, Omnes F, Maurel P, Acher O, Brillouet F, C-Fan JC, Salerno J. Appl Phys Lett 1988;53:2389.

[6] Sugo M, Mori H, Tachikawa M, Itoh Y. Appl Phys Lett 1990;57:593.

[7] Yablonovitch E, Hwang DM, Gmitter TJ, Florez LT, Harbison JP. Appl Phys Lett 1990;56:2419.

[8] Lo YH, Bhat R, Hwang DM, Chua C, Lin C-H. Appl Phys Lett 1993;62:1038.

[9] Wada H, Kamijoh T. IEEE Photon Technol Lett 1996;8:173.

[10] Tewkesbury ST, Hornak LA, Nariman HE, McGinnis SP. IEEE Trans Comp Hybrids Manuf Technol 1993;7:674.

[11] Shenoy KV, Fonstad CG, Mikkelson JM. IEEE Electron Dev Lett 1994;15:106.

[12] Shimbo M, Furukawa K, Fukuda K, Tanzawa K. J Appl Phys 1986;60:2987.

[13] Lasky JB. Appl Phys Lett 1986;48:78.

[14] Lo YH, Bhat R, Hwang DM, Koza MA, Lee TP. Appl Phys Lett 1991;58:1961.

[15] Wada H, Ogawa Y, Kamijoh T. Appl Phys Lett 1993;62:738.

[16] Babic DI, Streubel K, Mirin RP, Margalit NM, Bowers JE, Hu EL, Mars DE, Yang L, Carey K. IEEE Photon Technol Lett 1995;11:1225.

[17] Ersen A, Schnitzer I, Yablonovich E, Gmitter T. Solid-State Electron 1993;36:1731.

[18] Ejeckam FE, Chua CL, Zhu ZH, Lo YH, Hong M, Bhat R. Appl Phys Lett 1995;67:3936.

[19] Hawkins AR, Reynolds TE, England DR, Babic DI, Mondry MJ, Streubel K, Bowers JE. Appl Phys Lett 1996;68:3692.

[20] Chan WK, Yi-Yan A, Gmitter TJ, Florez LT, Jackel JL, Hwang DM, Yablonovitch E, Bhat R, Harbison JP. IEEE Photon Technol Lett 1990;2:194.

[21] Eda K, Sugimoto M, Tomita Y. Appl Phys Lett 1995;66:827.

[22] Spierings GACM, Haisma J. In: Proceedings of First International Symposium on Semiconductor Wafer Bonding: Science, Technology and Applications (Electrochemical Society, Pennington, NJ), Vol. 92-97. 1992. p. 18.

[23] Tong Q-Y, Schmidt E, Gosele U, Reiche M. Appl Phys Lett 1994;64:625.

[24] Bengtsson S. J Electron Mater 1992;21:841.

[25] Lehmann V, Mitani K, Stengl R, Mii T, Gösele U. Jpn J Appl Phys 1989;28:L2141.

[26] Wada H, Kamijoh T. IEEE J Selct Topics in Quantum Electron 1997;3:937.

[27] Wada H, Takamori T, Kamijoh T. IEEE Photon Technol Lett 1996;8:1426.

[28] Iga K, Koyama F, Kinoshita S. IEEE J Quantum Electron 1988;24:1845.

[29] Matsuo S, Tateno K, Kurokawa T. Technical Digest on Spatial Light Modulators in 1997 OSA Topical Meetings. Paper SMB2-1, p. 19.

AUTHOR INDEX

Adesida I., 1333
Agarwal A. K., 1459
Agarwal B., 1489
Asbeck P. M., 1555
Awano Y., 1513

Balakrishna V., 1459
Bayraktaroglu B., 1429
Berg M., 1607
Bito Y., 1419
Brandt C. D., 1459
Brar B., 1355
Broekaert T., 1355
Burk Jr. A. A., 1459
Bykhovski A., 1451

Camiade M., 1599
Chau F. H.-F., 1399
Chen Z., 1399
Chung H. Y., 1535
Clarke R. C., 1459
Comparini M. C., 1577
Cueva G., 1333

D'Avanzo D., 1437
Deufel R., 1607
Dickmann J., 1607
Dourlens C., 1599
Driad R., 1445

Emura K., 1613
Enoki T., 1527
Essilfie G., 1437

Fay P., 1333
Feudale M., 1577
Frazier G., 1355
Fujikura H., 1541
Fujimoto K., 1413
Furukawa H., 1405
Furuya K., 1395

Gaska R., 1451
Glück M., 1383
Grider D. E., 1473
Guthrie J., 1489

Hackbarth T., 1383
Harada N., 1513
Hasegawa H., 1483, 1541, 1561
Hashimoto S., 1571
Hashimoto T., 1389
Hashizume T., 1483
Hau G., 1419
Higgins J. A., 1633
Hikosaka K., 1513
Hirama A., 1541
Höck G., 1383
Hokazono A., 1465
Hongo S., 1347
Hsu S. S. H., 1429
Hutchinson C., 1437
Hwang J. C. M., 1325

Iiyama K., 1571
Inokuma T., 1571
Ishii Y., 1527
Ishikawa O., 1413
Iwata N., 1419

Jackson G., 1399
Jaganathan S., 1489

Kamijoh T., 1655
Kasahara K., 1339
Kasai Y., 1571
Katsui T., 1571
Kawarada H., 1465
Kawashima K., 1413
Kimura T., 1519
Kishimoto S., 1547
Kitatani K., 1465
Kohn E., 1535, 1547
König U., 1383
Koyama Y., 1483
Kudo Y., 1413
Kunihiro K., 1339
Kunze M., 1535

Laframboise S. R., 1445
Lee C. P., 1399
Lee L. H., 1535
Lee Q., 1489
Lin B., 1399
Lipka K. M., 1547
Long S., 1489
Low T., 1437
Lu Z. H., 1445
Lunardi L., 1627

MacMillan M. F., 1459
Maezawa K., 1547
Mahajan A., 1333
Martin S. C., 1489
Masato H., 1413
Mathew T., 1489
Matsunami K., 1547
McAlister S. P., 1445
McDermott B. T., 1633
McKinnon W. R., 1445
Mensa D., 1489
Miyamoto Y., 1395
Mizutani T., 1547
Mochizuki K., 1425
Moll N., 1645
Morizuka K., 1347
Morris F., 1355
Muraguchi M., 1591
Mutoh M., 1561

Nagata H., 1405
Nagy A., 1645
Nguyen C., 1473
Nguyen N. X., 1473
Nishimura T. B., 1419
Nishitsuji M., 1413
Nozu T., 1347

O'Loughlin M. J., 1459
Ohkubo S., 1339
Ohno Y., 1339
Ohshima T., 1519
Oka T., 1425
Oshima K., 1395
Ostermann T., 1383
Ouchi K., 1425

Pan N., 1633
Parker E. H. C., 1497
Pavlidis D., 1429, 1507
Pierson R. L., 1633
Pons D., 1599
Pullela R., 1489

Quentin P., 1599

Robbins V., 1645
Rodwell M., 1489
Rohdin H., 1645
Runge K., 1633

Sai H., 1541
Sakai T., 1561
Sasaki H., 1655
Sato F., 1389
Sato M., 1519
Sawdai D., 1507
Saxarra M., 1383
Schmid P., 1547
Schuermeyer F., 1555
Schulman J. N., 1367
Schüppen A., 1373
Seabaugh A., 1355
Shigemasa R., 1519
Shirley T., 1437
Shur M. S., 1451
Siergiej R. R., 1459
Smith R. P., 1489
Sriram S., 1459
Su C.-y., 1645
Suemitsu T., 1527
Sugimura T., 1571
Sugiyama T., 1347
Suriani A., 1577

Takahashi H., 1561
Takahashi Y., 1339
Takamiya S., 1571
Takata M., 1413
Takeyama T., 1547
Tanaka T., 1405
Tanoue T., 1425
Tashiro T., 1389
Tatsumi T., 1389
Tobita H., 1395
Tomizawa M., 1547
Tsugawa K., 1465
Tsunotani M., 1519
Tsuzuku T., 1571

Ueda D., 1405
Umeda Y., 1527
Umezana H., 1465
Usunami T., 1547

Wada H., 1655
van der Wagt P., 1355
Wakita A., 1645
Wang N.-L. L., 1399
Washio K., 1619
Whall T. E., 1497
Whiteley W., 1437
Wölk C., 1607

Xu D., 1527

Yamane Y., 1527
Yeats B., 1437
Yokoyama H., 1527
Yokoyama N., 1513
Yoshida T., 1561
Yu R., 1633

Zampardi P. J., 1555, 1633
Zeuner M., 1383
Ziegler V., 1607
Zolper J. C., 1479